

Jörn J. Hägele

Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten

Videobasierte Analysen zu Aktivitäten und
Vorstellungen von Schülerinnen und Schülern
der gymnasialen Oberstufe bei der Bearbeitung
von fachmethodischer Instruktion



λογος

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf, Hans Niedderer, Mathias Ropohl und Elke Sumfleth

Diese Reihe im Logos Verlag Berlin lädt Forscherinnen und Forscher ein, ihre neuen wissenschaftlichen Studien zum Physik- und Chemielernen im Kontext einer Vielzahl von bereits erschienenen Arbeiten zu quantitativen und qualitativen empirischen Untersuchungen sowie evaluativ begleiteten Konzeptionsentwicklungen zu veröffentlichen. Die in den bisherigen Studien erfassten Themen und Inhalte spiegeln das breite Spektrum der Einflussfaktoren wider, die in den Lehr- und Lernprozessen in Schule und Hochschule wirksam sind.

Die Herausgeber hoffen, mit der Förderung von Publikationen, die sich mit dem Physik- und Chemielernen befassen, einen Beitrag zur weiteren Stabilisierung der physik- und chemiedidaktischen Forschung und zur Verbesserung eines an den Ergebnissen fachdidaktischer Forschung orientierten Unterrichts in den beiden Fächern zu leisten.

Martin Hopf, Hans Niedderer, Mathias Ropohl und Elke Sumfleth

Studien zum Physik- und Chemielernen

Band 332

Jörn J. Hägele

**Kompetenzaufbau
zum experimentbezogenen
Denken und Arbeiten**

Videobasierte Analysen zu Aktivitäten und
Vorstellungen von Schülerinnen und
Schülern der gymnasialen Oberstufe bei der
Bearbeitung von fachmethodischer Instruktion

Logos Verlag Berlin



Studien zum Physik- und Chemielernen

Martin Hopf, Hans Niedderer, Mathias Ropohl und Elke Sumfleth [Hrsg.]

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.



© Copyright Logos Verlag Berlin GmbH 2022

Alle Rechte vorbehalten.

ISBN 978-3-8325-5476-7

ISSN 1614-8967

Logos Verlag Berlin GmbH
Georg-Knorr-Str. 4, Geb. 10
D-12681 Berlin

Tel.: +49 (0)30 / 42 85 10 90

Fax: +49 (0)30 / 42 85 10 92

<https://www.logos-verlag.de>

Dissertation zur Erlangung des Doktorgrades – Dr. rer. nat. – am Fachbereich 07
(Mathematik und Informatik, Physik, Geographie) der Justus-Liebig-Universität
Gießen

Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten

**Videobasierte Analysen zu Aktivitäten und Vorstellungen
von Schülerinnen und Schülern der gymnasialen Oberstufe
bei der Bearbeitung von fachmethodischer Instruktion**

vorgelegt von

Jörn Jonathan Hägele

geboren in Hamburg, wohnhaft in Hamburg

Institut für Didaktik der Physik

März 2021

Tag der Disputation: 01.06.2021

Erstgutachterin: Prof. Dr. Claudia von Aufschnaiter

Zweitgutachter: Jun.-Prof. Dr. Andreas Vorholzer

Prüferin: Prof. Dr. Claudia Höhne

Prüfer: Prof. Dr. Alexander Eitel

Auf, auf, gib deinem Schmerze
und Sorgen Gute Nacht!
Lass fahren, was das Herze
betrübt und traurig macht;
bist du doch nicht Regente,
der alles führen soll:
Gott sitzt im Regimente
und führet alles wohl.

Paul Gerhardt

I have seen an end of all perfection:
but Thy commandment is exceeding broad.

Psalm 119:96

Inhaltsverzeichnis

1	Einleitung	1
2	Theoretische und empirische Grundlagen	11
2.1	Kompetenzen zum experimentbezogenen Denken und Arbeiten als Ziele von Unterricht	11
2.2	Modellierung von Kompetenzen zum experimentbezogenen Denken und Arbeiten	13
2.2.1	Fähigkeiten zum experimentbez. Denken und Arbeiten	15
2.2.2	Verständnisse zum experimentbez. Denken und Arbeiten . . .	16
2.3	Modellierung von Kompetenzaufbau (zum experimentbezogenen Den- ken und Arbeiten)	21
2.4	Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten: Befunde und zugehörige Instruktionsmaßnahmen	25
2.4.1	Kompetenzentwicklung über längere Zeiträume	26
2.4.2	Förderung von Fähigkeiten und Verständnissen – Information, Praxis und Reflexion als (Design-)Merkmale von Instruktionen	28
2.4.2.1	Informationen: Welche, wie viele, wann?	30
2.4.2.2	Praxis: Explorieren, Üben und Wiederholen	34
2.4.2.3	Reflexion: Die eigene Praxis in den Blick nehmen . .	36
2.4.2.4	Verschiedene Personen	37
2.4.3	Befunde zu Verständnissen zum experimentbez. Denken u. Arbeiten	38
2.4.3.1	Vorstellungen zum experimentbezogenen Denken und Arbeiten	39
2.4.3.2	Vorstellungen zu naturwissenschaftlichen Fragen . .	40
2.4.3.3	Vorstellungen zu Beobachtungen und Deutungen . .	42
3	Ziele und Forschungsfragen	45
3.1	Fragen zu Aktivitäten der Lernenden	47
3.2	Fragen zu Vorstellungen der Lernenden	48
4	Zugrundeliegendes Design	51
4.1	Fokussierte Fähigkeiten und Konzepte	52
4.2	Grundsätzliche Konzeption der Instruktion	55
4.3	Explizit- und implizit-fachmethodische Instruktionsvarianten	57
4.4	Übersicht über die inhaltlichen Abschnitte der beiden Instruktionsva- rianten	59
4.5	Prä- und Post-Erhebungen	63
4.6	Videoaufzeichnungen	64

5	Aktivitäten von Lernenden während der Bearbeitung der beiden Instruktionsvarianten	67
5.1	Eventbasiertes Kodieren von verbalen und nonverbalen Aktivitäten der Lernenden	68
5.1.1	Kodiermanual	68
5.1.2	Methodisches zur Beurteilerübereinstimmungsberechnung . .	76
5.1.3	Kodiertrainings und Beurteilerübereinstimmungen	84
5.2	Wahl und Gruppierung der analysierten Lernenden	88
5.2.1	Ausgangslage	88
5.2.2	Vorgehen bei der Auswahl der Lernenden	90
5.2.3	Analysierte Teilstichproben	92
5.2.4	Gruppierung der analysierten Lernenden	95
5.2.4.1	Gruppierung einzelner Personen	95
5.2.4.2	Gruppierung von Teams	99
5.3	Überlegungen zum Vorgehen bei der Auswertung	101
5.3.1	Zum Auswerten der Kodierungen	101
5.3.2	Zu statistischen Methoden	102
5.4	Auswertungen und Ergebnisse zur explizit-fachmethodischen Instruk-tionsvariante	105
5.4.1	Bearbeitungsdauern von Lernenden mit unterschiedlichen Kom-petenzzuwächsen	106
5.4.1.1	Gesamtbearbeitungsdauern für die Einheiten	106
5.4.1.2	Bearbeitungsdauern von Abschnitten der Instrukti-onseinheiten	114
5.4.1.3	Bearbeitungsdauern der einzelnen Karten der expli-zit-fachmethodischen Instruktionsvariante	118
5.4.2	Aktivitätsprofile der Lernenden	123
5.4.2.1	Vorgehen: Clusteranalysen zur Bildung von Aktivi-tätsprofiltypen	126
5.4.2.2	Aktivitätsprofile für Personen bei Einheit 1: Ergeb-nisse und exemplarische ausführliche Darstellung des Vorgehens	131
5.4.2.3	Aktivitätsprofile für Teams bei Einheit 1: Ergeb-nisse und exemplarische ausführliche Darstellung des Vorgehens bei Teams	141
5.4.2.4	Aktivitätsprofiltypen für Personen bei Einheit 2 . .	146
5.4.2.5	Aktivitätsprofiltypen für Teams bei Einheit 2	151
5.4.2.6	Aktivitätsprofiltypen für Personen bei Einheit 3 . .	153
5.4.2.7	Aktivitätsprofiltypen für Teams bei Einheit 3	157
5.4.2.8	Gesamtschau auf Aktivitätsprofile	159
5.4.3	Aktivitätsprofile und Kompetenzzuwächse	163

5.4.4	Fachmethodische Beiträge in der Gesamtschau und Kompetenzzuwächse	173
5.4.4.1	Vergleich der Gesamtdauer fachmethodischer Beiträge für Personen mit unterschiedlichen Kompetenzzuwächsen	175
5.4.4.2	Vergleich der Gesamtdauer fachmethodischer Beiträge des gesamten Teams von Personen mit unterschiedlichen Kompetenzzuwächsen	181
5.4.5	Fachmethodische Beiträge für spezifische Ausschnitte der Instruktion und Kompetenzzuwächse	184
5.4.5.1	Fachmethodische Beiträge und inhaltliche Abschnitte der Instruktion	184
5.4.5.2	Fachmethodische Beiträge und konkrete Situationen (Karten der Instruktion)	188
5.4.6	Fachmethodische Beiträge im Kontext von Experimentiersituationen und Kompetenzzuwächse	190
5.4.6.1	Wahl der Situationen und Vorgehen beim Vergleich	190
5.4.6.2	Ergebnisse der Vergleiche	192
5.4.6.3	Detaillierte Betrachtung einer Experimentiersituation	195
5.5	Auswertungen und Ergebnisse zur implizit-fachmethodischen Instruktionsvariante sowie zum Vergleich der beiden Instruktionsvarianten .	207
5.5.1	Bearbeitungsdauern	208
5.5.1.1	Vergleich der Instruktionsvarianten	208
5.5.1.2	Lernende mit unterschiedlichen Kompetenzzuwächsen	210
5.5.2	Aktivitätsprofile	214
5.5.2.1	Vergleich der Instruktionsvarianten	214
5.5.2.2	Lernende mit unterschiedlichen Kompetenzzuwächsen	221
5.5.3	Fachmethodische Beiträge	228
5.5.3.1	Vergleich der Instruktionsvarianten	228
5.5.3.2	Lernende mit unterschiedlichen Kompetenzzuwächsen	231
5.5.3.3	Lernende in unterschiedlichen Instruktionsvarianten und mit unterschiedlichen Kompetenzzuwächsen . .	232
5.5.4	Inhalte der fachmethodischen Beiträge	237
6	Vorstellungen von Lernenden während der Bearbeitung der explizit-fachmethodischen Instruktionsvariante	241
6.1	Qualitative Inhaltsanalyse zur Rekonstruktion von Vorstellungen der Lernenden	242
6.1.1	Induktives Generieren von Ideen-Kategorien	245
6.1.2	Zuweisen von Kategorien der Formulierung	250
6.1.3	Zuweisen von Kategorien des Modus	254
6.1.4	Kodieren der zugehörigen (Teil-)Aufgabe	257
6.1.5	Kodieren der Korrektheit der Einschätzungen	259

6.1.6	Beurteilerübereinstimmung	261
6.1.7	Generieren von Ideen-Oberkategorien	264
6.1.8	Einschätzen der Angemessenheit der Ideen-Oberkategorien	266
6.2	Zusammenstellung der analysierten Instruktionsextrakte	269
6.2.1	Darstellung des Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen	269
6.2.2	Darstellung des Instruktionsextrakts zur Unterscheidung von Beobachtung und Deutung	272
6.2.3	Einschränkungen und einordnende Hinweise zu den erstellten Instruktionsextrakten	274
6.3	Wahl der analysierten Lernenden und zugehörigen Videosequenzen und Transkripte	278
6.4	Auswertungen und Ergebnisse zur inhaltlichen Vielfalt der Vorstell.	283
6.4.1	Kennzeichen von naturwissenschaftlichen Fragen	283
6.4.1.1	Vielfalt an Vorstellungen	285
6.4.1.2	Abgleich mit dokumentierten Vorstellungen	288
6.4.2	Unterscheidung von Beobachtung und Deutung	290
6.4.2.1	Vielfalt an Vorstellungen	291
6.4.2.2	Abgleich mit dokumentierten Vorstellungen	292
6.5	Auswertungen und Ergebnisse zur Allgemeinheit der Vorstellungen	295
6.5.1	Allgemeinheit der Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen	295
6.5.2	Allgemeinheit der Vorstellungen zur Unterscheidung von Beobachtung und Deutung	301
6.5.3	Gesamtschau zur Allgemeinheit der Vorstellungen	304
6.6	Auswertungen und Ergebnisse zur Angemessenheit der Vorstellungen	306
6.6.1	Angemessenheit der Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen	306
6.6.2	Angemessenheit der Vorstellungen zur Unterscheidung von Beobachtung und Deutung	311
6.7	Auswertungen und Ergebnisse zu Vorstellungen und zur Korrektheit von Entscheidungen	312
6.7.1	Korrektheit einzelner Entscheidungen und Angemessenheit zugehöriger Vorstellungen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen	314
6.7.1.1	Gesamtschau für Kennzeichen von naturwissenschaftlichen Fragen	314
6.7.1.2	Betrachtung der einzelnen Aufgaben des Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen	316
6.7.1.3	Entscheidungen und fachinhaltliche Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen	319

6.7.2	Vorstellungen und Entscheidungen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen auf Ebene der Personen und auf Ebene der Teams	320
6.7.3	Entscheidungen und Vorstellungen zur Unterscheidung von Beobachtung und Deutung	324
6.8	Auswertungen und Ergebnisse zur Stabilität und Variabilität der Vorstellungen	326
6.8.1	Vorstellungen der Kohorte über die Aufgaben hinweg	326
6.8.2	Über mehrere Aufgaben hinweg gehäuft auftretende Vorstellungen der Lernenden	330
6.8.3	Inhalte von über m. Aufg. hinweg gehäuft auftr. Vorstellungen	338
6.8.4	Variabilität v. Vorstellungen im Verlauf der Aufg.bearbeitung	343
6.8.5	Zusammenschau zu Stabilität und Variabilität der Vorstellungen über Aufgaben hinweg	351
6.9	Auswertungen und Ergebnisse zu Vorstellungen vor und nach instruktionalen Informationen	353
6.9.1	Vergleich der Phasen vor und nach den instruktionalen Informationen	354
6.9.1.1	Für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen	354
6.9.1.2	Für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung	367
6.9.1.3	Zusammenschau	378
6.9.2	Gehäuft auftretende Vorstellungen zur Beschreibung der individuellen Verläufe von Vorstellungen	380
6.9.2.1	Zwei Phasen: Vor und nach der instruktionalen Information	380
6.9.2.2	Differenziertere Betrachtung für die Phasen des Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen	388
6.9.3	Ketten von Vorstellungen zur Beschreibung der individuellen Verläufe von Vorstellungen	392
6.10	Auswertungen und Ergebnisse zu Vorstellungen für Lernende mit unterschiedlichen Kompetenzzuwächsen und Eingangsvoraussetzungen	397
6.10.1	Gruppierung der Lernenden für die einzelnen Testvariablen .	399
6.10.2	Untersuchte Prozessmerkmale aus den vorigen Forschungsfrag.	401
6.10.3	Übersicht über Gemeinsamkeiten und Unterschiede von Personen mit niedrig und hoch ausgeprägten Testwerten	403
6.10.4	Auf die instruktionale Information bezogene Prozessvariablen	406
6.10.5	Untersuchte Prozessmerkmale und Kompetenzen zum experimentbezogenen Denken und Arbeiten	408
6.10.6	Untersuchte Prozessmerkmale und nicht-fachmethodische Eingangsvoraussetzungen	417

7	Zusammenführungen, Diskussionen und Ausblicke	419
7.1	Vielfalt in den Aktivitäten und Vorstellungen von Lernenden bei der Bearbeitung von Instruktion zum experimentbezogenen Denken und Arbeiten	420
7.1.1	Aktivitäten sowie Verknüpfung von Aktivitäten und Kompetenzzuwächsen	422
7.1.2	Äußerungen und Vorstellungen zu ausgewählten Konzepten des experimentbezogenen Denkens und Arbeitens	423
7.1.3	Vielfalt im Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten	427
7.2	Bedeutung von Äußerungen mit fachmethodischem Gehalt für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten .	431
7.2.1	Charakteristika von Äußerungen mit fachmethod. Gehalt . .	435
7.2.2	Zusammenhang von Fähigkeiten und Verständnissen	447
7.2.2.1	Aufbau von Verständnissen bei den beiden Instruktionsvarianten	448
7.2.2.2	Fähigkeiten und Verständnisse während der Bearbeitung von Instruktion	449
7.2.3	Allgemeine oder fachmethodische kognitive Aktivität	451
7.2.3.1	Operationalisierung fachmethodischer kognitiver Aktivität	453
7.2.3.2	Allgemeine relevanter als fachmethodische kognitive Aktivität	454
7.2.3.3	Höhere Konzeptualisierungsniveaus als spezifische fachmethodische kognitive Aktivität	456
7.3	Weitere Ansätze zur Erklärung von Unterschieden zwischen Instruktionsvarianten	463
7.3.1	Bearbeitungsdauern	463
7.3.2	Kompetenzaufbau und Testsituationen	465
7.3.3	Teilkompetenzen	467
7.3.4	Eingangsvoraussetzungen	472
7.4	Instruktionale Informationen und Praxis	475
7.5	Prozesse des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten	478
7.5.1	Zeitliche Veränderungen in den Charakteristika von Äußerungen mit fachmethodischem Gehalt	478
7.5.1.1	Zeitliche Veränderung bezüglich der untersuchten Kategorien fachmethodischer Beiträge	478
7.5.1.2	Zeitliche Veränderung in der Generalität der Äußerungen mit fachmethodischem Gehalt	481
7.5.2	Zeitliche Veränderung von Vorstellungen	485
8	Schluss	487

Literaturverzeichnis	491
Anhang	515
A Prä- und Post-Erhebungen	517
A.1 Anlage der Erhebungen	517
A.2 Testwerte der untersuchten Personen	517
B Ausführlichere Dokumentation zu Aktivitäten	522
B.1 Kodiermanual	522
B.2 Gruppierung der Lernenden	532
B.3 Bearbeitungsdauern	533
B.4 Mittlere Anteile von Aktivitäten	535
B.5 Clusteranalysen zu Aktivitätsprofilen	536
C Für Vorstellungen der Lernenden untersuchte Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen	541
D Für Vorstellungen der Lernenden untersuchte Sequenz zur Unterscheidung von Beobachtung und Deutung	545
E Transkripte der Bearbeitungen zu den beiden Instruktionsextrakten (ad Kapitel 6)	549
E.1 Transkriptionsregeln	549
E.2 Transkripte zu Kennzeichen von naturwissenschaftlichen Fragen (ausgelagert)	549
E.3 Transkripte zu Unterscheidung von Beobachtung und Deutung (ausgelagert)	549
F Ausführlichere Dokumentationen für die Methoden und Ergebnisse zu den Vorstellungen von Lernenden im Bearbeitungsprozess (ad Kapitel 6)	551
F.1 Übersicht über Ideen-Oberkategorien und Ideen-Kategorien	551
F.2 Generierung der Ideen-Oberkategorien	555
F.3 Allgemeinheit der Äußerungen, aus denen die Vorstellungen rekonstruiert werden	561
F.4 Entscheidungen der Lernenden und zugehörige Vorstellungen	563
F.4.1 Ausführliche Diskussion von Teams für Kennzeichen von naturwissenschaftlichen Fragen	563
F.4.2 Ausführliche Diskussion zum zeitlichen Umfeld von ausgewählten Entscheidungen zur Unterscheidung von Beobachtung und Deutung	567
F.4.3 Ergänzende Abbildungen und Tabellen zum Textteil	572

Inhaltsverzeichnis

F.5	Diagramme zur zeitlichen Auflösung der Vorstellungen der einzelnen Personen (ausgelagert)	577
F.6	Stabilität der Vorstellungen	577
F.7	Vorstellungen im zeitlichen Verlauf	582
F.7.1	Instruktionale Information zu Kennzeichen von naturwissenschaftlichen Fragen	582
F.7.2	Instruktionale Information zur Unterscheidung von Beobachtung und Deutung	589
F.7.3	Kodierungen für die Karten der instruktionalen Information zur Unterscheidung von Beobachtung und Deutung	590
G	Weitere Auszüge aus dem Material der Instruktion	591
	Abbildungsverzeichnis	595
	Tabellenverzeichnis	605
	Danksagungen	615

1 Einleitung

Wie Lernende mit Lernmaterial umgehen, das ihnen gegeben wird, ist ein zentrales Thema für Lehrende: ohne entsprechende Reaktionen verfehlt selbst die noch so gut durchdachte Lernumgebung ihr Ziel. In der vorliegenden Arbeit werden Bearbeitungsprozesse von Lernenden beleuchtet, die fachmethodische Instruktion bearbeiten, sie sollen Kompetenzen zu naturwissenschaftlichen Fachmethoden aufbauen.

Verortung der Arbeit und Eingrenzung des Gegenstands

Naturwissenschaftliche Methoden hatten immer wieder eine wesentliche Bedeutung in Legitimationsüberlegungen für Physikunterricht (ersichtlich beispielsweise aus Kircher et al., 2015, S. 19–21; Litt, 1952; Tesch, 2005; Wagenschein, 1995, S. 98–111). Theodor Litt (1952) sah in der richtig verstandenen und eingeordneten naturwissenschaftlichen Methode beispielsweise ein Argument für die »Nichtigkeit aller der Einwände, die den Wahrheitswert und den Lebenswert der mathematischen Naturwissenschaft fraglich zu machen bestimmt sind« (S. 487). Auch wenn heutzutage weitgehend Konsens darüber vorliegt, dass es *die* naturwissenschaftliche Methode nicht gibt (z. B. Höttecke & Rieß, 2015; Osborne, 2014; oder Gehlen, 2016, S. 32, die insbesondere auf Lederman et al., 2002, S. 501, verweist),¹ stellt (fach)methodisches Vorgehen des naturwissenschaftlichen Denkens und Arbeitens auch weiterhin einen wesentlichen Teil von Legitimationsüberlegungen zum Lehren und Lernen von Naturwissenschaften dar (vgl. Crawford, 2014; Fischer, 1998; Gräber et al., 2013; KMK, 2005; Lehrer & Schauble, 2015; OECD, 2016; Rehm et al., 2008; Zimmerman & Klahr, 2018). Eine weit verbreitete Annahme ist beispielsweise, dass eine »kritische Auseinandersetzung mit naturwissenschaftlichen Erkenntnismethoden [uns ermöglicht], als mündige Bürger an unserer hoch technologisierten Gesellschaft zu partizipieren« (Schmidt, 2016, S. 1, die u. a. auf Bybee, 1997, verweist). Unterschiedliche Legitimationsüberlegungen

¹Die zitierte Literatur entstammt bewusst der fachdidaktischen Forschung. Insbesondere in der stärker aus Richtung der Psychologie kommenden Literatur ist es geläufig von »the scientific method« zu sprechen (z. B. Lorch et al., 2010, S. 90); inwiefern dadurch allerdings eine uniforme experimentelle Methode impliziert ist, ist im jeweiligen Einzelfall zu untersuchen und nicht Gegenstand der vorliegenden Arbeit. Auch die viel generellere Frage, inwiefern es theoretische oder praktische Uniformität im naturwissenschaftsmethodischen Vorgehen gibt, kann in der vorliegenden Arbeit nicht behandelt werden, da sie ein zu großes Feld darstellt (siehe etwa Chalmers, 2001).

1 Einleitung

rekurrieren häufig und auf verschiedene Weise auf naturwissenschaftliches Denken und Arbeiten; daher ist es kaum verwunderlich, dass es ein wichtiger Bestandteil von Bildungsvorgaben für naturwissenschaftliche Fächer ist (z. B. Abd-El-Khalick et al., 2004; Australian Curriculum, Assessment and Reporting Authority [ACARA], 2015; Department for Education [DfE], 2013; KMK, 2004, 2005; NGSS Lead States, 2013) sowie einen Schwerpunkt von naturwissenschaftsdidaktischer Forschung darstellt (z. B. Emden, 2011; Nehring, 2014; Rönnebeck et al., 2016; Zimmerman, 2007).

Experimentbezogenes naturwissenschaftliches Denken und Arbeiten (nachfolgend verkürzt: experimentbezogenes Denken und Arbeiten) stellt einen Teil des naturwissenschaftlichen Denkens und Arbeitens dar, der auf den Einsatz von Experimenten und Versuchen in den Naturwissenschaften bezogen ist (siehe bspw. Vorholzer, 2016; mehr in Kapitel 2). Das Experiment ist ein zentraler Erkenntnisweg, eine zentrale (Fach-)Methode, der Naturwissenschaften (vgl. Arnold et al., 2017, Fußnote 1; sowie Gut-Glanzmann & Mayer, 2018; Ostlund, 1998; Rieß et al., 2012). Außerdem sind viele Aspekte des experimentbezogenen Denkens und Arbeitens auch für andere naturwissenschaftliche Denk- und Arbeitsweisen relevant. Beispielsweise ist das Formulieren von präzisen naturwissenschaftlichen Fragen nicht nur für das Experimentieren von Bedeutung, sondern auch für die naturwissenschaftliche Arbeitsweise des Beobachtens zentral (Arnold et al., 2010; Kohlauf et al., 2010; Ostlund, 1998). Es liegt daher nahe, dass insbesondere experimentbezogenes Denken und Arbeiten häufig und auf vielfältige Weise hinsichtlich seiner Bedeutung im Naturwissenschaftsunterricht beleuchtet wird (Übersichten z. B. in Furtak et al., 2012; Gut-Glanzmann & Mayer, 2018; Lazonder & Harmsen, 2016; Rieß & Robin, 2012; Rönnebeck et al., 2016; Zimmerman, 2007; Zimmerman & Klahr, 2018). Neben der weltweit untersuchten Gestaltung und Wirksamkeit vom am Experimentieren orientierten sogenannten forschenden Lernen (*inquiry-based learning*, z. T. auch *inquiry-based teaching*, manchmal auch unter dem Namen entdeckendes Lernen; beispielhaft seien nur einige Quellen zitiert: Alfieri et al., 2011; Dobber et al., 2017; Furtak et al., 2012; OECD, 2016) ist dabei im deutschsprachigen Raum in den letzten Jahren vor allem Forschung zur Modellierung und Erfassung zugehöriger Kompetenzen dominant (z. B. Gehlen, 2016; Gut, 2012; Krell, 2018; Schecker et al., 2016; Schreiber et al., 2014; Vorholzer et al., 2016, z. T. sind umfassende Verweise auf weitere Literatur enthalten), es finden sich allerdings auch zunehmend Studien zu Förderung und Aufbau von entsprechenden Kompetenzen (z. B. Arnold et al., 2017; Hild et al., 2020; Nehring et al., 2016; Ropohl & Scheuermann, 2018; Schmidt, 2016; Strippel, 2017; Vorholzer, 2016).

Problemaufriss

Bezüglich des *Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten* sind bisher vorrangig Forschungsarbeiten zur Wirksamkeit von Lernumgebungen vorhanden (international wie national; systematische Übersichten finden sich nur sehr wenige, worauf bspw. Strippel, 2017, S. 19, hinweist; für weitere Literatur siehe einzelne Abschnitte der Quellen im vorigen Absatz, insbesondere Lazonder & Harmsen, 2016, und Rieß et al., 2012, sowie ausführlicher in Kapitel 2). Dabei wird gewissen Merkmalen von Instruktion – wie beispielsweise dem instruktionalen Informieren über Konzepte des experimentbezogenen Denkens und Arbeitens – (tentativ) zugeschrieben, dass sie für den Kompetenzaufbau ausschlaggebende Faktoren sind.² Derartige Schlussfolgerungen so empirisch abzusichern und zu kommunizieren, dass sie z. B. als allgemeine Basis für Unterrichtsplanung oder -durchführung dienen können, ist allerdings mit mannigfaltigen Schwierigkeiten verbunden, unter anderem:

Vergleichbarkeit von bereits untersuchten Instruktionen. In unterschiedlichen Studien eingesetzte Instruktionen, bei denen zunächst davon ausgegangen wird, dass sie hinsichtlich des als relevant angesehenen Merkmals vergleichbar seien, unterscheiden sich – aus guten Gründen sowie vermutlich auch aufgrund persönlicher Prägungen und Präferenzen derjenigen, die die Instruktion konstruiert haben – häufig hinsichtlich vielfältiger anderer Gestaltungsdimensionen.³ Diese Vielfalt kann selbst zum Gegenstand von Forschung oder gar brauchbar für Lehrkräfte gemacht werden; sie kann allerdings in der Forschung auch zu inkonklusiven oder widersprüchlichen Ergebnissen führen. Unter den Instruktionen,

²Etwa deuten Kalthoff et al. (2018) an, dass sie es als übliches Vorgehen ansehen, der Explizitheit explizit-fachmethodische Instruktion Effektivität zuzuschreiben, wenn sie schreiben: »Thus, the results of the meta-analysis concerning the efficacy of demonstrations and cognitive conflicts cannot be ascribed unambiguously to explicitness« (S. 1308). In ähnlicher Weise vermuten Rönnebeck et al. (2016) Einfluss expliziter Adressierung von Aspekten der Natur der Naturwissenschaften und metakognitiven Aspekten. Als weiteres Beispiel kann dienen, dass Kuhn et al. (2017) vermuten, dass die spezifische Anlage der eingesetzten Prompts das zentral relevante Merkmal der von ihnen vorgenommenen Intervention ist: »[I]t is impossible to be sure which specific components of a multicomponent experience were most critical to its outcomes. Our conjecture, however, is that a key factor was the emphasis during the intervention on counterargument and evidence to weaken claims (›Suppose someone disagrees with you ...‹)« (S. 248).

³Außerdem ist es durchaus möglich, dass Instruktionen, bei denen aufgrund der Beschreibungen in den Forschungsartikeln zunächst davon ausgegangen werden muss, dass diese nicht vergleichbar sind, vergleichbar sind (eventuell sogar besser als zunächst als vergleichbar angenommene Instruktionsansätze). Es ist beispielsweise gut denkbar, dass Studien aus dem Bereich der *Worked Examples*, die möglicherweise ohne Bezug zum Vergleich von explizit-fachmethodischen und implizit-fachmethodischen Instruktionsansätzen sind, stärkere Vergleichbarkeit mit einigen explizit-fachmethodischen Instruktionsansätzen aufweisen als andere als explizit-fachmethodisch designierte Instruktionsansätze. Vergleichbare Ausführungen zu einander ähnlichen Realisierungen von Prompts, Heuristiken etc. finden sich bei Lazonder und Harmsen (2016, S. 686–687).

1 Einleitung

die in der Forschungslandschaft den sogenannten *expliziten Instruktionsansätzen* zugeordnet werden, weil in ihnen die adressierten Konzepte instruktional expliziert werden, finden sich auch Studien zu Instruktionen, die auf die Variablenkontrollstrategie bezogen sind – und die in der Terminologie der vorliegenden Arbeit zu den *explizit-fachmethodischen Instruktionsansätzen* gezählt werden, weil die Variablenkontrolle ein Bestandteil von naturwissenschaftlichen Fachmethoden ist.⁴ Für diese explizit-fachmethodischen Instruktionen zur Variablenkontrolle zeigen Kalthoff et al. (2018) große Varianz in den vorzufindenden Realisationen auf; sie können mindestens hinsichtlich vier Aspekten differenziert werden: (a) Werden die genannten Konzepte auch erklärt? (b) Werden Positiv- und Negativbeispiele zu den Konzepten gegeben? (c) Sollen die Konzepte angewendet und im Nachgang reflektiert werden? (d) Wird versucht, kognitive Konflikte hervorzubringen? Die im zitierten Forschungsartikel vorgestellten Studien weichen bezüglich dieser Merkmale deutlich voneinander ab.

Neben den Abweichungen hinsichtlich anderer Gestaltungsdimensionen finden sich in eigenen Sichtungen davon, wie in explizit-fachmethodischen Instruktionsansätzen die Konzepte instruktional expliziert werden, d. h. wie die Informationen gegeben werden, (laut berichtenden Forschungsartikeln, die nur zu geringen Teilen konkrete Einblicke in die genaue instruktionale Information geben; vgl. u. a. Rönnebeck et al., 2016, S. 166) zudem deutliche Unterschiede darin, wie die instruktionalen Informationen realisiert werden. Diese weichen hinsichtlich vieler Aspekte voneinander ab, wie beispielsweise Textlänge, Visualisierung, Strukturiertheit, sprachliche Schwierigkeit. Ferner werden beispielsweise auch solche instruktionalen Informationen, die für sich genommen gut vergleichbar sind, in unterschiedlichen Realisationen explizit-fachmethodischer Instruktionsansätze zu unterschiedlichen Zeitpunkten der Instruktionen gegeben. Selbst eine konklusive Befundlage würde daher nicht notwendigerweise ausreichend Hinweise darauf liefern, inwiefern die genaue Umsetzung des Gestaltungsmerkmals für die Effektivität der je spezifischen Realisation des Instruktionsansatzes relevant ist.

⁴Explizite Instruktionsansätze können in Bezug auf verschiedene Inhalte explizit sein (z. B. das Konzept der Variablenkontrolle, das Ohmsche Gesetz, den Energieerhaltungssatz; siehe z. B. die verschiedenen Studien, die von Alfieri et al., 2011 in die Metaanalysen einbezogen werden). Im Kontext des experimentbezogenen Denkens und Arbeitens wird unter dem Begriff der expliziten Instruktion zumeist eine in Bezug auf spezifische Fachmethoden explizite Instruktion verstanden, was in der vorliegenden Arbeit durch die Begriffswahl *explizit-fachmethodische Instruktion* explizit gemacht wird. Allerdings ist auch dieser Begriff mit Einschränkungen verbunden, da zum einen Fachmethoden mehr umfassen als das experimentbezogene Denken und Arbeiten und zum anderen durch den Begriff nicht beschrieben wird, welche Konzepte des experimentbezogenen Denkens und Arbeitens instruktional expliziert werden. Für die in der vorliegenden Arbeit untersuchten Instruktionsvarianten finden sich daher in Kapitel 4 Listen aller instruktional explizierten Konzepte.

Es ist nachvollziehbar, dass ein hochkomplexes Feld wie das (der Förderung) des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten zunächst basierend auf allgemeineren Studien erkundet wird und erst nach und nach einzelne Merkmale isoliert werden (wie bspw. der Zeitpunkt des instruktionalen Informierens, siehe Matlen & Klahr, 2013);⁵ selbst bei optimistischem Ausblick scheint allerdings eine nachträgliche Analyse verschiedener in den Studien eingesetzter Instruktionen noch in weiter Ferne zu liegen (und von einer an einem gemeinsamen Instruktionsraster entlang forschenden Community soll hier gar nicht die Rede sein), mindestens aus dem Grund, dass die Realisationen der Instruktionsansätze grundsätzlich nur sehr knapp beschrieben werden und auch die eingesetzten Instruktionsmaterialien meistens nicht zugänglich sind oder sie kaum Aufschlüsse geben, weil die genaue instruktionale Realisation erst im Zusammenspiel mit der umsetzenden Lehrkraft entsteht.

Konstruktion von Instruktionsvarianten, die ein spezifisches Merkmal variieren. Selbst in einer dezidiert auf Forschung angelegten Instruktionskonstruktion ist es in manchen Fällen kaum möglich, nur eines der instruktionalen Merkmale zu variieren und alle anderen zu kontrollieren. Soll beispielsweise der Frage nachgegangen werden, ob Erläuterungen ein für den Kompetenzaufbau wesentlicher Bestandteil instruktionaler Informationen sind, ergeben sich mehrere Probleme. Zum einen scheint sich das bloße Nennen von Konzepten zum Teil kaum vom Erläutern unterscheiden zu lassen. Es ist beispielsweise nicht klar, ob der Satz »Alle anderen Faktoren müssen konstant gehalten werden« eine Erläuterung oder eine Nennung eines Konzepts ist. Und manche Konzepte können Erläuterungen anderer Konzepte sein: das soeben genannte Beispiel könnte eine Erläuterung der Variablenkontrollstrategie sein, die allgemeiner so als Konzept formuliert werden kann, dass in einer Untersuchung möglichst nur die Variable verändert werden sollte, deren Einfluss überprüft werden soll (für diese Formulierung siehe Vorholzer, 2016, basierend auf Gott et al., 2015). Zum anderen könnte es sein, dass die Relevanz von Erläuterungen darin besteht, dass sie eine notwendige Voraussetzung zur Anwendung und Übung in späteren Kontexten

⁵Die in Kapitel 2 gegebenen Einblicke in den Stand der Forschung illustrieren gut, dass bereits mit den hier in der Einleitung als eher allgemein bezeichneten Studien tiefgehende und weitreichende Erkenntnisse gewonnen werden konnten (beispielsweise, dass stärker vorstrukturierende Instruktionsansätze im Mittel auch für Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten effektiver sind als weniger stark vorstrukturierende Ansätze). Daher sind die Ausführungen in den Aufzählungspunkten nicht im Sinne eines Forschungspessimismus zu lesen und die bereits generierten Ergebnisse sollen nicht in Abrede gestellt werden. Vielmehr wird die Systematisierung von Ergebnissen aus verschiedenen Studien kritisch beleuchtet, um den Beitrag der vorliegenden Arbeit herauszustellen.

1 Einleitung

darstellen. Die Anwendungs- und Übungskontexte müssen aber notwendigerweise für Instruktionsvarianten ohne Erläuterungen zu den Konzepten anders konzipiert sein als es für Instruktionsvarianten mit Erläuterungen möglich und passend und vermutlich intuitiv naheliegend der Fall wäre. Die Entfaltung der Bedeutung von Erläuterungen ist daher ggf. durch andere Merkmale von Instruktion bedingt.

Im gegebenen Beispiel zur Konstruktion von Instruktionsvarianten, bei denen gezielt das Merkmal des Gebens von Erläuterungen variiert werden soll, lässt sich sicherlich ein besserer Weg zur Instruktionskonstruktion finden als dem Autor der vorliegenden Arbeit aktuell bewusst ist. Der generelle Punkt bleibt jedoch bestehen, dass nämlich eine Konstruktion von mehreren Instruktionsvarianten, aus denen sich durch bloßes Vergleichen der erzielten Kompetenzzuwächse auf die Bedeutung spezifischer instruktionaler Elemente schließen lässt, ein kompliziertes, vermutlich häufig nicht mögliches Unterfangen darstellt.⁶

(Potentieller) Beitrag der Arbeit

Ein naheliegender Umgang mit dem geschilderten Problem der Isolierung spezifischer Instruktionsmerkmale besteht darin, die Bearbeitung der Instruktion durch Lernende daraufhin zu untersuchen, inwiefern möglichst gut vergleichbar angelegte Instruktionen tatsächlich vergleichbar bearbeitet werden und inwiefern die gezielt variierten Merkmale zu unterschiedlichem Handeln und Denken der Lernenden bei der Bearbeitung führen. Selbst wenn es in einem konkreten Fall gelingen sollte, variablenkontrolliert ideal vergleichbare Instruktionen anzulegen (oder posthoc aus Studien zu extrahieren), wovon kaum auszugehen ist, blieben die Fragen nach der Vergleichbarkeit der Bearbeitungsprozesse und nach den durch die gezielt variierten Merkmale induzierten Unterschieden im Denken und Handeln bei der Bearbeitung der Instruktion relevant: Antworten darauf könnten einen Beitrag dazu leisten, aufzuklären, wie die untersuchten Merkmale zum Kompetenzaufbau beitragen und worin genau das den Kompetenzaufbau förderliche Element besteht (und worin nicht).

Anhand zweier Beispiele wird nachfolgend illustriert, dass durch die Untersuchung des in der Instruktionsbearbeitung vollzogenen Denkens und Handelns der Lernenden ein wesentlicher Beitrag bezüglich weiterer (zumindest implizit) in der Forschungs-

⁶Selbst wenn eine derartige Konstruktion logisch möglich ist, sind real von den Forschenden weitere Faktoren wie beispielsweise ethische Aspekte oder die Konsequenzen der (potentiellen) Interpretationen als Aspekt von Validität (vgl. Messick, 1995) zu bedenken, die dazu führen, dass Abstriche in der Vergleichbarkeit gemacht werden.

landschaft zum experimentbezogenen Denken und Arbeiten aufgeworfener Fragen geleistet werden kann:

- In vielen Studien zur Förderung von Kompetenzen zum experimentbezogenen Denken und Arbeiten werden *Prompts* oder Hinweise gegeben. Diese Prompts sollen Lernende unter anderem dazu anregen, (anders) über Fachmethoden nachzudenken und ggf. Selbsterklärungen zu generieren, denen lernförderliche Wirkung zugeschrieben wird (z. B. Koenen, 2014; Lin & Lehman, 1999; und vergleichbar, aber etwas allgemeiner gefasst Renkl, 2002; Renkl & Schworm, 2002). In empirischen Studien konnte gezeigt werden, dass derartige Unterstützungsmaßnahmen sowohl die Experimentierprozesse verbessern als auch zu größeren Kompetenzzuwächsen zum experimentbezogenen Denken und Arbeiten führen (z. B. Großmann & Wilde, 2019; Zhang et al., 2015). Grundsätzlich wird dabei auch in den Blick genommen, wie die Lernenden die Prompts und Unterstützungsmaßnahmen nutzen. So untersuchen Arnold et al. (2016) beispielsweise die schriftlich notierten Antworten der in Kleingruppen arbeitenden Lernenden und schließen daraus, dass nur in etwa 64 % der Fälle Bezug zu Gütekriterien des naturwissenschaftlichen Denkens und Arbeitens hergestellt wurde. Worüber Lernende bei der Bearbeitung von Prompts sprechen und wie die Lernförderlichkeit von Prompts mit der tatsächlichen Nutzung einhergeht, bleibt jedoch mit Bezug auf das experimentbezogene Denken und Arbeiten bisher häufig offen;⁷ beispielsweise schreibt Koenen (2014) in der Diskussion der Ergebnisse:

[Es zeigt sich, dass] Schülerinnen und Schüler, die die Prompts als hilfreich empfunden haben, im Post-Test höhere Mittelwerte erreichten als diejenigen, die angegeben haben, die Lösungshinweise als nicht hilfreich empfunden zu haben. Dies könnte darauf hinweisen, dass Schülerinnen und Schüler, die die Lernhinweise als hilfreich empfinden, diese auch tiefergehend bearbeiten, dadurch mehr Selbsterklärungen generieren und so einen höheren Lernerfolg erzielen. (S. 106)

Ob die jeweiligen Lernenden die Prompts tatsächlich tiefergehend bearbeitet haben und (dadurch) mehr Selbsterklärungen generiert haben, ist für die voll-

⁷Für fachinhaltliche Kompetenz siehe bspw. die richtungsweisende und fassettenreiche Publikation von Chi et al. (1994) zu Inhalten der Biologie und als dt. Überblick Rabe und Mikelskis (2007).

zogene Umsetzung der Fördermaßnahme allerdings unklar.⁸ Auch Arnold et al. (2016) erwähnen die Notwendigkeit weiterer Analysen, bspw. unter Hinzunahme von Videoaufzeichnungen. Analysen des Handelns und Denkens von Lernenden bei der Bearbeitung von Instruktion könnten genauere Aufschlüsse darüber geben, wie stark die Lernenden auf Fachmethoden eingehen und ob sie z. B. Selbsterklärungen generieren.⁹

- Eingangsvoraussetzungen der Lernenden wie fachmethodische Kompetenzen oder fachinhaltliche Kompetenzen scheinen einen Faktor für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten darzustellen. Beispielsweise finden sich Unterschiede in der Wirksamkeit explizit-fachmethodischer und implizit-fachmethodischer Instruktionsansätze für Lernende mit guten und schlechten fachmethodischen Eingangsvoraussetzungen (Details in Unterunterabschnitt 2.4.2.4, ab S. 37). Einige Überlegungen zu Eingangsvoraussetzungen scheinen durch Analysen von Bearbeitungsprozessen deutlich Erhellung erfahren zu können. Etwa formuliert Zimmerman (2007) die Vermutung, dass »domain knowledge may serve to draw attention to the functional effect of the experimental manipulation, and therefore influence the choice of experimental design.« (S. 179). Wie trägt die bereits vorhandene fachinhaltliche Kompetenz dazu bei (oder wie hindert sie), Kompetenzen zum experimentbezogenen Denken und Arbeiten situativ zu entfalten und/oder situationsübergreifend aufzubauen? Analysen zum Denken und Handeln von Lernenden bei der Bearbeitung von Instruktion könnten unter anderem Aufschluss darüber geben, wann und wie umfangreich auf Fachinhalte Bezug genommen wird.

Interessanterweise findet Vorholzer (2016) entgegen der sonst eher typisch erscheinenden Befunde keinen Einfluss der fachmethodischen und fachinhaltlichen Eingangsvoraussetzungen auf den fachmethodischen Kompetenzzuwachs. Er erklärt dies dadurch, dass die von ihm konzipierte und eingesetzte Instruktion viele bzw. ausreichend Gelegenheiten zum (Wieder-)Entdecken der adressierten fachmethodischen Konzepte gibt (Vorholzer, 2016, S. 136–137). Ob die

⁸Bei Koenen wird allerdings von einer angegliederten Hausarbeit berichtet, in der der Umgang mit den Prompts unter anderem mit der Methode des lauten Denkens untersucht wird. Trotzdem bleibt offen, *wie* die Lernenden bei der *tatsächlichen Bearbeitung der Fördermaßnahme* die Prompts genutzt haben (unter anderem, weil andere Lernende untersucht wurden und dafür zusätzlich die Methode des lauten Denkens genutzt wurde).

⁹Schmidt-Weigand et al. (2009) fragen in einem Ausblick einer Studie zu Prompts, die auf experimentelle Strategien bezogen sind, ebenfalls: »How strong do learning times vary across students? Do students adapt the progression in a reasonable way? How is learning in a dyad influenced by the learning partner? Those [subsequent] studies may also explore individual differences in the use of incremental presentation and prompts.« (S. 137). Weitere Beispiele aus der Forschungsliteratur könnten aufgezählt werden.

Konzepte von den Lernenden tatsächlich (wieder-)entdeckt wurden, lässt sich allerdings nur durch Analysen der Bearbeitungsprozesse nachgehen. Würden Lernende wiederholt Bezug auf fachmethodische Konzepte nehmen, so könnte die Vermutung von Vorholzer (2016) empirisch ein Stück besser abgesichert werden.

Erkenntnisse zum Denken und Handeln von Lernenden bei der Bearbeitung von Instruktion können nicht nur, wie bisher herausgestellt, mit Blick auf Wirkmechanismen von Fördermaßnahmen einen Beitrag im Forschungsfeld zum Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten leisten. Sie können auch erste Ausgangspunkte für Forschung zu Prozessen des Kompetenzaufbaus darstellen, also dazu, wie sich das (kompetenzbezogene) Denken und Handeln von Lernenden im zeitlichen Verlauf verändert, und somit Hinweise auf Entwicklung von beispielsweise Fähigkeiten und Vorstellungen im zeitlichen Verlauf liefern. Solche Erkenntnisse könnten dann wiederum einen Beitrag zur Formulierung von Modellen der Lernprogression und damit auch zur adaptiven (ggf. auf die Lernprogressionen angepassten) Förderung leisten.

Aufbau der Arbeit

Ausgehend vom Stand der Forschung (die Einleitung ergänzend und vertiefend in Kapitel 2 ausgeführt) wird das Ziel verfolgt, Hypothesen zum Kompetenzaufbau der Lernenden, zu Wirkmechanismen von Instruktionen und zu Prozessen des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten abzuleiten (detaillierter in Kapitel 3). In der vorliegenden explorativen Arbeit wird das Denken und Handeln anhand der Aktivitäten und Vorstellungen von Schülerinnen und Schülern der gymnasialen Oberstufe bei der Bearbeitung einer explizit-fachmethodischen und einer dazu parallel angelegten implizit-fachmethodischen Instruktionsvariante (dargestellt in Kapitel 4) in den Blick genommen. Dazu werden Videoaufzeichnungen mithilfe qualitativ-inhaltsanalytischer Verfahren untersucht sowie statistische und inhaltliche Verknüpfungen mit Vor- und Nach-Erhebungen zu Kompetenzen der Lernenden vorgenommen. Das methodische Vorgehen und die zugehörigen Ergebnisse sind jeweils in Kapitel 5 (Aktivitäten) und Kapitel 6 (Vorstellungen) dargestellt und im Detail diskutiert. In Kapitel 7 werden die Ergebnisse aus diesen beiden Kapiteln zusammengeführt und vor dem Hintergrund von ausgewählten Methoden, Befunden und Diskursen aus der Forschung diskutiert, um die Schlussfolgerungen der Arbeit zu bewerten und Hypothesen und Desiderata abzuleiten. Kapitel 8 beinhaltet ein kurzes Résumé der Arbeit.

2 Theoretische und empirische Grundlagen der Arbeit

Wie eine Untersuchung von Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten gestaltet wird, hängt wesentlich davon ab, welche Gründe es für die Untersuchung gibt und wie die Begrifflichkeiten *Kompetenzen zum experimentbezogenen Denken und Arbeiten* und *Kompetenzaufbau* gefasst werden. In dem vorliegenden Kapitel wird der Stand der Forschung zum Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten entlang dieser beiden Begrifflichkeiten zusammengefasst. Anhand der Begrifflichkeit *Kompetenzen zum experimentbezogenen Denken und Arbeiten* wird zunächst der Forschungsgegenstand eingegrenzt, indem eine Entscheidung für eine spezifische Ausdeutung von experimentbezogenem Denken und Arbeiten (Abschnitt 2.1) sowie für eine spezifische Modellierung zugehöriger Kompetenzen getroffen wird (Abschnitt 2.2, ab S. 13). Daran anschließend wird der Kompetenzaufbau in den Blick genommen, indem theoretische Überlegungen in eine Systematik integriert werden (Abschnitt 2.3, ab S. 21) und dann bisherige Befunde zum Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten zusammengefasst werden (Abschnitt 2.4, ab S. 25).

2.1 Kompetenzen zum experimentbezogenen Denken und Arbeiten als Ziele von Unterricht

Experimentbezogenes Denken und Arbeiten¹⁰ stellt sowohl eine umfangreich untersuchte *instruktionale Methode* zur Förderung von naturwissenschaftlichen Kompetenzen als auch einen in der Forschung und in Bildungsvorgaben betonten *Gegegenstand des Lernens* im Naturwissenschaftsunterricht dar (vgl. für die Unterscheidung die Zusammenführung bei Vorholzer, 2016; zudem Strippel, 2017, und systematisiert bei Teig, im Druck für 2021). Zum einen gibt es ein breites Feld fachdidaktischer Studien zum Einsatz des experimentbezogenen Denkens und Arbeitens als Unterrichtsmethode und zu zugehörigen Effekten (Übersichten finden sich z. B. in Alfieri et al., 2011; Börlin, 2012, S. 11; Furtak et al., 2012; Lazonder & Harmsen, 2016; Rönnebeck et al., 2016). Es geht dabei darum, die Lernenden zum experimentbezogenen Denken und Arbeiten aufzufordern und sie darin anzuleiten; untersucht wird unter anderem, wie

¹⁰Experimentbezogenes Denken und Arbeiten wird als Teil des naturwissenschaftlichen Denkens und Arbeitens, der auf den Einsatz von Experimenten und Versuchen in den Naturwissenschaften bezogen ist, aufgefasst (siehe Kapitel 1 und z. B. Vorholzer, 2016).

Lehrkräfte das experimentbezogene Denken und Arbeiten als Unterrichtsmethode einsetzen (z. B. Abrahams & Millar, 2008; Dobber et al., 2017; Tesch, 2005), aber auch, welche Effekte der Einsatz auf verschiedene Kompetenzen der Lernenden hat (fachinhaltliche Kompetenzen, z. B. Maiseyenko, 2014; Rizzo & Taylor, 2016; Wirth et al., 2008; verschiedene fachmethodische Kompetenzen, z. B. Lazonder & Harmsen, 2016; z. T. mehrere oder undifferenziert, z. B. Blanchard et al., 2010; Furtak et al., 2012; Hofstein & Lunetta, 2004; Wilson et al., 2010). Zum anderen wird in Bildungsvorgaben und in der Forschung betont, dass experimentbezogenes Denken und Arbeiten ein zentraler Gegenstand des Lernens im Naturwissenschaftsunterricht sein sollte. So weisen beispielsweise die deutschen Bildungsstandards für den mittleren Schulabschluss zu den drei naturwissenschaftlichen Fächern und auch die einheitlichen Prüfungsanforderungen für das Abitur im Fach Physik die Erkenntnisgewinnung als einen eigenständigen Bereich aus und betonen darin besonders das Experimentieren (KMK, 2004, 2005). Auch in der Forschung wird experimentbezogenes Denken und Arbeiten als ein zentrales Ziel des Lernens diskutiert und propagiert (z. B. Abd-El-Khalick et al., 2004; Kind & Osborne, 2017; Rieß et al., 2012; Schecker et al., 2016). Beispielsweise lassen sich viele der in der Einleitung erwähnten Überlegungen zur Legitimation von Naturwissenschaftsunterricht so interpretieren, dass die Schülerinnen und Schüler nicht nur experimentelles Denken und Arbeiten miterleben, sondern auch selbstständig ausführen können – und zwar nicht nur als Methode, um Fachinhalte zu erlernen, sondern auch als eigenständiges Ziel des Lernens. Auch in der Forschung zum *scientific thinking* werden sowohl Fachinhalte als auch Fachmethoden in den Blick genommen (vgl. Zimmerman & Klahr, 2018).

In der vorliegenden Arbeit geht es um Kompetenzen zum experimentbezogenen Denken und Arbeiten. *Das experimentbezogene Denken und Arbeiten wird also als Gegenstand und Ziel des Lernens betrachtet.* Es geht demnach darum, wie gut Schülerinnen und Schüler bereits experimentbezogenes Denken und Arbeiten ausführen und dass sie darin etwas dazulernen. Wie gut halten unterschiedliche Schülerinnen und Schüler die Variablenkontrolle zu unterschiedlichen Zeitpunkten ein? Was wissen die Lernenden zu verschiedenen Zeitpunkten darüber, was naturwissenschaftliche Fragen kennzeichnet? Die Arbeit ist auf die Untersuchung solcher und weiterer Fragen zu den vorliegenden Kompetenzen der Lernenden gerichtet.¹¹

¹¹Über die fokussierte Unterrichtsmethode ist damit noch nichts gesagt: Damit, dass *nicht* der Einsatz des experimentbezogenen Denkens und Arbeitens als Unterrichtsmethode im Fokus der Arbeit steht, ist nämlich nicht notwendigerweise ausgeschlossen, dass als instruktionale Methode experimentbezogenes Denken und Arbeiten genutzt wird. Vielmehr werden, wie in Kapitel 4 erläutert, die Kompetenzen der Lernenden sogar im Kontext zweier Instruktionsvarianten untersucht, die beide zum experimentbezogenen Denken und Arbeiten auffordern und die Lernenden dabei anleiten.

2.2 Modellierung von Kompetenzen zum experimentbezogenen Denken und Arbeiten

Kompetenzen zum experimentbezogenen Denken und Arbeiten meinen in der vorliegenden Arbeit das

Konglomerat aus *Fähigkeiten, Fertigkeiten, Verständnissen und Bereitschaften* zum *Formulieren von Fragen und Hypothesen*, zum *Planen und Durchführen von Untersuchungen* sowie zum *Analysieren und Interpretieren von experimentellen Daten*.

Diese Definition ist in Anlehnung an Klieme (2004) und Weinert (2001/2014) formuliert,¹² bei denen Kompetenzen charakterisiert werden als

funktional bestimmte, auf bestimmte Klassen von Situationen und Anforderungen bezogene kognitive Leistungsdispositionen, die sich psychologisch als Kenntnisse, Fertigkeiten, Strategien, Routinen oder auch bereichsspezifische Fähigkeiten beschreiben lassen (Klieme, 2004, S. 11),

oder die

bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können (Weinert, 2001, 2001/2014).

In der naturwissenschaftsdidaktischen Forschung wird typischerweise (z. B. Neumann et al., 2007; Schecker & Parchmann, 2006) auf die in den Bildungsstandards für den mittleren Schulabschluss zitierte Formulierung von Weinert verwiesen. Das (nicht überall so dezidiert vorgenommene) Inkludieren von Verständnissen in die eingerahmte Kompetenzdefinition rekurriert auf die in dem Zitat von Klieme (2004) aufgeführten Kenntnisse und wurde vor dem Hintergrund verschiedener Diskussionen

¹²Vor dem Hintergrund der Verortung der Arbeit im allgemeinbildend-schulischen Rahmen ist der dort mittlerweile als zentrale Orientierung herauskristalisierte (siehe z. B. Helmke & Schrader, 2014; Klieme & Hartig, 2007) und in den beiden aufgeführten Zitaten abgebildete Kompetenzbegriff maßgeblich. Die Diskussion deutlich abweichender Kompetenzbegriffe – wie etwa aus der Berufspädagogik (für einen Einstieg z. B. Erpenbeck & von Rosenstiehl, 2007b) oder aus dem Europäischen Qualifikationsrahmen (EQR; Europäische Kommission, 2008) – ist nicht Teil der Arbeit; stattdessen wird nur auf intradisziplinär relevante Aspekte zur Kompetenzdefinition eingegangen.

über die Relation von Kompetenz und Wissen vorgenommen, in denen die Wichtigkeit von Wissen bzw. Kenntnis für das Vorhandensein und den Aufbau von Kompetenzen herausgestellt wurde (vgl. z. B. Klieme et al., 2001; Klieme & Hartig, 2007; Kurtz, 2010; von Aufschnaiter & Hofmann, 2014; siehe im Kern schon bei Weinert, 2001). In der vorliegenden Arbeit wurde der Begriff *Verständnisse* gewählt, weil damit verschiedenste Formen des Wissens von einfachem Begriffswissen bis hin zu komplexen (z. B. situationsspezifischen), korrekten oder fehlerhaften, Vorstellungen eingeschlossen werden sollen (mehr dazu weiter unten in Unterabschnitt 2.2.2 ab S. 16).

Die eingerahmte Definition bezieht sich auf drei Teile des experimentbezogenen Denkens und Arbeitens und greift dadurch eine in der Forschungslandschaft häufig entlang von Teilprozessen getroffene Unterscheidungen auf. Sowohl national als auch international werden Unterscheidungen entlang von Teilprozessen des experimentbezogenen Denkens und Arbeitens in den Naturwissenschaftsdidaktiken an vielen Stellen vorgenommen (siehe z. B. Emden, 2011; Gut-Glanzmann & Mayer, 2018; Pedaste et al., 2015).¹³ Eine Einteilung von Kompetenzen zum experimentbezogenen Denken und Arbeiten entlang von Teilprozessen kann zwar nicht vollständig empirisch gerechtfertigt werden, weil Studien keine klare Favorisierung eines mehrdimensionalen Modells (und daher auch keine Anzahl und Beschreibung von zu unterscheidenden Teilkompetenzen) nahelegen (u. a. Nehring et al., 2016; Schmidt, 2016; Vorholzer et al., 2016). In der Studie, die der Arbeit zugrundeliegt, wird trotzdem zwischen

- dem Formulieren von Fragen und Hypothesen (FH),
- dem Planen und Durchführen von Untersuchungen (PU) und
- dem Auswerten und Interpretieren von Daten (AI)

unterschieden, was zusammenfassend in zwei Überlegungen begründet ist (ausführlicher bei Vorholzer, 2016). (a) Es kann aus fachdidaktischer Sicht sinnvoll sein, auch bei unklarer Psychometrie das experimentbezogene Denken und Arbeiten in Teilkonstrukte aufzuteilen, um Bezüge zwischen den Teilkonstrukten herzustellen oder die Konstruktion von Aufgaben oder Unterrichtsmaterial entlang der Teilkonstrukte vorzunehmen (Vorholzer, 2016, S. 19–21, 115–116; vgl. Kauertz, 2014; Wellnitz & Mayer, 2013). (b) Die Entscheidung für *drei* Teilprozesse lässt sich aus der zentralen

¹³Eine weitere Möglichkeit wäre beispielsweise eine Unterscheidung entlang von Problemtypen bzw. Arbeitsweisen (Gut-Glanzmann & Mayer, 2018; Nehring et al., 2016). Solche Unterscheidungen eignen sich aber eher für Untersuchungen von Kompetenzen des experimentbezogenen Denkens und Arbeitens in »wenig vorstrukturierten, authentischen, jedoch wohldefinierten experimentellen Problemsituationen« (Gut-Glanzmann & Mayer, 2018, S. 131), die nicht Gegenstand der Arbeit sind (vgl. Kapitel 4).

Rolle der drei Teilprozesse für naturwissenschaftliche Erkenntnisgewinnung begründen (beispielsweise sind naturwissenschaftliche Untersuchungen ohne das Formulieren von Fragen oder Hypothesen kaum denkbar; vgl. Vorholzer, 2016, S. 13–14, der u. a. auf Lederman et al., 2014, verweist; außerdem Kuhn & Dean, 2005; Lehrer & Schauble, 2015). Auch Vergleiche bildungspolitischer Vorgaben sowie nationaler und internationaler fachdidaktischer Forschungsarbeiten zeigen, dass sich die drei Teilprozesse in nahezu allen Beschreibungen und Definitionen identifizieren lassen (vgl. Vorholzer, 2016, S. 13, mit Verweis auf eine Übersicht bei Emden, 2011, und auf weitere Bildungsvorgaben und Forschungsvorhaben, oder Stiller, 2015).

Nachfolgend finden sich Erläuterungen zu den in der eingerahmten Definition (S. 13) benannten Komponenten von Kompetenzen zum experimentbezogenen Denken und Arbeiten. Auf Bereitschaften, die schon bei Weinert (2001/2014) von den anderen Komponenten separiert werden, wird dabei aus ähnlichen Gründen wie in der Klieme-Expertise (z. B. Klieme & Leutner, 2006, »Trennung kognitiver und motivationaler Faktoren«) nicht eingegangen; diese Komponente bleibt im Rahmen der vorliegenden Arbeit ein Desiderat.

2.2.1 Fähigkeiten und Fertigkeiten zum experimentbezogenen Denken und Arbeiten

In Bezug auf das experimentbezogene Denken und Arbeiten können die Begriffe *Fähigkeiten* und *Fertigkeiten* auf verschiedene Weisen differenziert oder gleichbedeutend genutzt werden. Bei Glug (2009) werden beispielsweise in Anlehnung an Frey (2006) Kompetenzen als aus Fähigkeiten zusammengesetzt angesehen, wobei Fähigkeiten wiederum aus Fertigkeiten zusammengesetzt sind. Für die Untersuchung von Kompetenz ist laut Glug eine Untersuchung des Konglomerats der zugehörigen Fähigkeiten notwendig, wobei erst durch »messbare Fertigkeiten [...] die theoretisch begründeten Fähigkeiten praktisch definiert« werden (2009, S. 22, vgl. S. 38–45). Fähigkeiten sind demnach latent, Fertigkeiten sind manifest (Glug, 2009, S. 39). Diesem Ansatz der hierarchischen Gliederung steht beispielsweise die Unterscheidung von Emden (2011) gegenüber, der Fähigkeiten und Fertigkeiten als komplementäre Regulatoren für Handeln beschreibt (und dabei auf arbeitspsychologische Überlegungen sowie den Gebrauch in der naturwissenschaftsdidaktischen Forschungslandschaft zurückgreift). Bei ihm werden die »Voraussetzungen zur händischen Durchführung laborspezifischer Handlungsabläufe als experimentelle Fertigkeiten« bezeichnet, Fähigkeiten beziehen sich auf »die planvolle Durchführung, Überwachung und Reflexion naturwissenschaftlicher Untersuchungen« (Emden, 2011, S. 15). Demnach beziehen sich Fähigkeiten

eher auf intellektuelle, Fertigkeiten eher auf sensomotorische Aspekte (Emden, 2011, S. 14). Eine weitere Möglichkeit, die beiden Begriffe zu verwenden, ist ein egalisieren-der Gebrauch, der in vielen naturwissenschaftsdidaktischen Publikationen vorhanden ist (Emden, 2011, S. 12).

Der *Schwerpunkt der vorliegenden Arbeit liegt auf intellektuellen Aspekten* des experimentbezogenen Denkens und Arbeitens (den Fähigkeiten im Sinne von Emden, 2011). Es ist an dieser Stelle jedoch anzumerken, dass diese nicht völlig von den motorischen Aspekten getrennt werden können, weil beispielsweise die Fähigkeit zur Kontrolle von Variablen auch mit Fertigkeiten des manuellen Umgangs mit Experimentiergegenständen zusammenhängt (vgl. u. a. Schreiber, 2012).

2.2.2 Verständnisse zum experimentbezogenen Denken und Arbeiten

Kompetenzen zum experimentbezogenen Denken und Arbeiten umfassen neben Fähigkeiten auch *Verständnisse* von Konzepten des experimentbezogenen Denkens und Arbeitens. Verständnisse von Lernenden zum experimentbezogenen Denken und Arbeiten beinhalten dabei mindestens (intraindividuell widerspruchsfreies und interindividuell konsensfähiges) Wissen dazu sowie (individuelle, ggf. z. T. widersprüchliche, ggf. situationsspezifische) Vorstellungen dazu, wie experimentbezogene Denk- und Arbeitsweisen ausgeführt werden (sollen) und was dabei zu beachten ist.¹⁴ Je fortgeschrittener die Lernenden sind, desto angemessenere Vorstellungen haben sie, was sich darin zeigt, dass die Vorstellungen näher an den aus fachlicher (d. h. in diesem Fall: fachmethodischer) Sicht angemessenen Konzepten des experimentbezogenen Denkens und Arbeitens sind.

Verständnisse zum naturwissenschaftlichen Denken und Arbeiten – unter das auch das experimentbezogene Denken und Arbeiten subsumiert ist – werden als *eine zentrale Zieldimension von naturwissenschaftlichem Unterricht* angesehen und als solche eigenständig untersucht. Sie finden sich prominent neben dem Verständnis von fachlichen Inhalten.¹⁵ Beispielsweise werden von vielen Autoren drei Formen von

¹⁴Nach Fertigstellung der entsprechenden Teile der Arbeit war auch eine Sichtung des neu erschienenen Werkes von Reinisch et al. (2020) möglich. Dort scheint eine ähnliche oder zumindest passende begriffliche Fassung vorgenommen zu werden.

¹⁵Die nachfolgend weiter ausdifferenzierte Unterscheidung von fachinhaltlichem Wissen und fachmethodischem, auf das experimentbezogene Denken und Arbeiten bezogene, Wissen ist nicht so zu lesen, dass fachinhaltliche Kompetenzen für Kompetenzen zum experimentbezogenen Denken und Arbeiten unbedeutend wären (die Forschung zeigt, dass – von Teilprozess zu Teilprozess unterschiedlich stark – eher das Gegenteil der Fall zu sein scheint, z. B. Zimmerman & Klahr, 2018, insb. S. 8 u. S. 16). In einem weitgefassten Sinn kann fachinhaltliches Wissen sogar zu den Verständnissen zum naturwissenschaftlichen Denken und Arbeit gezählt werden, weil ohne grundlegendes fachinhaltliches Wissen naturwissenschaftliche Denk- und Arbeitsweisen kaum

Wissen bzw. Verständnis (auf Englisch meist knowledge, manchmal understanding) unterschieden (eine Übersicht findet sich bspw. bei Koenen, 2014, Tabelle 2 auf S. 5): Inhaltswissen (engl. content knowledge), Prozesswissen (engl. procedural knowledge) und Erkenntniswissen (engl. epistemic knowledge).¹⁶ Kind und Osborne (2017) beklagen beispielsweise eine Überbetonung des Inhaltswissens in der Forschung und in der Schule und stellen als drei Ziele von Unterricht den Aufbau von Wissen der drei Arten dar:

- [•] Content knowledge of the appropriate domain-specific concepts—that is the ontological entities that science uses to reason with;
- [•] Procedural knowledge which is knowledge of the procedures and associated constructs that science uses to establish it [sic] claims to know; and
- [•] Epistemic knowledge which is knowledge of the epistemic constructs and values and how these are used to justify science’s claims to know. (Kind & Osborne, 2017, S. 3)

In den Veröffentlichungen zu PISA wird unter den Begriffen Vergleichbares verstanden und auch dort wird Wissen aus allen drei Bereichen als wichtiges Kennzeichen von in den Naturwissenschaften kompetenten Schülerinnen und Schülern angesehen (z. B. OECD, 2016).¹⁷

ausgeführt werden könnten (vgl. Koslowski, 1996; Zimmerman, 2007). In einem enger gefassten Sinn wird in der vorliegenden Arbeit jedoch unter *Verständnisse zum experimentbezogenen Denken und Arbeiten* nur solches Wissen gezählt, welches auf Fachmethoden gerichtet ist; dadurch wird die häufig vorgenommene Unterscheidung fachinhaltlicher Kompetenzen und fachmethodischer Kompetenzen aufrechterhalten (wenngleich diese zum Teil angezweifelt wird, z. B. Koslowski, 1996; Nehring & Schwichow, 2020; diskutiert auch bei Schauble, 1999), denn zu fachmethodischen Kompetenzen werden in der Verständniskomponente nur solche Verständnisse gezählt, die inhärent fachmethodisch sind.

¹⁶Diese Unterscheidung bezieht sich auf den Inhalt des Wissens. Insbesondere wird daher unter Prozesswissen solches Wissen subsumiert, dass auf (Handlungs-)Prozesse bezogen ist. Die Unterscheidung von Inhalts- und Prozesswissen kann daher nicht mit der Unterscheidung von deklarativem und prozeduralem Wissen gleichgesetzt werden (in ähnlicher, wenngleich nicht vollständig übereinstimmender Weise bei Arnold et al., 2017, die Fach- und sogenanntes Methodenwissen als zwei Arten deklarativen Wissens beschreiben). Beispielsweise lässt sich mit den Worten von Oberauer (1993) verdeutlichen, dass mit der Zuordnung eines Elementes zur Wissensform *Prozesswissen* zunächst keine Aussage darüber getroffen wird, welche Funktion das Wissen mental hat: »Diese Unterscheidung [von Sach- und Handlungswissen, die der hier aufgeführten Unterscheidung von Inhalts- und Prozesswissen weitgehend entspricht] betrifft den *Inhalt* des Wissens. Die Differenz von deklarativem und prozeduralem Wissen betrifft dagegen die *Rolle* des Wissens im Denken und Handeln« (S. 36).

¹⁷In der vorliegenden Arbeit wird nur Bezug auf Kompetenzen zum experimentbezogenen Denken und Arbeiten genommen. Allerdings finden sich auch für andere fachmethodische Kompetenzen ähnliche Differenzierungen, die die Wichtigkeit von Verständnissen als Ziele von Unterricht

Die vorgeschlagene Trennung von Prozesswissen und Erkenntniswissen wird nicht überall vorgenommen, wie das folgende Zitat aus dem Sammelbeitrag von Abd-El-Khalick et al. (2004, S. 413) zeigt, in dem zunächst eine zur obigen Unterscheidung passende Differenzierung zitiert, danach aber eine andere Unterscheidung propagiert wird:

Engagement in inquiry would seem to require the development of at least three, if not more, interconnected knowledge bases. According to Krajcik, Czerniak, and Berger (1998 [korrekt wäre wohl: 1999]), the three types of knowledge to be developed in project-based science are (a) content: knowing the central concepts and principles in a domain; (b) inquiry and problem solving: knowing how to problem-solve, design, and carry out investigations and employ metacognitive strategies; and (c) epistemic: knowing the “rules of the game,” what counts as evidence, and when to collect more evidence to support/refute a position.

In our work (Duschl, Ellenbogen, & Erduran, 1999) we see the need to coordinate inquiry learning around three goals: conceptual, social, and epistemic. Conceptual and epistemic goals are consistent with Krajcik et al.’s first and third types of knowledge (Krajcik, Czerniak, & Berger, 1998 [1999]).

Die aus Krajcik et al. (1999) zitierte Beschreibung der drei Wissensarten stimmt im Wesentlichen mit der in der vorliegenden Arbeit weiter oben zitierten Beschreibung von Kind und Osborne (2017) überein. Abd-El-Khalick et al. (2004) selbst nehmen aber im zweiten Absatz des Zitats eine andere Unterscheidung vor, ohne (im zitierten oder im weiteren Text) zu explizieren, wo sie die zweite Wissensart von Krajcik et al. (1999) verorten (oder ob sie diese als irrelevant ansehen). Dies könnte damit zusammenhängen, dass eine vollständige Abgrenzung des Wissens über das Vorgehen (also des Prozesswissens) von erkenntnistheoretischen Begründungen des Vorgehens (also vom Erkenntniswissen) aus Sicht einiger Autoren kaum möglich zu sein scheint (siehe z. B. Stiller, 2015, Abschnitt 2.3). Vielleicht subsumieren Abd-El-Khalick et al. (2004, S. 413) das Prozesswissen als Teilmenge des Erkenntniswissens.

ergänzend zu Fähigkeiten hervorheben. Für das Argumentieren schreiben Hofstein und Kind (2012) beispielsweise: »We sense two overlapping learning aims: [1: Verständnisse] students should understand [a] the scientific standards and [b] their guiding epistemologies; [2: Fähigkeiten] and students should be able to apply these standards in their own argumentation« (S. 201). Anmerken lässt sich, dass erstens sowohl Verständnisse [1] als auch Fähigkeiten [2] und zweitens sowohl Prozesswissen [a] als auch Erkenntniswissen [b] in diesem Zitat enthalten sind.

2.2 Modellierung von Kompetenzen zum experimentbez. Denken und Arbeiten

In der vorliegenden Arbeit werden Prozesswissen und Erkenntniswissen in Anlehnung an Kind und Osborne (2017) getrennt, die für das Experimentieren schreiben:

Experimentation, similarly, is not just a matter of knowing how to get reliable data—which is a procedural issue—but also why reliability is important—which is an epistemic issue. To conduct an experiment, undoubtedly, it is essential to have some understanding of what a variable is, and which variables might be controlled. Why it is necessary to control variables, however, is an epistemic construct. (Kind & Osborne, 2017, S. 16)

Gemäß den Autoren lassen sich also Antworten auf die Frage, *warum* (oder besser: für welche Ziele) Variablen kontrolliert werden *müssen*, dem Erkenntniswissen zuordnen; Wissen darüber, *wie* Variablenkontrolle durchgeführt wird und was alles dazugehört, ist Prozesswissen.

Gegenstand der Arbeit sind die Verständnisse der Lernenden zu dem *Wie* des experimentbezogenen Denkens und Arbeitens.¹⁸ Sie können gemäß obiger Überlegungen als Teil der Kompetenzen zum experimentbezogenen Denken und Arbeiten und damit als ein wesentliches Ziel von Naturwissenschaftsunterricht angesehen werden. Die Überlegungen lassen sich auch als Begründungen für die ausdrückliche Aufnahme der Verständnisse in die Kompetenzbegriff am Anfang des Abschnitts 2.2 lesen – denn Kompetenzen zum experimentbezogenen Denken und Arbeiten umfassen in der Forschung und in Bildungsvorgaben eben nicht nur Fähigkeiten (und Fertigkeiten), sondern auch ausdrücklich Verständnisse von zugehörigen prozessbezogenen Vorgehensweisen und Konstrukten. In gewisser Weise zeigen das auch die folgenden Ausführungen.

Die Schwerpunktsetzung der Arbeit auf Verständnisse lässt sich der von Gut-Glanzmann und Mayer (2018) als *conceptual understanding* bezeichneten Rahmung für das experimentbezogene Denken und Arbeiten zuordnen. Eine Integration theoretischer Überlegungen und empirischer Befunde aus dem Kontext anderer Rahmungen scheint angebracht, weil sich die von der conceptual-understanding-Rahmung bei Gut-Glanzmann und Mayer (2018) unterschiedenen Rahmungen (*scientific-reasoning*-,

¹⁸Eine weitere, innere, Differenzierung der Verständnisse zum Wie des experimentbezogenen Denkens und Arbeitens schlagen bspw. Zohar und Peled (2008) vor: »(a) procedural knowledge that has to do with the effective use of strategies [...] and (b) conditional knowledge that refers, among other things, to knowing when and why to use strategies« (S. 338). Die in der vorliegenden Arbeit vorgenommene Differenzierung in verschiedene Wissensbestandteile basiert allerdings auf der der Arbeit zugrundeliegenden Studie, die in Kapitel 4 vorgestellt wird; es wird daher an dieser Stelle darauf verzichtet, ausführlicher über Differenzierungsmöglichkeiten zum Prozesswissen zu berichten.

skill-orientierte Ansätze oder Rahmungen als Problemlösen) »aufgrund inhaltlicher Überschneidungen und gegenseitiger Abhängigkeiten in der Handlungsanalyse nur schwer trennen« lassen. Es scheint sogar der Fall zu sein, dass Verständnissen (conceptual understanding) in allen anderen Rahmungen eine nicht unwesentliche Bedeutung zugewiesen wird. Beispielsweise verweisen Gut-Glanzmann und Mayer (2018) in ihren Ausführungen zu verschiedenen Rahmungen des experimentbezogenen Denkens und Arbeitens (S. 127–132) jeweils auf die Wichtigkeit von Prozesswissen: Als ein Desiderat in der Rahmung als wissenschaftliches Denken ergibt sich die Frage nach dem Zusammenhang von Prozesswissen (dort: Methodenwissen) und experimentellen Fähigkeiten beim Erwerb experimenteller Kompetenz (Gut-Glanzmann & Mayer, 2018, S. 129); für jeden der Problemtypen in der Rahmung des experimentbezogenen Denkens und Arbeitens als Problemlösen wird ein zur Lösung erforderliches sogenanntes Problemverständnis aufgeführt (Gut-Glanzmann & Mayer, 2018, Tabelle 8.2 auf S. 132). Gleichzeitig sind Rahmungen, die ausdrücklich die Wichtigkeit der Verständnisse für das experimentbezogene Denken und Arbeiten betonen, wenig prominent in der Forschungslandschaft, obwohl sie einen nicht zu vernachlässigenden Beitrag zur Beschreibung von Kompetenzen zum experimentbezogenen Denken und Arbeiten leisten können (z. B. Gut-Glanzmann & Mayer, 2018, S. 124; siehe auch Millar et al., 1996; Roberts, 2001). Eine Betonung von Verständnissen zum experimentbezogenen Denken und Arbeiten in der vorliegenden Arbeit scheint demnach aus mehreren Gründen angemessen zu sein.

2.3 Modellierung von Kompetenzaufbau (zum experimentbezogenen Denken und Arbeiten)

Auf einer allgemeinen Ebene wird Kompetenzaufbau in der vorliegenden Arbeit von Kompetenzstatūs¹⁹ unterschieden: Während der Begriff Kompetenzaufbau darauf bezogen ist, dass sich die Kompetenz von Lernenden im zeitlichen Verlauf wandelt, beschreibt der Begriff Kompetenzstatus, welche Kompetenz bei Lernenden zu einem Zeitpunkt vorliegt (vgl. z. B. Erpenbeck & von Rosenstiehl, 2007a, S. XXVI).

Kompetenzstatūs. Eine genauere Betrachtung von Kompetenzstatūs zeigt – insbesondere mit Blick auf das experimentbezogene Denken und Arbeiten –, dass zur Beschreibung von Kompetenzen, die bei Lernenden zu einem Zeitpunkt bzw. in einem kleinen Zeitfenster vorliegen, zeitlich ausgedehnte Prozesse des Denkens und Handelns von Personen untersucht werden können. Dies wird auf mindestens zwei Ebenen deutlich.

Prozessdiagnostik. Neben der in der Forschung zu vorliegenden Kompetenzen häufig vorgenommenen Analyse von Aufgabenlösungen, Produkten, die Lernende erstellen, bzw. Ergebnissen von Handlungen der Lernenden (in Bezug auf experimentbezogene Kompetenzen z. B. Hübinger, 2008; Lazonder & Egberink, 2014; Muth, 2017; Vorholzer, 2016) können auch Analysen des Denkens und Handelns von Lernenden bei der Bearbeitung von Aufgaben oder dem Erstellen von Produkten Aufschluss über die bei den Lernenden vorliegenden Kompetenzen geben (in Bezug auf experimentbezogene Kompetenzen selten, meist zur Validierung schriftlicher Erfassungen, z. B. Beretz, 2013; Dickmann, 2016). Diese Unterscheidung wurde in Anlehnung an von Aufschnaiter et al. (2015) getroffen, die in analoger Weise verschiedene Arten von Diagnostik differenzieren; gemäß der dort vorzufindenden Ausführungen lässt sich beispielsweise für das Denken und Handeln bei der Aufgabenbearbeitung oder bei der Produkterstellung analysieren, »ob die generierte Lösung erst mit dem Aufsuchen verschiedener alternativer Lösungsansätze entstanden ist und (damit einhergehend) welche zeitliche Dauer der Bearbeitungsprozess hat« (von Aufschnaiter et al., 2015, S. 745). Dabei könnte sich beispielsweise »zeigen, dass inhaltlich richtige Ergebnisse [...] aus inhaltlich unangemessenen Prozessen generiert werden oder umgekehrt« (von Aufschnaiter et al., 2015, S. 745). In Bezug auf das experimentbezogenen Denken und Arbeiten ist beispielsweise denkbar, dass die Lernenden zu fachinhaltlich angemessenen Ergebnissen kommen, obwohl ihr Vorgehen beim Experimentieren aus fachmethodischer Sicht eher unangemessen ist oder sie ihr (ggf. angemessenes) Vorgehen fachmethodisch unangemessen begründen.

¹⁹ *Kompetenzstatūs* ist der Plural von Kompetenzstatus.

Experimentieren als ein Prozess. Für Kompetenzen zum experimentbezogenen Denken und Arbeiten kann über die bisherigen Ausführungen hinaus angemerkt werden, dass experimentbezogenes naturwissenschaftliches Denken und Arbeiten ein zeitlich ausgedehnter Prozess ist, der verschiedene Teilprozesse umfasst. Um gewisse Aspekte von Kompetenzen zum experimentbezogenen Denken und Arbeiten zu erfassen, kann es also ggf. ein adäquates Vorgehen sein, Abfolgen von experimentellen (Sprach-) Handlungen von Lernenden zu erfassen und diese eventuell mit (ideal)typischen Abläufen des Experimentierens abzugleichen (vgl. Arndt, 2016; Čančula et al., 2015; Sonnenschein, 2019; Strippel, 2017). Solch ein Vorgehen scheint vor allem dann angemessen, wenn Kompetenzen von Lernenden zum experimentbezogenen Denken und Arbeiten in wenig vorstrukturierten Umgebungen erfasst werden sollen, weil in diesen Fällen die eigenständige Strukturierung des Experimentierprozesses als Teil der Kompetenz angesehen werden kann (vgl. Gut, 2012). Wenngleich Prozesse des experimentbezogenen Denkens und Arbeitens zeitlich deutlich ausgedehnt sein können, ist in bisherigen Studien die Betrachtungen dieser Prozesse meist darauf ausgerichtet, die – als im analysierten Prozess einigermaßen stabil angenommene – vorliegende Kompetenz der Lernenden zu erfassen (Emden, 2011; Ruiz-Primo & Shavelson, 1996; Schreiber et al., 2014).

Kompetenzaufbau. Eine innere Differenzierung des Begriffs *Kompetenzaufbau* kann beispielsweise dadurch erfolgen, dass zwischen zeitlich *diskreten* Folgen von Kompetenzstatūs (auf Status 1 folgt Status 2 usw.) und zwischen einer *kontinuierlichen* Beschreibung des Wandels der Kompetenz unterschieden wird. Ferner kann unterschieden werden, wie stark der zeitliche Aspekt des Kompetenzaufbaus fokussiert wird: Neben Untersuchungen zur Veränderung der Kompetenz von einem Zeitpunkt zu einem zweiten (Kompetenzveränderung) sind auch Untersuchungen denkbar, bei denen mehrere Zeitpunkte verglichen werden (Kompetenzentwicklung); Untersuchungen können von Anfangs- bis Endzeitpunkt nicht nur einige Minuten bis wenige Stunden, sondern auch Wochen, Monate oder gar Schuljahre umfassen; zudem lässt sich auch unterscheiden, welchen zeitlichen Umfang die untersuchten Kompetenzstatūs bzw. Prozesse haben (bspw. ob Ergebnisse aus zwei Tests miteinander verglichen werden oder zwei jeweils halbstündige Videos).

Zur Beschreibung von Kompetenzaufbau als zeitliche Folge verschiedener Kompetenzstatūs können *Modelle* herangezogen werden, die differenzieren, wie Kompetenz von Lernenden (zu einem jeweils festen Zeitpunkt) ausgeprägt ist. Der erste und naheliegende Einsatz solcher Modelle besteht im interindividuellen Vergleich unterschiedlicher Lernender (zu demselben Zeitpunkt). Gleichzeitig können diese Modelle auch für den Kompetenzaufbau herangezogen werden – dieser wird in der vorliegen-

2.3 Modellierung von Kompetenzaufbau (zum experimentbez. Denken u. Arbeiten)

den Arbeit insbesondere auf den lernendenspezifischen, intraindividuellen Wandel der Kompetenz im Laufe der Zeit bezogen aufgefasst, häufig allerdings auch einfach nur mit einer allgemeineren über die untersuchte Kohorte gemittelten Änderung der modellierten Aspekte in Verbindung gebracht.

Strukturmodelle trennen verschiedene Dimensionen von Kompetenz (u. a. Schecker & Parchmann, 2006; für experimentbezogene Kompetenzen z. B. Gehlen, 2016; Nehring, 2014; Schreiber, 2012). Beispielsweise ist die für die vorliegende Arbeit vorgenommene Entscheidung, Kompetenzen zum experimentbezogenen Denken und Arbeiten als eigenständige Kompetenz von fachinhaltlicher Kompetenz zu unterscheiden, eine auf Kompetenzstrukturen bezogene Modellierung wie auch die Differenzierung der drei Teilprozesse des experimentbezogenen Denkens und Arbeitens (Formulieren von Fragen und Hypothesen; Planen und Durchführen von Untersuchungen; Auswerten und Interpretieren von Daten).

*Niveaumodelle*²⁰ beschreiben Unterschiede darin, wie eine Kompetenz ausgeprägt ist und welche Qualität diese Ausprägung relativ zu anderen hat (u. a. Hartig & Klieme, 2007; Schecker & Parchmann, 2006). Was eine höherwertige Kompetenz ausmacht, wird dabei sehr verschieden konzeptioniert. Für experimentelle Kompetenz schlägt Hammann (2004) beispielsweise vor, eigene Niveaus für jeden der drei folgenden Teilbereiche zu formulieren: Hypothesen-Suchraum (entspricht dem Formulieren von (Fragen und) Hypothesen), Experimentier-Suchraum (entspricht dem Planen von Untersuchungen) und Analyse von Daten. Für die als Teilkompetenz aufgefasste Suche im Experimentier-Suchraum beschreibt er beispielsweise vier Stufen vom unsystematischen Umgang mit Variablen (Variablenveränderung ohne System; u. a. sogenannte *change-all*-Ansätze) über teilweise systematischen Umgang und über systematischen Umgang in bekannten Domänen bis hin zum systematischen Umgang in unbekannten Domänen (nur die Ausprägung der unabhängigen Variablen wird verändert; gelingt auch in Domänen, für die Lernende geringes fachinhaltliches Vorwissen aufweisen).²¹ Unabhängig von der Konzeptionierung der Niveaus lassen

²⁰Häufig werden Niveaumodelle auch als Stufenmodelle bezeichnet. Der Begriff der Stufen kann jedoch so gelesen werden, dass er suggeriert, dass die Stufen nacheinander abgelaufen werden, was jedoch nicht explizit Teil dieser Modelle ist. Die zeitliche Abfolge von verschiedenen ausgeprägten Kompetenzen ist erst Gegenstand von den im weiteren Text beschriebenen Kompetenzentwicklungsmodellen.

²¹Gut (2012) zeigt auf, dass die von Hammann (2004) formulierten Niveaus nahelegen, dass »die Entwicklung in den unteren Primarklassen ausschliesslich [sic] in Form eines Zuwachs der Qualität, mit der experimentiert wird, stattfindet und erst in den oberen Klassen Aufgaben mit grösserem [sic] Umfang, Transfer oder höherer Komplexität bewältigt werden«, obwohl dies empirisch nicht gerechtfertigt und, laut Gut, vermutlich von Hammann auch nicht intendiert ist. In der Arbeit von Gut (2012) sind weitere Dimensionen zur Beschreibung von Niveau-Unterschieden dargestellt (sogenannte Progressionsdimensionen), mithilfe derer in der Arbeit von Gut (2012) konsistentere Beschreibungen von Niveaus angestrebt werden.

sich diese nicht per se im Sinne zeitlicher Abfolgen im individuellen Lernen auffassen (vgl. Prediger & von Aufschnaiter, 2017; Renkl, 2012).

Sowohl Struktur- als auch Niveaumodelle können allerdings genutzt werden, um Kompetenzentwicklung zu beschreiben, weil unterschiedliche Ausprägungen der Kompetenz von Lernenden aufgelöst und in eine zeitliche Reihenfolge gebracht werden können. In der Forschung wird teilweise als weitere Kategorie die Gruppe der *Kompetenzentwicklungsmodelle* eingeführt, welche dezidiert auf die Beschreibung davon ausgerichtet sind, in welcher zeitlichen Reihenfolge sich Kompetenzen herausbilden. Der Begriff wird zum Teil auch für Modelle genutzt, die anhand von querschnittlichen Studien entwickelt wurden und aus denen sich demnach nur eingeschränkt Aussagen dazu ableiten lassen, wie die verschieden ausgeprägten Kompetenzen lernendenspezifisch zeitlich aufeinander folgen.²²

Bei der Untersuchung von Kompetenzaufbau stellen sich neben Fragen nach zu verschiedenen Zeitpunkten vorliegenden Kompetenzen auch Fragen, die auf das Denken und Handeln, welches von einem Kompetenzstatus zum nächsten führt, gerichtet sind.²³ Während im ersten Fall die Kompetenzen (als Dispositionen) am Handeln der Lernenden festgemacht werden, sind die Analysen im zweiten Fall darauf bezogen, welche (kompetenzbezogenen) Handlungen von den Lernenden in der Zwischenzeit ausgeführt werden und welche (kompetenzbezogenen) Verständnisse der Lernenden durch diese Handlungen offenbar werden, durch die die Wege zu den Kompetenzstatus gekennzeichnet sind. Beispielsweise lässt sich nicht nur die Frage danach, welche Lernenden von dem vor einem Unterricht bearbeiteten Test zu dem nach dem Unterricht bearbeiteten Test besonders viel oder wenig dazugelernt haben, als Frage zum Kompetenzaufbau der Lernenden auffassen; auch die Frage danach, ob Lernende mit besonders großem Kompetenzaufbau besonders häufig oder besonders lange bestimmten Handlungen nachgehen oder besonders ausgeprägt spezifische Vorstellungen verbalisieren, ist eine auf den Kompetenzaufbau gerichtete Frage im Sinne der vorliegenden Arbeit.

Fragen zu dem von einem Kompetenzstatus zu einem nächsten Kompetenzstatus führenden Denken und Handeln der Lernenden stellen sich unter anderem deshalb,

²²Schecker und Parchmann (2006, S. 48) unterscheiden zwischen präskriptiv-normativen und empirisch-deskriptiven Modellen. Der der Fußnote vorausgehende Absatz der vorliegenden Arbeit ist als Ausführung über empirisch-deskriptive Modelle anzusehen, weil es um die zu verschiedenen Zeitpunkten *vorliegenden* Kompetenzen geht. Mit Blick auf Kompetenzniveau- und Kompetenzentwicklungsmodelle lässt sich allerdings anmerken, dass eine derartige Differenzierung für präskriptiv-normative Modelle vermutlich nicht sinnvoll aufrechterhalten werden kann, weil ein Vorschreiben von Niveaus ohne damit einhergehendes Anstreben davon, dass sich die höheren Niveaus auch später im Lernen zeigen, nur schlecht denkbar ist.

²³Ähnliche Überlegungen finden sich bei von Aufschnaiter et al. (2015).

weil auch die in Kompetenzentwicklungsmodellen angegebenen Abfolgen von durchlaufenen Kompetenzniveaus (selbst bei Auflösung für jedes einzelne Individuum) aufgrund der diskreten, nicht notwendigerweise organisch aufeinander aufbauenden Formulierung der Niveaus nur wenig Aufschluss darüber geben, durch welche Wege die (nur) anhand von isolierten Kompetenzstatus beschriebene Kompetenzentwicklung gekennzeichnet ist. Wie gelangen Lernende von einem Kompetenzniveau zu einem nächst höheren? Geschieht dies tatsächlich durch ein einfaches Hinzufügen der in der höheren Niveaubeschreibung hinzugekommenen Teilfähigkeit? Und gibt es ggf. typische Zwischenschritte (werden im Beispiel von Hammann oben bspw. zunächst oberflächliche Systematisierungen wie etwa »kleinkariertes, akribisches Messen« vorgenommen, bevor später auf die Variation der unabhängigen Variable und noch später auf die Konstanthaltung von Kontrollvariablen geachtet wird)?

Im Sinne von Angebots-Nutzungs-Modellen werden sinnvollerweise bei der Untersuchung von Fragen zum Kompetenzaufbau, die darauf bezogen sind, wie Lernende von einem Kompetenzstatus zu einem späteren gelangen, sowohl die instruktionale Angebotsseite als auch die lernenden-individuelle Nutzungsseite in den Blick genommen (vgl. u. a. Lersch, 2007; Seidel, 2014). Beispielsweise können Beschreibungen einer eingesetzten Instruktion zwar Aufschluss darüber geben, womit eventuell zustandekommender Kompetenzzuwachs zusammenhängen könnte; spezifischere Hinweise auf die tatsächlich zur Kompetenzveränderung führenden Denk- und Handlungsprozesse (sowie tatsächlich medierende Personenmerkmale) geben allerdings erst Analysen der Lernenden bei der Bearbeitung der Instruktion.

2.4 Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten: Befunde und zugehörige Instruktionsmaßnahmen

Nachfolgend werden als Ankerpunkte der vorliegenden Arbeit zentrale Befunde aus dem Forschungsstand abgebildet. Einige wenige Studien bezüglich des Aufbaus von Kompetenzen zum naturwissenschaftlichen Experimentieren beziehen sich auf mehrere aufeinanderfolgende Kompetenzstatus über einen längeren Zeitraum von mehreren Monaten bis hin zu mehreren Schul- oder Lebensjahren. Ein kurzer Überblick über solche Studien zur Kompetenzentwicklung wird in Unterabschnitt 2.4.1 gegeben, weil eine Untersuchung zum Kompetenzaufbau bei der Bearbeitung von Instruktion letztlich in die Verläufe der Kompetenzentwicklung auf großen Zeitskalen eingebettet ist. Die meisten Studien zum Kompetenzaufbau im in der vorliegenden Arbeit fokussierten Bereich nehmen Kompetenzveränderung von einem Zeitpunkt zu einem einzigen späteren oder die Wirkung von Instruktion auf eine (oder mehrere) nachfolgende Testung(en) in den Blick. Da in diesen Studien oft Vergleiche

von verschiedenen Instruktionsmaßnahmen und den zugehörigen Kompetenzzuwächsen vorgenommen werden, können viele der Studien erst vor dem Hintergrund der eingesetzten Instruktionsmaßnahmen angemessen eingeordnet werden. In Unterabschnitt 2.4.2 werden Befunde zur Kompetenzveränderung vor dem Hintergrund des Denkens und Handelns der Lernenden bei der Bearbeitung der Fördermaßnahmen zusammengefasst. Die Studien zu den Instruktionsmaßnahmen sind im Wesentlichen auf Fähigkeiten bezogen. Ein Schwerpunkt der vorliegenden Arbeit liegt – wie in Abschnitt 2.2 dargestellt – auf den Verständnissen der Lernenden. Soweit es möglich ist, wird der Forschungsstand mit diesem Fokus zusammengefasst. Ferner finden sich in Unterabschnitt 2.4.3 spezifische Befunde zu Vorstellungen von Lernenden zum experimentbezogenen Denken und Arbeiten.

2.4.1 Kompetenzentwicklung über längere Zeiträume

Zu längeren Prozessen der Kompetenzentwicklung zum experimentbezogenen Denken und Arbeiten finden sich wenige Studien, in denen meist methodisch nicht mit Längsschnitten gearbeitet wird. Beispielsweise wird in der Dissertation von Grube (2011) über eine Untersuchung der Kompetenzen zum experimentbezogenen Denken und Arbeiten über die Jahrgänge der Sekundarstufe I hinweg berichtet (Querschnitt für Klassen 5–10, Prä-Post-Vergleich für jede Jahrgangsstufe). Die dort modellierten Teilkompetenzen zu den Teilprozessen des experimentbezogenen Denkens und Arbeitens haben im Mittel über die Jahrgänge unterschiedliche Schwierigkeiten. Schülerinnen und Schülern fällt es am schwersten, naturwissenschaftliche Fragen zu formulieren und zu erkennen. In absteigender Reihenfolge nennt Grube (2011) dann die – etwas anders als in der vorliegenden Arbeit zugeschnittenen – Teilkompetenzen Planung, Hypothese, Deutung. Kompetenzen der Schülerinnen und Schüler nehmen in Bezug auf die Teilkompetenzen unterschiedlich stark zu, woraus Grube (2011) folgert, dass eine »gezielte Diagnostik und Förderung einzelner Teilfähigkeiten wissenschaftlichen Denkens [...] vor dem Hintergrund dieser Befunde von besonderer Bedeutung« ist (S. 91).

Ein weiteres Beispiel ist Ding (2018), der chinesische Schülerinnen und Schüler sowie Studentinnen und Studenten jahrgangsweise (Jahrgänge 4–16) in verschiedenen Teilfähigkeiten zum wissenschaftlichen Denken testet. Für verschiedene untersuchte Teilfähigkeiten zeigen sich unterschiedliche Prozesse der Kompetenzentwicklung. Beispielsweise wird die Fähigkeit zur Variablenkontrolle früher in ersten Ansätzen aufgebaut als die zum *probabilistic reasoning*, aber für letztere vollzieht sich ein steilerer Anstieg über die Jahrgänge hinweg. Und Piekny und Maehler (2013) zeigen (querschnittlich) für jüngere Kinder und Schülerinnen und Schüler auf, dass »the three cognitive components of domain-general scientific reasoning [nach dem SDDS-Modell

von Klahr und Dunbar (1988): Generieren von Hypothesen, experimentelles Testen von Hypothesen, Evaluieren von Evidenz] emerge asynchronously. The development of domain-general scientific reasoning begins with the ability to handle unambiguous data, progresses to the interpretation of ambiguous data, and leads to a flexible adaptation of hypotheses according to the sufficiency of evidence. When children understand the relation between the level of ambiguity of evidence and the level of confidence in hypotheses, the ability to differentiate conclusive from inconclusive experiments accompanies this development« (S. 1, im Abstract).

Mit Blick auf die *Kompetenzveränderung in den einzelnen Jahrgängen* berichtet Grube (2011), dass sich in höheren Jahrgängen geringere Zuwächse finden. Dieser Befund steht gewissermaßen im Widerspruch zu Ding (2018), bei dem sich in den entsprechenden Jahrgängen (9–11) die stärksten Zuwächse finden. Die von Grube (2011) genannten methodischen Gründe (Deckeneffekt des Tests in einigen Teilbereichen) für ihre Befunde scheinen daher plausibler als die von ihr vorgeschlagenen inhaltlichen Erklärungen. Dabei muss allerdings berücksichtigt werden, dass die Schulsysteme, Bildungsvorgaben usw. vermutlich nur in geringem Maße vergleichbar sind. Für die vorliegende Arbeit ist zunächst von Relevanz, dass in beiden Studien unterschiedlich starke Zuwächse für verschiedene Jahrgänge gefunden wurden. Bei der Generalisierung der Befunde der vorliegenden Arbeit muss daher die Einschränkung durch die Wahl eines spezifischen Jahrgangs beachtet werden.

Studien, die einzelne *Teilprozesse* in den Blick nehmen stützen die bisher genannten Befunde und Folgerungen (vgl. z.B. Amsel & Brock, 1996; Piekny & Maehler, 2013). Die Studie von Bullock und Ziegler (1999) ist laut Klahr (1999) die erste Längsschnittstudie zum experimentbezogenen Denken und Arbeiten (über einen Zeitraum von 4 Jahren im Alter von 9 bis 12). Ergänzend zu den bereits berichteten generellen Befunden, findet sich innerhalb des Teilprozesses der Planung in der Studie ein Unterschied zwischen den verschiedenen betrachteten Fähigkeiten. Schülerinnen und Schüler sind bereits in früheren Jahrgängen in der Lage, korrekt geplante Experimente zu identifizieren und *auszuwählen*; erst in späteren Jahrgängen sind die Schülerinnen und Schüler dazu in der Lage, eigenständig korrekte Experimente zu *planen* (Bullock & Ziegler, 1999).

In Bezug auf *Personenmerkmale*, die Prozesse der Kompetenzentwicklung beeinflussen, finden Bullock und Ziegler (1999), dass die Identifizierung von korrekt geplanten Experimenten, die zeitlich früher von den Lernenden geleistet werden kann, vor allem von allgemein-kognitiven Faktoren vorhergesagt werden kann. Für die korrekte Planung, die von den Lernenden erst zeitlich später bewältigt wird, spielen auch andere Personenmerkmale wie verbalisierbares Wissen über experimentbezogenes

Denken und Arbeiten eine Rolle. Bei Grube (2011) wird ein nicht-auffindbarer Einfluss von Eingangsvoraussetzungen in Bezug auf das experimentbezogene Denken und Arbeiten (bei Grube: Ausgangskompetenz) auf methodische Probleme zurückgeführt. Ferner findet sich bei Grube (2011) kein Einfluss von Geschlecht auf die Prozesse der Kompetenzentwicklung zum experimentbezogenen Denken und Arbeiten.

Ein möglicher *Einfluss der Lernumgebung* wurde bei Grube (2011) untersucht. Einige Schulklassen waren dem Projekt *Biologie im Kontext* zugeordnet. Die generelle Beschulung gemäß dem Projekt zeigte zwar keine Effekte auf die Prozesse der Kompetenzentwicklung der Schülerinnen und Schüler, allerdings konnte für Lehrkräfte im Projekt, die zudem angaben, kompetenzorientiert zu unterrichten, ein Effekt auf Kompetenzen zum experimentbezogenen Denken und Arbeiten gefunden werden. Die Befunde lassen sich so deuten, dass konkrete Merkmale der Lernumgebung (die bei Grube maßgeblich von der jeweiligen Lehrkraft beeinflusst werden) einen stärkeren Einfluss auf die Prozesse der Kompetenzentwicklung haben als die Zugehörigkeit zu einem Projekt, welches die Lernumgebung auf einem weiter entfernten Niveau rahmt. Für Prozesse des Denkens und Handelns beim Kompetenzaufbau lässt sich vermuten, dass diese ähnlich wie oder noch stärker als Prozesse der Kompetenzentwicklung von der Fokussierung der Lehrkraft oder spezifischen Anlage der Instruktion abhängen.

2.4.2 Förderung von Fähigkeiten und Verständnissen – Information, Praxis und Reflexion als (Design-)Merkmale von Instruktionen

Zur Förderung des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten liegt vergleichsweise umfassende und vielfältige Literatur (meist mit Fokus auf Fähigkeiten) vor (zum Überblick z. B. Alfieri et al., 2011; Furtak et al., 2012; Lazonder & Harmsen, 2016; Rieß et al., 2012; Rönnebeck et al., 2016; Zimmerman, 2007). In vielen der entsprechenden Studien wird vorrangig auf die instruktionale Seite des Kompetenzaufbaus fokussiert; es wird danach gefragt, welche Art und Beschaffenheit der Lernumgebung Kompetenzveränderung von Lernenden wie stark fördert. Nicht aus allen dieser Studien lassen sich gleichermaßen Hinweise auf das Denken und Handeln der Lernenden und ihre Bedeutung für den Kompetenzaufbau seitens der Lernenden ableiten. In einigen Fällen finden sich jedoch spezifische Aussagen (Vermutungen, Hypothesen, Desiderata, ...) zu den Aktivitäten der Lernenden oder den Prozessen des Kompetenzaufbaus. Darüber hinaus ist es möglich aus den grundlegenden Überlegungen dazu, wie (Prozesse des) Kompetenzaufbau(s) zum experimentbezogenen Denken und Arbeiten gefördert werden kann (können), Vermutungen und Hypothesen zum Denken und Handeln der Lernenden in den Prozessen des Kompetenzaufbaus abzuleiten.

2.4 Kompetenzaufbau zum experimentbez. Denken u. Arbeiten: Befunde

In Studien zum Aufbau von Kompetenzen zum experimentbezogenen Denken und Arbeiten werden vielfältige Schwerpunkte gesetzt. In ihrem Review unterscheidet Zimmerman (2007) die Interventionen unter anderem danach, ob sie *instructional* oder *practice interventions* sind – das heißt, ob das Designmerkmal der in der Intervention eingesetzten Instruktion darin besteht, den Lernenden Fragen zu stellen, Hinweise oder Informationen zu geben, die ihnen beim besseren experimentbezogenen Denken und Arbeiten helfen sollen, oder darin, dass die Lernenden experimentbezogenes Denken und Arbeiten in einer gewissen Weise praktizieren und dadurch (ein)üben sollen. In ähnlicher Weise unterscheiden Rönnebeck et al. (2016) zwischen *immersion in practice* und *explicit instruction*.²⁴ Eine ähnliche, häufig getroffene, Unterscheidung ist die Unterscheidung von expliziter und impliziter Instruktion (z. B. Holliday, 2006), basierend darauf, ob die zu den Lernzielen gehörenden Konzepte instruktional expliziert werden oder nicht. Verschiedene Autoren nennen allerdings unterschiedliche Merkmale und Unterscheidungskriterien (Kalthoff et al., 2018; auch z. B. Holliday, 2006; Lazonder & Egberink, 2014). Wesentlich bei der Unterscheidung scheinen die Fragen zu sein, ob bzw. welche (und wie viele) Informationen den Lernenden gegeben werden. Beispielsweise sortieren Vorholzer und von Aufschnaiter (2019b) explizite und implizite Instruktionsansätze sowie minimal-informierende Instruktionsansätze in absteigender Reihenfolge auf einer Achse zum Grad der konzeptuellen Informationsmitteilung. Die Zugehörigkeit anderer Elemente (wie beispielsweise Übungsphasen) zu expliziter oder impliziter Instruktion wird unterschiedlich umgesetzt (siehe Kalthoff et al., 2018). Die getroffenen Unterscheidungen sind nicht trennscharf, sondern scheinen eher Merkmale von Instruktion zu beschreiben, die auch im Kontext der Förderung von Kompetenzen zum experimentbezogenen Denken und Arbeiten relevant sind. Manche Autoren plädieren dafür, Elemente aus jeweils beiden Seiten in einem gemeinsamen Instruktionsansatz zu integrieren (z. B. Holliday, 2006, der explizite und implizite Ansätze auf einem Kontinuum verortet).

Ein weiterer Schwerpunkt bei der Beschreibung von Instruktionen ist, inwieweit die Lernenden aufgefordert und unterstützt werden, das (eigene) Vorgehen zu reflektieren (z. B. Kalthoff et al., 2018; Rönnebeck et al., 2016). Der Einsatz von nachgelagerter Fokussierung auf die (eigenen) Prozesse des experimentbezogenen Denkens und Arbeitens ist im Kontext der Förderung von Kompetenzen zum experimentbezogenen Denken und Arbeiten seltener als im Kontext der Förderung von Kompetenzen zu

²⁴Sie tun dies nur für Instruktionen zum Argumentieren ausdrücklich (Rönnebeck et al., 2016, S. 164) und für Instruktionen zum Identifizieren von Forschungsfragen beiläufig (»these studies focus entirely on an immersion approach of participating in inquiry-type laboratory activities or research-like projects. Hence, apart from repeated practice, little is known about instructional activities to develop students' ability to identify research questions«, S. 169); für die anderen naturwissenschaftlichen Denk- und Arbeitsweisen wird diese Unterscheidung nur implizit aufgegriffen.

Nature of Science, von woher er vermutlich adaptiert ist (z. B. Abd-El-Khalick et al., 2004; Scharmann et al., 2005).

Anhand der aufgeführten Unterscheidungen lassen sich die drei nachfolgend genauer diskutierten Aspekte ableiten: Mitteilung von Informationen an die Lernenden, Umfang des Praktizierens von experimentbezogenem Denken und Arbeiten durch die Lernenden und Gelegenheiten für Lernende zum Reflektieren des Vorgehens des experimentbezogenen Denkens und Arbeitens. Dabei wird auch diskutiert, dass Instruktionen gewöhnlich nicht nur mit Blick auf eines der Merkmale konzipiert sind.

2.4.2.1 Informationen: Welche, wie viele, wann?

Instruktionale Strategien wie Feedback, Prompts, (Lösungs-)Beispiele oder Erklärungen können als Platzierung von Informationen im Laufe der Instruktion angesehen werden. In Anlehnung an Lazonder und Harmsen (2016) sowie Alfieri et al. (2011) lassen sich

- *Prozessüberwachungshilfen* (Welche Schritte der Aufgabenbearbeitung oder des experimentbezogenen Denkens und Arbeitens wurden vom Lernern schon ausgeführt?),
- *Prompts* (Welcher Schritt soll beim experimentbezogenen Denken und Arbeiten oder beim Bearbeiten der aktuellen Aufgabe als nächstes ausgeführt werden?),
- *Anleitungen* (Wie soll der nächste Schritt ausgeführt werden?),
- *Teillösungen* (Anbieten einer Teillösung zum nächsten Schritt, mit der weitergearbeitet werden soll) und
- *Feedback* (im engeren Sinne: Wie richtig/schnell/... war das, was vom Lernenden gemacht wurde?; im weiteren Sinne auch: Was kann besser gemacht werden und wie?) sowie
- *Erklärungen* (Was soll getan werden bzw. was ist richtig und warum?)

unterscheiden.²⁵ Für Kompetenzen aller Art (Kompetenzen zum experimentbezogenen Denken und Arbeiten eingeschlossen) ergeben sich – grob zusammengefasst²⁶ – für stärker anleitende (in der obigen Aufzählung später genannte) instruktionale Strategien größere Effekte

²⁵Bei Lazonder und Harmsen (2016) wird auch noch *Zergliederungen* (Zerlegen in Teilaufgaben oder Teilprozesse) aufgeführt. Dies ist als eine Vorstrukturierung anzusehen, die allerdings nicht notwendigerweise mit dem Geben von Informationen verbunden sein muss.

²⁶Eine Ungenauigkeit in der Zusammenfassung findet sich z. B. in Bezug auf Lösungsbeispiele (engl. worked examples), die bei Alfieri et al. (2011) im Mittel die höchste Wirksamkeit haben. Dies könnte daran liegen, dass es Worked Examples sowohl mit als auch ohne Erklärungen der Beispiellösung gibt und die rezipierten Studien hier auf verschiedene Art zurückgegriffen haben dürften.

2.4 Kompetenzaufbau zum experimentbez. Denken u. Arbeiten: Befunde

- auf Lernergebnisse (meist erlernte Fachinhalte, z. T. auch fachmethodische Fähigkeiten; gemessen an Posttests, meist im Vergleich zu einem Prätest; *learning outcomes* bei Lazonder & Harmsen, 2016, S. 691, Ergebnisse ab S. 701; *post-tests* bei Alfieri et al., 2011, S. 5, Ergebnisse ab S. 7),
- auf Angemessenheit der Aktivitäten im Lernprozess (Fähigkeiten der Lernenden im Lernprozess; *learning activities* bei Lazonder & Harmsen, 2016, S. 691, Ergebnisse ab S. 700; fällt vermutlich unter *aquisition scores* bei Alfieri et al., 2011, S. 5, Ergebnisse ab S. 7) und
- auf Angemessenheit der Produkte, die im Lernprozess erstellt werden (z. B. Anzahl fehlerhafter und/oder korrekter Versuchsergebnisse oder Aufgabenlösungen, Qualität einer Concept Map; *performance success* bei Lazonder & Harmsen, 2016, S. 691, Ergebnisse ab S. 699; *aquisition scores* bei Alfieri et al., 2011, S. 5, Ergebnisse ab S. 7).

Für Kompetenzen zum experimentbezogenen Denken und Arbeiten als Gesamtkonstrukt liegt keine spezifische zusammenfassende (Meta-)Studie zu Instruktionseffekten vor; am nächsten kommt der Aspekt der Angemessenheit der Aktivitäten bei Lazonder und Harmsen (2016), denn diese schreiben:

Learning activities for example, was defined as a measurement of skill and could therefore only pertain to inquiry skills and regulative skills. Likewise, measures of performance success could only target domain knowledge whereas learning outcomes could refer to all three types of outcome focus. (S. 696)

Es wird also bezüglich der Instruktionsbearbeitung (»learning activities« bzw. »measures of performance«) in gewisser Hinsicht eine Trennung von Fachinhalten (»inquiry skills«) und Fachmethoden (»domain knowledge«) vorgenommen. Das Ergebnis von Lazonder und Harmsen (2016), dass die Angemessenheit der Aktivitäten im Prozess positiv mit dem Grad der *guidance* zusammenhängt,²⁷ kann so interpretiert werden, dass stärker anleitende Instruktionen das korrekte experimentbezogene Denken und Arbeiten im Prozess besser fördern. Es bleibt aufgrund der nicht-vorgenommenen Trennung von Fachinhalten und Fachmethoden bezüglich der Lernergebnisse (*learning outcomes*) allerdings offen, inwiefern der Befund zu den Lernprozessen auch auf die Lernergebnisse zum experimentbezogenen Denken und Arbeiten im Spezifischen

²⁷Im Text heißt es: »[G]uided inquiry learning led to a more proficient use of inquiry skills than did unguided inquiry learning, $z = 5.85$, $p < .001$. The d statistic and its confidence interval show that the true effect size is likely to fall in the range of 0.44 to 0.88, which [...] can be considered a medium to large effect« (S. 699–700).

übertragbar ist, ob also stärker informierende und stärker anleitende Instruktion auch hinsichtlich des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten effektiver ist. Ferner fallen die für den Kompetenzaufbau neben den Fähigkeiten ebenfalls relevanten Verständnisse zum experimentbezogenen Denken und Arbeiten in der Klassifikation vermutlich nur unter die Lernergebnisse, da sie in keiner der Dimensionen aus den beiden Studien explizit inkludiert sind.

Der am stärksten beforschte Teilprozess des experimentbezogenen Denkens und Arbeitens ist das Planen und Durchführen von Untersuchungen, wobei insbesondere die Variablenkontrollstrategie und zugehörige Fähigkeiten und Konzepte fokussiert werden. Die Metastudien von Ross (1988) und Schwichow et al. (2016) operationalisieren explizite Instruktion zwar jeweils anders und kommen daher zu unterschiedlichen Schlüssen darüber, wie effektiv explizite Instruktionsansätze sind. Insgesamt legen die beiden Metastudien allerdings – in Übereinstimmung mit den Befunden aus dem vorigen Absatz – nahe, dass mehr Informationen, durch die das Lernen stärker angeleitet wird, für das Lernen hilfreicher sind als weniger Anleitung durch weniger Informationen. Auch für die anderen Teilprozesse des experimentbezogenen Denkens und Arbeitens liegen Hinweise vor, dass Informationen zum experimentbezogenen Denken und Arbeiten förderlich für entsprechende Kompetenzen sind (z. B. Vorholzer, 2016; Zimmerman, 2007).

Wann Informationen gegeben werden sollten, ist jedoch mit Blick auf verschiedene Effekte nicht hinreichend geklärt. Lazonder und Egberink (2014) berichten, dass die Angemessenheit der Experimentierprozesse im Verlauf der Instruktion gleichermaßen gefördert werden kann, »when children are taught the procedure of sound experimental design prior to their inquiry as when their use of this procedure is scaffolded [during inquiry; Ergänzung durch JJH]« (S. 301). Matlen und Klahr (2013, S. 622) führen verschiedene theoretische Argumente und empirische Befunde auf, die teilweise für das Geben von Informationen im Experimentierprozess, teilweise für das nachträgliche Informieren sprechen, wenn Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten (insbesondere zur Variablenkontrolle) stattfinden soll. In ihrer eigenen Studie finden sie ähnlich wie Lazonder und Egberink, »that children learn and transfer the CVS [control-of-variables strategy, dt. Variablenkontrollstrategie; JJH] relatively well regardless of whether they receive direct instruction before or after practice« 2014, S. 301, solange sie nur »received high guidance at some point during instruction« (Matlen & Klahr, 2013, S. 631). In Bezug auf Informationen im Verlauf der Instruktion finden Wilhelm und Beishuizen (2004, zitiert nach Zimmerman, 2007) Effekte im Prozess, aber keine Effekte im Posttest, so dass die Frage, wann Informationen gegeben werden sollten, auch auf das verfolgte Ziel

bezogen werden sollte (z. B. Unterstützung von fachinhaltlichem Lernen durch korrekte Experimentierprozesse und -ergebnisse oder Förderung von fachmethodischen Lernen; vgl. z. B. Vorholzer & von Aufschnaiter, 2019b, oder auch Ausführungen zu produktiven Fehlern wie bei Lazonder & Egberink, 2014; Matlen & Klahr, 2013). Arnold et al. (2014) untersuchen die Schwierigkeiten von älteren Schülerinnen und Schülern während der Prozesse des Kompetenzaufbaus und stellen fest, dass trotz Informationsphase zu Begriffen und Strategien des experimentbezogenen Denkens und Arbeitens im Vorhinein in den Prozessen noch häufig Schwierigkeiten auftreten. Sie folgern, dass zum angemessenen experimentbezogenen Denken und Arbeiten bearbeitungsbegleitende Erklärungen und Prompts erforderlich sein dürften, beispielsweise mit fortschreitender Zeit auf zunehmend höherem Niveau (S. 22). Die zitierte empirische Absicherung für diesen Vorschlag bezieht sich allerdings nur auf das Zunehmen der Informationsmenge und regelmäßige Prompts (Schmidt-Weigand et al., 2009), nicht aber auf die angesprochenen Niveaus. Ferner stellt sich vermutlich in den meisten Studien, in denen Informationen gegeben werden, die Frage, wie die Lernenden mit diesen Informationen umgehen. Als Hypothese für das schlechtere Abschneiden einer expliziten Instruktion führen Kalthoff et al. (2018) beispielsweise an, dass die entsprechenden Lernenden »have read through the explicit explanations of concepts like CVS but have not reflected upon or applied them« (S. 1320). dies könnte eventuell damit zusammenhängen, dass die Informationen zu ungünstigen Zeitpunkten gegeben wurden.

Neben den Fragen, wie viele und wann Informationen gegeben werden sollten, wird auch untersucht, welche Inhalte die Informationen haben sollten, um verschiedene Effekte (Lernergebnisse, Angemessenheit der Aktivitäten im Lernprozess, Angemessenheit der Produkte aus dem Lernprozess) besonders gut zu ermöglichen. Dabei werden u. a. (1) Informationen über die einzusetzenden (experimentellen) Strategien selbst sowie (2) Informationen darüber, wie eine angemessene Strategie auszuwählen ist, (3) Informationen über die Fachinhalte, die der Instruktion zugrundeliegen sowie (4) Informationen über Ziele des experimentbezogenen Denkens und Arbeitens unterschieden (Klahr, 1999; Vorholzer & von Aufschnaiter, 2019b; Zimmerman & Klahr, 2018; z. B. ad 1: Kalthoff, 2019, die Informationen zum richtigen Notieren von Messwerten gibt, ad 2: Kuhn et al., 2000, die die Auswahl der Strategien betonen). Je nach Instruktion und zugehörigen Aufgaben können beispielsweise Informationen über zugehörige Fachinhalte Einfluss auf Kompetenzen zum experimentbezogenen Denken und Arbeiten haben. Chinn und Malhotra (2002) finden beispielsweise, dass fachinhaltliche Erklärungen von unerwarteten Ergebnissen vor dem experimentbezogenen Denken und Arbeiten den Lernenden helfen, ihre Beobachtungen auf die

relevanten Aspekte zu lenken und die tatsächlichen Geschehnisse zu beobachten (siehe auch Kuhn & Ho, 1980). Wenn es um Informationen zum experimentbezogenen Denken und Arbeiten geht, wird in der Forschungsliteratur allerdings zumeist nicht differenziert dargestellt, auf welche experimentbezogenen Konzepte die gegebenen Informationen sich beziehen (siehe Vorholzer, 2016; Vorholzer & von Aufschnaiter, 2019b). Vor dem Hintergrund von Unterabschnitt 2.2.2 (zur Rolle von Verständnissen für Kompetenzen zum experimentbezogenen Denken und Arbeiten) dürfte es für Forschung zum Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten z. B. interessant sein, verstärkt auf die einzelnen Konzepte zum experimentbezogenen Denken und Arbeiten zu achten und deren Bedeutung in Informationen durch die Instruktion sowie in den Verständnissen der Lernenden zu untersuchen.

2.4.2.2 Praxis: Explorieren, Üben und Wiederholen

Zur Förderung von Kompetenzen zum experimentbezogenen Denken und Arbeiten werden auch Instruktionsansätze eingesetzt, deren wesentliches Merkmal im wiederholten Praktizieren von experimentbezogenem Denken und Arbeiten liegt. Zum Teil werden diese Instruktionsansätze als Vergleichs- oder Kontrollgruppen in Studien zu Informationen (s. o.) genutzt. Gibt es eine weitere Gruppe, in der keine Beschäftigung mit experimentbezogenem Denken und Arbeiten stattfindet, zeigt sich für gewöhnlich, dass die Praxis-Gruppe besser als diese abschneidet (z. B. Dean & Kuhn, 2007; Lazonder & Egberink, 2014; Schalk et al., 2019). Studien, die auf die Praxis fokussieren, wie beispielsweise Studien zum *microgenetic change* von Kuhn und anderen legen die Annahme nahe, dass Kompetenzaufbau dadurch zustandekommt, dass neue Strategien durch Praxis gefestigt werden und irgendwann die alten Strategien vollständig verdrängen (Kuhn, 1995). In derartigen Studien findet sich, dass die Kompetenzveränderung über die einzelnen Messzeitpunkte hinweg, die zugleich die Praxisphasen sind, eher graduell als abrupt erfolgen (z. B. Kuhn & Phelps, 1982; zusammengefasst in Zimmerman, 2007). Das gilt beispielsweise für die Variablenkontrollstrategie sogar dann, wenn sie schon einmal komplett entdeckt wurde – sie wird dann nicht ab sofort exklusiv genutzt, sondern erst graduell gefestigt. Insgesamt lassen sich die zugehörigen Befunde laut Kuhn und Dean (2005) so zusammenfassen, dass »Children in the preadolescent age range, however, have also been shown to develop inquiry skills without direct instruction if they engage in repeated encounters with situations that require these skills« (S. 866).

In einer Studie von Hofstein et al. (2005) finden sich Effekte der intensiven Praxis des Planens und Durchführens auf Kompetenzen zum Formulieren von Fragen in einer Posttestung. Dies zeigt, dass die Teilprozesse des Experimentierens zwar sehr wohl

didaktisch sinnvoll voneinander unterschieden werden können, allerdings empirisch nicht in jeglicher Hinsicht trennbar sein müssen.

Tempel et al. (2018) finden, dass mehrfaches Testen zur Variablenkontrolle die zugehörigen Fähigkeiten fördert. Da die Schülerinnen und Schüler bei der Testbearbeitung Teile des experimentbezogenen Denkens und Arbeitens vollziehen, liegt in gewisser Hinsicht Praxis vor. Auch ohne Informationen im instruktionalen Sinne kann also Kompetenzaufbau stattfinden.

Für gewöhnlich finden sich allerdings Studien, in denen die Wichtigkeit von Praxis gegenüber rein informierenden Instruktionsansätzen hervorgehoben wird. In vielen Fällen wird die Praxis der Lernenden betont, gleichzeitig werden aber Informationen in Formen von bspw. Prompts oder Begriffstrainings während oder vor dem experimentbezogenen Denken und Arbeiten gegeben (das ist beispielsweise eine Kritik von Klahr & Carver, 1995, an Kuhn & Phelps, 1982, wie Zimmerman, 2007, berichtet). Diese Informationen sind weniger umfangreich als in Studien, in denen die Wichtigkeit expliziter Instruktion hervorgehoben wird, aber es sind doch Informationen. An manchen Stellen wird argumentiert, dass die Wirksamkeit von Praxis sich erst durch die Verbindung mit den gegebenen Informationen erklären lässt, weil diese beispielsweise eine Fokussierung der Lernenden auf die angestrebten Kompetenzen herstellen (z. B. Kuhn et al., 2017; Lorch et al., 2010). Ein anderes Argument ist, dass Praxis und Informationen unterschiedliche Aspekte zum Kompetenzaufbau beitragen, die erst gemeinsam zum Vorhandensein von Kompetenz führen (Zohar & David, 2008). In Studien, die Erklärungen, Lösungsbeispiele und Anleitungen mit Praxis verbinden, schneiden Vergleichsgruppen, in denen keine Informationen gegeben werden, in Erhebungen nach der Instruktionsbearbeitung schlechter ab (z. B. Keselman, 2003; Lorch et al., 2010; Vorholzer, 2016, sowie van Harmelen, 2017, bei dem allerdings die Effekte *während* der Instruktionsbearbeitung für *task segmentation* größer sind als für *direct instruction*).

Aus den oben bereits zitierten Studien zur Reihenfolge von Information und Praxis wird deutlich, dass es auch bei gedanklicher Fokussierung auf Praxis dasselbe Problem gibt. Begrifflich lässt es sich in der Frage fassen, ob die Praxisphase in einem Instruktionsansatz als *Exploration* (Erkundung ohne voriges Erhalten von Informationen), als *Üben* (direktes Beziehen der erhaltenen Informationen auf konkrete Aufgaben oder Situationen) oder als *Wiederholen* (Praxis, in der es erforderlich ist, frühere Informationen zu erinnern und zu notieren oder anzuwenden) angelegt ist. Für eine Untersuchung zu Prozessen des Kompetenzaufbaus stellt sich dann beispielsweise die Frage nach der unterschiedlichen Ausgestaltung der Praxis durch die Lernenden in den verschiedenen Phasen (d. h. u. a. in Abhängigkeit von vorhandenen Informationen).

2.4.2.3 *Reflexion: Die eigene Praxis in den Blick nehmen*

Ein weiteres Element, dessen Wirkung und Effektivität für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten diskutiert wird, ist Reflexion, im Sinne von retrospektivem Nachdenken über das eigene experimentbezogene Denken und über das eigene Vorgehen beim experimentellen Arbeiten. Generell lässt sich als Befund anführen, dass Reflexionen Kompetenzen zum experimentbezogenen Denken und Arbeiten fördern (z. B. Etkina et al., 2010; Pedaste et al., 2012; White et al., 1999). Da Reflexion nicht der Schwerpunkt der Arbeit ist, sei hier nur noch angeführt, dass Reflexion spontan auftreten und vollständig von den Lernenden gesteuert sein kann, aber auch durch Feedback oder Reflexionsfragen angeregt werden kann. In gewisser Weise lässt sich eine Anregung zum Reflektieren des eigenen experimentbezogenen Denkens und Arbeitens als eine spezielle Information auffassen. So legt beispielsweise der Untertitel von Lin und Lehman (1999) »Effects of Prompting College Students to Reflect on Their Own Thinking« die Verbindung zu oben diskutierten Informationen nahe. Auch Feedback, welches nachgelagert und auf das experimentbezogene Denken und Arbeiten bezogen ist (z. B. Wollenschläger, 2012), kann sowohl als Information als auch als Anregung zur Reflexion aufgefasst werden.

Als Ziel beim Anregen zur Reflexion kann angesehen werden, die Lernenden auf das eigene experimentbezogene Denken und Arbeiten aufmerksam zu machen und ihre Aufmerksamkeit auf diesen Gegenstand zu richten. Auch für einige der oben genannten Arten von Informationen (z. B. Erklärungen, Feedback, Anleitungen) kann dies als Ziel angesehen werden (vgl. Ellis, 2009, zitiert bei Kalthoff et al., 2018; vergleiche auch die Beschreibungen von impliziter Instruktion in der Literatur, z. B. bei Vorholzer, 2016, in denen nahegelegt wird, dass die Schülerinnen und Schüler nicht über Konzepte zum Experimentieren nachdenken müssen oder sollen, sondern über andere Inhalte). Ebenfalls kann als ein Mittel zur Aufmerksamkeitssteuerung angesehen werden, wenn Lernende aufgefordert werden, ihr experimentbezogenes Denken und Arbeiten zu rechtfertigen oder begründen. Beispielsweise vermuten Kuhn et al. (2017) dies als wesentlichen Katalysator für ihre Befunde bei einem sonst nur auf Praxis basierenden Instruktionsansatz, dessen Förderung deutlich besser als die des regelunterrichteten Kontrollansatzes war. Das damit verbundene bessere Lernen könnte auch damit zusammenhängen, dass eine Externalisierung sonst versteckter oder unbewusster Vorgänge stattfindet (vgl. Kuhn et al., 2000). Für Prozesse des Kompetenzaufbaus stellt sich folglich unter anderem die Frage, ob und wie Lernende über (ihr eigenes) experimentbezogenes Denken und Arbeiten sprechen (reflektieren) und dann auch inwiefern sie von diesem spezifischen Reflektieren profitieren.

2.4.2.4 Verschiedene Personen

Ein hervorstechendes Ergebnis, welches zur differenzierten Darstellung der oben berichteten Befunde beiträgt und im Rahmen der Arbeit von Bedeutung ist, liegt in der Unterscheidung verschiedener Personengruppen bei Untersuchungen zur Förderung von Kompetenzen zum experimentbezogenen Denken und Arbeiten. In der Zusammenschau der Studien zeichnet sich ab, dass Lernende mit unterschiedlichen fachmethodischen Eingangsvoraussetzungen bei Kompetenzen zum experimentbezogenen Denken und Arbeiten unterschiedlich stark von Informationen (oder auch Praxis) profitieren.

In der Tendenz findet sich bei der Betrachtung von Kompetenzen zum experimentbezogenen Denken und Arbeiten als ein Ziel der Förderung, dass Schülerinnen und Schüler mit niedrigeren Eingangsvoraussetzungen in Bezug auf das experimentbezogene Denken und Arbeiten (Fähigkeiten, Vorwissen/Verständnis, ...) mehr von stärker auf Informationen basierenden Instruktionsansätzen profitieren als solche mit höheren Eingangsvoraussetzungen (z. B. Lorch et al., 2010; White et al., 1999; Zohar & Peled, 2008). Für die Einordnung dieses Befunds relevant ist, dass in Übereinstimmung mit Aussagen zum allgemeinen Lernen (z. B. Ausubel, 1968) auch beim experimentbezogenen Denken und Arbeiten grundsätzlich gilt, dass Lernende mit mehr Vorwissen im Mittel von allen Instruktionen stärker profitieren (ergänzend zu eben zitierten Quellen z. B. Blanchard et al., 2010, für stärker auf Praxis orientierte Instruktion). Zusammengefasst sind die Befunde bezüglich der durch die Instruktion gegebenen Informationen demnach so zu verstehen, dass die Spanne zwischen Lernenden mit höheren vs. niedrigeren fachmethodischen Eingangsvoraussetzungen in Bezug auf Kompetenzen zum experimentbezogenen Denken und Arbeiten verkleinert wird (verstreut gibt es auch konträre Befunde z. B. Koenen, 2014).

In einigen Fällen finden sich für die stärker informierenden expliziten Instruktionen nahezu keine Unterschiede für Lernende mit unterschiedlichen Eingangsvoraussetzungen, während sich für weniger oder gar nicht informierende implizite Instruktionsansätze bei Lernenden mit schlechten Eingangsvoraussetzungen nur sehr geringe Kompetenzzuwächse finden (Lorch et al., 2010; Zohar & David, 2008; und Zohar & Peled, 2008, die zusätzlich zeigen, dass Lernende mit schlechten Eingangsvoraussetzungen zwar am Ende dieselben Kompetenzen erreichen, aber diese im Prozess langsamer aufbauen). Ferner stellen Klahr und Nigam (2004) anknüpfend an ihre Studie die Frage »whether there are specific features of some learners that render discovery learning effective« (S. 666). Dies ist vermutlich zunächst auf andere Per-

sonenmerkmale (wie fachinhaltliches Vorwissen, Motivation, ...) bezogen,²⁸ kann aber auch so gelesen werden, dass Aktivitäten der Lernenden aus den Prozessen des Kompetenzaufbaus möglicherweise solche spezifischen Features sind. Für Prozesse des Denkens und Handelns beim Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten stellen sich davon ausgehend die Fragen, wie sich Lernende mit höheren vs. niedrigeren Eingangsvoraussetzungen voneinander unterscheiden und ob es spezifische Elemente in den Prozessen des Kompetenzaufbaus gibt, die höhere Lernzuwächse von Personen erklären können. Außerdem ist mit Bezug auf die Kompetenzveränderung zu fragen, inwiefern das Auftreten dieser Elemente mit den Eingangsvoraussetzungen der Lernenden zusammenhängt.

2.4.3 Befunde zu Verständnissen zum experimentbezogenen Denken und Arbeiten

Verständnisse von Personen umfassen in der vorliegenden Arbeit ein Kontinuum von einfachem Faktenwissen bis hin zu Vorstellungen, die situationsspezifisch und vielschichtig sein können. Diese dispositionellen Größen haben gemeinsam, dass sie sich immer nur an einem Vollzug oder einem Produkt konkreter Handlungen einer Person festmachen lassen. Handlungen lassen demnach zunächst auf vorhandene Fähigkeiten und dann auch auf eventuell vorhandene Verständnisse schließen. Auch anhand anderer Fähigkeiten lassen sich Wissen und Vorstellungen von Personen rekonstruieren (dazu mehr in Kapitel 6). Das heißt allerdings nicht, dass für jede Fähigkeit, die sich aus Beobachtersicht mit vorhandenem Verständnis erklären lässt, die Person bereits über die zugehörigen Verständnisse verfügt.

Gemäß der Begriffsfassung aus Abschnitt 2.2 beziehen sich nur wenige Studien darauf, über welche prozessbezogenen Verständnisse Personen zum experimentbezogenen Denken und Arbeiten verfügen (Übersichten insgesamt z. B. bei Hammann et al., 2006; Stiller, 2015; siehe u. a. auch Arnold et al., 2014; Boudreaux et al., 2008). Im Rahmen der Schülervorstellungsforschung wird von Vorstellungen gesprochen, was so aufgefasst werden kann, dass es um die tatsächlichen (teilweise situationsspezifischen) Verständnisse der Personen zum experimentbezogenen Denken und Arbeiten geht. Die Vorstellungsforschung hat sich bisher jedoch fast nur auf Fachinhalte bezogen und berichtet im Bereich des naturwissenschaftlichen Denkens und Arbeitens vorrangig

²⁸Diese sind ebenfalls nur zu sehr geringen Teilen untersucht. Da dies nicht im Fokus der Arbeit steht, sei beispielhaft nur genannt, dass sich im Bereich des Teilprozesses *Auswerten und Interpretieren von Daten* Ergebnisse dazu finden, dass fachinhaltliche Eingangsvoraussetzungen (Vorwissen oder angemessene oder unangemessene Vorstellungen) Einfluss darauf haben, wie Evidenz interpretiert wird (z. B. Amsel & Brock, 1996; Koslowski, 1996).

über Vorstellungen, die sich dem Bereich des Erkenntniswissens (engl. epistemic knowledge) zuordnen lassen (Übersichten in Hammann & Asshoff, 2014; Schecker et al., 2018).

2.4.3.1 Vorstellungen zum experimentbezogenen Denken und Arbeiten

Eine in der Literatur berichtete Vorstellung von Schülerinnen und Schülern zum experimentbezogenen Denken und Arbeiten ist, dass Experimentieren aus Ausprobieren und Herausfinden besteht. Sie sehen demnach das experimentbezogene Denken und Arbeiten als ein Sammeln von Fakten an, die dann bereits die naturwissenschaftliche Erkenntnis darstellen. Diese Vorstellung findet sich unter älteren Schülern seltener als unter jüngeren (Höttecke & Hopf, 2018). In ähnlicher Weise wird darüber berichtet (Schauble et al., 1991; siehe auch Hammann et al., 2006), dass Schülerinnen und Schüler beim experimentbezogenen Denken und Arbeiten vor allem Effekte erzeugen wollen (anstatt zu betonen, dass sie einer Frage nachgehen oder eine Hypothese prüfen).

Heinicke berichtet an verschiedenen Stellen über Vorstellungen zum Messen (Heinicke, 2012; Heinicke & Heering, 2013). Zum einen berichtet sie von einem Modell, in dem die Vorstellungen der Lernenden in zwei Paradigmen eingeordnet werden: Point-Paradigma und Set-Paradigma. Demnach werden (hier vereinfacht notiert) Messwerte manchmal als voneinander unabhängige punktgenaue Zahlenwerte, manchmal als Näherungswerte für ein Werteintervall angesehen. Heinicke und Heering (2013) zitieren eine Studie von Allie et al. (1998), in der zwei Drittel der befragten Studienanfänger sich konsistent einem Paradigma zuordnen ließen. In einer ausführlichen Kritik des Modells spricht sich Heinicke vor dem Hintergrund historischer Überlegungen und empirischer Befunde dafür aus, Lernenden beide Paradigmen als Optionen darzulegen, die sich in verschiedenen Situationen in besonderer Weise bewähren, auch wenn in der Scientific Community für gewöhnlich das Set-Paradigma als das angemessenere angesehen wird (Heinicke, 2012, S. 475–477). Diese Studie zeigt zum einen, dass die Situatitivität von Vorstellungen berücksichtigt werden sollte, zum anderen, dass auch fortgeschrittene Lernende noch Probleme mit gewissen Vorstellungen zum experimentbezogenen Denken und Arbeiten haben.

Auch Studien zum Wissenschaftsverständnis (engl. nature of science) oder zum Erkenntniswissen (engl. epistemic knowledge) enthalten gelegentlich Hinweise auf (prozessbezogene) Vorstellungen zum experimentbezogenen Denken und Arbeiten. Vorrangig umfassen diese Studien allerdings Aspekte, die in der vorliegenden Arbeit nicht betrachtet werden, und die Verallgemeinerungen aus diesen Studien können

nicht ohne Weiteres als Befunde zu den hier in den Blick genommenen Vorstellungen zum experimentbezogenen Denken und Arbeiten aufgefasst werden. Eine Analyse der Teilbefunde erweist sich allerdings als aufwändig und scheint im Rahmen der Arbeit nicht notwendig. Nur in Bezug auf das Erkennen und Formulieren naturwissenschaftlicher Fragen und die Unterscheidung von Beobachtung und Deutung im Rahmen der Auswertung und Interpretation experimenteller Daten (siehe bereits Tabelle 4.2), die Gegenstände von Analysen in Kapitel 6 der vorliegenden Arbeit sind, werden nachfolgend einige Befunde dargestellt.

2.4.3.2 Vorstellungen zu naturwissenschaftlichen Fragen

Zu Vorstellungen zum Formulieren von Fragen kann beispielsweise auf die Studie von Grygier (2008) zurückgegriffen werden. Sie führt eine Interviewstudie mit Grundschülerinnen und Grundschülern der vierten Klasse durch und findet, dass diese in der Lage sind, zu beschreiben, dass naturwissenschaftliche Fragen grundsätzlich mit einem Experiment beantwortet werden können (z. B. S. 192, 194), auch wenn die entsprechenden Niveaus eher selten erreicht werden.

Bei Leach et al. (1996) wurden Schülerinnen und Schüler im Alter von 9, 12 und 16 Jahren vielfältige Fragen gegeben, für die sie entscheiden sollten, ob Naturwissenschaftlerinnen und Naturwissenschaftler (unabhängig davon, ob sie männlich oder weiblich sind) Interesse hätten, mehr darüber herauszufinden. Die Fragen wurden im Vorhinein so formuliert, dass Schülerinnen und Schüler als Argumente für ihre Entscheidungen auf verschiedene Aspekte rekurren konnten, wie beispielsweise die Art des Phänomens, auf das sich die Frage bezieht, auf den Bezug zum schulischen Naturwissenschaftsunterricht oder darauf, ob der Frage mit empirischen Untersuchungen nachgegangen werden kann. Die Antworten wurden zunächst danach kodiert, wie die Frage eingeschätzt wurde (interessant vs. uninteressant für Naturwissenschaftlerinnen und Naturwissenschaftler). Danach wurden innerhalb jeder Frage ähnliche Begründungen gruppiert. Zuletzt wurde über die Fragen hinweg untersucht, welche Gemeinsamkeiten und Unterschiede die Schülerinnen und Schüler erkannt haben und wie konsistent sie mit den von ihnen genutzten Aspekten (Art des Phänomens, Schulbezug, empirische Prüfbarkeit, ...) begründet haben.

Wenngleich die Frage, wofür sich Naturwissenschaftlerinnen und Naturwissenschaftler interessieren, sich aufgrund des Interessensbezugs nicht vollständig dem Prozesswissen zuordnen lässt, geben die Befunde von Leach et al. (1996) Aufschluss über Vorstellungen der Lernenden zu Kennzeichen von naturwissenschaftlichen Fra-

gen. Die drei zentralen Kategorien, in die die Begründungen der Schülerinnen und Schüler tatsächlich fielen, bezogen sich auf

- whether the question is open to investigation
- the domain of the phenomenon in question and
- the personal and institutional characteristics and interests of scientists (Leach et al., 1996, S. 118).

Die letzte Kategorie scheint den geringsten Bezug zum Prozesswissen aufzuweisen. Für jede Frage werden in dem Artikel die Anteile nach den drei Altersgruppen aufgelöst. Die Muster für verschiedene Fragen unterscheiden sich teilweise deutlich voneinander. Auch wurde über alle Fragen hinweg der Anteil für die drei zentralen Kategorien innerhalb der jeweiligen Altersgruppe an allen vergebenen Kategorien berichtet (Abbildung 2.1). Die Erwähnung der empirischen Prüfbarkeit ist bei 16-Jährigen prozentual häufiger als bei Zwölfjährigen, für die wiederum der Anteil höher ist als für Neunjährige. Für den Bezug zur Fachdisziplin (engl. domain) ist zunächst ein starker Anstieg und dann ein leichter Abfall des relativen Anteils verzeichnet. In Bezug auf die letzte Kategorie finden sich keine Unterschiede für die Altersgruppen, ein Gesamtanteil wird im Originaltext nicht angegeben.

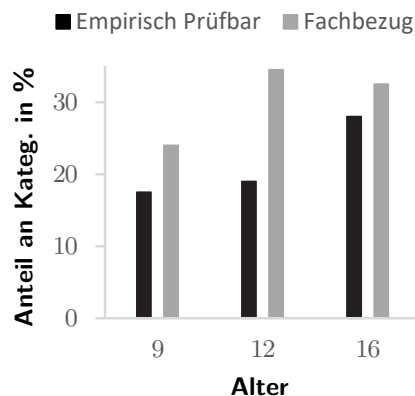


Abbildung 2.1: Befunde zu zwei Kategorien für Begründungen, die Schülerinnen und Schüler verschiedener Altersstufen für die Einschätzung von Fragen gegeben haben (Werte entnommen aus Millar et al., 1996).

Innerhalb der Kategorien von Leach et al. (1996) kann nach Subkategorien aufgelöst werden, für die sich teilweise deutlich voneinander verschiedene Muster über die Altersgruppen hinweg ergeben (hier nicht dargestellt). Ferner geben eine nicht zu vernachlässigende Anzahl von Schülerinnen und Schüler im Alter von 9 Jahren zu ihren

Einschätzungen keine Begründungen an. Laut Leach et al. (1996) scheint dies teilweise mit tautologischen Antworten (z. B. »Die Frage ist gut [d. h. Naturwissenschaftlerinnen und -wissenschaftler haben Interesse, der Frage nachzugehen], weil Wissenschaftler das interessant finden«) von Neunjährigen zusammenzuhängen.

2.4.3.3 Vorstellungen zu Beobachtungen und Deutungen

Für eine ausführliche Interviewstudie führt Petermann (2017) Befunde zu Schüler-
vorstellungen zum Beobachten und Deuten zusammen. Ihre Zusammenstellung wird
in Tabelle 2.1 abgebildet. Die Zuordnung in die Spalten zu Beobachtungen und
Deutungen lässt sich zwar hinterfragen, dennoch bildet die Übersicht den Stand der
Literatur gut ab.

Sowohl die dargestellte Literaturübersicht als auch die Ergebnisse der Arbeit
von Petermann (2017) lassen darauf schließen, dass sich das Feld zum Beobachten
und Deuten in vielfältiger Weise mit Gebieten wie beispielsweise den epistemolo-

Tabelle 2.1: Vorstellungen zum Beobachten und Deuten – Zusammenschau der Lite-
ratur entnommen aus Petermann (2017, S. 16–17).

Beobachtungen	Deutungen
Eine Sammlung von gewissenhaft erhobenen Daten führt zu sicherem Wissen. (1, 4, 5, 7)	Forscher interpretieren Ergebnisse neutral und objektiv. Auswertungen sind nicht durch Vorkenntnisse und Erwartungen beeinflusst. (1, 2, 5, 7, 8)
Eine Beobachtung (oder ein Experiment) ist ein Ausgangspunkt des Erkenntnisweges in den Naturwissenschaften. (3) [Dies stellt eine anknüpfungsfähige Vorstellung zu Beobachtungen dar.]	Ergebnisse aus Experimenten lassen sich verschieden interpretieren. (2,8) [Dies stellt eine angemessene Vorstellung zu Deutungen dar.]
Daten stellen offensichtliche Fakten dar. (6)	Eine Erklärung ist eine Veranschaulichung [und kein Rückgriff auf ein Aussagensystem]. (2)
Ergebnisse naturwissenschaftlicher Forschung sind in der Natur vorgebildet. Die Natur ist vor ihrer wissenschaftlichen Erfassung organisiert und ontologisch real. (3)	
Daten sind das Ziel einer Untersuchung und die Antwort auf die Untersuchungsfrage. (6)	
Daten sprechen für sich selbst. (9)	
Naturwissenschaftliche Ergebnisse werden grundsätzlich auf Richtigkeit geprüft. (7)	

Anmerkung. Der Inhalt ist aus Petermann (2017) entnommen, auch Text in eckigen Klammern. Literaturverweise in runden Klammern beziehen sich auf:

¹ Ertl, 2010; ² Höttecke, 2001a; ³ Höttecke, 2001b; ⁴ Kruse & Denz, 2016; ⁵ McComas, 1998; ⁶ McNeill & Berland, 2017; ⁷ Priemer, 2006; ⁸ Schaake, 2011; ⁹ Villanueva & Hand, 2011.

2.4 Kompetenzaufbau zum experimentbez. Denken u. Arbeiten: Befunde

gischen Überzeugungen oder auch dem Prozesswissen zum experimentbezogenen Denken und Arbeiten überschneidet, wie Petermann auch selbst erwähnt (2017, z. B. S. 13). Verständnisse zum experimentbezogenen Denken und Arbeiten, die sich dem Prozesswissen zuordnen lassen und sich auf das Beobachten und Deuten beziehen, umfassen beispielsweise die Unterscheidung von Beobachtungen und Deutungen. Für den Prozess des Auswertens und Interpretierens ist beispielsweise das Wissen darüber, *warum* diese Unterscheidung getroffen werden muss, dem Erkenntniswissen zuzuordnen und daher nicht im Fokus der vorliegenden Arbeit. Verständnisse dazu, *dass* Beobachtungen und Deutungen unterschieden werden müssen und *wie* sich Beobachtungen und Deutungen unterscheiden, lassen sich dem Prozesswissen zuordnen. In der Literatur finden sich scheinbar nur indirekte Hinweise auf Vorstellungen zu solchen prozessbezogenen Aspekten beim Beobachten und Deuten (vgl. Tabelle 2.1). Aus der empirischen Arbeit von Petermann (2017) lassen sich hingegen beispielsweise in Bezug auf die Unterscheidung von Beobachtung und Deutung die folgenden Vorstellungen extrahieren (siehe Abbildung 2.2).

Beobachtung = Wahrgenommenes: Beobachten ist das (fokussierte) Wahrnehmen/Sehen bzw. die Wiedergabe/Beschreibung dessen. Beobachten ist die Gewinnung von Daten (welche gegebenenfalls ganz objektiv und ohne Nachdenken passiert).

Deutung = Erklärung: Die Deutung ist die Erklärung/Begründung der Beobachtung bzw. dafür, warum etwas passiert.

Deutung = Verallgemeinerung: Deutungen sind Verallgemeinerungen (von Abhängigkeiten/ Zusammenhängen).

Deutung = Verknüpfung mit Bekanntem: Deutung ist die Verknüpfung Einordnung mit/in Bekanntem/Wissen/Rahmen.

Deutung = Interpretation: Deutungen sind Interpretationen/Schlussfolgerungen.

Deutung = Vermutung: Eine Deutung ist eine (begründete) Vermutung.

Beobachten = Deuten: Beobachtung und Deutung sind (fast) identisch/gleich.

Beobachtungen und Deutungen trennen, weil es verschiedene Prozesse sind.

Beobachtungen und Deutungen trennen, weil sich die Deutungen zu einer Beobachtung unterscheiden können.

Teilmenge: Beobachten ist in Deuten enthalten.

Reihenfolge: Nach der Beobachtung folgt die Deutung.

Wechselspiel: Beobachtung und Deutung hängen wechselseitig zusammen. Man kann von Beobachtungen zu Deutungen gelangen und umgekehrt.

Abbildung 2.2: Vorstellungen zum Unterscheiden von Beobachtungen und Deutungen (Zusammenstellung von Aussagen aus Petermann, 2017, Anhang B, ab S. 201).

3 Ziele und Forschungsfragen

Kompetenzen zum experimentbezogenen Denken und Arbeiten stellen eine wichtige Zieldimension von naturwissenschaftlichem Unterricht dar und sind (entsprechend) ein vielfältig untersuchter Gegenstand der naturwissenschaftsdidaktischen Forschung (siehe u. a. Kapitel 1). In der Forschung zu diesem Gegenstand finden sich viele punktuelle Kompetenzmessungen (siehe u. a. Kapitel 2). Diese können kaum Aussagen darüber generieren, wie die Lernenden die entsprechenden Kompetenzen aufgebaut haben bzw. in welcher Weise Instruktionen den Aufbau unterstützt haben – auch dann nicht, wenn für die punktuellen Messungen jeweils die Prozesse des Experimentierens betrachtet werden (Abschnitt 2.3, ab S. 21). Experimentell angelegte Vergleiche von Instruktionsansätzen liefern zwar Hypothesen zu relevanten Faktoren im Kompetenzaufbau; diese Hypothesen basieren allerdings nur in seltenen Fällen auf systematisch generierten Befunden zu den Prozessen der Bearbeitung der eingesetzten Instruktionen (vgl. Kapitel 1 sowie Unterabschnitt 2.4.2, ab S. 28).

Ein Ziel dieser auf das experimentbezogene Denken und Arbeiten gerichteten Arbeit besteht darin, basierend auf Analysen zum fachmethodischen Denken und Handeln der Lernenden Hinweise zu Wirkmechanismen von fachmethodischer Instruktion zu erlangen sowie erste Erkenntnisse zu Prozessen des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten zu generieren. Beispielsweise erscheint es vor dem Hintergrund der Forschung sinnvoll, zu untersuchen, in welchem Verhältnis unterschiedliche Aktivitäten zueinander und zum fachmethodischen Kompetenzaufbau stehen und ob dies für unterschiedliche Instruktionsansätze differiert, weil manche inhaltlichen bzw. kognitiven Aktivitäten für fachinhaltlichen Kompetenzaufbau relevant sind und sich für Fachmethoden u. a. die Frage nach der Relevanz experimenteller Aktivitäten für den Kompetenzaufbau aufdrängt. Aufgrund der in Kapitel 2 herausgearbeiteten Bedeutung der Personenspezifität in Analysen zum Kompetenzaufbau liegt es zudem nahe, interindividuelle Unterschiede und Gemeinsamkeiten im Denken und Handeln der einzelnen Lernenden in besonderer Weise in den Blick zu nehmen. Für die Herstellung von Bezügen zu Wirkmechanismen bietet es sich dabei unter anderem an, auf Merkmale von Instruktionen einzugehen, zu denen sich in Kapitel 2 umfassende Forschung diskutieren ließ; dazu gehören insbesondere die Frage nach der Bedeutung und Rolle instruktionaler Informationen im Kompetenzaufbau sowie die Frage nach der Relevanz der Fokussierung der Lernenden auf (ihr eigenes) experimentbezogenes Denken und Arbeiten. Ferner stellen sich Fragen nach der zeitlichen Veränderung bspw.

3 Ziele und Forschungsfragen

des sich in rekonstruierbaren Vorstellungen widerspiegelnden Denkens der Lernenden.

Ein weiteres Ziel der Arbeit besteht darin, einen Beitrag im Forschungsfeld der Verständnisse von Schülerinnen und Schülern zum experimentbezogenen Denken und Arbeiten zu leisten. Wie in Unterabschnitt 2.4.3 (ab S. 38) dargestellt, finden sich nur wenige Ergebnisse zu Verständnissen, die sich dem Bereich des Prozesswissens zuordnen lassen (der neben dem Inhaltswissen und dem Erkenntniswissen ein zentraler Zielbereich von naturwissenschaftlichem Unterricht ist; vgl. Unterabschnitt 2.2.2, ab S. 16). Zunächst kann der Beitrag der Arbeit also in einer differenzierteren Darstellung verschiedener Verständnisse, über die Lernende verfügen, bestehen. Ferner hat eine Untersuchung von Verständnissen, die auf Analysen von Bearbeitungsprozessen basiert, unter anderem den Vorteil, dass die Verständnisse mit ihren jeweiligen Aktualisierungen und Nutzungen in den Prozessen in Verbindung gebracht werden können. Da Verständnisse einen zentralen, allerdings wenig beleuchteten Bestandteil von Kompetenzen zum experimentbezogenen Denken und Arbeiten darstellen, wird insbesondere hierfür eine zeitliche Auflösung angestrebt, die als Grundlage für Hypothesen zu Prozessen des Kompetenzaufbaus dienen kann.

Für eine Erfassung des Denkens und Handelns von Lernenden im zeitlichen Verlauf bieten sich Videoaufzeichnungen von Lernenden bei der Arbeit in einer Lernumgebung an. Als Lernumgebung ist dabei für Analysen zum Denken und Handeln im zeitlichen Verlauf, die u. a. als Grundlage für Hypothesen zu Prozessen des Kompetenzaufbaus dienen sollen, eine Umgebung vorteilhaft, die Kompetenzveränderung im Bereich des experimentbezogenen Denkens und Arbeitens grundsätzlich möglich (besser noch: wahrscheinlich) macht. Basierend auf Befunden zur Effektivität von explizit-fachmethodischen Instruktionsansätzen (siehe Kapitel 1 und Kapitel 2) bietet sich beispielsweise eine Instruktion aus diesem Bereich an. Ferner liegt im Forschungsfeld zu expliziter Instruktion aufgrund der Kontrastierung mit impliziten Instruktionen ausreichend Literatur vor, vor deren Hintergrund in der Arbeit, basierend auf Analysen von Bearbeitungsprozessen, generierte Hypothesen zu Wirkmechanismen von Instruktion eingeordnet und diskutiert werden können. Daher wird in der vorliegenden Arbeit auf eine explizit-fachmethodische und eine implizit-fachmethodische Instruktionsvariante zurückgegriffen, die von Vorholzer (2016) für Schülerinnen und Schüler der elften Klassenstufe entwickelt wurden. Diese wurden bereits vor Beginn des die vorliegende Arbeit rahmenden Projekts²⁹ in einem Interventionsdesign mit Prä- und Post-Erhebungen verwendet; außerdem wurden Videoaufzeichnungen von Lernenden während der Bearbeitung der Instruktionsvarianten erstellt. Die für die vorliegende – vorrangig auf die Videoaufzeichnungen bezogene – Arbeit relevanten Details der zugrundeliegenden Studie sind in Kapitel 4 dargestellt.

²⁹DFG AU 155/11-1; Projektnummer 317314720.

3.1 Fragen zu Aktivitäten der Lernenden

In einem ersten Zugang zur Untersuchung des Denkens und Handelns der Lernenden bei der videografierten Bearbeitung der beiden untersuchten fachmethodischen Instruktionsvarianten wird der Frage nachgegangen, wie typische Aktivitäten und Bearbeitungsprozesse der Lernenden aussehen. So stellt sich beispielsweise die Frage, in welchem Umfang typische Aktivitäten (z.B. Lesen, Schreiben, auf Fachinhalte gerichtete Äußerungen, Off-Task etc.) bei der Bearbeitung auftreten. Analysen der Aktivitäten von Lernenden erlauben Rückschlüsse darauf, wie die Lernenden die Auseinandersetzung mit fachmethodischen Aufgaben (bewusst oder unbewusst) zeitlich rahmen und strukturieren. Sie erlauben zudem, zwischen fachinhaltlichen und fachmethodischen Fokussierungen und Orientierungen auf andere Inhaltsbereiche aus der Physik oder aus dem Alltag zu unterscheiden. In den Aktivitäten wird damit, zumindest in ersten Ansätzen, auch deutlich, welche Fähigkeiten zum experimentbezogenen Denken und Arbeiten bzw. Strategien der Aufgabenbearbeitung eingesetzt werden, ob also z.B. experimentbezogenes Denken und Arbeiten in den Blick genommen wird oder nicht. Unterschiede (und Gemeinsamkeiten) bezüglich der Bearbeitung der Instruktion zwischen verschiedenen Lernenden können Hinweise darauf liefern, welche Aktivitäten besonders relevant für den Kompetenzaufbau sind. Lernende können sich beispielsweise darin unterscheiden, wie hoch der erreichte Kompetenzzuwachs ist oder welche Instruktionsvariante sie bearbeiten.

Die Forschungsfragen sind auf die explizit-fachmethodische Instruktionsvariante (F-Akt1) sowie auf die implizit-fachmethodische Instruktionsvariante (F-Akt2) jeweils für sich genommen sowie auf den Vergleich der beiden Instruktionsvarianten (F-Akt3) gerichtet. Gemäß der vorgenommenen Betonung der Personenspezifität beinhalten die Forschungsfragen jeweils (zumindest implizit) auch die Frage danach, welche interindividuellen Unterschiede und Gemeinsamkeiten sich für die einzelnen Lernenden bezüglich der Aktivitäten finden lassen.

F-Akt1 Welche Unterschiede und Gemeinsamkeiten zeigen sich in den Aktivitäten von Lernenden – insbesondere mit hohen vs. niedrigen Kompetenzzuwächsen – bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante?

F-Akt2 Welche Unterschiede und Gemeinsamkeiten zeigen sich in den Aktivitäten von Lernenden – insbesondere mit hohen vs. niedrigen Kompetenzzuwächsen – bei der Bearbeitung der implizit-fachmethodischen Instruktionsvariante?

F-Akt3 Welche Unterschiede und Gemeinsamkeiten zeigen sich in den Aktivitäten von Lernenden, die unterschiedliche Instruktionsvarianten (explizit-fachmethodisch vs. implizit-fachmethodisch) bearbeiten?

3.2 Fragen zu Vorstellungen der Lernenden

Ein wesentlicher Teil der Arbeit ist auf die Untersuchung von Vorstellungen der Lernenden ausgerichtet. In diesem zweiten Zugang wird anhand der Sprechhandlungen der Lernenden auf das Denken der Lernenden geschlossen, genauer: anhand der Äußerungen mit fachmethodischem Gehalt wird auf die Vorstellungen der Lernenden geschlossen. Vorstellungen stellen gemäß Unterabschnitt 2.4.3 (ab S. 38) einen zentralen Teil des Denkens von Lernenden dar; in ihrer Vernetzung ergeben sich die Verständnisse der Lernenden. In den Analysen zu den Vorstellungen der Lernenden wird u. a. das Ziel verfolgt, aufzuschlüsseln, über welche fachmethodischen Inhalte die Lernenden bei der Bearbeitung der Instruktion sprechen und in welcher Weise sie dies tun.

Explizit-fachmethodische Instruktionsansätze zielen durch Informieren über die adressierten Konzepte und meist auch zusätzliche Arbeitsaufträge zu den Konzepten darauf ab, Lernende zum Denken und damit auch zum Sprechen über Fachmethoden anzuregen (siehe Kapitel 2). Aufgrund der Anlage der untersuchten Instruktionen (Details in Kapitel 4) ist davon auszugehen, dass die Lernenden bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante zeitlich insgesamt deutlich umfangreichere sowie inhaltlich umfassendere Äußerungen mit fachmethodischem Gehalt machen als bei der implizit-fachmethodischen Instruktionsvariante. Daher beziehen sich die Forschungsfragen zu den Vorstellungen der Lernenden zum experimentbezogenen Denken und Arbeiten in der vorliegenden Arbeit nur auf die ertragsversprechendere explizit-fachmethodische Instruktionsvariante. Gemäß der Schwerpunktsetzung der Arbeit (Details in den Kapiteln 2 und 6) werden exemplarisch Vorstellungen zu naturwissenschaftlichen Fragen und zum Beobachten und Deuten betrachtet. Zunächst stellt sich in einer allgemeinen Gesamtschau die Frage danach, welche Vorstellungen (in diesen beiden Bereichen) ganz grundsätzlich aufgefunden werden können:

F-Vor1 Welche Vorstellungen entwickeln Lernende zum experimentbezogenen Denken und Arbeiten während der Bearbeitung der explizit-fachmethodischen Instruktionsvariante?

Es kann davon ausgegangen werden, dass Vorstellungen nicht in einer statischen Form vorliegen, sondern Lernende ihre Vorstellungen bei der Artikulation (neu bzw.

erneut) entwickeln (z. B. von Aufschnaiter & von Aufschnaiter, 2003). Daher umfasst das Wort *entwickeln* in der Forschungsfrage sowohl das Aktivieren einer bereits vorliegenden Vorstellung für eine jeweilige Artikulation als auch das Generieren einer bisher nicht vorliegenden Vorstellung.

Hinweise darauf, wie die Lernenden die Vorstellungen bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante einsetzen, können u. a. daraus abgeleitet werden, wie generell bzw. allgemein die Äußerungen der Lernenden sind, aus denen auf ihre Vorstellungen geschlossen wird. Die Vorstellungen können dabei beispielsweise aus Äußerungen rekonstruiert werden, die losgelöst von konkreten Situationen oder auf eine Vielzahl von Situationen bezogen sind; sie können aber auch aus sehr spezifisch für eine konkrete Situation formulierten Äußerungen abgeleitet werden (vgl. u. a. von Aufschnaiter & Rogge, 2010a).

F-Vor2 Welche Vorstellungen zum experimentbezogenen Denken und Arbeiten sind bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante auf konkrete Situationen, welche auf Verallgemeinerungen bezogen?

Neben den Fragen nach den verschiedenen vorliegenden Vorstellungen der Lernenden und ihrer Allgemeinheit ist aus fachdidaktischer Perspektive auch interessant, wie angemessen diese Vorstellungen sind. Empirische Befunden deuten darauf hin, dass Lernenden das Formulieren von Fragen und Hypothesen in der Regel Schwierigkeiten bereitet (z. B. Grube, 2011). Es wird also zunächst davon ausgegangen, dass hier verstärkt unangemessene und wenig anschlussfähige Vorstellungen auftreten. Zum Beobachten und Deuten findet Petermann (2017) für Elftklässler, dass die Vorstellungen grundsätzlich angemessen oder zumindest anschlussfähig sind. Es wird daher davon ausgegangen, dass auch in den hier untersuchten Videoaufzeichnungen die Schülerinnen und Schüler vorrangig angemessene Vorstellungen vorweisen, weil sie ebenfalls Elftklässler sind. Anschließend an die dargestellten Befunde stellt sich auch die Frage, inwiefern unangemessene Vorstellungen und Schwierigkeiten beim experimentbezogenen Denken und Arbeiten bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante miteinander einhergehen bzw. ob angemessene Vorstellungen sich in korrekten Entscheidungen beim experimentbezogenen Denken und Arbeiten widerspiegeln (z. B. Entscheidungen dazu, ob eine Frage naturwissenschaftlich ist oder nicht).

3 Ziele und Forschungsfragen

F-Vor3 Wie angemessen sind die Vorstellungen der Lernenden zum experimentbezogenen Denken und Arbeiten?

F-Vor4 Mit welchen Vorstellungen gehen korrekte und fehlerhafte Entscheidungen zum experimentbezogenen Denken und Arbeiten einher?

Prozesse des Kompetenzaufbaus betreffen die Änderungen des Denkens und Handelns im zeitlichen Verlauf. Vorbereitend für die Generierung von Hypothesen zu Prozessen des Kompetenzaufbaus stellen sich unter anderem Fragen, die auf eine zeitliche Auflösung der Äußerungen, aus denen auf die Vorstellungen der Lernenden geschlossen wird, bezogen sind. Es ergibt sich beispielsweise die Frage, wie stabil bzw. variabel die Vorstellungen der Lernenden über die Situationen hinweg sind, weil damit Hinweise auf die Bedeutung der Stabilität von Vorstellungen in den Prozessen des Kompetenzaufbaus verbunden sind. Außerdem stellt sich im Kontext von explizit-fachmethodischer Instruktion die Frage, inwiefern die instruktional bereitgestellten Informationen über Konzepte des experimentbezogenen Denkens und Arbeitens mit den Vorstellungen der Lernenden zusammenhängen.

F-Vor5 Wie stabil bzw. variabel sind die Vorstellungen der Lernenden über die Situationen hinweg?

F-Vor6 Welche Unterschiede und Gemeinsamkeiten gibt es zwischen den Vorstellungen der Lernenden in Phasen vor und nach instruktionalen Informationen zu Konzepten des experimentbezogenen Denkens und Arbeitens?

Die detaillierten Analysen, die für die Beantwortungen der vorherigen Forschungsfragen zu den Vorstellungen der Lernenden notwendig sind, sollen in einem letzten Schritt mit den tatsächlich vorliegenden Kompetenzen sowie tatsächlich erreichten Kompetenzzuwächsen der Lernenden zum experimentbezogenen Denken und Arbeiten in Verbindung gebracht werden, um Fragen zum Zusammenhang von Vorstellungen während der Bearbeitung der explizit-fachmethodischen Instruktionsvariante mit dem Kompetenzaufbau zu beantworten. Zusammenfassend lässt sich die zu untersuchende Forschungsfrage so formulieren:

F-Vor7 Welche Unterschiede und Gemeinsamkeiten finden sich bezüglich der Forschungsfragen F-Vor1 – F-Vor6 für Lernende mit unterschiedlichen Kompetenzzuwächsen zum experimentbezogenen Denken und Arbeiten von der Prä- zur Post-Erhebung sowie für Lernende mit unterschiedlichen Eingangsvoraussetzungen?

4 Zugrundeliegendes Design

In der vorliegenden Arbeit wird das Denken und Handeln von Lernenden bei der Bearbeitung von fachmethodischer Instruktion in den Blick genommen. Dazu wird auf eine explizit-fachmethodische und eine implizit-fachmethodische Instruktionsvariante zurückgegriffen, die von Vorholzer (2016) entwickelt wurden. Die Instruktionsvarianten wurden bereits im Rahmen einer Interventionsstudie eingesetzt, die die Basis der vorliegenden Arbeit bildet. Die Struktur der der Arbeit zugrundeliegenden Studie ist (vereinfacht) in Abbildung 4.1 dargestellt: Die Instruktion zur Förderung von Kompetenzen zum experimentbezogenen Denken und Arbeiten wurde in einem mehrwöchigen Zeitraum eingesetzt (d. h., beide Instruktionsvarianten wurden eingesetzt) und von Prä- und Post-Erhebungen flankiert. Nach einer Darstellung der für die Instruktion maßgeblichen Fähigkeiten und Konzepte (Abschnitt 4.1) wird ausführlich auf die Konzeption der Instruktion eingegangen (Abschnitte 4.2, 4.3 und 4.4), weil das Denken und Handeln der Lernenden vor dem Hintergrund der eingesetzten Instruktion interpretiert wird. Abschnitt 4.5 enthält einen knappen einführenden Überblick zu den Prä-Post-Vergleichen der zugrundeliegenden Studie, weil auch die Testwerte der Lernenden eine Grundlage für die vorliegende Arbeit darstellen. Abschließend findet sich in Abschnitt 4.6 ein Überblick über die Videoaufzeichnungen von der Bearbeitung der Instruktion, die eine wesentliche Grundlage für die vorliegende Arbeit darstellen.

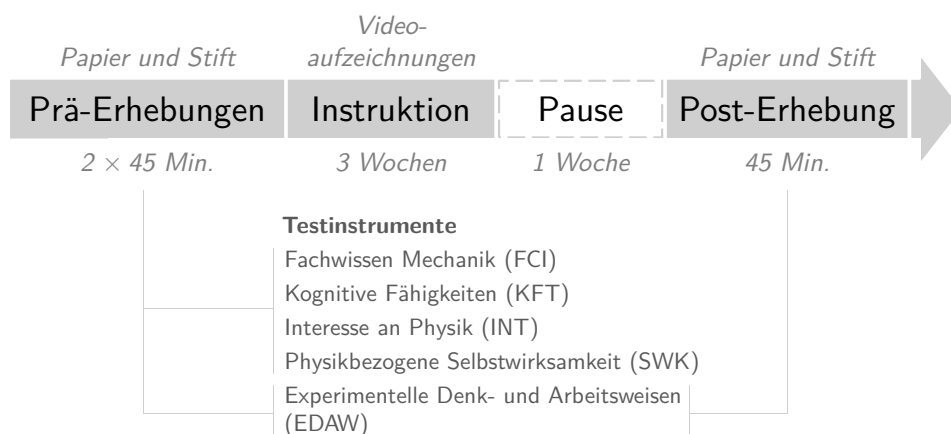


Abbildung 4.1: Überblick über die zeitliche Struktur der Erhebungen für die zugrundeliegende Studie.

4.1 Fokussierte Fähigkeiten und Konzepte

Die in der Studie, die der Arbeit zugrunde liegt, in den Blick genommenen Kompetenzen umfassen insbesondere *Fähigkeiten zum Formulieren von Fragen und Hypothesen (FH)*, *zum Planen und Durchführen von Untersuchungen (PU)* und *zum Auswerten und Interpretieren von experimentellen Daten (AI)*. Sie entstammen einer Zusammenfassung der Literatur durch Vorholzer (2016) und sind in Tabelle 4.1 (auf S. 53) zu finden. Die Reihung der Fähigkeiten innerhalb der Tabelle impliziert keine ordinale Skalierung im Sinne von Niveaustufen. Für Fähigkeiten, die auf denselben Kern bezogen sind und sich nur hinsichtlich des gewählten Operators unterscheiden (z. B. »eine inhaltlich passende Fragestellung identifizieren / zuordnen / formulieren«), ist der Übersichtlichkeit halber nur jeweils eine Formulierung in die Tabelle aufgenommen (Vorholzer bezeichnet diese als die »jeweils als am anspruchsvollsten angenommenen Formulierung«, 2016, S. 16). Die verschiedenen Möglichkeiten zur Wahl der Operatoren sind jedoch keineswegs unbedeutend für Analysen des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten. Es gibt beispielsweise Hinweise für Fähigkeiten zur Variablenkontrollstrategie, dass es einen Unterschied macht, ob Lernende ein kontrolliertes Experiment zu einer vorgegebenen Frage planen sollen oder sie die fehlende Aussagekraft von Experimenten, bei denen mehr als eine Variable verändert wurde, erkennen sollen (Bullock & Ziegler, 1999; Schwichow et al., 2016). Daher werden im Verlauf der Arbeit die verschiedenen genutzten Operatoren an den relevanten Stellen unterschieden, wenngleich in Tabelle 4.1 nur jeweils ein Operator genutzt wird.

Die in dem Projekt insgesamt in den Blick genommenen Kompetenzen umfassen neben den Fähigkeiten auch *Verständnisse zum Formulieren von Fragen und Hypothesen, zum Planen und Durchführen von Untersuchungen und zum Auswerten und Interpretieren von Daten*. Anhand von umfangreichen Zusammenstellungen wie beispielsweise bei Gott et al. (2015) lässt sich erkennen, dass für die Studie notwendigerweise eine Einschränkung auf wesentliche Konzepte vorgenommen werden musste. Eine Zusammenstellung der Konzepte, zu denen im Rahmen des Projekts Verständnisse aufgebaut werden sollten, wurde von Vorholzer (2016) vorgenommen und begründet. Dazu diente zum einen die Orientierung an den drei als zentral herausgearbeiteten Teilprozessen des experimentbezogenen Denkens und Arbeitens (FH, PU, AI); zum anderen wurde die Relevanz der Begriffe, Regeln und Strategien (also der Konzepte) für die einzelnen Teilprozesse und das experimentbezogene Denken und Arbeiten insgesamt beachtet (Vorholzer, 2016). Zudem liegt auch nicht immer Konsens in der Forschungslandschaft vor, wie Begriffe, Regeln und Strate-

gien für experimentbezogenes Denken und Arbeiten beschrieben werden sollen, so dass notwendigerweise eigene Formulierungen und Schwerpunktsetzungen vorgenommen wurden (Vorholzer, 2016). Tabelle 4.2 (auf S. 54) dient als Übersicht über die Konzepte, die der Arbeit zugrundeliegen. Im Wesentlichen sind die Überlegungen von Vorholzer (2016) abgebildet; zusätzlich wurde auf Analysen der von Vorholzer (2016) entwickelten Instruktionsvarianten durch Gabi (2018) zurückgegriffen, die für Untersuchungen mit direktem Bezug zur vorliegenden Arbeit erstellt wurden. Konzepte, die kursiv gesetzt wurden, stellen Ankerkonzepte für die im Gesamtprojekt eingesetzten Testungen und Instruktionsvarianten dar und wurden im Rahmen der Instruktion durch die anderen Konzepte ergänzt und unterstützt.

Tabelle 4.1: Fähigkeiten zu den drei Teilprozessen des experimentbezogenen Denkens und Arbeitens (Inhalt entnommen aus Vorholzer, 2016, Tabelle 2.2).

	Formulieren von Fragen und Hypothesen	Planen und Durchführen von Untersuchungen	Auswerten und Interpretieren von Daten
Die Schülerinnen oder Schüler ...			
Fähigkeiten	...erläutern den Unterschied zwischen naturwissenschaftlichen und nicht naturwissenschaftlichen Fragestellungen. ...formulieren eine naturwissenschaftliche Fragestellungen. ...formulieren eine inhaltlich^a zu einer Vermutung bzw. Hypothese passende Fragestellung (und umgekehrt).	...planen ein inhaltlich^a zu einer Fragestellung bzw. Vermutung bzw. Hypothese passendes Experiment ...identifizieren abhängige, unabhängige und Kontrollvariablen in einem Experiment. ...planen im Hinblick auf Variablenkontrolle angemessene Experimente.	...unterscheiden systematisch zwischen Beobachtungen und Deutungen. ...formulieren im Hinblick auf vorliegende Daten zulässige Interpretationen. ...formulieren eine inhaltlich^a zu einer Vermutung bzw. Hypothese passende Interpretation. ...erstellen sachangemessene Tabellen und Diagramme. ...bewerten den Nutzen eines Experiments im Hinblick auf den Erkenntnisgewinn (insbesondere bzgl. der Frage bzw. Vermutung bzw. Hypothese).

Anmerkungen. Literaturbelege finden sich bei Vorholzer (2016, Tabelle 2.2).

^a Inhaltliche Passung bedeutet, dass die Vermutung bzw. das Experiment bzw. die Interpretation grundsätzlich dazu geeignet sein muss, eine Antwort auf die Frage bzw. zur Vermutung bzw. Hypothese zu liefern (unabhängig davon, ob dabei z. B. Aspekten der Variablenkontrolle angemessen berücksichtigt wurden).

Tabelle 4.2: Konzepte zu den drei Teilprozessen des experimentbezogenen Denkens und Arbeitens basierend auf Vorholzer (2016, Tab. 2.2) mit Ergänzungen aus Gabi (2018, S. 10–11).

	Formulieren von Fragen und Hypothesen	Planen und Durchführen von Untersuchungen	Auswerten und Interpretieren von Daten
Konzepte	<i>Eine naturwissenschaftliche Fragestellung ist mit naturwissenschaftlichen Methoden^b untersuchbar und unterscheidet sich dadurch von einer sozial- oder geisteswissenschaftlichen Frage.</i> Auch unsinnig erscheinende Fragestellungen können naturwissenschaftliche Fragen sein, wenn sie mit naturwiss. Methoden^b untersucht werden können. <i>Die Fragestellung einer Untersuchung sollte immer so präzise wie möglich formuliert sein.</i> Eine allgemeine Fragestellung beschreibt was generell mit dem Versuch untersucht werden soll und lässt offen welche Merkmale beobachtet werden sollen. Eine präzise Fragestellung musst möglichst genau festlegen, welches Merkmal verändert und welches Merkmal beobachtet werden soll. <i>Vermutungen bzw. Hypothesen treffen eine Vorhersage über die im Rahmen einer Fragestellung zu untersuchenden Zusammenhänge.</i> Vorhersagen über den Versuchsaufbau sollten nur dann formuliert werden, wenn sie auch sinnvoll begründet werden können. Man unterscheidet bei Vorhersagen zwischen Vermutung (Begründung durch Phänomen) und Hypothesen (Begründung durch Theorie/Modell).	<i>Eine Untersuchung muss so geplant werden, dass die Fragestellung bzw. die dazugehörige Hypothese/Vermutung damit beantwortet bzw. verifiziert oder falsifiziert werden kann.</i> <i>In einer Untersuchung sollte möglichst nur die Variable verändert werden, deren Einfluss überprüft werden soll (sie heißt unabhängige Variable; UV).</i> <i>Die abhängige Variable (AV) ist die Variable, die während der Änderung der unabhängigen Variable beobachtet/gemessen wird, um zu überprüfen, ob es die UV einen Einfluss auf sie hat.</i> <i>Alle anderen Variablen, die einen Einfluss haben könnten, sollten möglichst konstant gehalten werden. (Variablenkontrolle)</i> Variablenkontrolle ist eine wichtige Voraussetzung dafür, dass aus einer Untersuchung richtige Schlussfolgerungen gezogen werden können. <i>Beim Planen einer Untersuchung ist es wichtig, darauf zu achten, dass die Untersuchung möglichst den gesamten zur Verfügung stehenden Messbereich abdeckt.</i>	<i>Eine Beobachtung beschreibt ausschließlich, was passiert; sie ist intersubjektiv überprüfbar. Eine Deutung enthält immer eine Verallgemeinerung oder Bewertung.</i> <i>Eine Interpretation bezieht sich auf die in der Untersuchung gewonnen Daten und soll sachangemessen aus diesen Daten begründet werden (z. B. darf sie nicht im Widerspruch zu den Daten stehen).</i> <i>Eine Interpretation soll direkten Bezug auf die Fragestellung und die Vermutung/Hypothese nehmen, diese verifizieren oder falsifizieren.</i> Deutungen, die keinen direkten Bezug zur Untersuchung haben, können zu neuen Fragen und Vermutungen führen. <i>In einem Diagramm wird die unabhängige Variable immer auf der Abszisse (horizontale Achse) aufgetragen. (Analog für AV mit Ordinate.)</i> <i>Die Achsen eines Diagramms sind so skaliert, dass die Messwerte auf den größten Teil der Achse entfallen.</i> <i>In einer Tabelle wird die unabhängige Variable immer in der ersten, die abhängige(n) in der (den) folgenden und die berechnete(n) Variable(n) in der (den) dann folgenden Spalte(n) notiert. (Analog bei zeilenweiser Darstellung.)</i> <i>In Tabellen und Diagrammen müssen Werte immer mit einer zugehörigen Maßeinheit angegeben werden. Eine Beschriftung der Zeilen/Spalten bzw. Achsen ist ausreichend.</i> <i>Naturwissenschaftliche Experimente werden durchgeführt, um eine Frage zu beantworten oder eine Ursache-Wirkungs-Beziehung zu prüfen (und nicht, um ein möglichst gutes Ergebnis oder einen Effekt zu erzielen).</i>

Anmerkung. Kursiv = Ankerkonzepte für Testungen. Literaturbelege bei Vorholzer (2016, Tabelle 2.2).

^a Inhaltliche Passung bedeutet, dass die Vermutung bzw. das Experiment bzw. die Interpretation grundsätzlich dazu geeignet sein muss, eine Antwort auf die Frage bzw. zur Vermutung bzw. Hypothese zu liefern (unabhängig davon, ob dabei z. B. Aspekte der Variablenkontrolle angemessen berücksichtigt wurden).

^b Naturw. Methoden beinhalten u. a. auch Beobachtung, Messung, Zählung objektivierbarer Ereignisse sowie Modellierung und Exploration und sind nicht als auf das variablenkontrollierte Experiment beschränkt.

4.2 Grundsätzliche Konzeption der Instruktion

Die Instruktion wurde für Schülerinnen und Schüler der elften Klasse zur Bearbeitung in Teams à 2–3 Personen entwickelt und berücksichtigt *zentrale Designprinzipien*. Erstens ist die Instruktion so konzipiert, dass die (angenommenen) Lernvoraussetzungen der Lernenden berücksichtigt sind. Zweitens werden darüber hinaus auch Annahmen zu Lernverläufen berücksichtigt (ausführlicher dargestellt in Vorholzer et al., 2020). Drittens fordert die Instruktion zur selbstbestimmten, intensiven Auseinandersetzung mit Aufgaben auf. Viertens liegt die gesamte Instruktion in Textform vor, wobei der Textumfang durch mehrfache Iterationen auf ein Minimum reduziert wurde.

Das Vorliegen der *Instruktion in Textform* ist so umgesetzt, dass Arbeitsaufträge und Informationen auf Aufgabenkarten bzw. Informationskarten notiert sind, wobei jede Karte für gewöhnlich eine Aufgabe oder mehrere inhaltlich direkt zusammengehörende Teilaufgaben (bzw. eine auf einen Aspekt gerichtete Information) beinhaltet. Diese Anlage wurde von Vorholzer (2016) gewählt, um Interaktionen von Lernenden mit Lehrkräften und Forschenden zu minimieren und damit auch die Vergleichbarkeit der Bearbeitungsprozesse von verschiedenen Teams zu unterstützen.

Die Instruktion bietet Gelegenheiten zum eigenständigen experimentbezogenen Denken und Arbeiten und besteht aus *Serien von jeweils auf das gleiche Konzept bezogenen Explorationsaufgaben, Informationstexten und Übungsaufgaben*. So ist beispielsweise (nachfolgend auch als Beispiel 1 referenziert) in Abbildung 4.2 (auf S. 56) ein Ausschnitt aus der Instruktion dargestellt, in dem vor einem Informationstext zu Kennzeichen von naturwissenschaftlichen Fragen einige Explorationsaufgaben vorhanden sind, bei denen ohne vorgegebene Kriterien entschieden werden soll, ob Beispielfragen naturwissenschaftlich oder nicht-naturwissenschaftlich sind (Karte 1-21); nach der instruktionalen Explizierung einiger Kriterien (Karte 1-23) wird dazu aufgefordert, sie auf verschiedene Fälle begründet anzuwenden (Karten 1-24 und folgende). Außerdem gibt es beispielsweise (nachfolgend als Beispiel 2 referenziert) Möglichkeiten, vor einer instruktionalen Explizierung des Zusammenhangs von dem Abwurfwinkel eines Projektils und seiner Wurfweite beim schrägen Wurf, den Zusammenhang eigenständig zu erkunden, der zudem in späteren Übungsaufgaben intensiver behandelt wird. Serien zu verschiedenen Konzepten können einander auch überlappen (schematisch dargestellt in Abbildung 4.3 auf S. 57) und einige Serien beinhalten auch nur zwei der drei Elemente.

4 Zugrundeliegendes Design

1-21 Eine naturwissenschaftliche Fragestellung?

AUFGABE

Nicht jede Fragestellung, die etwas mit Physik, Chemie oder Biologie zu tun hat, ist automatisch eine *naturwissenschaftliche* Fragestellung. Anders herum können auch alltägliche Fragen naturwissenschaftliche Fragen sein.

Bei welchen der folgenden Fragestellungen handelt es sich Ihrer Einschätzung nach um naturwissenschaftliche Fragestellungen?

- ☐ Ist naturwissenschaftliches Arbeiten ein spannendes Thema für Schüler/innen?
- ☐ Wird Kaffee mit Milch schneller kalt als Kaffee ohne Milch?
- ☐ Hat der Luftdruck einen Einfluss darauf, bei welcher Temperatur Wasser kocht?
- ☐ Beeinflusst die Größe eines Hundes, wie oft er im Durchschnitt pro Stunde bellt?
- ☐ Schmeckt Mineralwasser mit Kohlensäure besser als Leitungswasser?
- ☐ Von welchen Faktoren hängt die Endgeschwindigkeit eines Fallschirmspringes ab?
- ☐ Sollten Schüler/innen etwas darüber lernen, welche Faktoren die Periodendauer eines Fadenpendels beeinflussen?

1-22 Eine naturwissenschaftliche Fragestellung?

KONTROLLE

[Lösungen zur Karte 21]

Stellen Sie sich vor, Sie sollten jemandem erklären, was eine naturwissenschaftliche Fragestellung von einer nicht-naturwissenschaftlichen Fragestellung unterscheidet. Was würden Sie sagen? Die Beispiele oben können Ihnen bei Ihren Überlegungen helfen!

[Platz zum Notieren der Antwort]

1-23 Eine naturwissenschaftliche Fragestellung?

INFORMATION

Es kann nicht immer eindeutig entschieden werden, ob eine Fragestellung „naturwissenschaftlich“ ist oder nicht. **Ein zentrales Merkmal einer naturwissenschaftlichen Fragestellung ist jedoch, dass sie mit naturwissenschaftlichen Methoden untersucht werden kann.**

Naturwissenschaftliche Methoden sind Messungen/Beobachtungen/Zählungen von objektivierbaren Ereignissen. Nicht objektivierbare Ereignisse sind z. B. Geschmack, Meinung, Einstellung usw.

[Erläuterende Beispiele]

1-24 Eine naturwissenschaftliche Fragestellung?

AUFGABE

Warum ist „Wird Kaffee mit Milch schneller kalt als Kaffee ohne Milch?“ eine naturwissenschaftliche Fragestellung?

[Weitere Aufgaben bis Karte 1-26]

Abbildung 4.2: Sequenz von Explorationsaufgaben, Informationstext und Übungsaufgaben zu Kennzeichen von naturwissenschaftlichen Fragen aus der explizit-fachmethodischen Instruktionsvariante (Konzept ausformuliert auf Karte 1-23; graue Hinterlegungen kennzeichnen Anmerkungen).

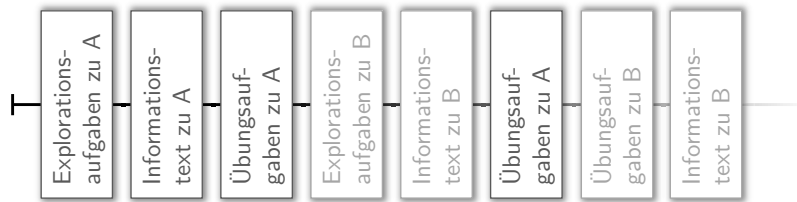


Abbildung 4.3: Exemplarische Struktur von teilweise überlappenden Serien aus Explorationsaufgaben, Informationstexten und Übungsaufgaben zu zwei Konzepten A und B. Der erste Teil der Serie zum Konzept A ist eine in sich abgeschlossene Serie mit allen drei Komponenten; für das Konzept B liegt hingegen keine in sich geschlossene Serie vor.

4.3 Explizit- und implizit-fachmethodische Instruktionsvarianten

Die Instruktion liegt in zwei Varianten vor – einer explizit-fachmethodischen Variante und einer implizit-fachmethodischen Variante. Die Varianten bestehen beide aus drei Einheiten, welche planmäßig dieselben Dauern haben (45 + 90 + 90 Minuten), und sie beziehen sich auf dieselben fachlichen Inhalte als fachliche Kontexte (siehe Tabelle 4.3). Die Instruktionsvarianten unterscheiden sich aber in der *Adressierung* der Fähigkeiten und Konzepte zu den drei Teilprozessen zum experimentbezogenen Denken und Arbeiten, welche in der Studie in den Blick genommen werden: Formulieren von Fragen und Hypothesen (FH), Planen von Untersuchungen (PU), Auswerten und Interpretieren von Daten (AI).

Die Fähigkeiten und Konzepte zu den Teilprozessen FH, PU und AI werden in der *explizit-fachmethodischen Instruktionsvariante* explizit adressiert, wobei jeweils eine Einheit schwerpunktmäßig einem Teilprozess zugeordnet ist (siehe Tabelle 4.3). Die *explizite Adressierung* der Fähigkeiten und Konzepte umfasst die zeitlich strukturierte Exploration, Mitteilung und Übung der Konzepte des experimentbezogenen Denkens und Arbeitens sowie eigenständiges Ausführen von zugehörigen Tätigkeiten. Das Beispiel 1 von oben lässt sich als eine Sequenz von Aufgaben verstehen, die zu den ersten beiden Konzepten aus Tabelle 4.2 gehört. Es ist zu betonen, dass die explizit-fachmethodische Instruktionsvariante nicht für den Aufbau fachinhaltlicher Kompetenzen angelegt ist – fachinhaltliche Konzepte werden daher nicht in strukturiert aufeinanderfolgenden Explorations, Informationen und Übungen adressiert.

Die *implizit-fachmethodische Instruktionsvariante* ist so konzipiert, dass experimentbezogenes Denken und Arbeiten *implizit adressiert* wird: Erstens sind – wie auch für die explizit-fachmethodische Instruktionsvariante – alle Texte, Diagramme und Tabellen der Instruktion (Aufgabenstellungen, Informationen, Hinweise, ...) so angelegt, dass sie den (jeweils anwendbaren) fachmethodischen Konzepten genügen.

Tabelle 4.3: Inhalte und Fokusse der Einheiten der expliziten und der impliziten Variante der Instruktion (basierend auf Vorholzer, 2016, S. 95).

		Einheit 1	Einheit 2	Einheit 3
Themenbereich		Schweredruck in Flüssigkeiten	Freier Fall und Luftreibung	Senkrechter, waagerechter und schiefer Wurf
Fokus der Instruktionsvariante	Explizit-fachmethodisch	Formulierung von naturwissenschaftlichen Fragen und Hypothesen (FH)	Planung von naturwissenschaftlichen Untersuchungen (PU)	Auswertung und Interpretation von Daten (AI)
	Implizit-fachmethodisch	Modellierung des Schweredrucks; Berechnung des Schweredrucks; Parameter, die den Schweredruck beeinflussen usw.	Parameter, von denen durchschnittliche Fallgeschwindigkeit abhängt; Existenz und Berechnung maximaler Fallgeschwindigkeit usw.	Parameter, von denen die maximale Wurfhöhe und Wurfweite abhängt; Unterschiede Wurf mit und ohne Reibung usw.
Planmäßige Dauer		ca. 45 Minuten	ca. 90 Minuten	ca. 90 Minuten

Anmerkung. Die Tabelle stammt aus Vorholzer (2016, S. 72) und wurde zur Darstellung transponiert und minimal angepasst. Die planmäßige Dauer umfasst in den beiden längeren Einheiten bereits organisatorische Vorbemerkungen, Hinweise zur Videografierung etc.; die Bearbeitungsdauer wird jeweils eher auf 75–80 Minuten angesetzt.

Beispielsweise sind alle aufgeführten naturwissenschaftlichen Fragen, zu denen die Lernenden weiterführende Überlegungen oder Untersuchungen anstellen sollen, unter Nennung der jeweiligen abhängigen und unabhängigen Variablen präzise formuliert (vergleiche Tabelle 4.2 auf S. 54). Zweitens ermöglichen die Arbeitsaufträge das eigenständige Ausführen experimentbezogener Denk- und Arbeitsweisen. Demnach kann sowohl für die explizit-fachmethodische als auch für die implizit-fachmethodische Instruktionsvariante in einer gewissen Weise davon gesprochen werden, dass diese entlang der Konzepte des experimentbezogenen Denkens und Arbeitens strukturiert sind. Es sind jedoch einige Unterschiede zwischen den Varianten hervorzuheben: Zum einen werden Konzepte des experimentbezogenen Denkens und Arbeitens in der implizit-fachmethodischen Instruktionsvariante nicht mitgeteilt, und Phasen zur

4.4 Übersicht über die inhaltlichen Abschnitte der beiden Instruktionsvarianten

(Ein-)Übung der Anwendung von vorher genannten fachmethodischen Konzepten sind demnach nicht im Material enthalten. Zum anderen werden in der implizit-fachmethodischen Variante im Kontrast zur explizit-fachmethodischen Instruktionsvariante fachinhaltliche Konzepte (zu dem in Tabelle 4.3 aufgeführten Fokus) fokussiert und *zeitlich strukturiert* (d. h. durch spezifisch auf fachmethodisches Lernen angelegte Aufgabenreihenfolgen) adressiert.³⁰ Beispiel 2 von oben (S. 55) beschreibt eine Sequenz von Aufgaben, die in der implizit-fachmethodischen (und nicht in der explizit-fachmethodischen) Instruktionsvariante vorhanden ist, da die Erarbeitung eines fachinhaltlichen Konzeptes (nämlich: »Der Abwurfwinkel hat einen [genauer spezifizierbaren] Einfluss auf die Wurfweite«) forciert wird.

Ausführlichere Gegenüberstellungen der Instruktionsvarianten finden sich in Vorholzer (2016) oder Vorholzer et al. (2020). Außerdem gewähren die in der vorliegenden Arbeit befindlichen Erläuterungen zu Methoden und Ergebnissen detaillierte Einblicke in die Instruktion. Im Anhang der vorliegenden Arbeit und im Anhang bei Vorholzer (2016) finden sich zusammenhängende Ausschnitte aus der Instruktion.

4.4 Übersicht über die inhaltlichen Abschnitte der beiden Instruktionsvarianten

Die Instruktionen sind innerhalb jeder der drei Einheiten zusätzlich in inhaltliche Abschnitte gegliedert, die den Lernenden durch Abschnittstitelblätter transparent gemacht sind. Da einige Analysen der Arbeit auf Ebene der (Instruktions-)Abschnitte erfolgen, sind diese in den Tabellen 4.4 und 4.5 kurz beschrieben. Für die explizit-fachmethodische Instruktionsvariante entsprechen die (Instruktions-)Abschnitte in etwa den Konzepten aus Tabelle 4.2 (auf S. 54). Für die implizit-fachmethodische Instruktionsvariante sind die (Instruktions-)Abschnitte auf Teilbereiche der jeweils adressierten Fachinhalte bezogen. Die in Abschnitt 4.2 geschilderten Serien zu Konzepten beginnen für beide Instruktionsvarianten zumeist in den entsprechenden (Instruktions-)Abschnitten, Wiederholungsaufgaben finden sich allerdings auch in späteren Abschnitten (und ggf. Einheiten) der Instruktion. Die (Instruktions-)Abschnitte sind nach der Einheit benannt und dann fortlaufend als A[i] nummeriert; Karten der Instruktion werden mit der Einheitennummer und einer fortlaufenden Zahl angegeben

³⁰In gewisser Hinsicht ist es möglich, die implizit-fachmethodische Instruktion zugleich als explizit-fachinhaltliche Instruktion zu bezeichnen. In diesem Sinne müssten also auch die fachinhaltlichen Fähigkeiten und Konzepte aufgelistet werden, die im Material adressiert werden. Da der Fokus der Arbeit allerdings sowohl auf explizit-fachmethodischer Instruktion als auch auf den Kompetenzen des experimentbezogenen Denkens und Arbeitens liegt, wird darauf verzichtet und stattdessen auf Vorholzer (2016, S. 94–96) verwiesen.

4 Zugrundeliegendes Design

(ohne Buchstaben). Der (Instruktions-)Abschnitt 1-A1 umfasst beispielsweise die Karten 3 bis 19 der ersten Einheit.

Tabelle 4.4: Abschnitte der explizit-fachmethodischen Instruktionsvariante.

Abschnitt (Karten)	Kurzbeschreibung
1-A0 (1-01 – 1-02)	Vorspann zu Einheit 1: »Naturwissenschaftliche Fragestellungen, Vermutungen und Hypothesen«
1-A1 (1-03 – 1-19)	Präzise und allgemeine Fragen unterscheiden; Vermutungen als begründete Vorhersagen
1-A2 (1-20 – 1-29)	Naturwissenschaftliche und nicht-naturwissenschaftliche Fragen
1-A3 (1-30 – 1-33)	Vermutungen und Hypothesen als unterschiedlich begründete Vorhersagen
1-A4 (1-34 – 1-35)	Lob; Zusammenfassungsblatt
2-A0 (2-01 – 2-02)	Vorspann zu Einheit 2: »Naturwissenschaftliche Untersuchungen planen«
2-A1 (2-03 – 2-12)	Begriffseinführung: Abhängige und unabhängige Variable, Kontrollvariablen
2-A2 (2-13 – 2-25)	Kontrollierte und unkontrollierte Versuchsdesigns unterscheiden
2-A3a (2-26 – 2-44)	Versuchsplanung unter Berücksichtigung der Variablenkontrolle
2-A3b (2-45 – 2-49)	[Fachinhaltliche Informationen zur Geschwindigkeitsabhängigkeit der Luftreibung]
2-A4 (2-50 – 2-56)	Bezug von Versuchsplanung und Fragestellungen herstellen; Messwiederholungen und Messgenauigkeit
2-A5 (2-57 – 2-58)	Lob; Zusammenfassungsblatt

(Fortsetzung auf der nächsten Seite)

4.4 Übersicht über die inhaltlichen Abschnitte der beiden Instruktionsvarianten

Tabelle 4.4 (Fortsetzung)

Abschnitt (Karten)	Kurzbeschreibung
3-A0 (3-01 – 3-02)	Vorspann zu Einheit 3: »Auswerten und Deuten naturwissenschaftlicher Untersuchungen«
3-A1 (3-03 – 3-18)	Tabellen und Diagramme geeignet anlegen
3-A2 (3-19 – 3-25)	Unterscheidung von Beobachtung und Deutung
3-A3 (3-26 – 3-43)	V Versuchsergebnisse angemessen deuten
3-A4 (3-44 – 3-54)	Erfolge und Misserfolge bei naturwissenschaftlichen Untersuchungen
3-A5 (3-55 – 3-56)	Lob; Zusammenfassungsblatt

Anmerkungen. Abschnitt 2-A3 enthält einen fakultativen Einschub zu Fachinhalten und wird daher in zwei Teilabschnitte getrennt, die im Material allerdings nicht durch ein Abschnittstitelblatt voneinander getrennt sind (daher die eckigen Klammern). Die Kurzbeschreibungen dienen dazu, einen groben Überblick über die gesamte Instruktion zu geben; das Instruktionsmaterial ist am Institut für Physikdidaktik der JLU Gießen einsehbar und in überarbeiteter Fassung auch publiziert (Vorholzer, 2017).

Tabelle 4.5: Abschnitte der implizit-fachmethodischen Instruktionsvariante.

Abschnitt (Karten)	Kurzbeschreibung
1-A0 (1-01 – 1-02)	Vorspann zu Einheit 1: »Luftblasen im Glas? Druck aus verschiedenen Perspektiven« (Schweredruck in Flüssigkeiten)
1-A1 (1-03 – 1-09)	[Luftblasen beobachten als Druckvergleich]
1-A2 (1-10 – 1-15)	[Weitere Ergebnisse zu Druck im Wasser]
1-A3 (1-16 – 1-27)	[Modellierung von Druck im Wasser durch Wasserwürfelmodell]
1-A5 (1-28 – 1-31)	[Druck berechnen]
1-A6 (1-32 – 1-33)	Lob; Zusammenfassungsblatt

(Fortsetzung auf der nächsten Seite)

Tabelle 4.5 (Fortsetzung)

Abschnitt (Karten)	Kurzbeschreibung
2-A0 (2-01 – 2-02)	Vorspann zu Einheit 2: »Physikalische Experimente: Reibung aus verschiedenen Perspektiven« (Freier Fall und Luftreibung)
2-A1 (2-03 – 2-06)	[Reibung auf einer Unterlage]
2-A2 (2-07 – 2-10)	[Einfluss der Masse]
2-A3 (2-11 – 2-17)	[Einfluss der Kontaktfläche und Reibungszahl]
2-A4 (2-18 – 2-27)	[Reibung in der Luft]
2-A5 (2-28 – 2-35)	[Geschwindigkeitsabhängigkeit der Luftreibung]
2-A6 (2-36 – 2-40)	[Anwendungen und eigenständige Planung anschließender Versuche]
2-A7 (2-41 – 2-44)	Lob; Zusammenfassingsblatt
2-A8 (2-45 – 2-54)	Zusatzmaterial: Messunsicherheiten; Abschließendes Lob
3-A0 (3-01 – 3-02)	Vorspann zu Einheit 3: »Der „Wasserwerfer“: Experimente zum schiefen Wurf« (Senk-, waagerechter u. schiefer Wurf)
3-A1 (3-03 – 3-07)	[Einführung in den Wasserwerfer: Relative Höhe, Wurfweite; Versuchsaufbau]
3-A2 (3-08 – 3-11)	[Austrittsgeschwindigkeit]
3-A3 (3-12 – 3-14)	[Abwurfwinkel]
3-A4 (3-15 – 3-26)	[Flugbahn beschreiben: Komponentenzerlegung]
3-A5 (3-27 – 3-33)	[Relative Höhe und optimaler Abwurfwinkel]
3-A6 (3-34 – 3-40)	[Reale Flugbahnen]
3-A7 (3-41 – 3-42)	Lob; Zusammenfassingsblatt

Anmerkung. Kurzbeschreibungen zu im Material nicht durch Abschnittstitelblätter ausgewiesenen Abschnitten stehen in eckigen Klammern. Die Kurzbeschreibungen dienen dazu, einen groben Überblick über die gesamte Instruktion zu geben; das Instruktionsmaterial ist am Institut für Physikdidaktik der JLU Gießen einsehbar und in überarbeiteter Fassung auch publiziert (Vorholzer, 2017).

4.5 Prä- und Post-Erhebungen

Zu den Prä- und Post-Erhebungszeitpunkten wurden verschiedene Testinstrumente eingesetzt (siehe bereits Abbildung 4.1 auf S. 51). Als Haupterhebungsinstrument diente der Test zu experimentellen Denk- und Arbeitsweisen (*EDAWT*; Vorholzer, 2016; Vorholzer et al., 2016). Personenmerkmale wurden vor der Intervention mit folgenden Instrumenten erhoben: Fachwissenstest zur Mechanik (eine ergänzte und für die elfte Klassenstufe adaptierte Form des Force Concepts Inventorys, siehe Vorholzer, 2016); Test zu allgemeinen kognitiven Fähigkeiten (Subskala N2 des KFTs; Heller & Perleth, 2000); Einschätzungsbogen zur physikbezogenen Selbstwirksamkeit (Steckenmesser-Sander, 2015). Die Güte der Instrumente wurde für die im Folgenden als Ausgangsbasis dienende Stichprobe von Vorholzer (2016) geprüft. Grob zusammengefasst finden sich für den *EDAWT* gute bis sehr gute Werte; für die anderen Instrumente ergeben sich zufriedenstellende bis gute Werte (für statistische Kennzahlen siehe Anhang A, Tabelle A.1; ausführlichere Diskussionen zur Validität finden sich bei Vorholzer, 2016).

Analysen der Prä- und Post-Erhebungen wurden von Vorholzer (2016) vorgenommen. An der Studie haben insgesamt 222 Personen aus 12 Parallelklassen derselben Schule (elfte Klassenstufe, ca. 16–17 Jahre alt) teilgenommen. Die explizit-fachmethodische Instruktion wurde von 111 Schülerinnen und Schülern (63 % weiblich) bearbeitet, die implizit-fachmethodische Instruktion ebenfalls von 111 Schülerinnen und Schülern (67 % weiblich). Mithilfe der Prä-Erhebungen wurde für eine Vergleichbarkeit der beiden Populationen gesorgt. Die Ergebnisse sind ausführlich dokumentiert (z. B. Vorholzer et al., 2018; Vorholzer et al., 2016) und werden im Rahmen der vorliegenden Arbeit dort dargestellt, wo sie zur Hypothesenformulierung beziehungsweise zur Diskussion oder Einordnung der Ergebnisse notwendig sind. Grundsätzlich wurde gefunden, dass es für beide Varianten einen signifikanten Kompetenzzuwachs gibt (explizit: $t = 7.506$, $p < .001$; $\eta^2 = 0.46$; implizit: $t = 4.806$, $p < .001$; $\eta^2 = 0.26$; entspricht jeweils einem großen Effekt). Eine Kovarianzanalyse zeigt zudem, dass der Kompetenzzuwachs bei der explizit-fachmethodischen Instruktion signifikant größer ist als bei der implizit-fachmethodischen Instruktion ($F(1, 127) = 9.77$, $p = 0.002$, dieser Unterschied entspricht einem mittleren Effekt: $\eta_p^2 = 0.07$).

4.6 Videoaufzeichnungen

Als Datenquellen für die vorliegende Arbeit existieren neben den für die einzelnen Personen vorliegenden Testwerten aus den im vorigen Unterabschnitt dargestellten Prä- und Post-Erhebungen *Videoaufzeichnungen der Bearbeitungsprozesse* von einer (basierend auf freiwilliger Zustimmung generierten) Teilstichprobe der getesteten Lernenden. Insgesamt wurden 71 Personen bei der Bearbeitung von mindestens einer Einheit der explizit-fachmethodischen Instruktionsvariante und 50 Personen bei der Bearbeitung von mindestens einer Einheit der implizit-fachmethodischen Instruktionsvariante videografiert. Über 90 % der Personen wurden für mindestens zwei Einheiten aufgezeichnet, über 65 % für alle drei. Eine Übersicht über die Personen- und Team-Anzahlen für die jeweiligen Instruktionsvarianten findet sich in Tabelle 4.6.

Die Instruktion wurde in der Regel von festen Teams à 2–3 Personen bearbeitet, die gemeinsam an einem Kartensatz arbeiteten, der die schriftlichen Anweisungen und Information enthält. Sie saßen (bzw. standen für einige Experimente) dabei in etwa wie in Abbildung 4.4 dargestellt gemeinsam an einem Tisch. Daher liegen die Videoaufzeichnungen der Personen in einer Teamstruktur vor. Wie in Tabelle 4.6 dargestellt wurden 24 Teams für die explizit-fachmethodische und 18 Teams für die implizit-fachmethodische Variante der Instruktion videografiert.³¹ Genauere Angaben zu den in den jeweiligen Analysen berücksichtigten Personen und Teams finden sich bei den entsprechenden Analysen in den Kapiteln 5 und 6.

Tabelle 4.6: Überblick über Personenanzahlen nach Instruktionsvariante sowie Geschlecht.

Instruktionsvariante	Teams	S	S _w	S _m
<i>Gesamtstudie</i>	83	222	144	78
Explizit-fachmeth.	41	111	70	41
Implizit-fachmeth.	42	111	74	37
<i>Videoaufzeichnungen</i>	42	121	75	46
Explizit-fachmeth.	24	71	46	25
Implizit-fachmeth.	18	50	29	21

Anmerkung. Die Anzahlen für vorliegende schriftliche Aufzeichnungen der Schülerinnen und Schüler entsprechen den Anzahlen für die Gesamtstudie.

S = Schülerinnen und Schüler,

S_w = Schülerinnen, S_m = Schüler.

³¹Die Ungleichheit der Stichprobengrößen erklärt sich durch die freiwillige Zustimmung zur Videoaufzeichnung. Teams, in denen mindestens eine Person nicht mit der Videoaufzeichnung einverstanden war, wurden nicht aufgezeichnet.



(a) Beim Experimentieren



(b) Beim Beantworten von Fragen

Abbildung 4.4: Typische Kameraperspektiven, aus denen die Schülerinnen und Schüler bei der Bearbeitung der Instruktion gefilmt wurden (Bilder bereits veröffentlicht in Vorholzer et al., 2020, Reproduced with permission from Springer Nature).

5 Aktivitäten von Lernenden während der Bearbeitung der beiden Instruktionsvarianten

Analysen zu den verbalen und nonverbalen Aktivitäten von Lernenden bei der Bearbeitung der beiden eingesetzten Instruktionsvarianten erlauben, die Auseinandersetzung auf der Verhaltensebene zu charakterisieren sowie zu beschreiben, wie sich die Art und Weise der Auseinandersetzung im Verlauf der Zeit und in Verbindung mit der Instruktion bzw. den bearbeiteten Aufgaben verändert. Aussagen zum (Sprach-)Handeln der Lernenden ermöglichen zum Teil Rückschlüsse auf das damit verbundene Denken der Lernenden. Verknüpfungen der Aktivitäten mit den von Lernenden erreichten Kompetenzzuwächsen zum experimentbezogenen Denken und Arbeiten ermöglichen die Generierung von Hypothesen dazu, welche Bedeutung die jeweiligen Aktivitäten für den Kompetenzaufbau haben. Dabei wird gemäß der in der Forschung hervorgehobenen Relevanz der von Lernenden vorgenommenen Fokussierung auf Fachmethoden (Kapitel 2) insbesondere auf die Äußerungen der Lernenden mit fachmethodischem Gehalt eingegangen, weil diese als sprachliche Fokussierung auf Fachmethoden aufgefasst werden.

In Abschnitt 5.1 wird dargestellt, welche Aktivitäten untersucht werden und welche Herausforderungen im Rahmen der Entwicklung eines Kodiermanuals für eventbasiertes Kodieren zu bewältigen waren. Da die Verknüpfung von Aktivitäten und Kompetenzzuwächsen im wesentlichen durch Vergleiche von Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen realisiert wird, wird in Abschnitt 5.2 die für die Kontrastierung vorgenommene Gruppierung der Lernenden geschildert. Die für die Auswertung grundlegenden methodischen Überlegungen und Vorgehensweisen werden in Abschnitt 5.3 vorgestellt. Spezifische Auswertungen, Ergebnisse und erste zugehörige Diskussionen sind in Abschnitt 5.4 (zur Forschungsfrage F-Akt1, die auf den Vergleich von Lernenden mit unterschiedlichen Kompetenzzuwächsen innerhalb der explizit-fachmethodischen Instruktionsvariante bezogen ist) und Abschnitt 5.5 (zu F-Akt2, bezogen auf die implizit-fachmethodische Instruktionsvariante, sowie zu F-Akt3, bezogen auf den Vergleich von Lernenden aus den beiden Instruktionsvarianten) zu finden. Eine zusammenführende Diskussion und die Generierung von Hypothesen erfolgt in Kapitel 7.

5.1 Eventbasiertes Kodieren von verbalen und nonverbalen Aktivitäten der Lernenden

Zur Erfassung ausgewählter Aktivitäten der Lernenden bei der Bearbeitung der untersuchten Instruktionsvarianten wird ein eventbasiertes Kodierverfahren genutzt, welches auf einem ausführlichen Kodiermanual basiert. Eventbasiertes Kodieren bedeutet, dass jedem Event (z. B. jeder Äußerung zu Fachmethoden) eine Kategorie aus dem Kategoriensystem der Kodiermanuals zugewiesen wird. Bei intervallbasiertem Kodieren würde jedem Zeitintervall (von vorher festgelegter Länge) eine Kategorie zugewiesen werden. Hofmann (2015) diskutiert ausführlich die Vor- und Nachteile eventbasierten und intervallbasierten Kodierens und sich aus den Verfahren jeweils ergebende Konsequenzen. Die Entscheidung für eventbasiertes Kodieren erfolgte, weil das Vorgehen sich auch bei großen Streuung der zeitlichen Länge der untersuchten Äußerungen gut anwenden ließ und weil davon ausgegangen wurde, dass Reihenfolgen von Aktivitäten sich hiermit besser untersuchen lassen, was für die Ziele des der Arbeit zugrundeliegenden Projektes relevant war.³²

5.1.1 Kodiermanual

Für die Entwicklung Kodiermanuals – die Kategorien sind übersichtsartig in Abbildung 5.1 dargestellt, ein Auszug des Manuals findet sich in Abbildung 5.3 (auf S. 72), das gesamte Manual findet sich in Abschnitt B.1 (ab S. 522) – wurden Literatursichtungen vorgenommen und Kodiermanuals aus der Arbeitsgruppe des Instituts

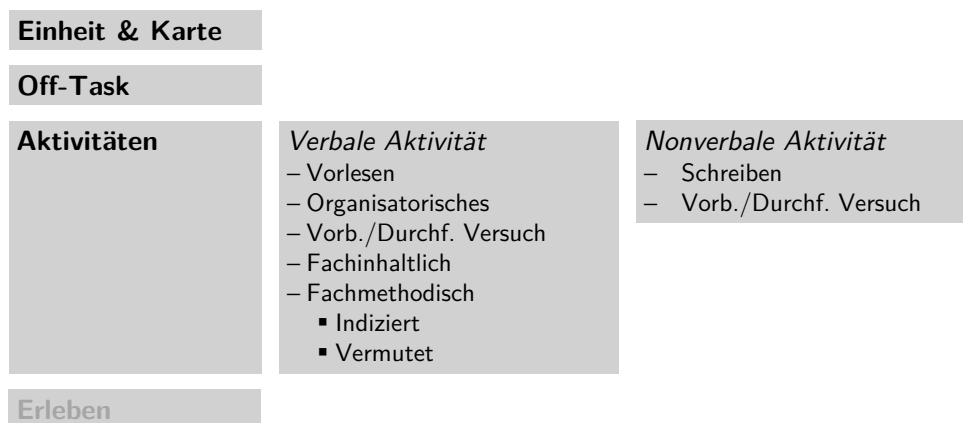


Abbildung 5.1: Übersicht über die Ebenen und Kategorien des Kodiermanuals.

³²Nach getroffener Entscheidung für das Verfahren, die auch mit der Entscheidung für die Software MAXQDA verbunden war, stellte sich allerdings unter anderem heraus, dass zur *statistischen* Analyse von zeitlichen Serien oftmals intervallbasiertes Kodieren eine bessere Basis dargestellt hätte. Es bleibt allerdings offen, ob damit zu validieren Deutungen gelangt hätte werden können.

für Physikdidaktik an der Justus-Liebig-Universität Gießen als Vorlagen genutzt. Insbesondere die bei Rogge (2010) und Steckenmesser-Sander (2015) veröffentlichten Manuale dienten als Ausgangspunkte für das Kodiermanual der vorliegenden Arbeit. Während der Laufzeit des der Arbeit zugrundeliegenden Projektes wurden weitere Studien mit Kodiermanualen veröffentlicht, die Ähnlichkeiten zum entwickelten Manual aufweisen, obwohl nicht auf sie zurückgegriffen wurde (u. a. Kalthoff, 2019).

Inhaltlich war dem Fokus der vorliegenden Arbeit entsprechend eine Weiterentwicklung der gesichteten Manuale erforderlich, um fachmethodischen Kompetenzaufbau zu fokussieren. Dazu wurden drei Kategorien *verbaler Aktivitäten* eingeführt (Ebene 3 des Kategoriensystems), die fachliche (d. h. fachbezogene) Äußerungen der Lernenden differenzieren:

- **Fachinhaltlich:** Äußerungen dazu, welche Gegenstände und Objekte genutzt werden, wie diese beschaffen sind, welche naturwissenschaftlichen (insbesondere physikalischen) Erkenntnisse sich (vermutlich) über diese generieren lassen, welche empirischen oder mathematischen Zusammenhänge es zwischen (insb. physikalischen) Größen gibt und Ähnliches. Beiträge werden unabhängig davon kodiert, ob sie aus fachlicher Sicht angemessen sind (für eine fachlich unangemessene Aussage siehe letztes Beispiel in der nachfolgenden Auflistung).

Beispielsweise werden die folgenden Aussagen als fachinhaltliche Aktivitäten angesehen (alle Beispiele bezogen auf Versuche, die an einen Aufbau angelehnt sind, wie er im Anhang in Abbildung 5.2 dargestellt ist): »Das ging aber schnell [dass beim Pusten durch einen Strohhalm Luftblasen aus dem umgebenden Wasser aufsteigen]«; »Ich denke, in dem Glas mit der niedrigeren Wasserhöhe steigen zuerst Luftblasen aus dem Strohhalm auf«; »Je höher die Wassersäule über dem Strohhalm ist, desto geringer ist der Druck am Ende des Strohhalms«.

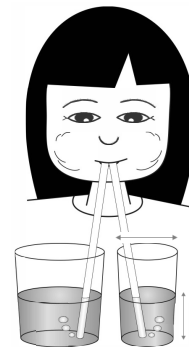


Abbildung 5.2: Skizze zu einem Versuchsaufbau aus der ersten Instruktionseinheit beider Varianten.

- **Fachmethodisch:** Äußerungen dazu, was methodische Entitäten (wie bspw. naturwissenschaftliche Fragen, Hypothesen, Variablen, Beobachtungen, Deutungen, Tabellen(strukturen)) kennzeichnet, wie sie beschrieben oder unterschieden werden können und wie methodische Entitäten und mit ihnen verbundene Untersuchungsplanungen und -durchführungen aus methodischer Sicht bewertet werden (können). Dabei wird gemäß Tabelle 4.2 (auf S. 54) zwischen den Teilprozessen (FH: Formulieren von Fragen und Hypothesen, PU: Planen und

Durchführen von Untersuchungen, AI: Auswerten und Interpretieren von Daten) unterschieden, weil die Formulierung der Indikatoren auf dieser Ebene besser gelangen (siehe Abschnitt B.1, ab S. 522). Beiträge werden unabhängig davon kodiert, ob sie aus fachmethodischer Sicht angemessen sind (nachfolgend nur Beispiele für eher angemessene Beiträge).

Beispielsweise werden die folgenden Aussagen als fachmethodische Aktivitäten angesehen: »Weil die Strohhalme nicht gleich tief im Wasser eingetaucht waren, ist der Vergleich für zwei verschiedene Strohhalm Durchmesser unzulässig« (PU); »Das ist doch keine naturwissenschaftliche Frage, ob es dir Freude macht, in die Strohhalme zu pusten« (FH); »Den Strohhalm Durchmesser musst du in die erste Tabellenspalte schreiben, weil das die unabhängige Variable ist« (AI); »Dass der Druck am Ende des Strohhalmes größer ist, ist eine Deutung« (AI).

- **Vorbereiten/Durchführen/Verbal:** Äußerungen dazu, was bei der Vorbereitung oder Durchführung eines Versuches praktisch oder technisch zu beachten ist.

Beispielsweise werden die folgenden Aussagen als verbale Aktivitäten zur Vorbereitung und Durchführung von Versuchen angesehen: »Du musst in beide Gläser gleich viel Wasser schütten«; »Ich nehme jetzt die beiden Gläser hier«; »Da stand doch eben, dass du beide Trinkhalme gleich tief eintauchen musst«; »Drei, zwei, eins, los – pusten«.

Überprüfungen des Kodiermanuals fanden in mehreren Durchgängen statt. Dabei wurde das Kodiermanual für alle drei Einheiten und für beide Instruktionsvarianten eingesetzt. Auch Diskussionen in der Forschungscommunity haben zum Ausschärfen der Kategorien beigetragen. Eine wesentliche Weiterentwicklung, die vor allem basierend auf den aufgezeichneten Äußerungen der Schülerinnen und Schüler stattgefunden hat, war die Unterscheidung fachmethodisch-indizierter und fachmethodisch-vermuteter Beiträge, die darauf bezogen ist, wie deutlich die Äußerungen von Lernenden im Zusammenhang mit Fachmethoden und insbesondere fachmethodischen Konzepten stehen:

- **Fachmethodisch-Indiziert:** Die auf den Seiten 69 und 70 als Beispiele für fachmethodische Beiträge aufgeführten Äußerungen enthalten (1) entweder einzelne Worte oder Begründungselemente, anhand derer sich recht deutlich zeigt, dass die Äußerung auf Fachmethoden bezogen ist (sogenannte – kontextabhängig festgelegte – Indizien; beispielsweise: »naturwissenschaftliche Frage«, »Deutung«, »unzulässiger Vergleich«). Auch Äußerungen von Lernenden, die (2) als fachmethodische Regeln oder Definitionen formuliert sind oder eindeutig auf eine solche verweisen, werden als fachmethodisch-indizierte Beiträge kodiert

5.1 Eventbasiertes Kodieren von verbalen u. nonverbalen Aktivitäten der Lernenden

(beispielsweise: »Du musst die Variablenkontrolle einhalten«; »Variablenkontrolle heißt, alle Variablen möglichst genau zu messen, also zu kontrollieren«; wie für die Gesamtkategorie **Fachmethodisch** werden auch bei der Subkategorie **fachmethodisch-indizierter** Beiträge die Kodierungen unabhängig von der aus fachmethodischer Sicht gegebenen Angemessenheit der Äußerung vorgenommen, wie das zweite Beispiel verdeutlicht).

- **Fachmethodisch-Vermutet:** Für zwei der oben gegebenen Beispiele zum verbalen Vorbereiten und Durchführen von Versuchen lässt sich vermuten, dass die Äußerungen auf Fachmethoden bezogen sind: »Du musst in beide Gläser gleich viel Wasser schütten« und »Da stand doch eben, dass du beide Trinkhalme gleich tief eintauchen musst«. Abhängig von der spezifischen Art (wie bspw. Betonung) und dem spezifischen (sehr eng umrissenen) Kontext der Äußerung könnte vermutet werden, dass die Äußerungen zum Konstanthalten der Wasserhöhe bzw. der Eintauchtiefe einen Bezug zu dem fachmethodischen Prinzip der Variablenkontrolle herstellen. In als fachmethodisch-vermutet kodierten Beiträgen finden sich keine Indizien (anders also als bei fachmethodisch-indizierten Beiträgen), weshalb der Art und die Kontext der Äußerung – auch bei den beiden gegebenen Beispielen – zur Kodierentscheidung hinzugezogen werden müssen.

Von den Kategorien **Fachinhaltlich**, **Fachmethodisch-Indiziert** und **VorbereitenDurchführenVerbal** wird den oben gegebenen Beschreibungen gemäß jeder Lernendenäußerung nur genau eine Kategorie zugewiesen. Die Kategorie **Fachmethodisch-Vermutet** ist jedoch ausdrücklich zeitgleich mit einer der beiden Kategorien **Fachinhaltlich** oder **VorbereitenDurchführenVerbal** zu kodieren (weil der fachmethodische Charakter einer fachlichen Äußerung parallel zum nahezu eindeutig identifizierbaren Bezug auf Fachinhalte oder auf Praktisch-Technisches *vermutet* wird).

Das Manual umfasst weitere Kategorien verbaler Aktivitäten auf der dritten Ebene des Kategoriensystems (**Vorlesen** von Text, **Organisatorisches:** Allgemeine Arbeitsprozessorganisation, **Sonstiges**; ausführliche Beschreibungen im Anhang) sowie außerdem Kategorien zu *nonverbalen Aktivitäten* (ebenfalls Ebene 3 des Kodiersystems), die unabhängig von den verbalen Aktivitäten (aber untereinander disjunkt) kodiert werden:

- **VorbereitenDurchführenNonverbal:** Zielgerichtete Manipulation von Versuchsmaterial, durch die ein selbst ausgedachter oder im Lernmaterial angelegter Versuch vorbereitet oder durchgeführt wird. Dazu gehören beispielsweise das Herstellen eines Versuchsaufbaus oder das vermutlich mit Erkenntnisinteresse

verbundene Pusten in einen Strohhalm. Herumspielen mit Versuchsmaterial wird nicht mit dieser Kategorie kodiert.

- **Schreiben:** Worte, Zeichnungen, Rechnungen etc. werden auf eigenen oder bereitgestellten Zetteln in Reinschrift oder als Kladde notiert.

Ferner findet sich eine Kategorie, die abbildet, in welchen Zeiträumen die Personen offensichtlich nicht-instruktionsbezogenen Aktivitäten nachgehen, weil sie beispielsweise über das Wetter oder die Farbe des Physiklehrerschals reden. Diese Kategorie trägt den Namen **OffTask**; sie schließt aus, dass zeitgleich andere Kategorien zu verbalen oder nonverbalen Aktivitäten für dieselbe Person kodiert werden (Ebene 2 des Kategoriensystems). Außerdem wurden Kategorien zum *Erleben* der Lernenden kodiert, diese stellen aber keinen Gegenstand der vorliegenden Arbeit dar und sind daher nicht erläutert (Ebene 4 des Kategoriensystems).

In der Übersicht über die Kategorien des Kodiermanuals in Abbildung 5.1 ist eine weitere Ebene aufgeführt (Ebene 1), innerhalb der die aktuell vom Team bearbeitete Karte abbildet wird (jede Karte eine Kategorie), so dass Zuordnungen der Aktivitäten der Lernenden zu den Karten der Instruktion möglich sind. Abschnitt B.1 enthält eine vollständige Darstellung des Manuals. Darin sind Beschreibungen, Indikatoren sowie reale und fiktive Beispiele angegeben. Beispielhaft ist ein Auszug aus dem Kodiermanual für die Kategorie **Fachinhaltlich** in Abbildung 5.3 aufgeführt.

Codes	Beschreibung	Indikatoren Die Schülerin / der Schüler ...	Beispiele
>Fachinhaltlich<	Äußerung oder Frage zu / Diskussion über fachinhaltsbezogene Themen, Konzepte, Erkenntnisse, Beispiele usw. Fachinhaltsbezogene Auseinandersetzung mit einem (Gedanken-)Versuch, z. B. Beobachtungen, Deutungen, ...	a) ... stellt eine inhaltsbezogene (nicht organisatorische) Frage.	a) – „Das hat doch was mit der Masse des Wassers zu tun, oder?“ – „Hängt die Fallgeschwindigkeit vom Gewicht ab?“
		b) ... trifft eine Vorhersage über den Ausgang eines Versuchs (mit und ohne Begründung).	b) – „Ich glaube, im linken Glas steigen zuerst Luftblasen auf“ – „Die Blasen werden größer sein, weil der Durchmesser vom Halm größer ist“
		c) ... äußert / diskutiert über eine Beobachtung (auch Messwerte) zu einem gerade durchgeführten oder erinnerten Versuch (ggf. mit Deutung).	c) – „Im linken Glas sind zuerst Blasen aufgestiegen!“ – „Da sind doch vorhin in beiden Gläsern gleichzeitig Blasen aufgestiegen“ – „Die Stoppuhr zeigt 1,2 Sekunden“
		d) ... überträgt ein inhaltsbezogenes Konzept auf einen konkreten Fall (z. B. zur Begründung einer Vorhersage oder Deutung einer Beobachtung).	d) – „Der große Kegel wird langsamer fallen, weil die Luftreibung größer ist“ – „In der Unterwasserhöhle ist der Druck kleiner, weil dort die Wassersäule kleiner ist“
		e) ... äußert / diskutiert ein inhaltsbezogenes Konzept.	e) – „Der Schweredruck hängt immer von der Höhe der Wassersäule ab“ – „Die Wassermenge im Topf hat Einfluss auf die Kochzeit“

Abbildung 5.3: Exemplarischer Auszug aus dem Kodiermanual für die Kategorie **Fachinhaltlich** mit Beschreibung, Indikatoren und Beispielen. Komplettes Manual in Anhang B.1 (ab S. 522).

5.1 Eventbasiertes Kodieren von verbalen u. nonverbalen Aktivitäten der Lernenden

Während für die explizit-fachmethodische Instruktionsvariante die Arbeitsaufträge und Informationen auf spezifische Aspekte von Fachmethoden bezogen sind (und somit durch den instruktionalen Kontext von Äußerungen erste Hinweise auf inhaltliche Bezüge dieser möglich sind), ist dies für die implizit-fachmethodische Variante nicht der Fall. Für die implizit-fachmethodische Instruktionsvariante wurde daher eine ergänzende Ebene (zwischen Ebene 3 und Ebene 4, nicht dargestellt in Abbildung 4.1) im Kodiermanual eingeführt, um die die Inhalte und Bezüge der fachmethodischen Beiträge spezifischer abzubilden. Jeder fachmethodische Beitrag (fachmethodisch-indiziert oder fachmethodisch-vermutet) wurde zusätzlich mit einer der in Tabelle 5.1 aufgelisteten Kategorien kodiert. Die Kategorien entsprechen im Wesentlichen den inhaltlichen Abschnitten der explizit-fachmethodischen Instruktionsvariante (siehe Tabelle 4.4) und wurden eingeführt, um grobe inhaltliche Vergleiche zwischen beiden Instruktionsvarianten hinsichtlich der Inhalte der fachmethodischen Beiträge zu ermöglichen. Beispielsweise wird der Äußerung »Das ist keine Antwort auf die hier gestellte Frage« die Kategorie **Passung von Versuch und Frage** zugewiesen.

Tabelle 5.1: Auf einer ergänzenden Ebene des Kategoriensystems (nicht dargestellt in Abbildung 5.1) für die implizit-fachmethodische Instruktionsvariante eingesetzte Kategorien zum Inhalt fachmethodischer Beiträge.

Kategorie	Beschreibung
Naturwissenschaftliche Fragen	Was macht naturwissenschaftliche Fragen aus? – Klassifizieren von und Kennzeichen von naturwissenschaftlichen Fragestellungen (auch in Abgrenzung zu nicht-naturwissenschaftlichen Fragen), ...
Präzise und allgemeine Fragen	Wie können sich naturwissenschaftliche Fragen voneinander unterscheiden? – Unterscheidung von präzisen und allgemeinen (naturwissenschaftlichen) Fragestellungen, von Wie- und Was-Fragen, Kennzeichen der jeweiligen Fragestellungstypen, ...
Vorhersagen: Vermutungen und Hypothesen	Wie können sich Vorhersagen voneinander unterscheiden? – Vorhersagen nach irgendwelchen Kriterien klassifizieren (bspw. vorerfahrungsbasiert, theoriebasiert, ...), ...
Passung von Frage und Vorhersage	Wie sind Frage und Vorhersage aufeinander zu beziehen? – Passung von Frage und Vermutung/Hypothese thematisiert, ...

(Fortsetzung auf der nächsten Seite)

Tabelle 5.1 (Fortsetzung)

Kategorie	Beschreibung
Variablen und ihre Kontrolle	Wie sind Messgrößen im Versuch zu behandeln? – Definitionen von Variablen, Zuweisung von Rollen für bestimmte Messgrößen (z. B. »Du musst das hier immer auf der gleichen Höhe lassen«), Prinzip der Variablenkontrolle, ...
Passung von Versuch und Frage	Wie sind Versuch und Frage aufeinander zu beziehen? – Passung von Frage und Versuch(saufbau) thematisiert, ...
Passung von Versuch und Vorhersage	Wie sind Versuch und Vorhersage aufeinander zu beziehen? – Passung von Frage und Vermutung/Hypothese, ...
Messen und Messbereiche	Was sollte man beim Messen beachten? – Messbereich thematisiert, ...
Tabellen und Diagramme	Was sollte man beim Anlegen von Tabellen und Diagrammen aus fachmethodischer Sicht beachten? – Einheiten in Tabellen/Diagrammen, Spalten/Zeilen in Tabellen, Achsenwahl in Diagrammen, ...
Beobachtungen und Deutungen	Unterscheiden von Sichtbarem und Unsichtbarem, von Subjektivem und Objektivem, ...
Passung von Deutungen zu Ergebnissen	Was muss man beachten, wenn man Ergebnisse zu einem Versuch berichtet? – Deutungen müssen zu dem Versuch, zur Hypothese, zur Fragestellung passen, ...
Erfolg und Sinnhaftigkeit von Untersuchungen	Wann werden Ergebnisse zu einem Versuch als Erfolg gewertet?

Neben der bereits in den Kategorienbeschreibungen thematisierten Frage, ob Kategorien disjunkt vergeben werden (grundsätzlich innerhalb jeder Ebene disjunkt, abgesehen von fachmethodisch-vermuteten Beiträgen), wurden für das eventbasierte Kodieren weitere *Festlegungen zum Vorgehen* vorgenommen (ausführlicher in Anhang B.1, insbesondere auf S. ??):

- Da grundsätzlich Personen einzeln (wenngleich in ihrer Teamkonstellation verortet) untersucht werden sollen, werden die bisher dargestellten Kategorien zu verbalen und nonverbalen Aktivitäten sowie zum OffTask (und zum Erleben) *für jede Person einzeln* kodiert (nur die teamweise bearbeiteten Karten werden auf Ebene des Teams kodiert). Da Teams aus maximal drei Personen bestehen, wurde dies technisch durch drei separate Kategoriensysteme realisiert. Im

5.1 Eventbasiertes Kodieren von verbalen u. nonverbalen Aktivitäten der Lernenden

Kodierprozess wurde jedoch davon abgesehen, die Videos dreifach zu betrachten; in den ersten Probedurchläufen stellte sich das einfache Durchgehen der Videos (mit entsprechend vielen Wiederholungsschleifen an kurzen Ausschnitten mit sehr vielen Kodierungen) dem dreifachen Durchgehen der Videos gegenüber ähnlich aufmerksamkeitsfordernd, aber weniger zeitintensiv dar.

- Für verbale Aktivitäten gilt (in Anlehnung an bei Rogge, 2010, dargestellte Befunde) eine *3-Sekunden-Regel*, die eine zeitliche Spezifikation dessen darstellt, was als *ein* zu kodierendes Segment angesehen wird: Haben aufeinanderfolgende Äußerungen derselben Person, denen die gleiche Kategorie zugewiesen wird, einen Abstand von weniger als 3 Sekunden, so werden sie als ein Segment kodiert. Haben Äußerungen, die zum gleichen Satz (oder zur gleichen Kategorie) gehören, einen Abstand von mehr als 3 Sekunden, so werden sie einzeln (und ggf. mit der Kategorie des gesamten Satzes) kodiert. Eine Folge der Regel ist, dass direkt aufeinanderfolgende Äußerungen, die verschiedene Aspekte in den Blick nehmen als ein Segment kodiert werden; ein Auszählen der kodierten Segmente nach verschiedenen Inhalten ist daher nur sehr eingeschränkt möglich. Für die Berechnung zeitlicher Anteile ergibt sich, dass kodierte Segmente mit Längen größer als 3 Sekunden im Durchschnitt vermutlich eher etwas zu lang sind, was für verbale Aktivitäten, die einzig von der Regel betroffen sind, aufgrund der meist kürzeren Dauer nur eine geringe Einschränkung darstellt (in Abbildung 5.4 beispielhaft ein Histogramm für die Länge fachmethodisch-indizierte Beiträge bei Einheit 1).

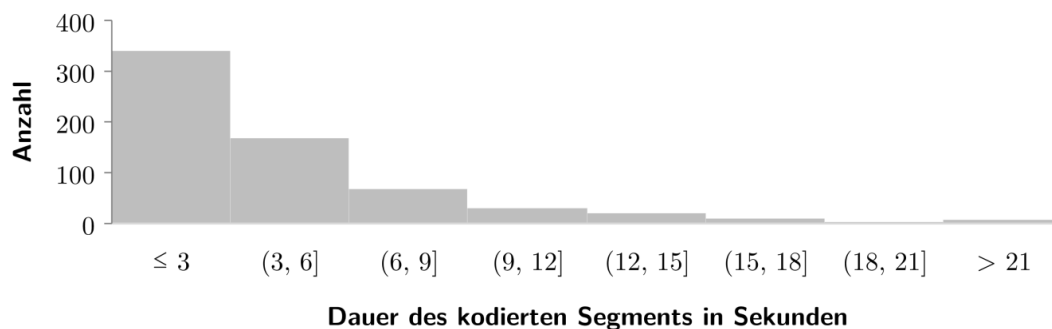


Abbildung 5.4: Histogramm zur Länge der mit Fachmethodisch-Indiziert kodierten Segmente in Einheit 1 der explizit-fachmethodischen Instruktionsvariante.

5.1.2 Methodisches zur Beurteilerübereinstimmungsberechnung

Um einschätzen zu können, welche Interpretationen anhand der mit dem Kodiermanual erzeugten Kodierungen (Details dazu in Unterabschnitt 5.3.1) und darauf basierenden statistischen Auswertungen (Details dazu in Unterabschnitt 5.3.2) mit welcher Sicherheit abgeleitet werden können, kann neben inhaltlichen, argumentativen Rechtfertigungen für das methodische Vorgehen und die darauf basierenden Interpretationen auch auf die sogenannte Beurteilerübereinstimmung (auch Inter-coderreliabilität) Bezug genommen werden. In der vorliegenden Arbeit ist allein schon deshalb auf die Beurteilerübereinstimmung einzugehen, weil mehr als eine Person die Kodierungen erzeugt hat (Details dazu in Unterabschnitt 5.1.3). Zwar setzen alle Personen denselben Kodierleitfaden ein, ob sie dies allerdings in derselben (oder in gut vergleichbarer) Weise tun, ist eine u. a. durch die Bestimmung der Beurteilerübereinstimmung adressierte Frage.

In der vorliegenden Arbeit wird die Beurteilerübereinstimmung mithilfe absoluter, relativer sowie zufallsbereinigter Maße angegeben. Alle genutzten Maße sind für den Vergleich von zwei Kodierungen (z. B. von zwei verschiedenen Personen) zu demselben Einschätzungsgegenstand (Videoausschnitt/Transkript) konzipiert. Seien also zwei Kodierungen gegeben: Kodierung 1 und Kodierung 2. Alle genutzten Beurteilerübereinstimmungsmaße basieren dann auf der Bestimmung der Übereinstimmung von einzelnen Items der Kodierung, wie nachfolgend beschrieben.

Übereinstimmung von Items bestimmen. Als ein *zu kodierendes Item* kann ein Zeitintervall eines Videos oder ein Textstück aus einem Transkript angesehen werden. Nachfolgende Ausführungen beziehen sich vor allem auf Zeitintervalle, lassen sich aber analog auch für Intervalle von Zeichen in einem Transkript lesen. Für die Berechnung aller genutzten Beurteilerübereinstimmungsmaße ist zu entscheiden, ob die Beurteiler für ein solches zu beurteilendes Item (d. h. Intervall) in derselben Weise entschieden haben oder voneinander abweichen. Diese Entscheidung erfolgt dichotom: ja/nein.

Als *ein auf Übereinstimmung zu prüfendes Item* wird im genutzten Programm MAXQDA ein Zeitintervall angesehen, für welches in einer der Kodierungen eine Kategorie zugewiesen ist. In MAXQDA kann eingestellt werden, ob eine der Kodierungen als Ausgangspunkt gewählt wird oder beide Kodierungen gleichberechtigt verglichen werden. Der zweite Fall wird in der vorliegenden Arbeit gewählt, weil nicht davon ausgegangen wird, dass eine der Kodierungen als korrekte Kodierung angesehen werden kann.³³ Wird die Übereinstimmung nun ausgehend von beiden Kodierungen

³³Für formative Beurteilerübereinstimmungsbetrachtungen im Training der Hilfskräfte (nicht in der Arbeit berichtet) wurde bspw. die Kodierung des Autors der Arbeit als Ausgangspunkt gewählt, weil davon ausgegangen wurde, dass diese einen angemessenen Orientierungspunkt darstellt.

beurteilt, wird für jedes in Kodierung 1 kodierte Segment überprüft, ob es mit einem kodierten Segment aus Kodierung 2 übereinstimmt, und umgekehrt für jedes in Kodierung 2 kodierte Segment, ob es mit einem kodierten Segment aus Kodierung 1 übereinstimmt. Kodierte Segmente wiegen bei dem geschilderten Vorgehen schwerer als unkodierte, und in beiden Kodierungen kodierte Segmente, die sich unterscheiden, wiegen schwerer als Nicht-Übereinstimmungen aufgrund von Nicht-Kodieren durch eine Person. Ausführlicher wird dies erläutert bei Klug (2017) sowie in der Dokumentation von MAXQDA (VERBI Software, 2019).

Da aufgrund des eventbasierten Vorgehens beim Kodieren *die Grenzen der zu kodierenden Intervalle nicht a priori festgelegt* sind und verschiedene Kodierungen demnach nicht in den kodierten zeitlichen Segmenten übereinstimmen müssen (bspw., weil zwei Kodierer nicht dieselben Anfangs- und Endpunkte für Kodierungen wählen), wird die dichotome Entscheidung bezüglich der Übereinstimmung von kodierten Items durch Rückgriff auf einen *Schwellwert* getroffen.³⁴ Im für die Arbeit genutzten Programm MAXQDA wird die Berechnung auf folgende Weise vorgenommen:

Sei, wie in Abbildung 5.5 visualisiert, $S_1 = [a, b]_1^c$ ein kodierte Segment aus Kodierung 1 mit Anfangspunkt a und Endpunkt b , dem die Kategorie c zugewiesen ist. Ob für dieses kodierte Segment eine Übereinstimmung vorliegt, wird daran festgemacht, ob die andere Kodierung ein kodierte Segment S_2 mit derselben Kategorie c enthält, welches mindestens einen Anteil $u > w$ mit S_1 überlappt, wobei w ein vorher festgelegter Schwellwert ist (bspw. 50 %). Der Anteil u wird als relativer Anteil der Zeitdauer des sogenannten Überlapps $S_1 \cap S_2$ (Schnittmenge der beiden Intervalle) an der von beiden Intervallen gemeinsam überdecken Zeit $S_1 \cup S_2$ (Vereinigung der beiden Intervalle) berechnet: $|S_1 \cap S_2| / |S_1 \cup S_2|$, wobei $|\cdot|$ die Gesamtlänge in Sekunden

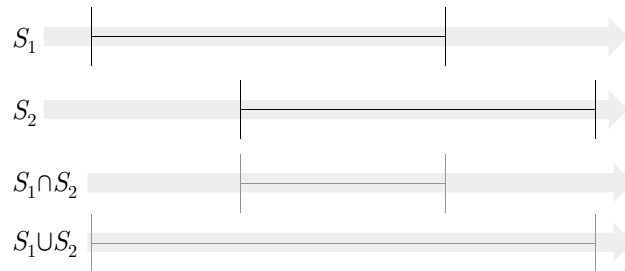


Abbildung 5.5: Visualisierung von zwei zu vergleichenden kodierten Segmenten sowie deren Schnittmenge und deren Vereinigung.

³⁴Prinzipiell könnte auch anhand einer mathematischen Funktion, die ein Übereinstimmungsmaß in Abhängigkeit vom gewählten Schwellwert beschreibt, auf weitere Aspekte der Übereinstimmung zweier Kodierungen eingegangen werden. Dazu würde beispielsweise zählen, wie gut Kodierer Intervallgrenzen bestimmen, oder auch, für welche Kategorien unterschiedliche Kodierer besonders häufig unterschiedliche oder unterschiedlich große Zeitintervalle kodieren. Derartige Überlegungen sind dem Autor der vorliegenden Arbeit nicht aus der Literatur bekannt; sie führen allerdings über den Rahmen der Arbeit hinaus.

angibt.³⁵ In Abbildung 5.5 ist der Anteil u beispielsweise geringer als 50 %, so dass die beiden Kodierungen 1 und 2 hinsichtlich des auf Übereinstimmung zu prüfenden Items nicht als übereinstimmend angesehen würden, wenn als Schwellwert ein w von 50 % gesetzt wäre.

Klug (2017) weist darauf hin, dass für kurze kodierte Segmente mit Dauern von wenigen Sekunden aufgrund von Toleranz (bspw. wegen Reaktionszeit) entsprechend geringe Schwellwerte w für als Übereinstimmung zu wertende Anteile u ($> w$) zu wählen sind. In der Fußnote auf S. 86 gibt er für zwei kodierte Segmente von jeweils 3 Sekunden Länge an, dass ein Schwellwert von $w = 33\%$ gewählt werden müsste, wenn ein Versatz von 1.5 Sekunden toleriert werden soll. In der vorliegenden Arbeit werden für unterschiedliche Kategorien unterschiedliche Schwellwerte genutzt. Für Kategorien zu verbalen Beiträgen und zum Off-Task wird ein Schwellwert von $u_{\text{verb}} = 30\%$ genutzt, weil die Beiträge grundsätzlich kurz ausfallen (siehe z. B. Abbildung 5.4). Ein Grund für einen niedrigen Schwellwert besteht darin, dass es für die vorliegende Arbeit vor allem auf die relative Lage der kodierten Segmente zueinander ankommt und nicht auf den genauen absoluten Zeitpunkt, allerdings besteht durch den niedrigen Schwellwert die Gefahr, zu einem kodierten Segment aus Kodierung 1 ein übereinstimmendes Segment in Kodierung 2 zu finden, obwohl die Kategorie in Kodierung 2 für eine andere Äußerung zugewiesen wurde. Aufgrund des getrennten Kodierens verschiedener Personen innerhalb desselben Videos hält sich das Risiko allerdings in Grenzen. Für Kategorien zu nonverbalen Aktivitäten der Lernenden wird ein Schwellenwert von $u_{\text{nv}} = 50\%$, für Kategorien zu aktuell bearbeiteten Karten der Instruktion ein Schwellenwert von $u_{\text{Karten}} = 90\%$ genutzt, weil diese Ereignisse typischerweise längere zusammenhängende Zeitsegmente abdecken, bei denen Toleranzen weniger bedeutend sind und vielmehr Abweichungen bezüglich der Grenzen identifiziert werden sollen, wenn sie entsprechend groß sind.³⁶

Absolutes Übereinstimmungsmaß: Anzahl der Items. Die absolute Übereinstimmung der beiden Kodierungen 1 und 2 ergibt sich als *Anzahl aller auf Übereinstimmung zu prüfenden Items, für die gemäß des jeweils festgelegten Schwellenwerts eine Übereinstimmung vorliegt*. Auch die Anzahl der (gemäß des festgelegten Schwel-

³⁵Bei intervallbasiertem Kodieren könnte jedes einzelne Intervall als ein zu beurteilendes Item angesehen werden. Ob eine Übereinstimmung vorliegt, wäre dann deutlich einfacher auszumachen.

³⁶Angenommen, eine Karte wird 60 Sekunden bearbeitet und in beiden Kodierungen 60 Sekunden kodiert. Dann würde ab einem Versatz von etwa 3.5 Sekunden nicht mehr von einer Übereinstimmung gesprochen. Ist der Beginn in beiden Kodierungen gleich, aber die Länge beider einen Kodierung kürzer, müsste das Ende in der einen Kodierung mehr als 6 Sekunden früher liegen als das in der anderen. Ferner ist anzumerken, dass ein zunächst angesetzter Wert von $u_{\text{nv}} = 70\%$ zu intolerant für kleine Abweichungen bezüglich des Beginns oder Endes dieser Aktivitäten war, weil sie deutlich kürzer andauern als die Kartenbearbeitungen.

lenwerts) nicht übereinstimmenden Items oder die Anzahl der auf Übereinstimmung zu prüfenden Items sollte in diesem Kontext möglichst angegeben werden.

Grundsätzlich werden aufgrund verschiedener Schwellwerte für die unterschiedlichen Ebenen des Kategoriensystems (Karten, Off-Task, verbale Aktivitäten, nonverbale Aktivitäten) eigenständige Werte für die Übereinstimmung angegeben, obwohl es durch einfache Addition der Anzahlen (die für sich genommen ja von dem jeweiligen Schwellwert der Entscheidungsregel unabhängig sind) möglich wäre, ein Gesamtmaß zu bilden und nur dieses anzugeben. Gründe für diese Unterscheidung liegen in der Verschiedenheit der Ebenen des Kategoriensystems auch hinsichtlich der im Kodiermanual beschriebenen Regeln und hinsichtlich der für Übereinstimmungen festgelegten Schwellwerte bezüglich des zeitlichen Überlapps.

Relatives Übereinstimmungsmaß: Prozentuale Übereinstimmung. Um verschiedene Beurteilerübereinstimmungen zu vergleichen, etwa zu verschiedenen analysierte Einheiten der Instruktion oder im Vergleich mit anderen Studien, eignet sich ein normiertes Maß besser als die Angabe absoluter Übereinstimmung. Ein üblicherweise angegebenes Maß ist die prozentuale Übereinstimmung

$$P_o = \frac{\text{Anzahl aller übereinstimmenden Items}}{\text{Anzahl aller zu prüfenden Items}}, \quad (5.1)$$

die zumeist als Wert vom Hundert angegeben wird (Wirtz & Caspar, 2002). Dabei werden in der vorliegenden Arbeit als übereinstimmende Items alle auf Übereinstimmung zu prüfenden Items angesehen, für die gemäß des jeweils festgelegten Schwellenwerts eine Übereinstimmung vorliegt. Als zu prüfende Items werden, wie oben erläutert, alle Zeitintervalle angesehen, für die in mindestens einer der Kodierungen ein kodiertes Segment vorliegt. Ähnlich wie bei den absoluten Übereinstimmungsmaßen werden die relativen Übereinstimmungsmaße für die Ebenen des Kategoriensystems einzeln angegeben.

Zufallsbereinigtes Übereinstimmungsmaß: $\kappa_B(\mathbf{n})$. Es ist möglich, dass Items aus Zufall übereinstimmen, weil einer der Kodierer zufällig dieselbe Kodierung gewählt hat wie der andere Kodierer. Ein häufig genutztes zufallsbereinigtes Maß für Beurteilerübereinstimmung bei nominalskalierten Kodierungen ist

$$\kappa = \frac{P_o - P_c}{1 - P_c}, \quad (5.2)$$

wobei P_o die in Gleichung 5.1 angegebene prozentuale Übereinstimmung (Index o für observed) und P_c die zufällige Übereinstimmung (Index c für chance) ist.

Die Berechnung von P_c , welches angegeben soll, mit welcher Wahrscheinlichkeit ein Kodierer zufällig dieselbe Kategorie wählt wie der andere oder zufällig die »richtige« Kategorie wählt, erfolgt auf verschiedene Weise. Der am häufigsten rezipierte Vorschlag kommt von Cohen (1960), der die Wahrscheinlichkeit anhand der in den realen Kodierungen vorliegenden relativen Häufigkeiten der Kategorien schätzt:

$$P_c = \sum_{j=1}^n h_{j.} \cdot h_{.j}, \quad (5.3)$$

wobei $h_{xy} = a_{xy}/N$ ist und a_{xy} = Anzahl der Items, denen in Kodierung 1 die Kategorie x und in Kodierung 2 die Kategorie y zugewiesen ist, und N = Anzahl aller zu prüfenden Items und n die Anzahl aller Kategorien des Kategoriensystems.

Einen Gegenvorschlag, der auf Probleme von Cohens Vorschlag bei ungleichen Randsummen eingeht, kommt von Brennan und Prediger (1981). Sie kritisieren an Cohens κ , dass zwei Kodierungen bei denen die Randverteilungen (d. h. die Anzahlen der einzelnen Kategorien für jede der beiden Kodierungen) einigermaßen übereinstimmen schwerer einen guten κ -Wert erreichen als zwei Kodierungen mit sehr verschiedenen Randverteilungen (S. 692). Stattdessen schreiben sie:

[W]hen marginals are free (i.e., not fixed a priori), it does not seem reasonable to allow the index of chance agreement to depend on the level of actual agreement in the marginals produced by the assigners.

When the marginal distributions are not specified a priori, each assigner might produce any set of marginals, consistent with the constraint that there are N objects to be classified into n categories. If each assigner randomly (in the sense of »blindly«) allocates objects to categories, then, for each assigner, the expected marginal proportion for each category is $1/n$. (Brennan & Prediger, 1981)

Daraus ergibt sich dann eine zufällige Übereinstimmung von

$$P_c(n) = \frac{1}{n} \quad (5.4)$$

und somit ein von der Anzahl der Kategorien abhängiges

$$\kappa_B(n) = \frac{P_o - P_c(n)}{1 - P_c(n)} \quad (\text{mit Index B für Brennan}).^{37} \quad (5.5)$$

³⁷Typischerweise wird das κ nach Brennan und Prediger (1981) mit κ_n bezeichnet. In der vorliegenden Arbeit wird allerdings ein eindeutigerer Name gewählt, um sowohl den Unterschied zu Cohens κ zu verdeutlichen als auch auf das genutzte n hinzuweisen.

Beim Vorschlag von Brennan und Prediger wird nicht berücksichtigt, dass gewisse Kategorien von sich aus bereits häufiger auftreten als andere. Kodierer könnten aber in einem solchen Fall (zumindest mit etwas Vorerfahrung) wissen, welche Randverteilung ungefähr zu erwarten sein dürfte. Das rigoroseste Vorgehen dürfte sein, Wirtz und Caspar (2002, S. 62–67) folgend zunächst die Randverteilungen für die zu vergleichenden Kodierungen miteinander abzugleichen, um den sogenannten Effekt unterschiedlicher Grundwahrscheinlichkeiten zu kontrollieren und anschließend auf den Effekt der mangelnden Konsistenz (ungleiches Verständnis der Kategorien oder abweichende Einschätzung der für das Zuweisen relevanten Merkmale) einzugehen, indem Cohens κ betrachtet wird.

In der Diskussion um Cohens κ und Alternativen spielt es unter anderem nur selten eine Rolle, dass aufgrund eventbasierten Vorgehens in einigen Fällen (wie in der vorliegenden Arbeit) keine Intervallgrenzen festgelegt sind (und bspw. nur durch den Schwellwert eine dichotome Entscheidung herbeigeführt wird). Im Manual zu MAXQDA (2020) wird tendenziell eher von zufallsbereinigten Übereinstimmungsmaßen abgeraten (vgl. S. 31–32), allerdings finden sich dort keine fundierten Studien dazu.

Da in der Entscheidung für ein Übereinstimmungsmaß viele Faktoren zu berücksichtigen sind, wird für die vorliegende Arbeit auch aus Zeitgründen das in MAXQDA bereits implementierte³⁸ Vorgehen zur Berechnung von $\kappa_B(n)$ von Brennan und Prediger (1981) genutzt, obwohl tendenziell eine Nutzung des häufig eingesetzten Maßes von Cohen durchaus plausibel erscheint. Eine Vergleichbarkeit der angegebenen κ -Werte innerhalb der Arbeit ist so gegeben; im Vergleich zu anderen Studien, die Cohens κ nutzen, sind die in der vorliegenden Arbeit angegebenen zufallsbereinigten Übereinstimmungen eher niedriger (vgl. Klug, 2017). In der Literatur vorhandene Vergleichswerte für Cohens κ (z. B. bei Wirtz & Caspar, 2002, S. 59, die ein $\kappa > .6$ als gut bezeichnen) sind ohnehin auf die spezifische Situation zu beziehen, in dem konkreten Fall des Gebrauchs von $\kappa_B(n)$ allerdings mit zusätzlicher Vorsicht heranzuziehen.

Ähnlich wie bei den absoluten und relativen Übereinstimmungsmaßen werden die zufallsbereinigten Übereinstimmungsmaße für die Ebenen des Kategoriensystems einzeln angegeben; entsprechend können auch unterschiedliche Interpretationen hinsichtlich der Vergleichswerte vorgenommen werden. Mit dieser Entscheidung ist eine Reduktion des jeweils zugrundeliegenden n verbunden, weil nicht durch die Anzahl aller verfügbaren, sondern durch die Anzahl aller auf der jeweiligen Ebene

³⁸ Anders als Brennan und Prediger (1981) vorschlagen, werden nicht-genutzte Kategorien von MAXQDA aus der Anzahl aller Kategorien ausgeschlossen, so dass das berechnete P_c ggf. größer ausfällt und das berechnete κ kleiner. Da der Rechenweg von MAXQDA eingesehen werden kann, wird eingegriffen und die Rechenvorschrift von Brennan und Prediger (1981) genutzt. (Da der Rechenweg für Cohens κ deutlich abweicht, wäre ein analoges Vorgehen mithilfe des Rechenwegs von MAXQDA nicht ohne deutlichen zusätzlichen Aufwand möglich.)

des Kategoriensystems vorhandenen Kategorien geteilt wird. Die genauen Anzahlen sind am Ende des vorliegenden Unterabschnitts vermerkt.

Weitere Entscheidungen für die vorliegende Arbeit. Aufgrund des mehrschichtigen Kodiermanuals werden für die Berechnung der Beurteilerübereinstimmung neben der oben dargestellten Wahl des jeweiligen Maßes weitere konkrete Entscheidungen getroffen.

Die Maße für die Beurteilerübereinstimmung werden, wie bereits erläutert, *für die Ebenen* des Kategoriensystems (Karten, Off-Task, verbale Aktivitäten, nonverbale Aktivitäten) *einzelnen* angegeben. Dies entspricht den Beschreibungen und Regeln im Kodiermanual, die die Ebenen deutlich voneinander trennen.

Sowohl auf der Ebene nonverbaler als auch auf der Ebene verbaler Aktivitäten die Sonstiges-Kategorien aus der Beurteilerübereinstimmung ausgenommen. Dies liegt darin begründet, dass als relevant angesehene Nicht-Übereinstimmungen bereits durch das Fehlen einer ggf. anderen Kategorie abgedeckt sind. Außerdem sollen Nicht-Übereinstimmungen aufgrund eines ggf. überflüssigen Sonstiges-Segments nicht gewertet werden.

Wie bei der Vorstellung der Kategorie **Fachmethodisch** erwähnt wurde (siehe ab S. 69; vgl. auch Anhang B.1, ab S. 522), sind die Beschreibungen für fachmethodische Beiträge in drei inhaltliche Blöcke zu den drei Teilprozessen des experimentbezogenen Denkens und Arbeitens gegliedert. Daraus ergeben sich für den Kodiervorgang *sechs fachmethodische Kategorien*, weil jeweils zwischen indizierten und vermuteten fachmethodischen Beiträgen unterschieden wird. Allerdings ist im Kodierprozess die Entscheidung zwischen fachmethodisch-indizierten und fachmethodisch-vermuteten Beiträgen prominenter und die Entscheidung, welchem Teilprozess ein fachmethodischer Beitrag zugeordnet wird, eher nachrangig. Auch im weiteren Verlauf der Arbeit wird vorrangig auf fachmethodisch-indizierte und fachmethodisch-vermutete Beiträge Bezug genommen, ohne die jeweilige Zuordnung zum Teilprozess vorzunehmen. Den Ausführungen entsprechend wird auch die Beurteilerübereinstimmung berechnet: es werden nur Kategorien fachmethodischer Beiträge unterschieden, die Teilprozesse werden nicht aufgelöst.

Mit fachmethodisch-vermuteten Beiträgen (FMv), die *ausdrücklich parallel* zu einer einer der Kategorien **VorbereitenDurchführenVerbal** (VDv) oder **Fachinhaltlich** (FI) *kodiert werden* sollen, wird Zwecks Einfachheit so verfahren, als würden eigenständige Segmente kodiert. Wenn also eine Person einen Beitrag mit FI und FMv kodiert, die andere aber nur mit FI, finden sich demnach zwei übereinstimmende und ein nicht-übereinstimmendes Item. Alternativen wären (a) die Einführung zusätzlicher

5.1 Eventbasiertes Kodieren von verbalen u. nonverbalen Aktivitäten der Lernenden

Kategorien FI+FMv und VDv+FMv *vor* (nicht mehr möglich) bzw. *nach* (nicht dem realen Vorgehen entsprechend) Anwendung durch die Kodierer oder (b) die Nutzung eines anderen Übereinstimmungsmaßes (nicht ohne Weiteres auffindbar). Das nun gewählte Vorgehen zur Bestimmung von Übereinstimmung bildet die realen Entscheidungsprozesse (zwei separate Entscheidungen: FI?, FMv?) am ehesten ab.

Eine letzte Erläuterung betrifft die Kodierung von (zwei bis) *drei Personen je Team*. Die Ebenen des Kategoriensystem zu Off-Task und verbalen und nonverbalen Aktivitäten werden bei Anwesenheit von drei Personen verdreifacht (jede inhaltliche Kategorie wird mit jeder Person gepaart), bei zwei Personen verdoppelt. Karten werden als teamweise bearbeitet angesehen und daher pro Video nur einmal kodiert. Beurteilerübereinstimmungen ergeben sich dann natürlicherweise aus dem kompletten, alle Lernenden umfassenden Kategoriensystem, weil auch Zuweisungen unterschiedlicher Personen als abweichende Kodierungen gewertet werden.^{39,40}

Vor dem Hintergrund der obigen Erläuterungen ergeben sich für die Ebenen des Kategoriensystems die folgenden Anzahlen von Kategorien:

$n_{\text{Karten, E1, expl}} = 31$	$n_{\text{Karten, E1, impl}} = 32$
$n_{\text{Karten, E2, expl}} = 54$	$n_{\text{Karten, E2, impl}} = 54$
$n_{\text{Karten, E3, expl}} = 53$	$n_{\text{Karten, E3, impl}} = 41$
$n_{\text{Off-Task}} = 3$	
$n_{\text{verbal}} = 18$	
$n_{\text{nonverbal}} = 6$	

³⁹Im Kodiertraining wurde bei Nicht-Übereinstimmung darauf geachtet, ob *nur* die zugewiesene Person abweichend war und (sofern keine allzu großen Fallzahlen auftraten) als entsprechend unproblematisch angesehen, weil dieser Fehler sehr wahrscheinlich auf ungenügender Audio- bzw. Video-Qualität oder ggf. auf mangelnder Konzentration beruhte. Für die endgültige Berechnung der Beurteilerübereinstimmung wird allerdings streng nach Personen unterschieden, weil die Auswertungen klare Trennungen der Kodierungen verschiedener Lernender erfordern. Für gemeinsam von mehreren Lernenden getroffene Aussagen liegen klare Regeln im Kodiermanual vor.

⁴⁰An dieser Stelle lässt sich zusammenfassend anmerken, dass die Kodierentscheidung, für die eine Übereinstimmung geprüft wird, nicht wie idealerweise angenommen *nur* die Entscheidung für eine der Kategorien umfasst, sondern insgesamt fünf Entscheidungen: 1. Anfangspunkt, 2. Endpunkt, 3. Person, 4. Kategorie, 5. (Anhand der 3-Sekunden-Regel mit anderer Kodierung zusammenziehen?).

5.1.3 Kodiertrainings und Beurteilerübereinstimmungen

Zur Kodierung des zeitlich umfangreichen Videomaterials (Details in Abschnitt 5.2, ab S. 88) wurden Hilfskräfte eingesetzt. Ein Teil des Videomaterials wurde von einer Hilfskraft kodiert, die inzwischen im Institut für Physikdidaktik doktoriert und stärker in die Entwicklung des Manuals eingebunden war (Hilfskraft 1). Aufgrund der Vorerfahrungen, die Hilfskraft 1 aus anderen Projekten mitbrachte, kodierten der Autor der Arbeit, einer der Betreuer und die Hilfskraft exemplarisch Auszüge sowie vollständige Videodateien und führten darauf basierend Diskussionen zur Weiterentwicklung des Manuals, die jeweils in einer modifizierten Version des Kodiermanuals mündeten. Als eine Passung der Kategorien zum videomaterial aus Sicht aller drei Personen grundsätzlich hergestellt war, erfolgte eine Beurteilerübereinstimmungsprüfung. Dafür wurde für jede Einheit ein komplettes Video eines jeweils nicht für die weiteren Auswertungen genutzten Teams von beiden Kodierern kodiert (abgesehen von den Karten, für die davon ausgegangen wurde, dass der Zeiteinsatz einer Person gespart werden kann) und die Beurteilerübereinstimmung gemäß Unterabschnitt 5.1.2 berechnet. Da sich die zweite Einheit der explizit-fachmethodischen Instruktionsvariante im Training als besonders interessant herausstellte und die Anwendung des Kodiermanuals für diese Einheit aus Sicht aller drei Personen gut gelang, wurde anschließend mit der finalen Kodierung für diese Einheit begonnen. Die Werte der für diese Einheit ermittelten Beurteilerübereinstimmungsmaße finden sich in Tabelle 5.2 (Einheit 2). Sie sind als gut einzuschätzen. Im folgenden kodierte Hilfskraft 1 weitere 4 Videos, die Grundlage der Analysen der Arbeit darstellen, der Autor der Arbeit kodierte 2 weitere Videos. Für einzelne schwierige Zuweisungsentscheidungen wurden dabei Absprachen vorgenommen, um einen Konsens zu erzielen.

Für die restlichen Kodierungen wurden zwei weitere Hilfskräfte (Hilfskräfte 2 und 3) eingesetzt, die die Grundlagen der physikdidaktischen Ausbildung erfolgreich durchlaufen haben. Diese wurden vom Autor der Arbeit intensiv geschult: Zunächst machten sie sich mit der eingesetzten Kodiersoftware (MAXQDA 2018; siehe VERBI Software, 2018) vertraut; danach lernten sie initial anhand einer schriftlichen Einführung die Kategorien des Kodiermanuals und anschließend die technischen Kodierregeln für das eingesetzte eventbasierte Kodierverfahren kennen (siehe Abbildung 5.6 auf S. 86); anhand von Probekodierungen, die die beiden Hilfskräfte sowie der Autor der Arbeit vornahmen, wurden ferner ausführliche Trainingsgespräche geführt und Ausschärfungen in den Formulierungen des Kodiermanuals vorgenommen.

Zwischendurch erfolgten Beurteilerübereinstimmungsprüfungen im Sinne formativer Reliabilitätsprüfungen (vgl. Mayring, 2013), anhand derer Abweichungen zwischen

Tabelle 5.2: Übersicht über die Beurteilerübereinstimmungsberechnung für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante.

Ebene	Anzahl d.Kat. n	Schwell- wert u	Absolut		Relativ	Zufallsber.
			Ja	Nein	P_o	$\kappa_B(n)$
<i>Einheit 1</i>						
Karten	31	90 %	54	3	94,74 %	0,946
Off-Task	3	30 %	0	1	0 %	–
Nonverbal	6	50 %	26	4	86,67 %	0,84
Verbal	18	30 %	398	122	76,54 %	0,752
Gesamt	58	divers	478	130	78,62 %	0,782
Mittelwerte					85,98 %	0,846
<i>Einheit 2</i>						
Karten	54	90 %	–	–	–	–
Off-Task	3	30 %	0	3	0 %	–
Nonverbal	6	50 %	52	17	75,36 %	0,704
Verbal	18	30 %	343	150	69,57 %	0,678
Gesamt	82	divers	395	170	69,91 %	0,695
Mittelwerte					72,47 %	0,691
<i>Einheit 3</i>						
Karten	53	90 %	84	18	82,35 %	0,82
Off-Task	3	30 %	36	16	69,23 %	0,538
Nonverbal	6	50 %	76	55	58,02 %	0,496
Verbal	18	30 %	605	232	72,28 %	0,706
Gesamt	80	divers	801	321	71,39 %	0,71
Mittelwerte					70,47 %	0,64

Anmerkungen. Angegeben sind die Anzahlen der einbezogenen Kategorien auf der jeweiligen Ebene, der Schwellwert für den minimalen Anteil des Überlapps als Notwendigkeit für Übereinstimmung von einzelnen Items sowie die absolute, relative und zufallsbereinigte Übereinstimmung auf jeder jeweiligen Ebene. Die Werte für Kategorien und absolute Übereinstimmungen in der Zeile *Gesamt* werden als Summen berechnet, daraus ergeben sich dann die Werte für die prozentuale Übereinstimmung und die zufallsbereinigte Übereinstimmung. Die *Mittelwerte* basieren auf allen Werten, die nicht Null sind.

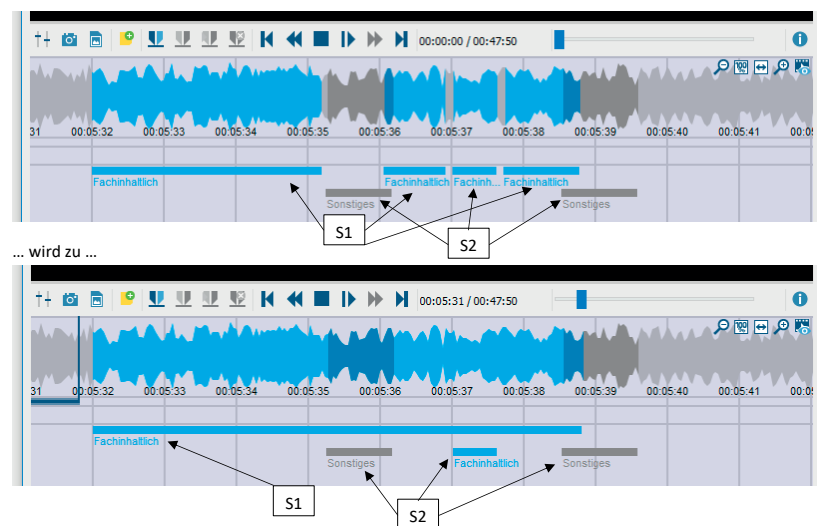
Mit der Unterscheidung von ›Methodisch-PU|vermutet‹ und ›Methodisch-PU|indiziert‹ haben Sie eine der wichtigsten Unterscheidungen des Kategoriensystems kennengelernt.

Prüfen Sie mit dem Auszug aus dem Manual erneut, ob Sie die Lösung (siehe unten) nachvollziehen können:

Jeder Beitrag muss außerdem einer der Kategorien rechts zugeordnet werden	>vermuteter Konzeptbezug< Art oder Kontext der Äußerung lässt vermuten, dass die Schülerin / der Schüler sich auf ein fachmethodisches Konzept (Kriterium, Definition, Kategorie, Regel) bezieht, es gibt jedoch keinen konkreten (indizierten) Hinweis.
	>indizierter Konzeptbezug< Einzelnes Wort oder Begründungselement im Kontext der Äußerung zeigt vergleichsweise deutlich, dass die Schülerin/der Schüler sich auf ein fachmethodisches Konzept (Kriterium, Definition, Kategorie, Regel) bezieht. Wird immer kodiert, wenn SuS selbst eine fachmethodische Regel formulieren!
– Kann parallel mit ›Vorb. / Durchf. (verbal)‹ kodiert werden, z. B. wenn Anweisung mit Konzept aus dem Bereich PU begründet wird	
– Kann parallel mit ›Fachinhaltlich‹ kodiert werden, wenn der Konzeptbezug, der die Aussage zu „Methodisch“ werden lässt, nur ein ›vermuteter Konzeptbezug‹ ist.	

Code	Aussage
M-PU indiziert	1: „Die Fallzeit ist hier die abhängige Variable.“
M-PU vermutet	2: „Dann lassen wir zwei gleich schwere Kegel aus gleicher Höhe fallen.“
M-PU indiziert	3: „Hier wird die Variablenkontrolle verletzt.“
M-PU indiziert	4: „Du musst beide Trinkhalme gleich tief eintauchen, sonst ist das kein fairer Versuch.“
Vorb./Durchf.(V)	5: „Dann lassen wir zwei Kegel runterfallen.“
M-PU indiziert	6: „Und dann verändern wir immer nur eine Variable gleichzeitig.“
M-PU vermutet	7: „Wir können ja mal die Höhe verändern – bei gleichem Durchmesser.“

(a)



(b)

Abbildung 5.6: Auszüge aus dem initialen Kodiertraining: (a) Zur Unterscheidung von fachmethodisch-indizierten und fachmethodisch-vermuteten Beiträgen (oben), (b) zur Anwendung der 3-Sekunden-Regel bei hochfrequentem Sprecherwechsel (unten); die genutzten Bezeichnungen weichen leicht von den in der Arbeit genutzten Bezeichnungen ab.

5.1 Eventbasiertes Kodieren von verbalen u. nonverbalen Aktivitäten der Lernenden

den Kodierungen der verschiedenen Personen ausführlich diskutiert wurden. Die abschließenden Beurteilerüberstimmungsprüfungen im Sinne summativer Reliabilitätsprüfungen (vgl. Mayring, 2013) erfolgten nach der Fertigstellung der Kodierungen, weil zu Beginn bereits ausreichend Hinweise aus den formativen Prüfungen vorlagen. Für die Einheiten 1 und 3 der explizit-fachmethodischen Instruktionsvariante wurde jeweils ein komplettes, in die weiteren Analysen der Arbeit nicht einbezogenes, Video von beiden Hilfskräften kodiert (inkl. Karten). Für die Beurteilerübereinstimmung ergaben sich gute bis sehr gute Werte (siehe Tabelle 5.2 auf S. 85). Der nur mäßig gute Wert für das κ der nonverbalen Aktivitäten bei Einheit 3 kommt u. a. durch die geringe Zahl verschiedener Kategorien (das n) zustande; die prozentuale Übereinstimmung ist als hinreichend anzusehen. Eine Sichtung der einzelnen Items zeigt, dass Nicht-Übereinstimmungen vor allem dadurch zustande kommen, dass die Kodierer an verschiedenen Stellen die Kategorie zum Schreiben zuweisen – eine Hilfskraft etwas häufiger als die andere. Hilfskraft 2 kodierte 11 Videos zu Einheit 1 und 8 Videos zu Einheit 3, Hilfskraft 3 kodierte 9 Videos zu Einheit 1 und 6 Videos zu Einheit 3. Die finalen Kodierungen wurden vom Autor der Arbeit gesichtet und in wenigen Einzelfällen angepasst.

Für die Einheiten der implizit-fachmethodischen Instruktionsvariante wurden keine gesonderten Prüfungen der Beurteilerübereinstimmung vorgenommen, weil die beiden Hilfskräfte 2 und 3 bereits anhand der explizit-fachmethodischen Instruktionsvariante umfassend trainiert wurden und nicht davon auszugehen war, dass die Anwendung des Kodiermanuals substantiell abweichende Herausforderungen mit sich bringen würde. Aus mündlichen Berichten der Hilfskräfte ging hervor, dass beim Anwenden des Kodiermanuals für die Videos zur implizit-fachmethodischen Instruktionsvariante keine Schwierigkeiten auftraten. Die Berichte enthielten zudem erste Hinweise auf Ergebnisse, weil bei der implizit-fachmethodischen Instruktionsvariante nur sehr wenige fachmethodische Beiträge kodiert wurden. Damit verbunden ist, dass die Hilfskräfte nur wenige Gelegenheiten hatten, die ergänzende Ebene zu den Inhalten der fachmethodischen Beiträge einzusetzen. Sichtungen aller zugehörigen Kodierungen durch den Autor der vorliegenden Arbeit zeigen, dass die Zuweisungen der auf die Inhalte der fachmethodischen Beiträge bezogenen Kategorien grundsätzlich in angemessener Weise erfolgt.

5.2 Wahl und Gruppierung der analysierten Lernenden

Die Videostichprobe enthält alle Lernenden aus der Gesamtstichprobe, die bei mindestens einer der drei Einheiten von einer der beiden analysierten Instruktionsvarianten videografiert wurden. Wie in Tabelle 4.6 berichtet, wurden nicht für alle Lernenden, die in der Gesamtstichprobe sind, auch Videoaufzeichnungen erzeugt. Dies geschah aufgrund von (inhaltlichen und forschungsökonomischen) Überlegungen zur Anzahl von Kameras pro Klassenraum und aufgrund von fehlenden Einverständniserklärungen (ausführlicher bei Vorholzer, 2016). Dennoch ist auch die Videostichprobe mit 71 videografierten Lernenden (in 24 Teams à 2–3 Personen) für die explizit-fachmethodische und 50 videografierten Lernenden (in 18 Teams à 2–3 Personen) für die implizit-fachmethodische Instruktionsvariante so umfangreich, dass für die Auswahl der Lernenden, die im mit der Arbeit verbundenen Projekt analysiert werden sollten, Kriterien festzulegen waren, entlang derer mit dem Kodieren der zugehörigen Videos begonnen wurde.

5.2.1 Ausgangslage

Wie in Vorholzer (2016) dargestellt, waren nicht alle Personen der Gesamtstichprobe zu allen Erhebungszeitpunkten anwesend. Ferner liegen nicht für alle Personen Werte für den Kompetenzzuwachs zum experimentbezogenen Denken und Arbeiten vor. Die Menge der Personen, für welche für den Kompetenzzuwachs Werte vorliegen, ist nicht deckungsgleich mit der Videostichprobe, weil von Vorholzer bewusst nur ein Teil der insgesamt teilnehmenden Population videografiert wurde. Dieser Sachverhalt ist in Abbildung 5.7 dadurch veranschaulicht, dass die beiden Kreise für *Erhebungen* und *Video* nicht deckungsgleich sind. In der Schnittmenge (der Personen, die sowohl bei den Erhebungen anwesend waren als auch mindestens einmal videografiert wurden) finden sich real 53 Personen, die die explizit-fachmethodische Instruktionsvariante bearbeitet haben, und 41 Personen, die die implizit-fachmethodische Instruktionsvariante bearbeitet haben.

Da die Instruktion aus drei Einheiten besteht, die an verschiedenen Tagen eingesetzt wurden, ist es möglich, dass Personen der Gesamtstichprobe bei manchen Instruktionseinheiten nicht anwesend sind. Die demnach vorliegende Situation ist in Abbildung 5.7 schematisch dargestellt: Personen, die bei Einheit 1 anwesend waren, sind durch die Ellipse *Einheit 1* abgebildet. Analog ist dies für Einheit 2 der Fall, während Einheit 3 zur Vereinfachung des Diagramms nicht dargestellt ist.⁴¹ Für

⁴¹Ferner sind manche Schnittmengen im Diagramm leer, dadurch dass manche Erhebungen in derselben Doppelstunde stattfanden wie die Bearbeitung von Einheit 1 der Instruktion. Diese

5.2 Wahl und Gruppierung der analysierten Lernenden

die Analysen, die in dem vorliegenden Kapitel berichtet werden, werden vorrangig Personen berücksichtigt, die bei allen drei Einheiten anwesend waren (innerhalb der Gesamtstichprobe sind dies 75 Personen für die explizit-fachmethodische, 72 Personen für die implizit-fachmethodische Instruktionsvariante; innerhalb der Videostichprobe sind es 57 Personen für die explizit-fachmethodische, 34 Personen für die implizit-fachmethodische Instruktionsvariante). In allen (Teil-)Stichproben, die in dem vorliegenden Kapitel berichtet/genutzt werden, befinden sich keine Personen, die nur bei *einer* Einheit anwesend waren (selbst dann nicht, wenn Videoaufzeichnungen nur für eine Einheit analysiert wurden); aufgrund der Teamstruktur finden sich jedoch einige Personen, die nur zwei Einheiten bearbeitet haben (genaue Aufschlüsselung in Unterabschnitt 5.2.3).

Die Teams, in denen die Schülerinnen und Schüler bei der Bearbeitung der Instruktions gearbeitet haben, wurden vor der Bearbeitung von Einheit 1 festgelegt und möglichst beibehalten. Daher ist es möglich, dass sich in einem Team eine Person befindet, die nur bei Einheit 1 (und Einheit 3, in der Abbildung nicht dargestellt) anwesend war, während eine zweite Person sowohl bei Einheit 1 als auch Einheit 2 (und ggf. auch Einheit 3) anwesend war. Personen, die bei Einheit 1 nicht anwesend

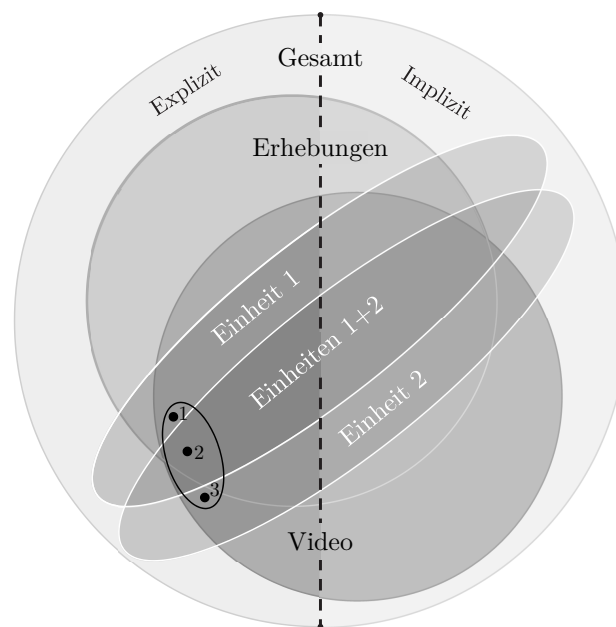


Abbildung 5.7: Schematisches Mengen-Diagramm zur Darstellung der Ausgangslage bezüglich der für Personen vorliegenden Daten (Drei Personen eines exemplarischen fiktiven Teams in schwarz eingetragen).

Randbemerkung ist aber für die folgenden Überlegungen zur Wahl und Gruppierung der Personen nicht von Bedeutung.

waren, wurden nur solchen Teams zugeordnet, die von vornherein aus zwei Personen bestanden oder bei denen in Einheit 2 oder 3 eine der Personen fehlte. Personen wurden in den Teams belassen, denen sie einmal zugeordnet wurden.

Da Teams natürlicherweise gemeinsam die Instruktion bearbeiten, gibt es keine Teams, die über die gestrichelte Linie in Abbildung 5.7 hinweg gehen (die Instruktionsvarianten-Gesamtstichproben sind disjunkt). Für die anderen Schnittlinien hingegen ist es möglich und auch häufig der Fall, dass Teams Personen aus voneinander getrennten Teilmengen enthalten. Das beispielhaft in schwarz eingezeichnete fiktiv-überspitzte Team enthält drei Personen, bearbeitet die explizit-fachmethodische Instruktionsvariante und es liegen Videoaufzeichnungen vor; die Personen unterscheiden sich allerdings in ihren Anwesenheiten:

- Person 1 Anwesend bei allen Erhebungen;
Anwesend bei Einheit 1, aber nicht Einheit 2
- Person 2 Anwesend bei allen Erhebungen;
Anwesend bei Einheit 1 und 2
- Person 3 Nicht bei allen Erhebungen anwesend;
Nicht bei Einheit 1 anwesend, aber bei Einheit 2

5.2.2 Vorgehen bei der Auswahl der Lernenden

Bei der Auswahl der analysierten Lernenden wurde teamweise vorgegangen, da Teams immer gemeinsam videografiert sind. Aufgrund des zeitlichen Ablaufs der Schulung der Hilfskräfte und des Kodierens (siehe Abschnitt 5.1) fand die endgültige Auswahl – nach grundsätzlicher Planung und Vorsortierung – zunächst für die Einheiten der explizit-fachmethodischen Instruktionsvariante, angefangen bei Einheit 2, gefolgt von Einheit 1 und Einheit 3, sowie zuletzt für die Einheiten der implizit-fachmethodischen Instruktionsvariante (gebündelt) statt. Dabei wurde grundsätzlich das Ziel verfolgt, Teams zu analysieren, in denen alle Lernenden bei allen drei Einheiten videografiert sind und möglichst viele Ergebnisse aus den Prä- und Post-Erhebungen vorliegen. Dazu wurde jedem Team, das in der Video-Stichprobe ist, basierend auf einem groben Rating ein Güte-Score zwischen 1 und 4 zugewiesen (4: Alle Werte aus den Erhebungen vorhanden und alle Personen für alle Einheiten anwesend; 1: Nur wenige Werte aus den Erhebungen vorhanden und nur wenige Personen für wenige Einheiten anwesend). Teams mit höheren Güte-Scores wurden immer bevorzugt gewählt.⁴²

⁴²Es war möglich und kam vor, dass bei der für die Güte-Score-Vergabe erfolgten sehr kurzen Video-Sichtung (insb. Prüfung, ob tatsächlich alle drei Personen da sind und der Ton funktioniert) andere

5.2 Wahl und Gruppierung der analysierten Lernenden

Nicht immer wurden allerdings alle für die jeweilige Einheit möglichen und für alle drei Einheiten analysierbaren Teams gewählt, weil auch weitere Kriterien zugrunde gelegt wurden, wie nachfolgend geschildert.

Aufgrund der im zugrundeliegenden Projekt verfolgten Ziele wurden jeweils Teams mit relativ hohen sowie mit relativ niedrigen (über die einzelnen Personen im Team gemittelten) Kompetenzzuwächsen ausgewählt, um Maximalkontraste hinsichtlich des Kompetenzzuwachses zu erlauben. Als relativ hoch bzw. relativ niedrig wurden der Orientierung halber dabei über die Personen innerhalb des jeweiligen Teams gemittelte Kompetenzzuwächse angesehen, die sich in Bezug auf die Videostichprobe im 35-Perzentil bzw. über dem 65-Perzentil aller teamweise gemittelten Kompetenzzuwächse befanden (dies entspricht etwa einer Drittelung aller Teams, die auf Video aufgezeichnet wurden, wobei auf das mittlere Drittel nachrangig zurückgegriffen wurde).⁴³

Es ist nicht davon auszugehen, dass deutliche Abnahmen der Kompetenzen zum experimentbezogenen Denken und Arbeiten während der Bearbeitung einer der eingesetzten Instruktionsvarianten auftreten. Wenn allerdings davon ausgegangen wird, dass plausiblerweise einige Personen Kompetenzzuwächse von etwa Null aufweisen, sind aufgrund von Messunsicherheiten des EDAW-Tests Streuungen im Bereich von Zuwächsen um die Null herum denkbar. Für Personen mit stark negativen Kompetenzzuwächsen ist allerdings eher davon auszugehen, dass die Werte für die Kompetenzzuwächse nicht den tatsächlichen Veränderungen in den Kompetenzen der Lernenden entsprechen (beispielsweise aufgrund von geringer Testmotivation o. Ä.). Für die teamweise stattfindende Auswahl der analysierten Lernenden wurde daher darauf geachtet, dass sich maximal eine Person mit negativem Kompetenzzuwachs in analysierten Teams befindet (wodurch zwei Teams ausgeschlossen wurden).

Als nachrangiges Kriterium wurde auch auf andere Ergebnisse aus den Erhebungen geachtet. Insbesondere wurde auf unterschiedliche Heterogenität innerhalb der Teams (also zwischen den Personen) hinsichtlich der Zuwächse und der Kompetenzen zum experimentbezogenen Denken und Arbeiten bei der Prä-Erhebung geachtet, um eine möglichst vielfältige Auswahl vorzunehmen. Bei Einheit 1 bzw. 3 wurde zudem jeweils beachtet, ob für dasselbe Team bereits Analysen für Einheit 2 bzw. die

Faktoren übersehen wurden, die letztlich dazu führten, dass Videos nicht analysiert wurden (bspw. zeitweise defekte Tonspuren oder fehlerhafte Test-Video-Zuordnung).

⁴³In der vorliegenden Arbeit werden für alle auf Erhebungen von Vorholzer (2016) basierenden Variablen (z. B. Kompetenzzuwächse, Prä-Kompetenzen zum experimentbezogenen Denken und Arbeiten) die dort berechneten Rasch-Measures (in logits) verwendet. Zur Normierung (Mittelwert 500, Standardabweichung 100) der entsprechenden Skalen und für weitere Informationen siehe Vorholzer (2016).

Einheiten 1 und 2 vorlagen. Obwohl vorausschauend ausgewählt wurde, war es aufgrund der Abwesenheit von Personen und der nachfolgend geschilderten weiteren Auswahlkriterien leider nicht immer möglich, diese Teams tatsächlich zu wählen, so dass in einigen Fällen auf alternative, hinsichtlich der Erhebungen vergleichbare, Teams zurückgegriffen werden musste.

Bevor die Kodierungen vorgenommen wurden, wurde geprüft, ob das Video der jeweiligen Einheit für das Team nutzbar ist. Teilweise lagen stark rauschende Audio-spuren vor, weshalb Teams aussortiert werden mussten. Bei kurzzeitigen Problemen in der Videospur wurde keine Aussortierung vorgenommen, sofern davon auszugehen war, dass in dieser Zeit wenig bis keine substantiellen nonverbalen instruktionsbezogenen Handlungen stattfanden (bspw. hakt in einem Fall das Video, während eine Person neues Material holt und die anderen beiden Personen über Dinge des Alltags sprechen).

Zusammenhängend mit dem geschilderten Vorgehen und basierend auf der Überlegung, dass die erste Einheit der Instruktion vermutlich gewissermaßen mitbestimmt, wie die anderen Einheiten bearbeitet und verstanden werden, wurde bei Einheit 1 eine größere Stichprobe angestrebt (was sich auch organisatorisch mit dem Beginn der Arbeit der Hilfskräfte 2 und 3 gut vereinbaren ließ). Die Stichprobengröße bei Einheit 2 wurde – obwohl eine Ergänzung durch weitere Lernende hilfreich schien – nicht mehr nachträglich angepasst, um nicht zusätzlich weitere Personen in die Kodierungen für diese Einheit einzubeziehen.

5.2.3 Analyisierte Teilstichproben

Aus dem geschilderten Vorgehen gehen fünf Video-Teilstichproben hervor (E_i steht dabei für die i te Einheit der jeweiligen Instruktionsvariante):

- Explizit-E1-Video: 33 Lernende
(in 11 Teams à 3 Personen)
- Explizit-E2-Video: 19 Lernende
(in 7 Teams: 5 Teams à 3 Personen und 2 Teams à 2 Personen)
- Explizit-E3-Video: 22 Lernende
(in 8 Teams: 6 Teams à 3 Personen und 2 Teams à 2 Personen)
- Implizit-E1bE3-Video (E1 bis E3): 15 Lernende
(in 6 Teams: 3 Teams à 3 Personen und 3 Teams à 2 Personen)

Die Schnittmengen der Video-Teilstichproben zur explizit-fachmethodische Instruktionsvariante stellen sich so dar: Es wurden 14 Personen (in 5 Teams) für $E1$ und

5.2 Wahl und Gruppierung der analysierten Lernenden

E2, 22 Personen (in 8 Teams) für *E1* und *E3* sowie 9 Personen (in 4 Teams) für *E2* und *E3* analysiert. Wenngleich die meisten der analysierten Personen alle drei Einheiten *bearbeitet* haben (54 Personen⁴⁴ von insgesamt 66 Personen), finden sich in der Schnittmenge der drei *Video*-Teilstichproben zur explizit-fachmethodischen Instruktionsvariante nur 9 Personen (ein Dreierteam und dreimal zwei Personen aus jeweils Dreierteams). Für eine Person dieser neun liegen keine Erhebungen zu fachinhaltlichen Kompetenzen, zu allgemein-kognitiven Fähigkeiten und zu emotional-motivationalen Aspekten vor. Bei Betrachtung aller *Video*-Teilstichproben liegen für alle Personen bis auf eine Person Werte zum Kompetenzzuwachs vor; Werte zu fachinhaltlichen Kompetenzen, zu allgemein-kognitiven Fähigkeiten und zu emotional-motivationalen Aspekten fehlen für 4 Personen. Für die implizit-fachmethodische Instruktionsvariante wurden alle analysierten Lernenden für die ersten beiden Einheiten analysiert, zusätzlich waren 3 Lernende (aus drei verschiedenen Teams) bei den ersten beiden Einheiten, aber nicht bei Einheit 3 anwesend. Zur Bezeichnung der 18 Lernenden, die allesamt die Einheiten 1 und 2 der implizit-fachmethodischen Instruktionsvariante bearbeitet haben (aber nur zum Teil Einheit 3), wird auch der Name *Implizit-E1bE2-Video* genutzt (18 Lernende in 6 Teams à 3 Personen). Eine Übersicht darüber, welche Personen für welche Einheiten kodiert sind und welche Daten für sie vorliegen, findet sich in Tabelle 5.3 auf Seite 94.

⁴⁴Bezogen auf beide Instruktionsvarianten.

Tabelle 5.3: Übersicht zu den Video-Teilstichproben: Anwesenheiten und Analysen.
(Testwerte aus den Erhebungen finden sich im Anhang in Tabelle A.2.)

Person	Bearb. Einh.	Analysiert für			Erhebungen vollständig
		E1	E2	E3	
Explizit-fachmethodische Instruktionsvariante					
A1	3	x	x		x
A2	3	x	x		(x)
A3	3	x	x		x
B4	3	x	x	x	x
B5	3	x	x	x	x
B6	2	x	x		x
C7	3		x		(x)
C8	2				
C9	3		x		x
D10	2	x		x	x
D11	3	x	x	x	x
D12	3	x	x	x	(x)
E13	2		x		
E14	2		x		x
E15	2		x		x
F16	3	x	x	x	x
F17	2	x	x		x
F18	3	x	x	x	x
G19	3	x	x	x	x
G20	3	x	x	x	x
G21	3	x	x	x	x
Fortsetzung rechts.					

Person	Bearb. Einh.	Analysiert für			Erhebungen vollständig
		E1	E2	E3	
Explizit-fachmethodische Instruktionsvariante (Forts.)					
H22	3	x		x	x
H23	3	x		x	x
H24	3	x		x	x
I25	3	x			x
I26	3	x			(x)
I27	3	x			x
J28	3	x		x	x
J29	3	x		x	x
J30	3	x		x	x
K31	3	x			x
K32	3	x			x
K33	3	x			x
L34	3	x		x	x
L35	3	x		x	x
L36	3	x		x	x
M37	3	x		x	x
M38	3	x		x	x
M39	3	x		x	x
Implizit-fachmethodische Instruktionsvariante					
Q50	3	x	x	x	x
Q51	2	x		x	x
Q52	3	x	x	x	x
R53	2	x		x	x
R54	3	x	x	x	x
R55	3	x	x	x	x
S56	2	x		x	x
S57	3	x	x	x	x
S58	3	x	x	x	x
T59	3	x	x	x	x
T60	3	x	x	x	x
T61	3	x	x	x	x
U62	3	x	x	x	x
U63	3	x	x	x	x
U64	3	x	x	x	x
V65	3	x	x	x	x
V66	3	x	x	x	x
V67	3	x	x	x	x

Anmerkungen. Bearb. Einh. = Anzahl bearbeiteter (aber nicht notwendigerweise auch analysierter) Einheiten. Leeres Feld = Nicht analysiert bzw. nicht anwesend bei Erhebung; x = Anwesend bzw. Werte für alle Erhebungen liegen vor; (x) = Werte zu EDAW-Erhebungen liegen vor (insb. also der Kompetenzzuwachs), aber andere Erhebungen sind nicht vollständig. Personenkürzel enthalten einen

5.2.4 Gruppierung der analysierten Lernenden

In den Forschungsfragen zu den Aktivitäten wird insbesondere nach Unterschieden und Gemeinsamkeiten für Lernende mit hohem vs. niedrigem Kompetenzzuwachs gefragt. Welche Kompetenzzuwächse als hoch oder niedrig angesehen werden, hängt jedoch vom gewählten Vergleichspunkt ab. Prinzipiell sind verschiedene Alternativen als Vergleichspunkte denkbar. Zum einen können verschiedene statistische Maße genutzt werden. So könnte *hoch* beispielsweise »über dem Mittelwert« bedeuten. Aber auch Perzentile könnten zur Klassifizierung genutzt werden. Zum anderen ist zu entscheiden, welche Stichprobe als Vergleichsgruppe genommen wird. Ist der Kompetenzzuwachs einer Person beispielsweise dann hoch, wenn er über dem Mittelwert der Gesamtstichprobe liegt oder über dem der Instruktionsvarianten-Gesamtstichprobe (also dem Mittelwert über alle Lernenden, die eine der beiden eingesetzten Instruktionsvarianten bearbeitet haben) oder über dem der Video-Teilstichprobe zu der in der jeweiligen Analyse betrachteten Einheit? Eine Person kann beispielsweise im Vergleich zur Gesamtstichprobe einen hohen Kompetenzzuwachs aufweisen, innerhalb der Instruktionsvarianten-Teilstichprobe zur explizit-fachmethodischen Instruktion liegt dieser Kompetenzzuwachs jedoch ggf. in der Nähe des Mittelwerts (was ein plausibles Beispiel ist, weil die Bearbeitung der explizit-fachmethodischen Instruktionsvariante im Mittel zu höheren Kompetenzzuwächsen führt als die Bearbeitung der implizit-fachmethodischen Instruktionsvariante).

Für die vorliegende Arbeit werden Lernende sowohl personen- als auch teamweise nach ihren Kompetenzzuwächsen gruppiert, weil es zur Beantwortung der Forschungsfragen unter anderem relevant sein könnte, ob Lernende für sich genommen hohe Kompetenzzuwächse erreichen oder ob sie in Teams arbeiten, in denen insgesamt gemittelt hohe Kompetenzzuwächse erreicht werden. Die jeweils genutzten Vergleichspunkte und die sich ergebenden Gruppierungen sind nachfolgend dargestellt.

5.2.4.1 Gruppierung einzelner Personen

Für die Analysen zu den Aktivitäten wird die Gruppierung basierend auf dem Mittelwert der Kompetenzzuwächse aller Personen vorgenommen, die alle drei Einheiten der jeweiligen Instruktionsvariante bearbeitet haben (Instruktionsvarianten-Gesamtstichprobe; hellgrau eingetragen in Abbildung 5.8). Der genutzte Mittelwert beträgt für die explizit-fachmethodische Instruktionsvariante 71.75 (Standardabweichung: 75.42), für die implizit-fachmethodische Instruktionsvariante 33.46 (Standardabweichung: 59.12). Grund für die bei der Mittelwertberechnung vorgenommene Beschränkung auf Personen, die alle drei Einheiten bearbeitet haben, ist, dass Personen, die nur

an einer oder zwei Einheiten teilgenommen haben, im Mittel deutlich geringere Kompetenzzuwächse aufweisen, was sich (zumindest für die explizit-fachmethodische Instruktionsvariante) anhand der Gesamtanlage von Instruktion und Test plausibel auf ihr Fehlen in manchen Einheiten zurückführen lässt (vgl. Vorholzer, 2016; Vorholzer & von Aufschnaiter, 2019a). Weil die Gruppierung für alle Lernenden vorgenommen wird, die alle drei Einheiten der jeweiligen Instruktionsvariante bearbeitet haben, ergibt sich, dass die Gruppierungsgrenzen unabhängig davon sind, welche Einheit der jeweiligen Instruktionsvariante betrachtet wird – in Abbildung 5.8 sind daher die hellgrauen Histogramme sowie die eingetragenen Trennlinien für alle Teilstichproben zur jeweiligen Instruktionsvariante gleich, während aufgrund der unterschiedlichen Stichproben(größen) die dunkelgrauen Histogramme von Teilstichprobe zu Teilstichprobe verschieden sind.

Da die analysierten Lernenden grundsätzlich so ausgewählt sind, dass sich im Mittelfeld der Kompetenzzuwächse nur wenige Personen befinden (siehe Unterabschnitt 5.2.2), wurde für die Video-Teilstichproben zur explizit-fachmethodischen Instruktionsvariante nach den Personen in der Nähe des mittleren Kompetenzzuwachses (71.75) gesucht. Der nächst niedrigere Kompetenzzuwachs findet sich für H24 mit 56.08. Um die kleinen Fallzahlen in den Gruppierungen von Personen mit hohem bzw. niedrigen Kompetenzzuwächsen nicht allzu stark zu beschränken, wurde diese Person (Abstand 15.67 vom Mittelwert) noch zur Gruppe der Personen mit niedrigen Kompetenzzuwächsen gezählt (markiert durch die linke gestrichelte Linie in Abbildung 5.8). Die Person mit dem nächst niedrigeren Kompetenzzuwachs ist F17 mit 43.49 (Abstand 28.26 vom Mittelwert). Im Bereich oberhalb des mittleren Kompetenzzuwachses werden Personen, die weniger als eine halbe Standardabweichung über dem Mittelwert liegen (bis 109.46, also mit einem Abstand von 37.71) zur Gruppe der Personen mit mittleren Kompetenzzuwächsen gezählt (markiert durch die rechte gestrichelte Linie in Abbildung 5.8).⁴⁵ Personen mit negativem Kompetenzzuwachs sind gesondert gruppiert, sofern der Kompetenzzuwachs kleiner als Null minus eine halbe Standardabweichung (also kleiner als -37.71) ist; andere Personen mit (nur leicht) negativem Kompetenzzuwachs sind in die Gruppe der Personen mit niedrigem Kompetenzzuwachs eingeordnet, da aufgrund von Messunsicherheiten eher davon auszugehen ist, dass für diese Personen weder substantielle Verbesserungen noch Verschlechterungen vorliegen.

⁴⁵Eine halbe Standardabweichung wurde gewählt, weil etwa dieser Wert (genauer: 0.46 Standardabweichungen) für normalverteilte Daten zu einer Drittelung der Stichprobe führt. Die Obergrenze der Gruppe von Personen mit niedrigen Kompetenzzuwächsen ist zwar nicht durch die Subtraktion der halben Standardabweichung vom Mittelwert definiert, liegt aber in der Nähe des entsprechenden Wertes.

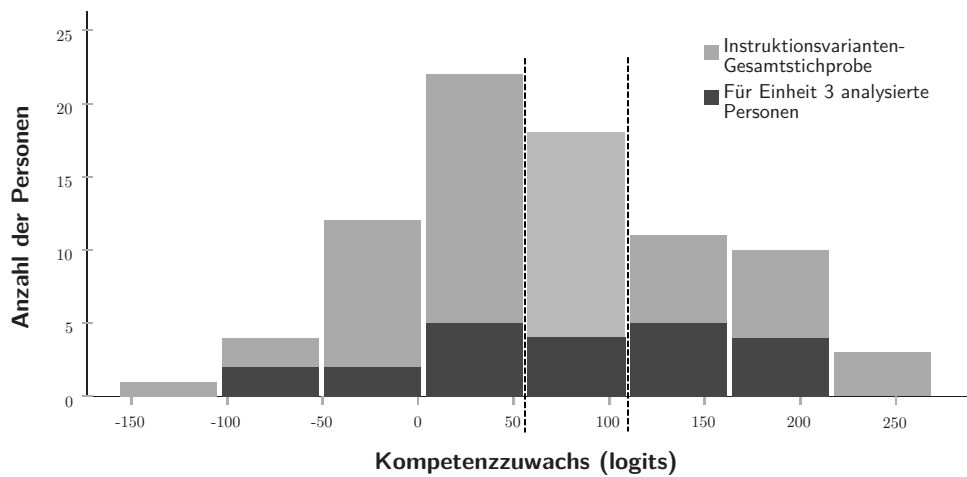
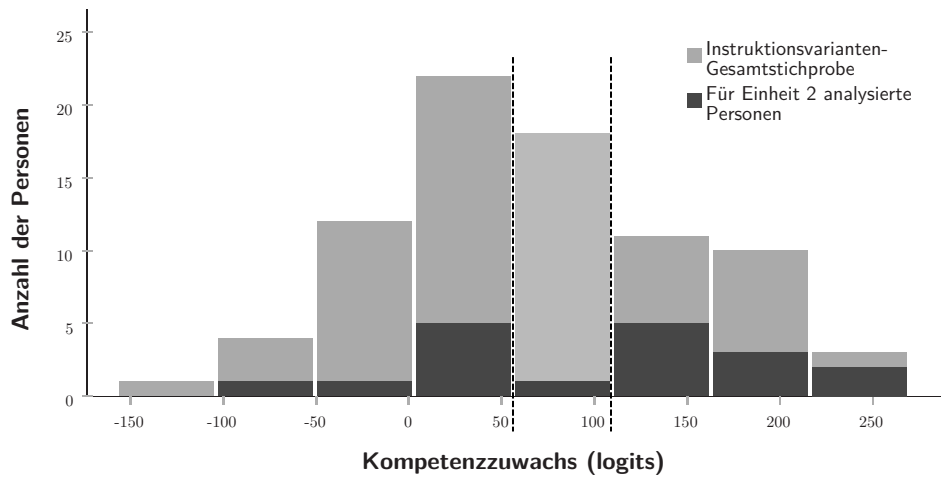
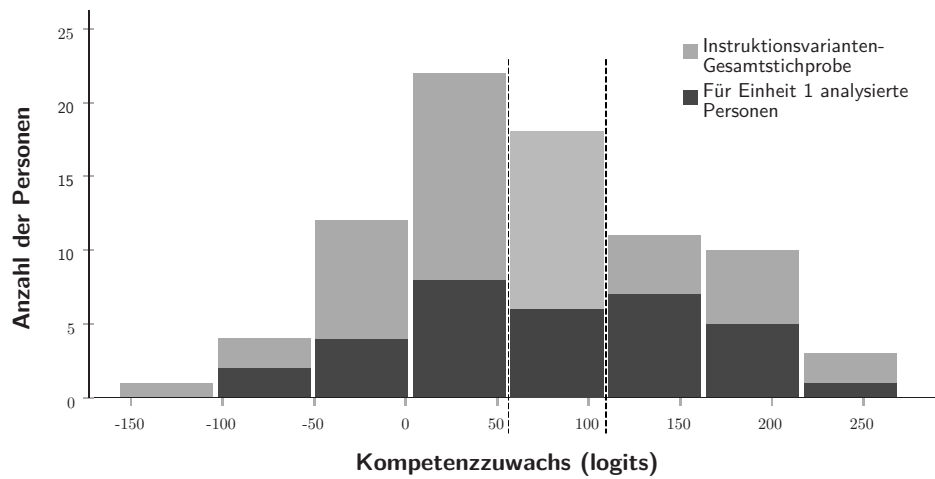


Abbildung 5.8: Gruppierung der Video-Teilstichproben hinsichtlich der vorliegenden Kompetenzzuwächse und grafischer Vergleich mit den jeweiligen Instruktionsvarianten-Gesamtstichproben. *Teil 1: Explizit-fachmethodische Instruktionsvariante.*

Personen mit negativen und niedrigen Kompetenzzuwächsen finden sich links von der ersten, Personen mit höheren Kompetenzzuwächsen rechts von der zweiten gestrichelten Linie. Dazwischen finden sich Personen mit mittleren Kompetenzzuwächsen.

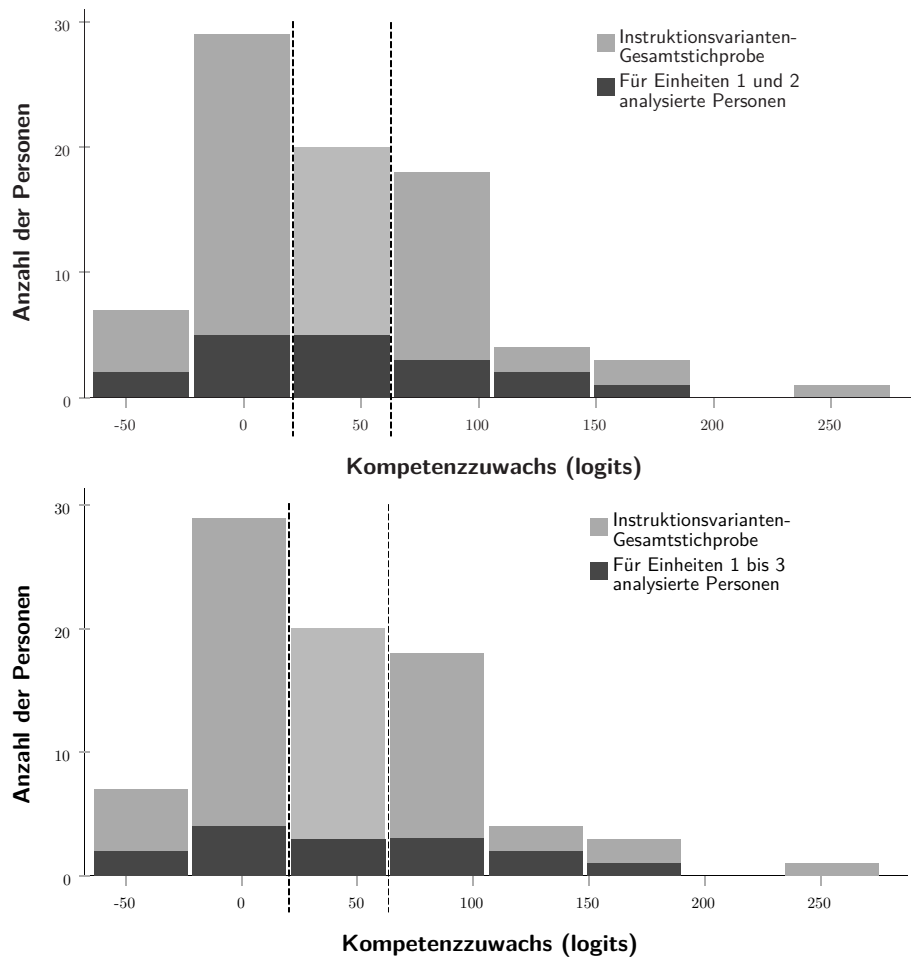


Abbildung 5.8: Gruppierung der Video-Teilstichproben hinsichtlich der vorliegenden Kompetenzzuwächse und grafischer Vergleich mit den jeweiligen Instruktionsvarianten-Gesamtstichproben. *Teil 2: Implizit-fachmethodische Instruktion*.

Personen mit negativen und niedrigen Kompetenzzuwächsen finden sich links von der ersten, Personen mit höheren Kompetenzzuwächsen rechts von der zweiten gestrichelten Linie. Dazwischen finden sich Personen mit mittleren Kompetenzzuwächsen.

5.2 Wahl und Gruppierung der analysierten Lernenden

Das Vorgehen zur Gruppierung für Lernende, die die implizit-fachmethodische Instruktionsvariante bearbeitet haben, ist vergleichbar mit dem Vorgehen für die Gruppierungen bei der explizit-fachmethodischen Instruktionsvariante. Die Grenzen sind in Relation zu all den Lernenden festgelegt, die alle drei Einheiten der implizit-fachmethodischen Instruktionsvariante bearbeitet haben. Als untere Grenze für hohe Kompetenzzuwächse wird eine halbe Standardabweichung über dem Mittelwert ($33.46 + 29.56 = 63.02$) gewählt (markiert durch die rechte gestrichelte Linie in Abbildung 5.8), während als niedrige Kompetenzzuwächse solche eingruppiert werden, die mehr als 10 logits unter dem Mittelwert liegen (der höchste Wert in dieser Gruppe beträgt 20.39; markiert durch die linke gestrichelte Linie in Abbildung 5.8). Als stark negative Kompetenzzuwächse, die aus der Gruppe der niedrigen Kompetenzzuwächse ausgeschlossen werden, gelten Werte, die unter demselben absoluten Wert wie für die explizit-fachmethodische Instruktionsvariante liegen (also unter -37.71).

In Abbildung 5.8 ist dargestellt, wie die Verteilungen der Kompetenzzuwächse für die jeweiligen Video-Teilstichproben aussehen und wie sie sich zur jeweiligen Instruktionsvarianten-Gesamtstichprobe verhalten. Die genauen Kompetenzzuwächse der Personen, anhand derer die Gruppierung vorgenommen wurde, finden sich in Anhang A.2 (ab S. 517).

5.2.4.2 Gruppierung von Teams

Da die Lernenden in Teams arbeiten und die Daten somit eine Teamstruktur haben, werden nicht nur einzelne Personen, sondern auch Teams gemäß der zugehörigen Kompetenzzuwächse gruppiert. Dafür wird gezählt, wie viele Personen eines Teams einen hohen, wie viele einen mittleren, wie viele einen niedrigen und wie viele einen (stark) negativen Kompetenzzuwachs aufweisen. Überwiegt die Zahl der Personen mit hohem (mittlerem/niedrigem) Kompetenzzuwachs gemäß der Personen-Gruppierung, wird das gesamte Team der Gruppe der Teams mit *gemittelt hohem* (mittlerem/niedrigem) Kompetenzzuwachs zugewiesen.

Für die Gruppierung der Teams, die die explizit-fachmethodische Instruktionsvariante bearbeitet haben, wurden drei Sonderfälle gesondert entschieden: Das Team F (alles außer mittel jeweils einmal vertreten) wird aufgrund der breiten Streuung als gemittelt mittlerer Team-Kompetenzzuwachs gruppiert und in den meisten Analysen daher nicht berücksichtigt (weil gemäß der Forschungsfragen zumeist Vergleiche für Teams mit hohen vs. niedrigen gemittelten Kompetenzzuwächsen vorgenommen werden). Das Team L (alles außer hoch vertreten) wird aufgrund zweier niedriger Werte (einer davon stark negativ) als gemittelt niedriger Team-Kompetenzzuwachs

gruppiert. Das Team C wurde nur für Einheit 2 kodiert, ferner sind nur zwei Personen anwesend. Da die Gruppe der Teams mit niedrigen Kompetenzzuwächsen für Einheit 2 sonst nur ein Team umfassen würde, wurde Team C dieser Gruppe zugeordnet, wenngleich eine Person einen mittleren und eine Personen einen niedrigen Kompetenzzuwachs aufweist. Die vorgenommenen Gruppierungen sind in Anhang B.2 (ab S. 532) aufgeführt.

Vergleiche von Teams mit gemittelt niedrigen Kompetenzzuwächsen und gemittelt hohen Kompetenzzuwächsen basieren auf sehr kleinen Fallzahlen. Für die Einheit 1 der explizit-fachmethodischen Instruktionsvariante können vier Teams mit vier Teams verglichen werden, für Einheit 2 zwei Teams mit drei Teams und für Einheit 3 drei Teams mit drei Teams (Teams mit gemittelt mittleren Kompetenzzuwächsen nicht erwähnt).

Für die Gruppierung der Teams, die die implizit-fachmethodische Instruktionsvariante bearbeitet haben, sind die Entscheidungen nicht so eindeutig wie für die explizit-fachmethodische Instruktionsvariante, weil die Heterogenität innerhalb der Team größer ist. Um sinnvoll mit den Gruppierungen weiterarbeiten zu können, wurde daher darauf geachtet, die sechs Teams gleichmäßig auf die Gruppen zu verteilen. Die beiden Teams, für die eine Zuweisung gemittelt hoher oder gemittelt niedriger Kompetenzzuwächse nicht plausibel möglich ist, finden sich in der Gruppe der Teams mit gemittelt mittleren Kompetenzzuwächsen. Es bleiben jeweils zwei Teams mit gemittelt hohen sowie mit gemittelt niedrigen Kompetenzzuwächsen übrig. Die vorgenommenen Gruppierungen sind in Anhang B.2 (ab S. 532) aufgeführt.

Für größere Datenmengen bietet sich vermutlich eine differenziertere Gruppierung der Teams an, die auch die Heterogenität in den Kompetenzzuwächsen expliziter abbildet (ggf. durch weitere Variablen). Darauf wird hier allerdings verzichtet, weil die Anzahlen der Teams in den Video-Teilstichproben so klein sind, dass Beschreibungen vorliegender Unterschiede vorgenommen werden können, sofern diese von Bedeutung für die jeweilige Analyse oder deren Diskussion zu sein scheinen.

5.3 Überlegungen zum Vorgehen bei der Auswertung

Das Einsetzen des Kodiermanuals ist ein qualitatives Vorgehen, dass daraufhin ausgelegt ist, quantitative Auswertungen der Daten zu ermöglichen, die Hinweise darauf liefern, welche Aspekte spezifischer und qualitativer untersucht werden können. Daraus folgt unter anderem, dass Statistik genutzt wird, um die Aktivitäten der Lernenden zu untersuchen und die Gruppen von Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen zu kontrastieren. Nachfolgend wird dargestellt, welche Daten sich aus der Anwendung des Kodiermanuals ergeben, welche statistischen Methoden in Abschnitt 5.4 eingesetzt werden und was dabei grundsätzlich berücksichtigt wird.

5.3.1 Zum Auswerten der Kodierungen

Die Anwendung des Kodiermanuals liefert kodierte Videoaufzeichnungen, die zunächst im Format von MAXQDA vorliegen. Die kodierte Videoaufzeichnungen bestehen aufgrund des eventbasierten Vorgehens aus kodierten Segmenten und unkodierten Abschnitten, wobei kodierte Segmente einander teilweise überlappen können. Jedes kodierte Segment hat einen Anfangs- und einen Endpunkt (und damit auch eine Länge). Außerdem gehört zu jedem kodierten Segment (mindestens) eine Kategorie des Kategoriensystems. Diese Rohdaten zu den kodierten Segmenten lassen sich aus MAXQDA exportieren und dann beispielsweise in Tabellenkalkulationsprogrammen weiterverarbeiten. Zusätzlich stellt auch MAXQDA interne Analysefunktionen bereit, die in der vorliegenden Arbeit dort erläutert sind, wo sie eingesetzt werden.

Analysen zu den Aktivitäten der Lernenden können sich – sehr allgemein betrachtet – auf das Vorliegen (überhaupt vorhanden?), die Lage (wo vorhanden?) oder die Länge (wie lange vorhanden?) von kodierten Segmenten für die Kategorien des Kategoriensystems beziehen. Dabei kann das Vorliegen, die Lage oder die Länge auf das gesamte Videomaterial oder auf (anhand von Abschnitten der von den Lernenden bearbeitete Instruktion oder anhand von anderen kodierten Segmenten ausgewählte) Ausschnitte des Videomaterials bezogen (vor diesen, zeitgleich zu diesen oder nach diesen?) untersucht werden. Beispielsweise kann gezählt werden, wie *häufig* eine Person im gesamten für sie analysierten Videomaterial fachmethodische Beiträge macht (wobei es ums Vorliegen geht) oder *wie lange* sie *zeitgleich* zum eigenen (nonverbalen) Durchführen von Experimenten über Fachinhalte spricht (wobei es um die Länge bezogen auf spezifische Ausschnitte der Instruktion geht).

Aufgrund der Gestaltung des eingesetzten Kodiermanuals ist für die vorliegende Arbeit die Analyse von *Längen* im Allgemeinen der Analyse von Häufigkeiten vorzuziehen. Insbesondere die Entscheidung für die 3-Sekunden-Regel (siehe Seite 75)

geht mit einer Schwerpunktsetzung auf Längen einher, weil durch das zeit- statt z. B. inhaltsbasierte Zusammenführen die Anzahl der kodierten Segmente nur von geringer inhaltlicher Bedeutung ist. Die *Länge eines kodierten Segments* wird in der vorliegenden Arbeit in Sekunden angegeben und ist aus Anfangs- und Endzeitpunkt des Segments berechnet. Wird von der *Länge einer Kategorie* (oder dem zeitlichen Anteil einer Kategorie) gesprochen, so ist damit die Summe aller Längen von mit dieser Kategorie kodierten Segmenten gemeint, also die Gesamtdauer der kodierten Aktivität. Dabei bezieht sich das Wort *Gesamtdauer* auf die jeweils betrachtete(n) Person(en) und den jeweils betrachteten Abschnitt des Videomaterials; falls dieser nicht genauer spezifiziert ist, ist die Gesamtdauer der Aktivität der jeweils betrachteten Person in der jeweils betrachteten Einheit der Instruktion gemeint. Die Längen von Kategorien (d. h. deren zeitliche Anteile) werden in der vorliegenden Arbeit im Allgemeinen *relativ zur Gesamtbearbeitungsdauer* (der betrachteten Einheit, des betrachteten Abschnitts der Instruktion, u. Ä. – sofern nicht näher spezifiziert entsprechend dem vorigen Satz) betrachtet.

Die *Lage* der Kodierungen ist in der vorliegenden Arbeit immer relativ zur einem Bezugspunkt angegeben, weil Vergleiche für – unterschiedlich schnell arbeitende – Personen bzw. Teams sonst nur schwerlich möglich wären (und ferner absolute Werte kaum inhaltlich Bedeutung haben dürften). Beispielsweise kann ein kodiertes Segment *während* der Bearbeitung einer Karte der Instruktion auftreten. Ein weiteres Beispiel wäre eine Analyse dazu, ob *nach* Segmenten, die mit *VorbereitenDurchführenVerbal* kodiert sind, häufig (was unter häufig verstanden wird, wäre dann genauer zu spezifizieren) ein mit *Fachmethodisch-Indiziert* kodiertes Segment folgt.

5.3.2 Zu statistischen Methoden

Für Vergleiche von Stichproben, wie sie zur Beantwortung der beiden Forschungsfragen zu den Aktivitäten eingesetzt werden, liegt es bei statistischer Verarbeitung der Daten nahe, dass *Lageparameter* der Populationen wie Mittelwerte (bspw. arithmetische Mittelwerte oder Mediane) verglichen werden. Die Wahl des betrachteten Lageparameters sollte vor allem vom Skalenniveau der zu vergleichenden Variablen sowie von der Erwartbarkeit großer Ausreißer abhängen (z. B. Rasch et al., 2006). Bei der Betrachtung der zeitlichen Anteile von Aktivitäten wird vorrangig davon ausgegangen, dass die genauen Verhältnisse der zeitlichen Anteile durchaus von Bedeutung sind. Daher wird in der Arbeit grundsätzlich mit Mittelwerten gearbeitet, auch wenn Ausreißer durchaus vorkommen dürften. Rangfolgen (und der damit assoziierte Median als Lageparameter der Verteilung) werden nur zur Diskussion einiger Befunde herangezogen, sofern die Notwendigkeit begründet werden kann.

5.3 Überlegungen zum Vorgehen bei der Auswertung

Als Streumaß zum Mittelwert wird die Standardabweichung angegeben (sofern sie sinnvoll interpretiert werden kann, ansonsten werden alle einzelnen Werte oder die Spannweite angegeben).

Innerhalb der explorativ anlegten Arbeit erfüllen die statistischen Methoden weniger eine prüfende, sondern stärker eine aufdeckende Funktion. Ergänzend zu umfassenden grafischen Auswertungen, in denen deskriptive Werte visuell aufbereitet werden, werden vor allem statistische Tests zum Vergleich von Lageparametern durchgeführt, um die Bedeutsamkeit (Signifikanz) und die Deutlichkeit (Effekt) genauer und in einer vergleichbaren Weise beschreiben zu können. Einer gut begründeten und inzwischen etablierten Leitlinie entsprechend werden grundsätzlich sowohl p -Werte (Signifikanzniveaus) als auch standardisierte Effektstärken anhand der Pearson-Korrelation r angegeben (z. B. Field, 2013; Fritz et al., 2012). Um die realen Effekte hinsichtlich ihrer praktischen fachdidaktischen Bedeutung einschätzen zu können (in der Medizin auch unter den Stichworten praktischer und klinischer Relevanz diskutiert; z. B. Stulz, 2012), wird zudem der Unterschied der Lageparameter für die verglichenen Gruppen angegeben, indem beispielsweise die Mittelwertdifferenz absolut und auch relativ zum größeren der beiden Mittelwerte aufgeführt wird. Aufgrund der Vielzahl durchgeführter Vergleiche werden diese häufig in tabellarischer Form zusammengefasst wie in Tabelle 5.4.

Tabelle 5.4: Beispiel für Übersicht über die Ergebnisse eines statistischen Tests zum Vergleich zweier Gruppen hinsichtlich zweier Variablen.

Vergleich	Gruppe 1			Gruppe 2			df	$t(df)$	p	r	Unterschied	
	N_1	MW	SD	N_2	MW	SD					Absolut	Relativ
Variable 1												
Variable 2												

Anmerkungen. [Art des Durchgeführten statistischen Tests, z. B. Welch-Test]; $N_{1/2}$ =Stichprobengröße für Gruppe 1 bzw. 2; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Signifikanz. Auch wenn Signifikanzen grundsätzlich durch auf drei Stellen genaue Werte angegeben werden, wird sprachlich bei der Beschreibung auf typische Signifikanzniveaus zurückgegriffen ($p < .01$: hochsignifikant, $p < .05$: signifikant, $p < .1$: vorsignifikant). In vielen Fällen werden vorsignifikante Unterschiede mitberichtet, so dass der Fehler erster Art (Unterschiede als bedeutsam zu berichten, die es real nicht sind) meist bei 10 % liegt. Basierend auf den berichteten Stichprobengrößen ergeben sich für Personenvergleiche dann Teststärken von ca. 30–50 % (berechnet mit R und dem Code von <http://powerandsamplesize.com/Calculators/Compare-2-Means/2-Sample-Equality>) abhängig von verglichenen Merkmalen (Aktivitätendauern usw.) und verglichenen Stichproben (für Einheit 2 liegt die Teststärke teilweise noch unter dem Wert von 30 %). Demnach ist die Wahrscheinlichkeit, aus einem nicht-signifikan-

ten Effekt korrekterweise zu folgern, dass kein Unterschied vorliegt, etwa 30–50 %, was sehr gering ist. Nicht-signifikante Ergebnisse werden daher in der Arbeit nur bei sehr deutlichen (bspw. externen, aufgrund anderer Vergleiche, oder theoretischen) Evidenzen als nicht-vorliegende Unterschiede interpretiert.

Effektstärke. In der Arbeit wird grundsätzlich Pearsons Korrelationskoeffizient r als Maß für die Effektstärke genutzt, welches eine gute Vergleichbarkeit für die eingesetzten Varianten des t-Tests erlaubt.⁴⁶ Auch wenn bei der Interpretation der Effektstärken der Kontext der untersuchten Effekte berücksichtigt werden sollte (z.B. Fritz et al., 2012) und demnach in der Arbeit bei der Formulierung von Deutungen berücksichtigt wird, werden bei der verbalen Beschreibung der Ergebnisse die Effektstärken in Anlehnung an die von Cohen (1988) vorgeschlagene, häufig rezipierte, Einteilung klassifiziert:

$ r < .1$	verschwindender Effekt
$.1 \leq r < .3$	kleiner Effekt
$.3 \leq r < .5$	mittlerer Effekt
$.5 \leq r $	großer Effekt

Eingesetzte statistische Verfahren. Generell werden in der vorliegenden Arbeit als *Welch-Tests* bezeichnete t-Tests mit Varianzkorrektur vorgenommen (vgl. Delacre et al., 2017; mittels SPSS, Version 27). Diese Entscheidung basiert zum einen darauf, dass Mittelwerte verglichen werden, zum anderen auf gehäuften Empfehlungen, grundsätzlich die varianzkorrigierten t-Tests den klassischen t-Tests nach Student gegenüber zu bevorzugen (z.B. Delacre et al., 2017; Ruxton, 2006). Für sehr kleine Stichproben oder deutlich unterschiedlich große Stichproben wird allerdings auf t-Tests nach Student zurückgegriffen, weil die Empfehlung sich allem Anschein nach nicht hinreichend gut übertragen lässt (vgl. z.B. Adusah & Brooks, 2011; de Winter, 2013), die Robustheit des t-Tests nach Student allerdings für kleine Stichproben gut belegt ist (vgl. zusätzlich Bortz & Weber, 2005; Heeren & D’Agostino, 1987). In der Arbeit werden dort, wo Welch-Tests eingesetzt werden, die Freiheitsgrade mit drei Nachkommastellen angegeben. Werden t-Tests nach Student eingesetzt, sind ganzzahlige Freiheitsgrade vermerkt.

⁴⁶Es ist aufgrund der vorgenommenen Gruppierungen, die dafür sorgen, dass etwa gleich große Gruppen von Personen bzw. Teams miteinander verglichen werden, in den meisten Fällen gut geeignet. Auch wenn in einigen anderen Fällen eher Cohens d gewählt werden sollte (Field et al., 2013, S. 58), wurde aus Gründen der Konsistenz, Vergleichbarkeit und Übersichtlichkeit darauf verzichtet, das Maß für die Effektstärke zu wechseln oder mehrere Maße anzugeben (wenngleich für hypothesentestende Arbeiten durchaus eine Angabe mehrere Effektstärkenmaße empfohlen wird, vgl. Fritz et al., 2012).

5.4 Auswertungen und Ergebnisse zur explizit-fachmethodischen Instruktionsvariante

Die erste Forschungsfrage zu den Aktivitäten der Lernenden bezieht sich auf die Unterschiede und Gemeinsamkeiten bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante, die für Lernende mit hohen vs. niedrigen Kompetenzzuwächsen vorliegen. Die Frage lautet:

F-Akt1 Welche Unterschiede und Gemeinsamkeiten zeigen sich in den Aktivitäten von Lernenden – insbesondere mit hohen vs. niedrigen Kompetenzzuwächsen – bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante?

Die folgenden Unterabschnitte entsprechen drei Bereichen, für die Analysen berichtet werden. Erstens werden die *Bearbeitungsdauern* der Einheiten insgesamt sowie der einzelnen Karten betrachtet. Derartige Analysen können eine Grundlage für weitere Analysen darstellen, weil sie aufklären können, ob und wo Unterschiede in den Aktivitäten oder Kompetenzzuwächsen ggf. durch unterschiedliche Bearbeitungsdauern bedingt sind. Ferner ist bekannt, dass time on task ein relevanter Faktor für das Lernen ist (z. B. Hattie, 2009; Stender et al., 2013). Zweitens werden die *Aktivitätsprofile* der Lernenden analysiert, die sich aus dem gleichzeitigen Betrachten ausgewählter kodierter Aktivitäten ergeben. Mithilfe von Clusteranalysen werden ähnliche Profile geclustert und die Cluster werden mit den Gruppen zu den Kompetenzzuwächsen verglichen. Diese Analysen dienen einer ganzheitlichen Betrachtung der Aktivitäten. Zum einen Rahmen sie somit die nachfolgenden Analysen zu den fachmethodischen Beiträgen, zum anderen leisten sie einen Beitrag dazu, Zusammenhänge von Aktivitäten insgesamt und Kompetenzzuwächsen zu identifizieren. Somit leisten sie einen spezifischen Beitrag zur Generierung von Hypothesen zu Prozessen des Kompetenzaufbaus, weil die Relevanz der Aktivitäten für Kompetenzzuwächse im Zusammenspiel untersucht wird. Drittens werden die *fachmethodischen Beiträge* der Lernenden genauer in den Blick genommen und unter anderem auf Bearbeitungsdauern und andere Aktivitäten bezogen analysiert. Wie in Abschnitt 3.1 (ab S. 47) ausgeführt, ist die Bedeutung der expliziten Beschäftigung mit Fachmethoden ein wichtiges Desiderat zum Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten. Analysen zu den fachmethodischen Beiträgen leisten beispielsweise einen Beitrag dazu, Hypothesen zur Relevanz dieser beim Kompetenzaufbau zu generieren.

Für die Analysen kann grundsätzlich auf die Gruppierung der (einzelnen) Personen und die Gruppierung der Teams zurückgegriffen werden. Sofern möglich und sinnvoll

wird grundsätzlich sowohl berichtet, welche Unterschiede und Gemeinsamkeiten sich für Personen mit hohen vs. niedrigen Kompetenzzuwächsen finden, als auch, welche sich für Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen finden. Dieses Vorgehen wird gewählt, um die Datenstruktur (teamweise Bearbeitung) zu berücksichtigen, und kann ferner dazu beitragen, die Kompetenzzuwächse der Lernenden nicht nur vor dem Hintergrund der materiellen Umgebung des Lernens (d. h. der Instruktion), sondern auch der personellen Umgebung des Lernens zu deuten.

5.4.1 Bearbeitungsdauern von Lernenden mit unterschiedlichen Kompetenzzuwächsen

Wird danach gefragt, welche Unterschiede und Gemeinsamkeiten es für Lernende mit hohen vs. niedrigen Kompetenzzuwächsen gibt, ist es naheliegend, zunächst die Bearbeitungsdauern zur Instruktion zu betrachten. Zugehörige Analysen können unter anderem Aufschluss darüber geben, ob Lernende mit unterschiedlichen Kompetenzzuwächsen auch unterschiedlich viel *time on task* sind, was vor dem Hintergrund der großen Bedeutung von eingesetzter Lernzeit für das Lernen einen wichtigen Aspekt von Kompetenzaufbau darstellt.

5.4.1.1 Gesamtbearbeitungsdauern für die Einheiten

Eine Betrachtung der Gesamtbearbeitungsdauern für die einzelnen Einheiten liefert Ergebnisse dazu, ob Lernende vergleichbar lange gearbeitet haben. Sie stellt damit eine Grundlage für die Interpretation weiterer Ergebnisse dar.

Vorbemerkung: Die Instruktion ist so angelegt, dass Lernende vergleichbar lange arbeiten. Dies entspringt der Konstruktion der Lerneinheiten (siehe Kapitel 4). Anzumerken ist, dass diese Tatsache für Einheit 2 der explizit-fachmethodischen Instruktionsvariante nur eingeschränkt gilt, weil ein fakultativer Abschnitt von Karte 45 bis Karte 49 enthalten ist. Dieser Abschnitt enthält fachinhaltliche Zusatzinformationen, die für die fachmethodischen Lernziele nicht relevant sind.

Die Verteilung der Gesamtbearbeitungsdauern ist nachvollziehbar. Die Gesamtbearbeitungsdauern der Einheiten ergeben sich jeweils als Summe der Längen aller Segmente, die mit einer Karten-Kategorie kodiert sind. (Das heißt, es wird hier, und auch nachfolgend, nicht die Dauer der Videoaufnahme zugrunde gelegt.) Gesamtbearbeitungsdauern sind für alle (zwei bis) drei Mitglieder eines Teams jeweils gleich; daher erfolgt diese Analyse nur auf Team-Ebene. Die Verteilungen der Gesamtbearbeitungsdauern für die drei Einheiten der explizit-fachmethodischen

Instruktionsvariante entstammen höchstwahrscheinlich aus normalverteilten Grundgesamtheiten (Shapiro-Wilk-Tests ergeben für E1: $W(11) = .961$; $p = .784$; für E2: $W(7) = .917$; $p = .445$; für E3: $W(11) = .925$; $p = .468$). Die absoluten Bearbeitungsdauern liegen für Einheit 1 etwa im Bereich der angesetzten Bearbeitungsdauer (ca. 40 Minuten statt angesetzten 45 Minuten, Details in Tabelle 5.5 usw.), für die Einheiten 2 und 3 deutlich niedriger (E2: ca. 67 Minuten statt angesetzten 90 Minuten; E3: ca. 64 Minuten statt angesetzten 90 Minuten); ein großer Teil der Differenz lässt sich durch die Stundenorganisation (organisatorische Vorbemerkungen etc.) erklären. Die Streuungen in den absoluten Gesamtbearbeitungsdauern über die unterschiedlichen Teams hinweg ist in einem erwartbaren Bereich und lässt sich durch Standardabweichungen von ca. 10 % der jeweils angesetzten Bearbeitungsdauer beschreiben (Einheit 1: ca. 5 Minuten bei angesetzten 45 Minuten, Einheiten 2+3: jeweils ca. 9 Minuten bei angesetzten 90 Minuten).

Die Lernenden setzen ihre Zeit grundsätzlich für die Arbeitsaufträge ein. Anhand der Kategorie OffTask aus dem Kodiermanual kann für jede Person analysiert werden, wie viel Zeit sie instruktionsfremden Beschäftigungen nachgeht. Die Anteile der Kategorie sind insgesamt sehr niedrig. In Abbildung 5.9 ist exemplarisch dargestellt, dass für Einheit 1 für den Großteil der Personen keine Phasen mit OffTask kodiert sind. Die beiden Personen mit den hohen zeitlichen Anteilen

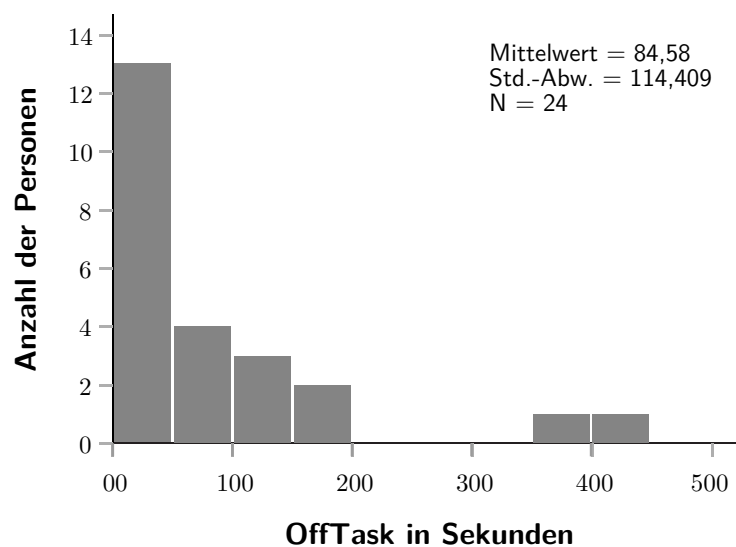


Abbildung 5.9: Histogramme für die absoluten Anteile der Kategorie OffTask in Einheit 1 der explizit-fachmethodischen Instruktionsvariante.

instruktionsfremder Beschäftigung sind D11 (7 Minuten 20 Sekunden; entsprechen ca. 16 % der angesetzten Einheitendauer von 45 Minuten für Einheit 1) und D12 (6 Minuten 6 Sekunden; ca. 13%). In allen anderen Einheiten und für alle anderen Personen sind die zeitlichen relativen Anteile deutlich geringer. In den nachfolgenden Analysen wird aufgrund der insgesamt geringen Anteile von Off-Task nicht weiter auf diese Kategorie eingegangen, es sei denn die Interpretation von Ergebnissen hängt substantiell mit ihr zusammen (bspw. für D11 oder D12 oder das Team D als Ganzes).

Die Daten deuten zum Teil darauf hin, dass Lernenden mit hohen Kompetenzzuwächsen systematisch größere Gesamtbearbeitungsdauern aufweisen als Lernende mit niedrigen Kompetenzzuwächsen. Da die Instruktion in Teams bearbeitet wurde, werden zunächst die Bearbeitungsdauern für *Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen* miteinander verglichen. In Tabelle 5.5 sind die Mittelwerte der Gesamtbearbeitungsdauern sowie Angaben zum statistischen Vergleich der Gruppen von Teams mit hohen vs. niedrigen Kompetenzzuwächsen dargestellt. Die Vergleiche der Teams zeigen, dass für Einheit 1 ein nicht-signifikanter Unterschied von kleiner Effektstärke existiert (Mittelwertunterschied von ca. 2 Minuten und 47 Sekunden, entspricht etwa 6.19 % der 45-minütigen angesetzten Bearbeitungsdauer). Für Einheit 2 liegt ein signifikanter Unterschied großer Effektstärke vor, mit einer Mittelwertdifferenz von ca. 10 Minuten und 9 Sekunden (11.28 % der angesetzten Bearbeitungsdauer von 90 Minuten). Für Einheit 3 liegt ein Unterschied mittlerer Effektstärke vor, der nicht-signifikant ist, aber einem Mittelwertunterschied von 8 Minuten und 26 Sekunden entspricht (9.37 % der angesetzten Bearbeitungsdauer von 90 Minuten). Es ist zu berücksichtigen, dass für Einheit 2 die Varianz in der größeren Stichprobe größer ist, so dass der p -Wert eher zu groß und das Ergebnis daher noch deutlicher ist, während in Einheit 3 die reale

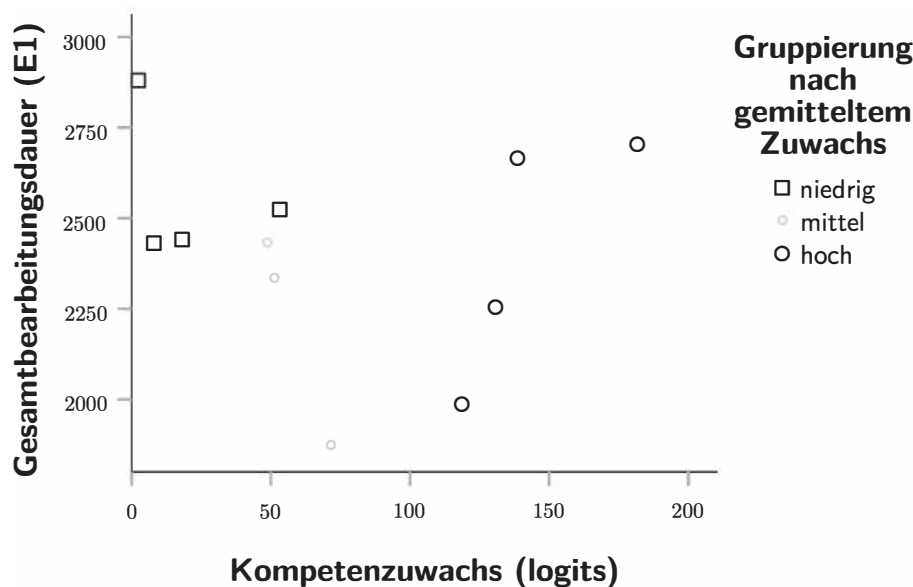
Tabelle 5.5: Übersicht über die Vergleiche der Gesamtbearbeitungsdauern der drei Einheiten (Teams mit gemittelt hohen vs. niedrigen Kompetenzzuwächsen).

Einheit	Zuwachs Niedrig			Zuwachs Hoch			df	$t(df)$	p	r	Unterschied	
	N_N	MW	SD	N_H	MW	SD					Absolut	Relativ
1	4	2569.0	211.3	4	2402.9	344.7	6	0.822	.443	-.318	-167,0	-6,19%
2	2	3723.3	183.7	3	4332.6	240.3	3	-3.631	.036	.902	609,0	11,28%
3	3	3585.4	671.6	3	4091.4	518.8	4	-1.033	.360	.458	506,0	9,37%

Anmerkungen. Students t-Test; $N_{H/N}$ = Stichprobengröße für Gruppe der Teams mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

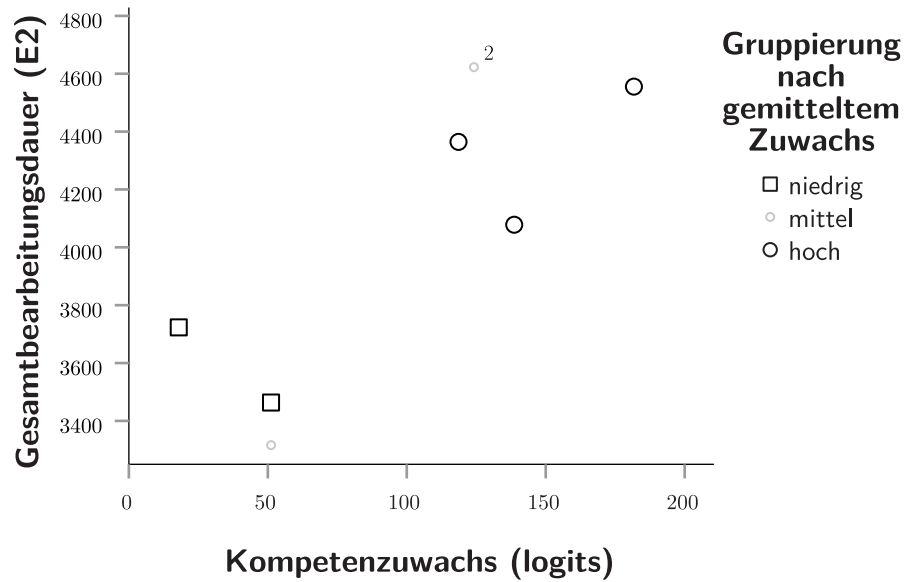
Irrtumswahrscheinlichkeit eher größer als der angegebene p -Wert sein dürfte. Durch die kleinen Stichprobenzahlen und ungleichen Varianzen dürften die angegebenen Effektstärken für die Team-Vergleiche bei allen drei Einheiten betragsmäßig eher überschätzt werden.

Wird im Detail aufgelöst, wie sich die Gesamtbearbeitungsdauern zu den mittleren Kompetenzzuwächsen der Teams verhalten, ergeben sich die Streudiagramme in Abbildung 5.10. Daraus geht hervor, dass die Unterschiede nur für Einheit 2 systematischer Natur sein dürften; für diese Einheit ergibt sich ein klarer ansteigender Trend. Für die anderen beiden Einheiten ergeben sich eher breite Bänder, die horizontal über das Diagramm verlaufen (mit ggf. minimal positivem Trend). Der Unterschied zwischen Einheit 2 und Einheit 3, die in der statistischen Beschreibung durch Tabelle 5.5 nicht allzu verschieden wirkten, wird besonders daran deutlich, dass Teams, die der Gruppe von Teams mit gemittelt mittleren Kompetenzzuwächsen zugewiesen sind (bspw., weil eine große Binnenvarianz vorliegt), bei Aufträgen gemäß der über die Personen gebildeten Mittelwerte von Kompetenzzuwächsen das Bild jeweils stimmig ergänzen.

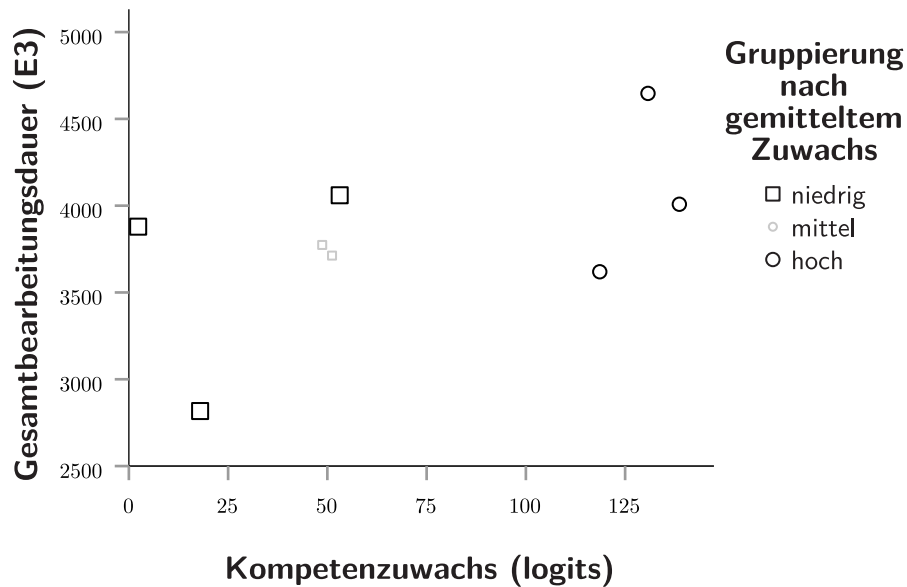


(a) Einheit 1 (11 Teams)

Abbildung 5.10: Informationen auf der nächsten Seite.



(b) Einheit 2 (7 Teams)



(c) Einheit 3 (8 Teams)

Abbildung 5.10: Gesamtbearbeitungsdauer aufgetragen über dem teamweise gemittelten Kompetenzzuwachs (eingetragene Punkte entsprechen jeweils einem Mittelwert der Zuwächse der jeweils drei Personen des jeweiligen Teams). Durch Symbole vermerkte Gruppierungszugehörigkeiten gemäß Team-Kompetenzzuwächsen entstammen Abschnitt 5.2 (wo nicht basierend auf arithmetischen Mittelwerten über Personen vorgegangen wird, so dass sich kleine Abweichungen ergeben). Die kleine Zwei bei einem Team für Einheit 2 zeigt an, dass dieses Team nur zwei Einheiten der Instruktion bearbeitet hat. Alle anderen Teams haben alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante bearbeitet, auch wenn sie nicht für alle analysiert wurden.

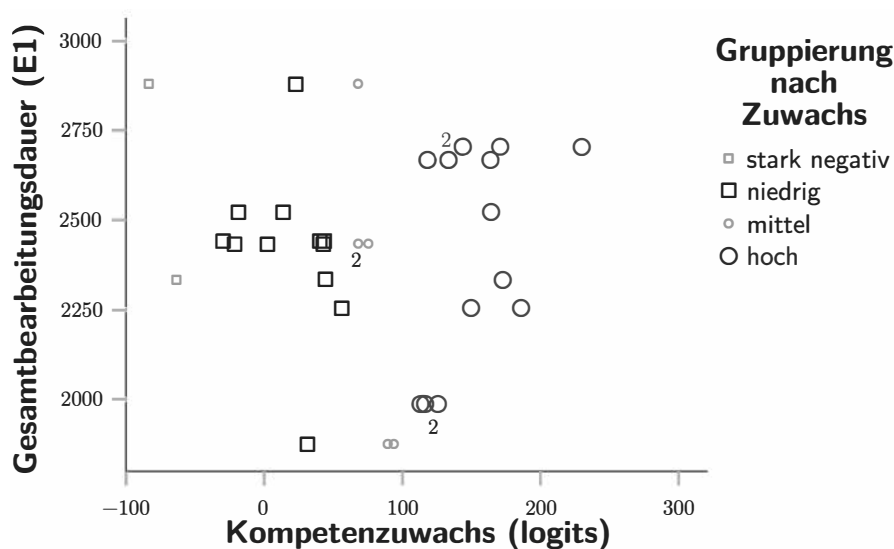
5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

Der Frage, ob sich Personen mit hohem Zuwachs eher in Teams befinden, die insgesamt länger arbeiten, kann nachgegangen werden, indem die Personen-Gruppierungen genutzt werden. Werden *Personen mit hohen vs. niedrigen Kompetenzzuwächsen* verglichen, ergibt sich Tabelle 5.6. Die Ergebnisse sind vergleichbar mit den oben zu den Team-Vergleichen dargestellten Ergebnissen; allerdings sind sie weniger stark ausgeprägt (relative und absolute Maße für Effekte), was darauf hindeutet, dass Gesamtbearbeitungsdauern weniger stark mit dem Kompetenzaufbau von Individuen zusammenhängen als mit dem über alle Teammitglieder gemittelten Kompetenzaufbau. Diese Deutung wird auch durch die in Abbildung 5.11 dargestellten Streudiagramme auf Personen-Ebene gestützt, in denen sich in ähnlicher Weise Trends bzw. keine Trends wie auf Team-Ebene finden.

Tabelle 5.6: Übersicht über die Vergleiche der Gesamtbearbeitungsdauern der drei Einheiten (Personen mit hohen vs. niedrigen Kompetenzzuwächsen).

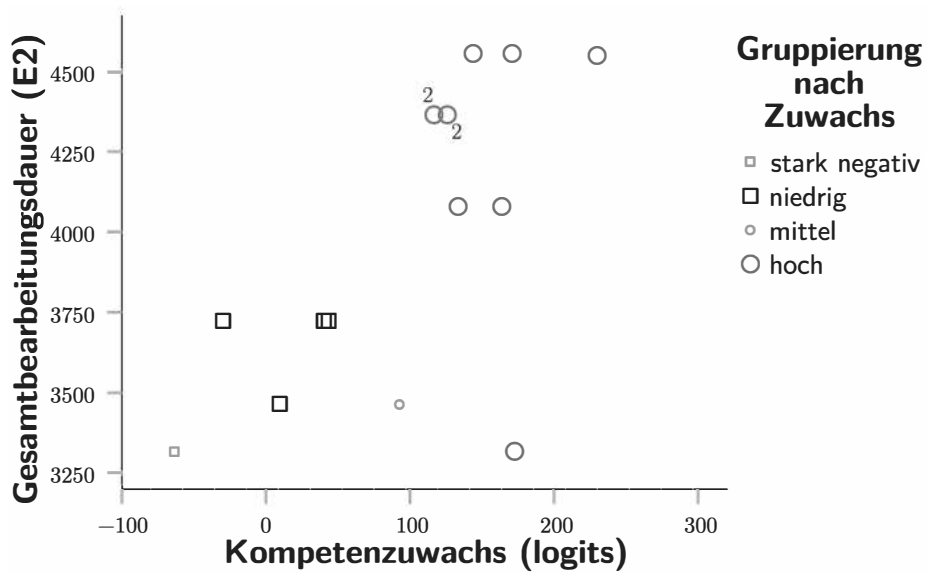
Einheit	Zuwachs Niedrig			Zuwachs Hoch			korr. df	t (df)	p	r	Unterschied	
	N_N	MW	SD	N_H	MW	SD					Absolut	Relativ
1	13	2418.5	271.4	13	2418.2	297.7	24.000	0.003	.997	<.001	-0,3	-0,01%
2	6	3762.1	454.6	11	4265.9	375.7	8.809	-2.318	.046	.615	503,0	9,31%
3	9	3620.4	659.1	10	3994.8	388.4	12.682	-1.488	.161	.385	374,0	6,93%

Anmerkungen. Welch-Test mit korrigierten Freiheitsgraden (korr. df); $N_{H/N}$ =Stichprobengröße für Gruppe der Teams mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

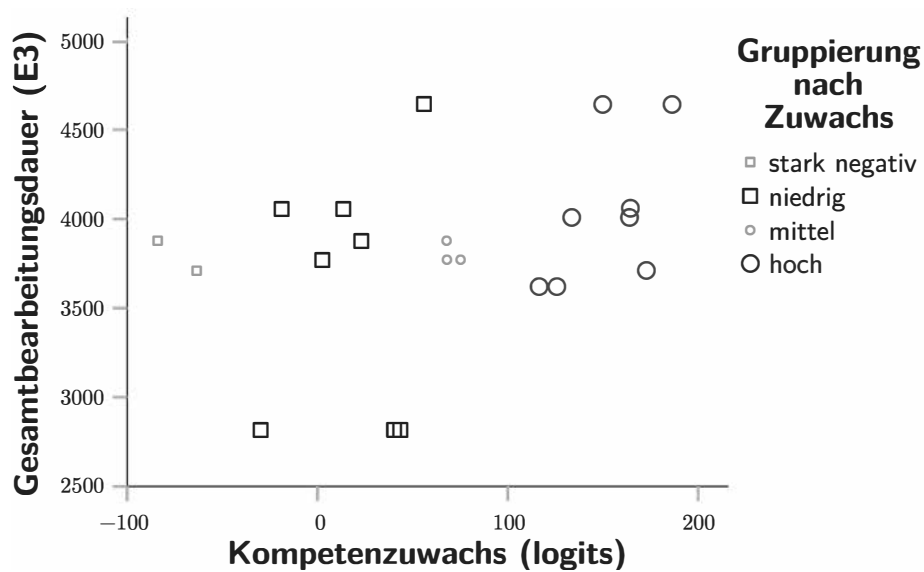


(a) Einheit 1 (33 Personen)

Abbildung 5.11: Informationen auf der nächsten Seite.



(b) Einheit 2 (19 Personen)



(c) Einheit 3 (24 Personen)

Abbildung 5.11: Gesamtbearbeitungsdauer des Teams, in dem sich die Person befindet, aufgetragen über dem Kompetenzzuwachs der Personen. (Die kleine Zwei an einigen Symbolen zeigt an, dass die Personen jeweils nur zwei Einheiten der Instruktion bearbeitet haben. Alle anderen Personen haben alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante bearbeitet, auch wenn sie nicht für alle analysiert wurden.)

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

Bei einer *zusammenfassenden Deutung der Gesamtbearbeitungsdauern* ist zu berücksichtigen, dass die Instruktion grundsätzlich so angelegt ist, dass die Bearbeitungsdauern verschiedener Teams etwa gleich sein sollten. Gleichzeitig scheinen Unterschiede von etwa zehn Prozent in der Bearbeitungsdauer aufgrund von Binnenvarianzen durchaus plausible Abweichungen zwischen unterschiedlichen Teams darzustellen. Insgesamt werden die Ergebnisse daher zunächst so gedeutet, dass die Gesamtbearbeitungsdauern nicht in einem bedeutsamen Zusammenhang mit den Kompetenzzuwächsen stehen.⁴⁷ Dass für Einheit 2 ein *systematischer* signifikanter Unterschied vorliegt, ist vor dem Hintergrund der fakultativen Einschubs (siehe erstes Ergebnis zu Gesamtbearbeitungsdauern) zunächst plausibel. (Gleichzeitig bleibt allerdings die Frage, weshalb Teams, die *fachinhaltliche* Zusatzinformationen bearbeiten, *höhere* Kompetenzzuwächse zum experimentbezogenen Denken und Arbeiten erreichen; eine mögliche Ursache könnte darin liegen, dass sie weniger Schwierigkeiten in den Teilen zu Fachmethoden hatten und daher mehr Zeit für den Einschub einsetzen können, während andere Teams aus Zeitgründen den Einschub überspringen müssen.) Die Ergebnisse zu den Abschnitten der Instruktion (unten in Unterunterabschnitt 5.4.1.2) zeigen jedoch, dass dieser fakultative Einschub nur einen sehr geringen zeitlichen Umfang hat und daher nur einen kleinen Anteil der Unterschiede in den Gesamtbearbeitungsdauern hervorgerufen haben kann. Dass die Analysen zu Einheit 3 einen systematischen Unterschied von mehr als 8 Minuten (mittlere Effektstärke) für die untersuchten Teams (ca. 6 Minuten für die Personenvergleiche) nahelegen, ist unerwartet. Dass der Unterschied so groß ist, dürfte vor dem Hintergrund des großen p -Wertes ein Artefakt der kleinen Stichprobengrößen sein und darf daher vermutlich nicht auf die Bearbeitung der explizit-fachmethodischen Instruktionsvariante von Lernenden der elften Klasse im Allgemeinen generalisiert werden. Sollte der Unterschied tatsächlich systematisch sein, müssten sich in den nachfolgenden Analysen zu Bearbeitungsdauern von Abschnitten und Karten der Instruktion Hinweise darauf finden. Vorrangig dienen die nachfolgenden Analysen zu Abschnitten und Karten der Instruktion dazu, Gemeinsamkeiten und Unterschiede in den Aufteilungen der Gesamtbearbeitungsdauern auf verschiedene Aufgaben und Inhalte der Instruktion für Lernende mit unterschiedlichen Kompetenzzuwächsen zu untersuchen.

⁴⁷Basierend auf den für die Einheiten 2 und 3 vorliegenden absoluten Unterschieden sowie den Streudiagrammen in den Abbildungen 5.10 und 5.11 wird in Unterunterabschnitt 5.5.1.2, ab S. 210, allerdings im Abgleich mit der implizit-fachmethodischen Instruktionsvariante geschlossen, dass die tendenziell positiven Zusammenhänge doch auf potentielle Zusammenhänge hinweisen könnten.

5.4.1.2 Bearbeitungsdauern von Abschnitten der Instruktionseinheiten

Die Einheiten der Instruktion sind in Abschnitte eingeteilt, die durch Abschnittstitelblätter auch den Lernenden transparent gemacht sind. Die Abschnitte beziehen sich jeweils auf spezifische Inhalte, für die explizit-fachmethodische Instruktionsvariante z. B. das Formulieren von Fragen oder die Unterscheidung von Beobachtungen und Deutungen (Inhalte aller Abschnitte sind in Tabelle 4.4 aufgeführt). Da der zweite Teil des dritten Abschnitts von Einheit 2 als fakultativ ausgewiesen ist, wird er im folgenden gesondert betrachtet (und als 2-A3b bezeichnet, während 2-A3a den davor liegenden Rest des dritten Abschnitts bezeichnet).

Analysen zu den Abschnitten der Einheiten der Instruktion ermöglichen zu identifizieren, welche inhaltlichen Bereiche der Instruktion für Lernende mit unterschiedlichen Kompetenzzuwächsen mit verschiedenen Bearbeitungsdauern einhergehen. Die Ergebnisse ermöglichen ferner unter anderem zu untersuchen, wodurch Unterschiede in den Gesamtbearbeitungsdauern verursacht sind bzw. ob trotz gleicher Bearbeitungsdauern inhaltlich spezifische Abschnitte unterschiedlich lange bearbeitet werden.

Bei absoluter Betrachtung: Lernende mit hohen Kompetenzzuwächsen arbeiten an allen Abschnitten (außer 1-A1) länger als Lernende mit niedrigen Kompetenzzuwächsen. Die Abschnittsbearbeitungsdauern der Lernenden werden sowohl für *Teams* mit unterschiedlichen gemittelten Kompetenzzuwächsen als auch für *Personen* mit unterschiedlichen Kompetenzzuwächsen verglichen. In Abbildung 5.12 ist dargestellt, dass Teams mit gemittelt hohen Kompetenzzuwächsen für jeden Abschnitt außer 1-A1 im Mittel mehr Zeit aufwenden als Teams mit gemittelt niedrigen Kompetenzzuwächsen. Aus Tabelle 5.7 geht hervor, dass die Unterschiede zwar nur zum Teil große Effektstärken haben und größtenteils nicht signifikant vorliegen, die absoluten oder relativen Unterschiede der Mittelwerte (absolut: Differenz der Mittelwerte für die beiden Gruppen; relativ: absoluter Unterschied geteilt durch größeren der beiden Mittelwerte) sind allerdings in den meisten Fällen nicht vernachlässigbar. Die jeweils ersten Abschnitte der Einheiten fallen etwas aus dem Muster heraus: Für 1-A1 findet sich ein umgekehrter Zusammenhang, für 2-A1 und 3-A1 finden sich nur sehr kleine Mittelwertunterschiede. Die Ergebnisse der Vergleiche von Personen mit unterschiedlichen Kompetenzzuwächsen sind im Anhang in Abbildung B.2 und Tabelle B.1 aufgeführt und analog zu den hier auf Ebene der Teams berichteten Ergebnissen; es finden sich allerdings etwas kleinere Unterschiede (Effektstärken sowie absolute und relative Mittelwertunterschiede sind geringer).

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

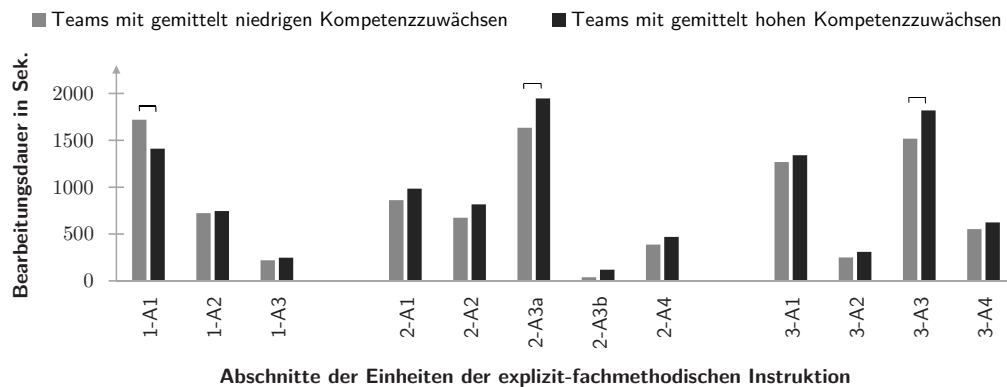


Abbildung 5.12: Mittelwerte der *absoluten* Abschnittsbearbeitungsdauern der explizit-fachmethodischen Instruktionsvariante für Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen. (Fehlerbalken entsprechen einer Standardabweichung des jeweiligen Mittelwerts; Abschnitte mit jeweils deutlichsten Unterschieden der Einheit sind mit einer Klammer markiert.)

Tabelle 5.7: Übersicht über die Vergleiche der *absoluten* Abschnittsbearbeitungsdauern für alle drei Einheiten (Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen).

Vergleich		Zuwachs Niedrig			Zuwachs Hoch			df	$t(df)$	p	r	Unterschied	
Einheit	Abschnitt	N_N	MW	SD	N_H	MW	SD					Absolut	Relativ
1	1	2	1718,4	280,8	4	1410,6	254,6	4	1,36	0,246	0,562	-307,8	-17,91%
	2		722,5	68,4		745,2	207,1		-0,144	0,893	0,072	22,7	3,05%
	3		219,6	98,2		247,1	34,1		-0,555	0,608	0,267	27,6	11,15%
2	1	2	860,5	138,3	3	983,2	115,4	3	138,31	0,356	0,532	122,7	12,48%
	2		673,9	238,6		815,8	80,6		238,64	0,384	0,507	141,9	17,40%
	3a		1633,2	181,3		1946,7	13,9		181,30	0,047	0,883	313,5	16,10%
	3b		38,5	36,8		118,6	87,8		36,77	0,325	0,561	80,1	67,55%
	4		387,4	337,8		468,3	232,0		337,78	0,766	0,185	81,0	17,29%
3	1	3	1267,3	268,2	3	1340,5	344,1	4	-0,29	0,786	0,144	73,1	5,46%
	2		249,6	114,2		309,6	142,3		-0,57	0,599	0,274	60,1	19,40%
	3		1516,7	153,1		1818,1	188,0		-2,153	0,098	0,733	301,4	16,58%
	4		551,8	187,8		623,2	284,1		-0,363	0,735	0,179	71,4	11,46%

Anmerkungen. Students t-Test; $N_{H/N}$ =Stichprobengröße für Gruppe der Teams mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Bei relativer Betrachtung: Lernende mit hohen Kompetenzzuwächsen setzen für die meisten Abschnitte ähnliche Anteile ihrer Gesamtbearbeitungsdauer ein wie Lernende mit niedrigen Kompetenzzuwächsen. Die relativen Anteile der Abschnittsbearbeitungsdauern an der tatsächlichen Gesamtbearbeitungsdauer des jeweiligen Teams werden sowohl für *Teams* mit unterschiedlichen gemittelten Kompetenzzuwächsen als auch für *Personen* mit unterschiedlichen Kompetenzzuwächsen verglichen. Aus Tabelle 5.8 geht hervor, dass die Unterschiede zwischen den relativen Bearbeitungsdauern von Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen in den meisten Fällen relativ zum jeweiligen Abschnittsbearbeitungsdaueranteil betrachtet meist nur sehr gering sind (letzte Spalte) – einzig für Einheit 1 und Abschnitt 2-3b finden sich größere relative Unterschiede zwischen den Mittelwerten der beiden Gruppen. Das Ergebnis ist in Abbildung 5.13 visualisiert. Da die meisten Unterschiede klein sind, werden die wechselhaften Tendenzen (an manchen Abschnitten arbeiten Teams mit gemittelt hohen Kompetenzzuwächsen länger, an anderen Teams mit gemittelt niedrigen) als unbedeutend interpretiert. Vielmehr investieren für Einheit 2 und Einheit 3 Lernende mit verschiedenen Kompetenzzuwächsen *ähnliche Zeitanteile für die Abschnitte*, obwohl es Unterschiede in den *absoluten* Abschnittsbearbeitungsdauern gibt.

Tabelle 5.8: Übersicht über die Vergleiche der *relativen* Anteile der Abschnittsbearbeitungsdauern an der Gesamtbearbeitungsdauer der jeweiligen Einheit (Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen).

Vergleich		Zuwachs Niedrig			Zuwachs Hoch			df	$t(df)$	p	r	Unterschied	
Einheit	Abschnitt	N_N	MW	SD	N_H	MW	SD					Absolut	Relativ
1	1	2	64,29%	3,21%	4	58,63%	6,73%	4	1,08	0,341	0,475	-5,66%	-8,80%
	2		27,43%	5,69%		30,77%	6,42%		-0,617	0,571	-0,295	3,34%	10,84%
	3		8,08%	2,76%		10,35%	1,58%		-1,347	0,249	-0,559	2,27%	21,93%
2	1	2	23,88%	2,63%	3	22,73%	2,78%	3	0,026	0,675	0,258	-1,15%	-4,83%
	2		18,61%	5,69%		18,88%	2,38%		0,057	0,942	-0,046	0,28%	1,47%
	3a		45,38%	2,73%		45,04%	2,83%		0,027	0,901	0,078	-0,34%	-0,76%
	3b		1,10%	1,08%		2,68%	1,97%		0,011	0,389	-0,501	1,58%	59,03%
	4		11,03%	9,96%		10,67%	5,06%		0,1	0,959	0,032	-0,36%	-3,29%
3	1	3	35,22%	1,01%	3	32,76%	7,78%	4	0,545	0,615	0,263	-2,47%	-7,01%
	2		6,78%	2,14%		7,44%	2,68%		-0,331	0,757	-0,163	0,66%	8,82%
	3		42,85%	4,50%		44,53%	1,76%		-0,604	0,578	-0,289	1,69%	3,79%
	4		15,14%	3,20%		15,27%	6,92%		-0,028	0,979	-0,014	0,12%	0,82%

Anmerkungen. Students t-Test; $N_{H/N}$ =Stichprobengröße für Gruppe der Teams mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

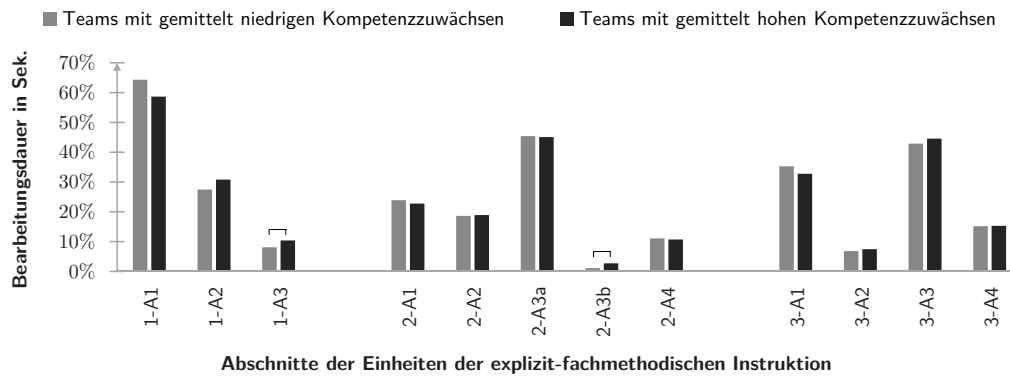


Abbildung 5.13: Mittelwerte der *relativen* Anteile der Abschnittsbearbeitungsdauern an den jeweiligen Gesamtbearbeitungsdauern der Einheiten der explizit-fachmethodischen Instruktionsvariante für Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen. (Fehlerbalken entsprechen einer Standardabweichung des jeweiligen Mittelwerts; Abschnitte mit Unterschieden mit jeweils deutlichsten Unterschieden sind mit einer Klammer markiert.)

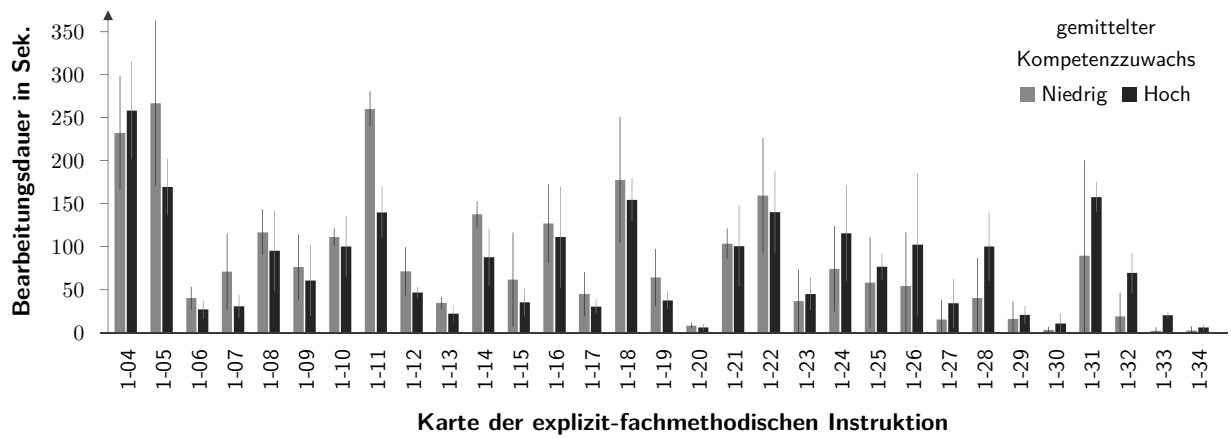
Die abweichenden Ergebnisse bei der ersten Einheit der explizit-fachmethodischen Instruktionsvariante deuten darauf hin, dass es Unterschiede hinsichtlich des Umgangs mit der zur Verfügung stehenden Zeit zwischen Lernenden mit unterschiedlichen Kompetenzzuwächsen geben könnte. Zunächst fällt auf, dass bei der relativen Betrachtung Lernende mit niedrigen Kompetenzzuwächsen an allen ersten Abschnitten der Einheiten (1-A1, 2-A1, 3-A1) tendenziell länger arbeiten als Lernende mit hohen Kompetenzzuwächsen. Zudem sind die Gesamtbearbeitungsdauern für Lernende mit niedrigen Kompetenzzuwächsen für die beiden zeitlich umfangreicher angesetzten Einheiten 2 und 3 kürzer als für Lernende mit hohen Kompetenzzuwächsen. In Einheit 1 arbeiten Lernende mit niedrigen Kompetenzzuwächsen deutlich kürzer am letzten Abschnitt (1-A3). Für die Deutung scheint insbesondere relevant, dass Einheit 1 zeitlich weniger umfangreich angesetzt als die beiden anderen Einheiten (45 Minuten vs. 90 Minuten) und dass bei Einheit 1 fast alle Teams bis zum Ende der Unterrichtsstunde arbeiten, während bei anderen Einheiten Teams auch nach Beendigung des Instruktionsmaterials frühzeitig die Bearbeitung beenden (dies ergibt sich aus Videosichtungen durch den Autor und allgemeinen Notizen der kodierenden Hilfskräfte sowie vorläufigen Analysen von Kartenbearbeitungsdauern, bei denen sich für einige Lernende mit niedrigen Kompetenzzuwächsen keine von Null verschiedenen Bearbeitungsdauern für die letzten Karten der Instruktion ergeben). Zusammenge-

nommen könnten diese Ergebnisse darauf hindeuten, dass Lernende mit niedrigen Kompetenzzuwächsen die zur Verfügung stehende Zeit weniger schlecht im Blick haben als Lernende mit hohen Kompetenzzuwächsen, weil sie in Einheit 1 nicht fertig werden und in den anderen (zeitlich deutlich umfangreicheren) Einheiten eher zu kurz arbeiten. Andere Deutungsmöglichkeiten bestehen darin, dass es Lernenden mit niedrigeren Kompetenzzuwächsen schwerer fällt, sich zu Beginn der Instruktionseinheit auf den fachmethodischen Gegenstand einzulassen, oder dass Lernende mit höheren Kompetenzzuwächsen besser mit den Arbeitsaufträgen umzugehen wissen und daher grundsätzlich kürzer brauchen und darauf erst im weiteren Verlauf durch langsames (oder gründlicheres) Arbeiten reagieren. Analysen zu einzelnen Karten erlauben einen Einblick, für welche spezifischen Arbeitsaufträge und Inhalte die Lernenden ihre Zeit verwenden, und werden daher im nächsten Unterunterabschnitt berichtet.

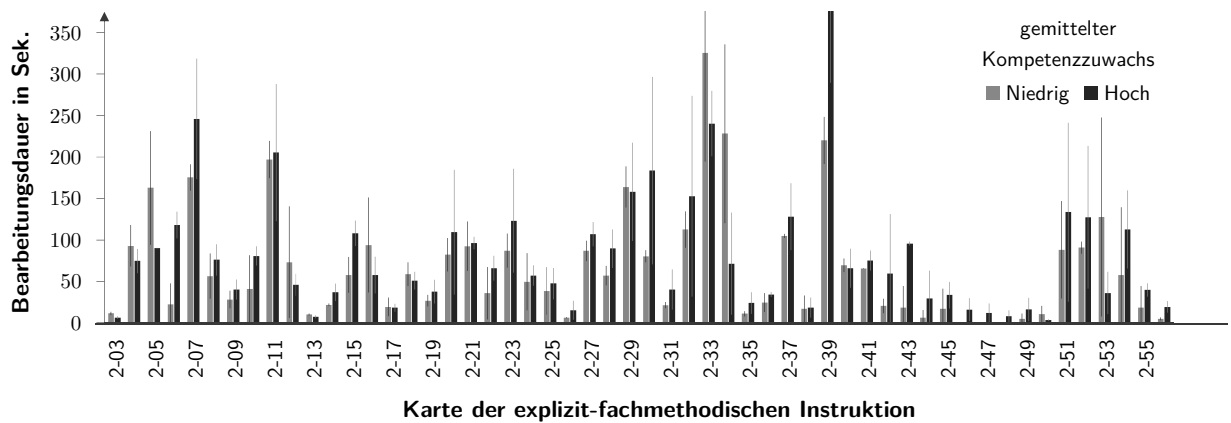
5.4.1.3 Bearbeitungsdauern der einzelnen Karten der explizit-fachmethodischen Instruktionsvariante

Analysen zu den Bearbeitungsdauern von Karten der Instruktion, auf denen die Arbeitsaufträge und Informationen notiert sind und auf denen Lernende Notizen machen können und sollen, können dem Identifizieren von spezifischen Inhalten und Arbeitsaufträgen dienen, für die Lernende mit unterschiedlichen Kompetenzzuwächsen unterschiedlich viel Zeit aufwenden. Dabei wird so vorgegangen, dass zunächst alle Teams aussortiert werden, die nicht bis zum Ende der jeweiligen Einheit gearbeitet haben, um Verzerrungen im Vergleich zu vermeiden. Dies betrifft nur die beiden Teams J und C bei Einheit 1, die beide niedrige Kompetenzzuwächse erreichen. Danach werden für alle verbliebenen Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen die Kartenbearbeitungsdauern verglichen.⁴⁸ Ein *Vergleich der Mittelwerte* ist in Abbildung 5.14 dargestellt. Da die Karten unterschiedlich angelegt sind (bspw. enthalten manche kurze Informationstexte, andere enthalten Experimentieraufträge, für die längere Bearbeitungsdauern angesetzt sind), entspricht es der Erwartung, dass von Karte zu Karte große Schwankungen in den Bearbeitungsdauern vorliegen. Für den Vergleich von Lernenden mit unterschiedlichen Kompetenzzuwächsen sind jeweils die beiden Säulen zu einer Karte zu betrachten.

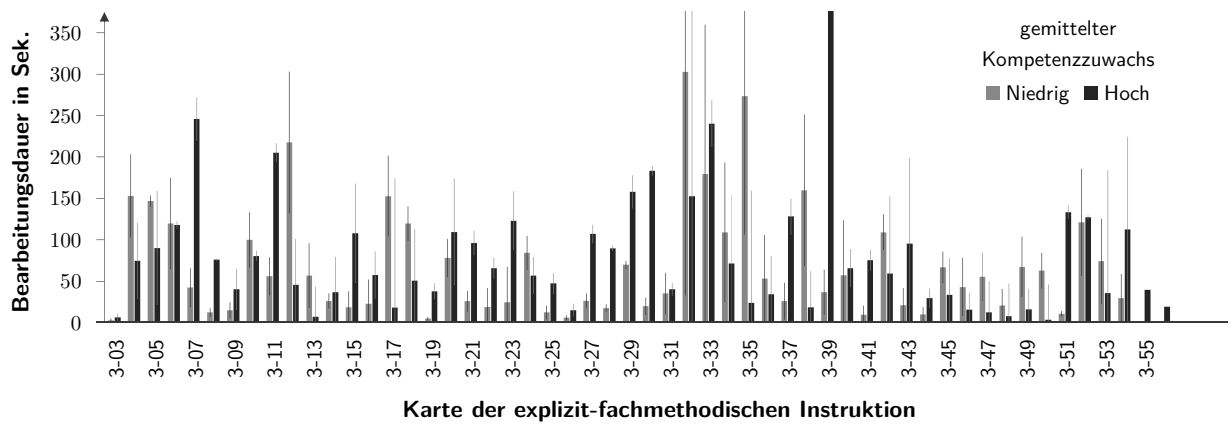
⁴⁸Ein Vergleich von *Personen* mit hohen vs. niedrigen Kompetenzzuwächsen liefert nahezu identische Ergebnisse wie nachfolgend basierend auf Vergleichen von Teams mit unterschiedlichen Kompetenzzuwächsen berichtet.



(a) Einheit 1 (9 Teams)



(b) Einheit 2 (7 Teams)



(c) Einheit 3 (8 Teams)

Abbildung 5.14: Mittelwerte der Bearbeitungsdauern zu den einzelnen Karten der drei Einheiten der explizit-fachmethodischen Instruktionsvariante für Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen. (Fehlerbalken entsprechen einer Standardabweichung zum jeweiligen Mittelwert.)

Werden die *Inhalte aller Karten* aller drei Einheiten betrachtet, für die sich in t-Tests Unterschiede mit großen Effektstärken finden, die mindestens auf einem Signifikanzniveau von $p < .2$ sind,⁴⁹ ergibt sich Tabelle 5.9. Zum einen finden sich mehrere Karten mit deutlichen Unterschieden, die eigenständiges Formulieren von Fragen (Dauer im Mittel größer für Gruppe der Teams mit hohen Kompetenzzuwächsen) oder eigenständiges Durchführen von Versuchen (von Karte zu Karte unterschiedlich) beinhalten. Zum anderen finden sich auch mehrere Karten, auf denen Lernende sich mit einer Versuchssituation vertraut machen können und sollen (Dauer im Mittel größer für Gruppe der Teams mit hohen Kompetenzzuwächsen). Deutliche Unterschiede finden sich auch für inhaltlich irrelevante Karten (wie Abschnittstitelblätter, auf denen nur eine Überschrift steht) – die tatsächlichen Aktivitäten der Lernenden könnten ggf. Aufschluss darüber gegeben, warum hier Unterschiede vorliegen; beispielsweise bespricht eins der analysierten Teams bei einem Abschnittstitelblatt angeregt durch die Abschnittsüberschrift eine zuvor bearbeitete Aufgabe nach.

Die Ergebnisse, dass Zeiteinsatz bei der Durchführung von naturwissenschaftlichen Versuchen nicht grundsätzlich mit höheren Kompetenzzuwächsen einhergeht, während Zeiteinsatz für das Formulieren von naturwissenschaftlichen Fragestellungen mit höheren Zuwächsen zusammenhängt, könnten so gedeutet werden, dass Zeiteinsatz beim Experimentieren leicht auf praktische Tätigkeiten verwendet werden kann, die ohne kognitive Aktivität zu den Konzepten des experimentbezogenen Denkens und Arbeitens stattfinden können, während Zeiteinsatz beim Formulieren von Fragen typischerweise gering sein dürfte und nur bei ausführlichen Überlegungen dazu, was bei der Formulierung der Frage zu beachten ist, die Bearbeitung deutlich länger dauern dürfte. Derartige Überlegungen könnten allerdings auf Fachinhalte (z. B. »Wenn es in der Frage um den Druck geht, was ist damit gemeint? Sagt es mir wirklich etwas über den Druck, wann die Luftblasen aus einem eingetauchten Strohhalm auftauchen?«) oder Fachmethoden (z. B. »Du musst doch immer in die Frage reinstecken, *was* einen Einfluss auf den Druck hat. Wir sollten also schreiben: Hat die Wasserhöhe im Glas einen Einfluss darauf, wo zuerst Luftblasen aufsteigen?«) gerichtet sein. Detailliertere Analysen der (verbalen) Aktivitäten können Aufschluss darüber geben, worauf die Überlegungen der Lernenden gerichtet sind (siehe insbesondere Unterabschnitte 5.4.4–5.4.6).

Eine *Einschränkung* der berichteten Ergebnisse zu den Kartenbearbeitungsdauern

⁴⁹ Aufgrund verschiedener Faktoren, u. a. der geringen Stichprobengrößen in Vergleichen von Teams, würden auf den typischen Signifikanzniveaus (etwa $p < .1$, für welches Unterschiede als vorsignifikant bezeichnet würden) einige aus inhaltlicher Sicht relevante Karten ausgeschlossen, obwohl große absolute Unterschiede in den Bearbeitungsdauern vorliegen.

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

Tabelle 5.9: Übersicht über Karten der explizit-fachmethodischen Instruktionsvariante, für die Unterschiede mit $p < .2$ und $r > .5$ im Vergleich der Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen auftreten.

Karte	Kurzbeschreibung des Inhalts	Länger arbeitet
1-07	Informationstext zu präzisen und allgemeinen Fragestellungen	N
1-11	Hypothese auswählen und Begründung verschriftlichen	N
1-14	Frage einschätzen (präzise/allgemein), Versuch kurz durchführen	N
1-18	Fragestellungen formulieren, die vorgegebenen Kriterien (präzise/allgemein) entsprechen	H
1-25	Fragestellungen einschätzen (naturwissenschaftlich oder nicht) und Begründung verschriftlichen	N
1-33	Informationstext zu Hypothesen und Vermutungen	H
2-03	Titelblatt der Einheit (nachdem Anonymisierungscode notiert)	N
2-05	Präzise Fragestellung zu gegebener Situation formulieren	N
2-06	Zustände von Variablen benennen	H
2-10	Informationsschaubild zu (un)abhängigen Variablen und Kontrollvariablen sowie einer präzisen Fragestellung	H
2-13	Titelblatt des zweiten Abschnitts	N
2-14	Einleitungstext zu einer Versuchsreihe	H
2-15	Variablen (un-/abhängige und Kontrollvariablen) benennen	H
2-28	Fachinhaltliche Frage zur durchschnittlichen Fallgeschwindigkeit	H
2-33	Versuch zu Kegeln planen und durchführen	N
2-34	Eigene Versuchsdurchführung zu Kegeln prüfen und bei Fehlern erneut durchführen	N
2-39	Zu Versuch passende Frage überlegen, Versuch durchführen	H
2-43	Vorgegebene tabellierte Werte zu Versuchsreihen interpretieren	H
2-56	Zusammenfassender Informationstext zu Messwiederholungen	H
3-05	Versuch zügig durchführen, Variablenkontrolle kurz diskutieren	N
3-29	Mit Versuch vertraut machen (Skizze zu Parametern ausfüllen)	H
3-31	Mit Versuch vert. machen (Parameter in Realsituation erkennen)	H
3-35	Versuchsdurchführung, Messwerte notieren, Antwort auf Fragestellung formulieren	H
3-37	Regeln zum Anlegen von Tabellen in Bsp. auf Einhaltung prüfen	H
3-38	Deutungen auf Pasung zu Messwerten überprüfen	N

Anmerkungen. H/N = Der Mittelwert für die Gruppe der Teams mit gemittelt hohen/niedrigen Kompetenzzuwächsen ist höher als der jeweils andere Mittelwert.

ergibt sich aus dem Vorgehen beim Kodieren. Die Karten wurden zwar so kodiert, dass sie auch dann erneut zugewiesen wurden, wenn sie später erneut von den Lernenden hervorgeholt wurden. Allerdings ist durch das Vorgehen nicht abgedeckt, wenn Lernende parallel zu einer Bearbeitung eine früher in der Instruktion verortete Informationskarte nutzen. Ferner ist es auch möglich, dass sich Lernende eigenständig bei früheren Karten Fragen stellen, die erst später durch die Instruktion aufgeworfen werden. Beispielsweise führen einige Lernende Experimente zu Beispielen aus der Instruktion durch, obwohl keine Aufforderung dazu gegeben wird. Später, wenn eine Experimentierserie durchgeführt werden soll, nutzen die Lernenden dann teilweise ihre vorigen Erfahrungen und verkürzen somit die Experimentierzeit. Mit dieser Einschränkung könnte das Ergebnis verbunden sein, dass sich im Umfeld der in Tabelle 5.9 berichteten Karten teilweise Karten finden, bei denen die jeweils andere Gruppe (Teams mit gemittelt niedrigen bzw. Teams mit gemittelt hohen Kompetenzzuwächsen) einen höheren Mittelwert für die Bearbeitungsdauer aufweisen (Beispielsweise Karten 2-05 und 2-06, aber teilweise auch nicht in der Tabelle berichtete Karten, bei denen die Effekte zwar nicht so groß sind, dass sie berichtet werden, aber die absoluten Unterschiede durchaus nicht unbedeutend sind). Das könnte nahelegen, dass teilweise von Karte zu Karte zeitlicher Ausgleich stattfindet, beispielsweise, weil Experimentiertätigkeiten auf nachfolgenden Karten weitergeführt werden. Diesem Umstand können und müssen kartenspezifische und kartenübergreifende Analysen von Aktivitäten Rechnung tragen.

5.4.2 Aktivitätsprofile der Lernenden

Wird danach gefragt, welche Bedeutung Aktivitäten für den Kompetenzaufbau haben, kann insbesondere auf das Zusammenspiel der Aktivitäten geachtet werden. Dafür kann in grober Näherung eine Betrachtung von Aktivitäten in ihrer Gesamtheit stattfinden. Lernende, die deutlich weniger über Fachinhalte sprechen, könnten beispielsweise mehr Zeit für fachmethodische Beiträge nutzen und daher höhere Kompetenzzuwächse erreichen. Andererseits wäre auch denkbar, dass nur Lernende, die mehr Zeit für beide Aktivitäten einsetzen als andere Lernende, hohe Kompetenzzuwächse erreichen.

In der Arbeit werden für die Lernenden ihre Aktivitätsprofile betrachtet, wobei als Aktivitätsprofil die – bildlich durch ein Balkendiagramm (illustrativ in Abbildung 5.15) darstellbare – Gesamtheit der zeitlichen Anteile von ausgewählten Kategorien an der Gesamtbearbeitungsdauer (der jeweils betrachteten Einheit) bezeichnet wird. Aktivitätsprofile werden in der Arbeit für jede einzelne Person sowie für jedes Team (Addition der Werte der Teammitglieder) betrachtet. Abbildung 5.15 zeigt beispielsweise, dass die dargestellte Person 8.84 % ihrer Gesamtbearbeitungsdauer bei der ersten Einheit der explizit-fachmethodischen Instruktionsvariante mit der nonverbalen Vorbereitung und/oder Durchführung von naturwissenschaftlichen Versuchen verbringt. Die Summe der zeitlichen Anteile ergibt keine 100 %, weil Aktivitäten einander zeitlich überlappen können, nicht alle Aktivitäten kodiert wurden, nicht alle kodierten Aktivitäten im Aktivitätsprofil abgebildet sind und aufgrund des eventbasierten Vorgehens vielen Zeitsegmenten überhaupt keine Kategorie zugewiesen

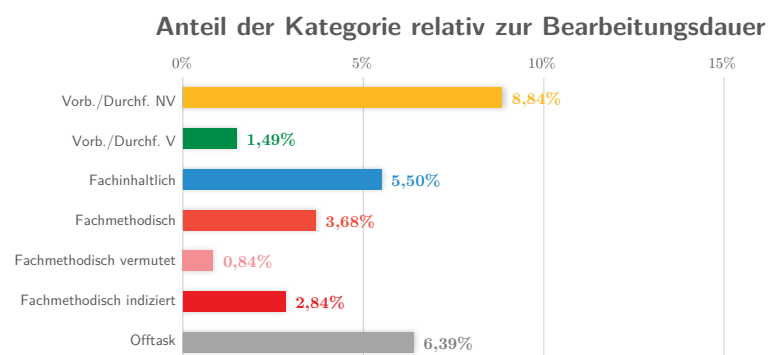


Abbildung 5.15: Illustrative Darstellung vom Aktivitätsprofil der Schülerin B6 als Balkendiagramm. (Der Anteil von OffTask ist untypisch im Vergleich zur Gesamtheit der Personen.)

ist. Der Mittelwert über den Anteil, den alle verbalen Aktivitäten einer Person zusammengekommen an der Gesamtbearbeitungsdauer für diese Person einnehmen, beträgt 14.56 %. Der Mittelwert über den Anteil, den die in den Aktivitätsprofilen analysierten verbalen Aktivitäten (Fachinhaltlich, Fachmethodisch und VorbereitenDurchführenVerbal) einer Person zusammengekommen an der Gesamtbearbeitungsdauern für die jeweilige Person einnehmen, beträgt 9 %. Die restlichen Prozente entfallen auf verbale Aktivitäten des Vorlesens oder die Kategorie Sonstiges und machen einen erwartbar großen Anteil der Bearbeitungsdauer aus. Der Mittelwert über den Anteil, den die verbalen Aktivitäten aller (zwei bis) drei Personen derselben jeweiligen Teams an der Gesamtbearbeitungsdauer des Teams einnehmen, beträgt 41.43 % (ist also etwas niedriger als das Dreifache des Personenmittelwerts, das 43.68 % beträgt). Der Mittelwert über den Anteil, den die in den Aktivitätsprofilen analysierten verbalen Aktivitäten (Fachinhaltlich, Fachmethodisch und VorbereitenDurchführenVerbal) aller (zwei bis) drei Personen derselben jeweiligen Teams an der Gesamtbearbeitungsdauer des Teams einnehmen, beträgt 26.61 % (und ist damit ebenfalls etwas niedriger als das Dreifache des Personenmittelwerts, das 26.99 % beträgt). Genauere Angaben zu den mittleren Dauern aller verbalen Aktivitäten finden sich in Tabelle 5.10; zu den Dauern der in den Aktivitätsprofilen analysierten verbalen Aktivitäten (Fachinhaltlich, Fachmethodisch und VorbereitenDurchführenVerbal) finden sich Details in Tabelle 5.11. Der mittlere Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführend von Versuchen beträgt 11.72 % (Standardabweichung: 7.08 %). Weitere Angaben zu nonverbalen Aktivitäten des Vorbereitens und Durchführens finden sich bei den nachfolgend dargestellten Ergebnissen zu Aktivitätsprofilen.

Tabelle 5.10: Mittlerer Anteil *aller* verbalen Aktivitäten an der Gesamtbearbeitungsdauer bei den einzelnen Einheiten aufgelöst auf Ebene der Personen und auf Ebene der Teams.

Einheit	Personen				Teams			
	MW	SD	Min	Max	MW	SD	Min	Max
1	15,36%	6,35%	3,98%	26,56%	46,09%	7,79%	31,73%	53,48%
2	13,85%	6,50%	4,97%	25,89%	38,09%	3,57%	33,37%	42,43%
3	13,96%	5,63%	5,09%	24,39%	38,97%	6,16%	32,26%	48,16%
Gesamt	14,56%	6,23%	3,98%	26,56%	41,43%	7,35%	31,73%	53,48%

Anmerkung. In der Gesamt-Zeile ist der Mittelwert über alle existierenden Paarungen von Einheiten mit Personen bzw. Teams aufgeführt.

Tabelle 5.11: Mittlerer Anteil *der in den Aktivitätsprofilen betrachteten* verbalen Aktivitäten an der Gesamtbearbeitungsdauer bei den einzelnen Einheiten aufgelöst auf Ebene der Personen und auf Ebene der Teams.

Einheit	Personen				Teams			
	MW	SD	Min	Max	MW	SD	Min	Max
1	8,72%	4,14%	2,15%	16,89%	26,16%	6,83%	15,04%	34,86%
2	10,63%	5,73%	3,58%	21,78%	29,96%	4,55%	21,40%	35,69%
3	8,00%	4,58%	2,00%	21,07%	22,48%	4,47%	18,28%	29,19%
Gesamt	9,00%	4,83%	2,00%	21,78%	25,61%	6,16%	15,04%	35,69%

Anmerkung. In der Gesamt-Zeile ist der Mittelwert über alle existierenden Paarungen von Einheiten mit Personen bzw. Teams aufgeführt.

In dem vorliegenden Unterabschnitt werden die Analysen zu Aktivitätsprofilen berichtet, die sich anhand aller für die explizit-fachmethodische Instruktionsvariante analysierten Lernenden ergeben. Da für die verschiedenen Einheiten der Instruktion unterschiedliche Stichproben zugrunde liegen, werden die Analysen für jede der Einheiten einzeln vorgenommen. Als Ergebnisse finden sich, zum einen große Vielfalt in den Aktivitätsprofilen (es finden sich viele verschiedene Aktivitätsprofiltypen) und zum anderen, dass die Aktivitätsprofile für die verschiedenen Einheiten der Instruktion grundsätzlich vergleichbar sind, allerdings zentrale Aktivitäten wie das nonverbale Vorbereiten und Durchführen oder das fachmethodisch-indizierte Beitragen durchaus von Einheit zu Einheit verschieden sind. Im nächsten Unterabschnitt (5.4.3, ab S. 163) werden dann Bezüge zu den Kompetenzzuwächsen der Lernenden hergestellt.

5.4.2.1 Vorgehen: Clusteranalysen zur Bildung von Aktivitätsprofiltypen

Für die Analysen zu den Aktivitätsprofilen werden fünf als potentiell für den Kompetenzaufbau relevant erachtete Aktivitäten von Lernenden berücksichtigt (zur Auswahl siehe Abschnitt 5.1): Nonverbale Durchführung von Versuchen (abgebildet durch die Kategorie *VorbereitenDurchführenNonverbal* aus dem Kategoriensystem), technische Äußerungen zur Vorbereitung und Durchführung von Versuchen (*VorbereitenDurchführenVerbal*), Äußerungen zu Fachinhalten (*Fachinhaltlich*) und indizierte und vermutete Äußerungen zu Fachmethoden (*Fachmethodisch-Indiziert* und *Fachmethodisch-Vermutet*). Der Unterschied zwischen den beiden Kategorien für fachmethodische Beiträge besteht darin, dass fachmethodisch-indizierte Beiträge einen deutlicheren Bezug zu Fachmethoden oder fachmethodischen Konzepten aufweisen als fachmethodisch-vermutete Beiträge, bei denen eine Beschäftigung mit Fachmethoden vom Beobachter vermutet wird, aber nicht an sogenannten Indizien festgemacht werden kann (siehe Abschnitt 5.1). Die Kategorie *OffTask* wird nicht in die Analysen einbezogen, weil sie nur mit geringem zeitlichen Anteil auftritt (siehe Unterabschnitt 5.4.1); die Kategorie *Fachmethodisch* hat als Summe der beiden Kategorien zu fachmethodischen Beiträgen keinen statistischen Mehrwert und wird daher ebenfalls nicht dezidiert in den statistischen Clusteranalysen betrachtet.

Die zeitlichen Anteile der Aktivitäten werden als *relative* Anteile der Längen der Kategorien *an der Gesamtbearbeitungsdauer* der jeweiligen Lernenden für die betrachtete Einheit angegeben und verwendet. Die Entscheidung basiert darauf, dass die Analysen zu den Aktivitätsprofilen auf die Relationen unterschiedlicher Aktivitäten zueinander gerichtet sind und dass die Gesamtbearbeitungsdauern der Einheiten und Abschnitte sich für unterschiedliche Teams zum Teil deutlich unterscheiden, während die relativen Anteile der Abschnittsbearbeitungsdauern sehr ähnlich sind (siehe Unterabschnitt 5.4.1).

Für die Aktivitäten ist davon auszugehen, dass diese in Prozessen des Kompetenzaufbaus zu verschiedenen Teilprozessen des experimentbezogenen Denkens und Arbeitens unterschiedlich stark ausgeprägt sind. Dass die Analysen zu Aktivitätsprofilen für jede Einheit einzeln durchgeführt werden, trägt diesem Umstand Rechnung. Ferner kann ein Vergleich der einheitenweisen Analysen einen Beitrag dazu leisten, Vermutungen zur Unterschiedlichkeit der Aktivitäten bei der Bearbeitung von expliziter Instruktion zu verschiedenen Teilprozessen des experimentbezogenen Denkens und Arbeitens genauer zu untersuchen.

Ähnliche Aktivitätsprofile werden zu *Aktivitätsprofiltypen* zusammengefasst, während unterschiedliche Aktivitätsprofile unterschiedlichen Aktivitätsprofiltypen zu-

geordnet werden. Basis der Identifikation von Aktivitätsprofiltypen bilden Clusteranalysen zu den zeitlichen Anteilen der Aktivitäten. Die nachfolgend berichteten Clusteranalysen gehen den Fragen nach, welche Aktivitätsprofiltypen für einzelne Personen identifiziert werden können und welche sich für Teams finden lassen.

Für Clusteranalysen ist grundsätzlich zu entscheiden, welches *Distanz bzw. Ähnlichkeitsmaß* eingesetzt wird und welcher *Clustering-Algorithmus* verwendet wird. Als Maß für die Ähnlichkeit bietet sich unter anderem das Korrelationsmaß nach Pearson an, weil – wie oben erläutert – vor allem die Verhältnisse der zeitlichen Anteile von Kategorien zueinander für die Bildung von Aktivitätsprofiltypen berücksichtigt werden sollen. Clustering-Algorithmen sind je nach betrachtetem Gegenstand, Ziel und Zweck auszuwählen (Romesburg, 2004). Für die aus den Kodierungen entnommenen Daten wurde das folgend erläuterte Verfahren zunächst an einzelnen Datensätzen erprobt und mit den Ergebnissen anderer Verfahren, die andere Algorithmen nutzen, abgeglichen. Das Verfahren hat sich als angemessen erwiesen, weil es viele der inhaltlich bedeutsamen Unterschiede abbilden kann und Aktivitätsprofiltypen nahelegt, die dem Ziel der Beschreibung von Unterschieden und Gemeinsamkeiten in den Aktivitätsprofilen dienlich sind. Das eingesetzte Verfahren zur Clusterung besteht aus drei Schritten, mit denen jeweils unterschiedliche Teilziele verfolgt werden.

Schritt 1: Ausreißer identifizieren und gegebenenfalls Partitionen erkennen. In einem ersten Schritt sollen Ausreißer identifiziert werden. Das heißt, dass Aktivitätsprofile (von Personen oder Teams), die sich hinsichtlich der Verhältnisse der Aktivitäten untereinander sehr stark von anderen Aktivitätsprofilen (von anderen Personen oder anderen Teams) unterscheiden, identifiziert werden. Diese können dann für den zweiten Schritt ausgeschlossen werden und im dritten Schritt ggf. wieder aufgenommen und inhaltlich ausgedeutet werden. Ein Nebenprodukt des ersten Schrittes kann darin bestehen, dass Partitionen erkannt werden, das heißt, dass zwei (oder drei) größere Gruppen von Aktivitätsprofilen identifiziert werden, die sich deutlich voneinander unterscheiden.

Als Clustering-Algorithmus⁵⁰ bietet sich die *Analyse nächstgelegener Nachbarn* an (auch single-linkage/SLINK genannt, z.B. Wentura & Pospeschill, 2015 oder Romesburg, 2004). Der Algorithmus lässt sich so beschreiben, dass zunächst alle Aktivitätsprofile als eigene Cluster oder Fälle aufgefasst werden. Basierend auf einer

⁵⁰Es werden nur hierarchische agglomerierende Clustering-Algorithmen genutzt, weil nicht-hierarchische und divisive Verfahren weder dem Gegenstand angemessen sein dürften noch häufig genutzt werden (Methodenberatung Zürich, 2021; Wentura & Pospeschill, 2015). Ferner werden Varianzmethoden, wie bspw. der Ward-Algorithmus, nicht genutzt, weil davon auszugehen ist, dass die Aktivitäten höher als dafür wünschenswert (angemerkt bspw. von Wentura & Pospeschill, 2015) miteinander korreliert sind. Alle Analysen werden in SPSS durchgeführt.

paarweisen Berechnung der Ähnlichkeiten aller Fälle (anhand der Pearson-Korrelation für die fünf betrachteten Variablen der Aktivitäten-Kategorien) werden dann in einem zweiten Step des Algorithmus die beiden ähnlichsten Cluster zusammengelegt. In jedem weiteren Step des Algorithmus müssen Ähnlichkeiten von Clustern, die aus mehreren Aktivitätsprofilen bestehen, berechnet werden. Die Ähnlichkeit zweier Cluster ergibt sich dabei aus der Ähnlichkeit der beiden am nächsten beieinanderliegenden Fälle (daher der Name: *nächstgelegener Nachbar*), von denen jeweils einer in einem der beiden zu vergleichenden Cluster ist. Darauf basierend werden in jedem weiteren Step des Algorithmus die beiden ähnlichsten Cluster oder Fälle zusammengeführt. Finden sich mehrere Paare mit gleichen Ähnlichkeiten, werden all diese (jeweils) zusammengeführt. Anschließend beginnt der nächste Step des Algorithmus und es werden erneut die Ähnlichkeiten berechnet und Cluster zusammengeführt. Der Algorithmus stoppt, sobald sich alle Fälle in einem Cluster befinden. Mithilfe der Reihenfolge der Zusammenführung von Fällen und Clustern sowie den in jedem Step des Algorithmus berechneten Distanzmaßen können die am besten passenden Cluster identifiziert werden (genauer weiter unten bei Schritt 3).

Das Verfahren der Analyse nächstgelegener Nachbarn kann dafür eingesetzt werden, Aktivitätsprofile mit sehr geringen Ähnlichkeiten voneinander zu trennen. Es hat eine hohe Sensibilität für Ausreißer und ermöglicht zudem das Erkennen von Gefahren der Kettenbildung (Handl, 2010, S. 404). Kettenbildung ist ein Phänomen bei Clusteranalysen, das darin besteht, dass *ein* Cluster mit jedem weiteren Step des Algorithmus mehr und mehr Fälle umfasst (weil es weitere Fälle quasi anzieht), während sich andere eigenständige Cluster nicht oder nur vereinzelt bilden (im oberen Drittel von Abbildung 5.16 auf S. 130 visualisiert, Erläuterung der Abbildung folgt erst später im Text). Das Phänomen kann nicht grundsätzlich vermieden werden, weil zum Teil durch ein angemessenes Abbilden der Struktur der Daten begründet ist (Romesburg, 2004, S. 139). Kettenbildung erschwert aber die Deutung von Clustern (Romesburg, 2004), beispielsweise, weil – dadurch, dass Cluster in sich recht heterogen sein können, ohne dass dadurch die Distanz zu anderen Clustern gefährdet ist – Ähnlichkeiten suggeriert werden, die faktisch nicht vorliegen. Ferner eignet sich der Algorithmus nächstgelegener Nachbarn dazu, grundsätzliche Partitionen zu erkennen (Wentura & Pospeschill, 2015, S. 171), beispielsweise weil es zwei grundsätzlich verschiedene Arten von Aktivitätsprofiltypen gibt (inhaltlich wäre beispielsweise denkbar, dass es Personen gibt, die relativ betrachtet deutlich mehr Zeit mit nonverbaler Versuchsdurchführung verbringen, wohingegen andere Personen darauf keine oder nur sehr wenig Zeit verwenden).

Schritt 2: Ähnlichkeiten der Aktivitätsprofile messen. In einem zweiten Schritt soll die eigentliche Analyse zu Aktivitätsprofiltypen vorgenommen werden, bei der besonders die Ähnlichkeit der Aktivitätsprofile *innerhalb* eines Clusters gesichert werden soll.⁵¹ Dazu werden die im ersten Schritt identifizierten Ausreißer aussortiert; anschließend wird (ggf. für jede in Schritt 1 identifizierte Partition) eine (eigene) erneute Clusteranalyse vorgenommen, ohne dass die Cluster aus Schritt 1 weiter berücksichtigt werden.

Als Clustering-Algorithmus wird eine Analyse der *Verlinkung innerhalb der Cluster* durchgeführt. Dieser Algorithmus ist ähnlich zum oben beschriebenen Algorithmus, bei dem eine Analyse nächster Nachbarn vorgenommen wird (siehe Schritt 1). Er weicht nur dahingehend ab, dass die Ähnlichkeiten von Clustern, die mehr als einen Fall beinhalten, anders berechnet wird: Es werden alle Fälle der beiden Cluster zusammengenommen und alle möglichen Paare aus ihnen gebildet. Dabei können auch zwei Fälle desselben Clusters ein Paar bilden. Für jedes der Paare von Fällen wird die Ähnlichkeit berechnet; aus allen Ähnlichkeiten wird dann der Durchschnitt ermittelt;⁵² dieser Durchschnitt wird als Distanz der beiden Cluster angesehen. Der gewählte Algorithmus führt dazu, dass die durchschnittliche Ähnlichkeit innerhalb des neuen Clusters möglichst hoch ist (Wentura & Pospeschill, 2015).

Schritt 3: Entscheidung für Cluster und Beschreibung der Aktivitätsprofiltypen. In einem dritten Schritt werden die Aktivitätsprofile der durch den zweiten Schritt vorgeschlagenen Cluster als Aktivitätsprofiltypen interpretiert und grafisch verglichen, um Gemeinsamkeiten und Unterschiede zu beschreiben. Dabei ist es möglich, dass Unterschiede oder Gemeinsamkeiten erkannt werden, die vom Algorithmus in Schritt 2 (zugunsten anderer Aspekte) übergangen wurden. Beispielsweise kann es sein, dass innerhalb eines Clusters der Anteil fachmethodischer Beiträge den Anteil fachinhaltlicher Beiträge in manchen Aktivitätsprofilen überwiegt, in manchen aber nicht; aus inhaltlicher Sicht, ist hier aber auf Konsistenz zu achten, weil die Kategorien (und vermutlich auch deren Verhältnisse zueinander) zentral für die Arbeit sind. Daher werden in derartigen Fällen händisch Anpassungen der Aktivitätsprofiltypen vorgenommen, so dass diese nicht vollständig mit den vorgeschlagenen Clustern übereinstimmen. Jeder der in Schritt 1 aussortierten Aktivitätsprofile wird, sofern inhaltlich plausibel, einem der gebildeten Aktivitätsprofiltypen hinzugefügt.

Vorgeschlagene Cluster können prinzipiell aus jedem Step des Algorithmus von

⁵¹Im ersten Schritt wird die Ähnlichkeit *verschiedener* Cluster *minimiert*, im zweiten Schritt wird die Ähnlichkeit *innerhalb* jedes einzelnen Clusters *maximiert*.

⁵²Die Sätze ab dem Doppelpunkt sind durch Wortersetzungen auf die hier vorliegende Situation angepasst, stammen aber ansonsten nahezu wörtlich von Wentura und Pospeschill (2015, S. 170).

Schritt 2 hervorgehen. Aus Clusteranalysen ergibt sich keine fest vorgegebene Clusterung. Es ergibt sich aber in jedem Step des Algorithmus eine Gesamtheterogenität der gebildeten Clusterung, welche als Orientierung für die Entscheidung für die Cluster genutzt werden kann. Vereinfacht gesprochen, bieten sich die Clusterungen, nach denen die Gesamtheterogenität einen großen Sprung macht, besonders gut an (Methodenberatung Zürich, 2021).

Die Gesamtheterogenität der Clusterung wird häufig in Dendrogrammen abgebildet. Ein Beispiel für ein Dendrogramm findet sich in Abbildung 5.16; eine Zwei-Cluster-Lösung wäre zu bevorzugen, weil hier der größte Heterogenitätssprung stattfindet (von 16 zu 25), aber auch der markierte Heterogenitätssprung von der Drei-Cluster-zur Zwei-Cluster-Lösung ist sehr groß (von 8 zu 16), weshalb auch eine Drei-Cluster-Lösung statistisch sehr gut denkbar ist. Aus dem Dendrogramm gehen ferner alle Zusammenführungen hervor, die in den Steps des Algorithmus vorgenommen werden: Vertikale Linien heißen, dass die verbundenen Cluster zusammengeführt werden; horizontale Linien heißen, dass ein Cluster in einem Step eigenständig bleibt und nicht mit einem anderen zusammengeführt wird. Im Dendrogramm ist auf der horizontalen Achse nicht aufgeführt, in welchem Step des Algorithmus die Zusammenführung

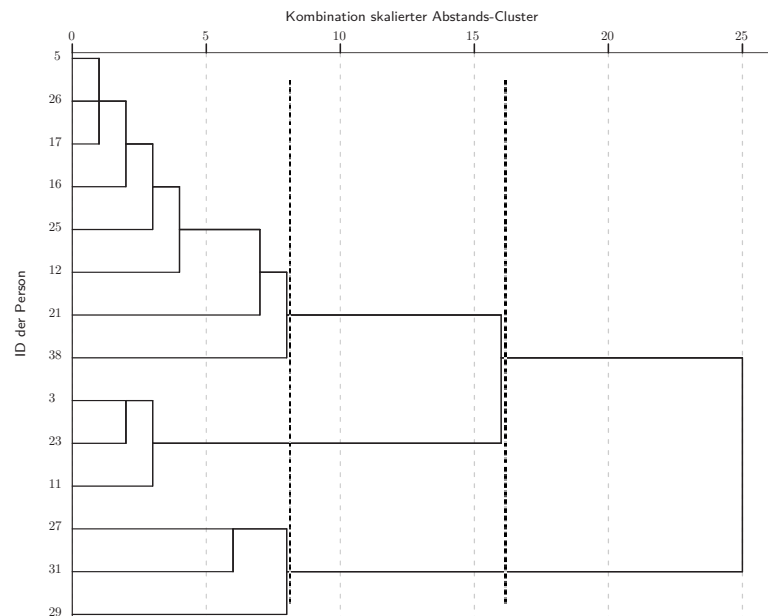


Abbildung 5.16: Exemplarisches Dendrogramm, welches den Prozess des Clusterings darstellt. (Zwei-Cluster-Lösung statistisch der Drei-Cluster-Lösung etwas überlegen; gegenstandsbezogene Überlegungen sind zur Entscheidung notwendig.)

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

erfolgt, sondern welches Gesamtheterogenitätsmaß in diesem Step erreicht wird. Im genutzten Statistikprogramm werden die Werte für die Gesamtheterogenität automatisch so skaliert, dass die höchste Gesamtheterogenität (erreicht im letzten Step des Algorithmus) 25 entspricht (erklärt bspw. bei Schendera, 2011).

Für die Entscheidung, welche der vorgeschlagenen Cluster gewählt werden, sollten grundsätzlich auch gegenstandsbezogene Überlegungen hinzugezogen werden (Wentura & Pospeschill, 2015, S. 172; Methodenberatung Zürich, 2021). Für die Arbeit wird – entlang der Zusammenführungspfade des Dendrogramms – unter zwei Gesichtspunkten entschieden, welche Cluster berichtet werden, auch wenn diese ggf. nicht statistisch zu bevorzugen wären:

- Es sollen möglichst drei oder mehr Personen (bzw. zwei oder mehr Teams) in einem Cluster sein, allerdings nicht mehr als $\frac{2}{3}$ der Stichprobe.
- Es soll in jedem Cluster eine inhaltliche Gemeinsamkeit identifiziert werden können, die in anderen Clustern nicht auftritt.

5.4.2.2 Aktivitätsprofile für Personen bei Einheit 1: Ergebnisse und exemplarische ausführliche Darstellung des Vorgehens

Die Analyse der Aktivitätsprofile der Personen, die Einheit 1 bearbeitet haben, wird an dieser Stelle ausführlich dargestellt, während für weitere Analysen zu Personen nur die Ergebnisse berichtet werden und auf Besonderheiten eingegangen wird. Zunächst werden die Ergebnisse präsentiert, dann folgen ausführliche Erläuterungen zu den Schritten des Verfahrens. Alle Ausführungen sind darauf bezogen, welche zeitlichen Anteile die Aktivitäten an der Gesamtbearbeitungsdauer von Einheit 1 ausmachen.

Vorlaufend zu der Aufteilung der Aktivitätsprofile der Personen in Aktivitätsprofiltypen wird eine Betrachtung der mittleren Anteile aller betrachteten Aktivitäten für die erste Einheit der explizit-fachmethodischen Instruktionsvariante dargestellt. In Abbildung 5.17 ist neben den mittleren Anteilen der Aktivitäten auch die Standardabweichung bezüglich der analysierten 33 Personen eingetragen. Die große Streuung der Werte über die Personen hinweg spiegelt sich in den großen Standardabweichungen wider, die in der Abbildung eingetragen sind. Die Schwankungen gehen für gewöhnlich nicht in einfacher Weise miteinander einher (etwa so, dass Lernende, die gering ausgeprägte fachmethodische Aktivitäten zeigen, deutlich höhere fachinhaltliche Aktivitäten aufweisen würden), was sich in den nachfolgend dargestellten Aktivitätsprofiltypen widerspiegelt.

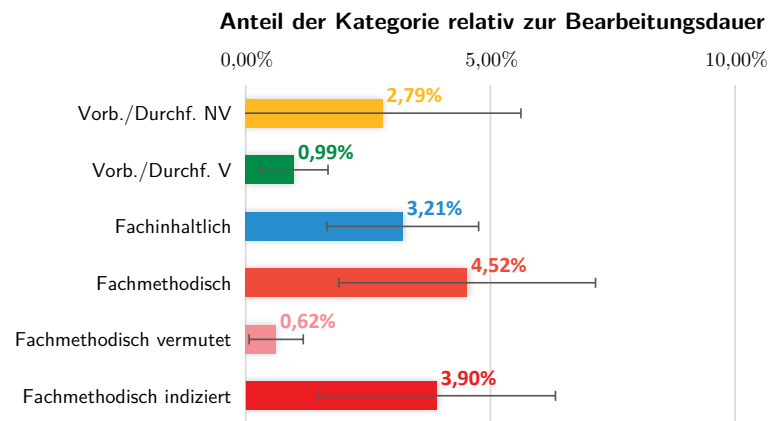


Abbildung 5.17: Mittlere Anteile der betrachteten Aktivitäten von 33 Personen bei der ersten Einheit der explizit-fachmethodischen Instruktionsvariante. Zusätzlich sind jeweils eine Standardabweichung des Mittelwerts nach oben und nach unten eingetragen.

Die Aktivitätsprofiltypen der einzelnen Personen finden sich in zwei Partitionen (siehe Erläuterungen zu Schritt 1 weiter unten; die Kürzel für die Partitionen ergeben sich daraus, dass die Aktivitätsprofile von Personen für die erste Einheit betrachtet werden und die Partitionen mit P1 und P2 bezeichnet werden):

E1-P-P1 Personen, die nonverbale Aktivitäten des Vorbereitens und Durchführens von naturwissenschaftlichen Versuchen zu einem (innerhalb von Einheit 1) *großen zeitlichen Anteil* der Gesamtbearbeitungsdauer (in etwa mehr als 5 %) tätigen. (Kurznotation: $VD_n \gg 0$.)

E1-P-P2 Personen, für die die Aktivität des Vorbereitens und Durchführens von naturwissenschaftlichen Versuchen einen *geringen zeitlichen Anteil* der Gesamtbearbeitungsdauer einnimmt. (Kurznotation: $VD_n \geq 0$.)

Genauer betrachtet kommt es nicht auf die absoluten Werte der zeitlichen Anteile an, sondern darauf, ob die Personen im Verhältnis zu ihren anderen Aktivitäten viel oder wenig Zeit mit **VorbereitenDurchführenNonverbal** verbringen. Der als Grenze angegebene Wert von 5 % ist als Schätzung allerdings eine gute Beschreibung der Datenlage und zeigt, dass es sich nicht einfach um eine Teilung direkt in der Nähe des Mittelwerts (2.79 %; siehe Abbildung 5.17) handelt.

Für die Partition mit großen Anteilen nonverbaler Aktivitäten des Vorbereitens und Durchführens von naturwissenschaftlichen Versuchen finden sich vier Aktivitätsprofiltypen. Zu jedem Aktivitätsprofiltyp ist in Abbildung 5.18 ein Repräsentant abgebildet, d. h. ein Aktivitätsprofil einer Person, das dem jeweiligen Typ zugeordnet ist. Die

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

Kürzel für die Aktivitätstypen ergeben sich analog zu den Partitionskürzeln, wobei für die Cluster ein weiterer Teil angehängt wird.

E1-P-P1-C1 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *deutlich größer* als der zeitliche Anteil fachinhaltlicher Beiträge. Der zeitliche Anteil fachinhaltlicher Beiträge hat in etwa die Größe des zeitlichen Anteils von VorbereitenDurchführenNonverbal. (Kurznotation: $FMi \gg FI \approx VDn$.)
Personen: A3, D11, H23.

E1-P-P1-C2 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist zwar in allen zugehörigen Aktivitätsprofilen größer als der zeitliche Anteil fachinhaltlicher Beiträge, aber *nur wenig größer* (Verhältnis liegt deutlich näher bei 1 als im Aktivitätstyp E1-P1-C2). Der zeitliche Anteil fachinhaltlicher Beiträge hat in etwa die Größe des zeitlichen Anteils von VorbereitenDurchführenNonverbal. (Kurznotation: $FMi > FI \approx VDn$.)
Personen: B5, D12, F16, F17, G21, I25, I26, M38.

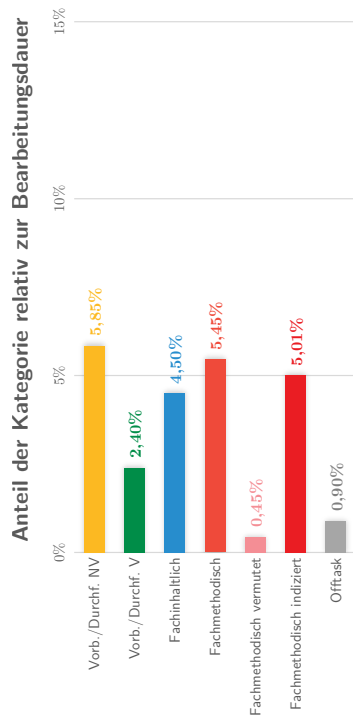
E1-P-P1-C3 Der zeitliche Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von naturwissenschaftlichen Versuchen ist *deutlich größer* als die zeitlichen Anteile aller anderen Aktivitäten (jeweils einzeln betrachtet). (Kurznotation: $VDn \gg$ alles andere.)
Personen: I27, J29, K31.

E1-P-P1-C4 Der zeitliche Anteil der fachmethodisch-indizierter Beiträge ist in allen zugehörigen Aktivitätsprofilen *kleiner* als der zeitliche Anteil fachinhaltlicher Beiträge. (Kurznotation: $FI > FMi$.)
Personen: L34.

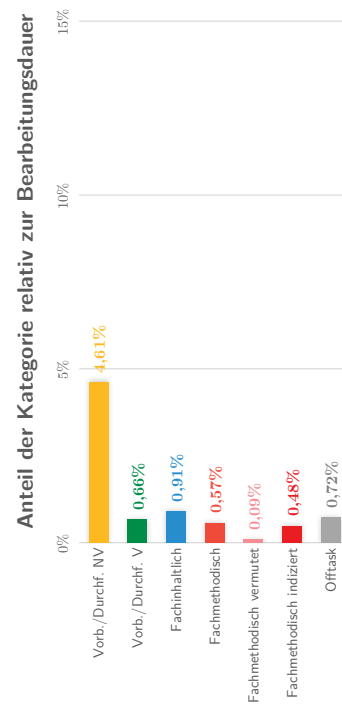
Für die Partition mit geringen Anteilen nonverbaler Aktivitäten des Vorbereitens und Durchführens von naturwissenschaftlichen Versuchen finden sich drei Aktivitätstypen:

E1-P-P2-C1 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *deutlich kleiner* als der zeitliche Anteil fachinhaltlicher Beiträge. (Kurznotation: $FI \gg FMi$.)
Personen: A1, A2, G20, J28, L36.

E1-P-P2-C2 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *deutlich größer* als der zeitliche Anteil fachinhaltlicher Beiträge. (Kurznotation: $FMi \gg FI$.)
Personen: B4, D10, G19, H22, J30, K32, L35, M39.

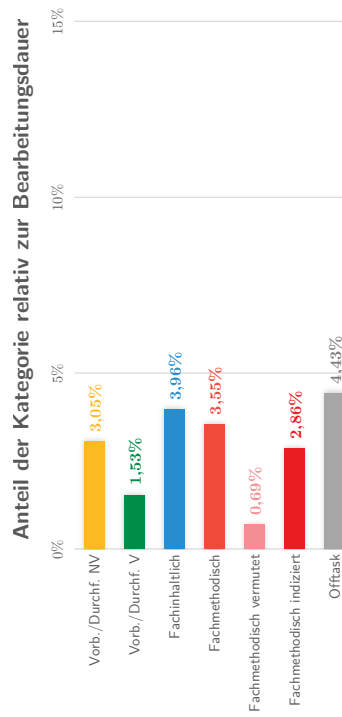


(a) E1-P-P1-C1 ($FMI \gg FI \approx VDN$): A3



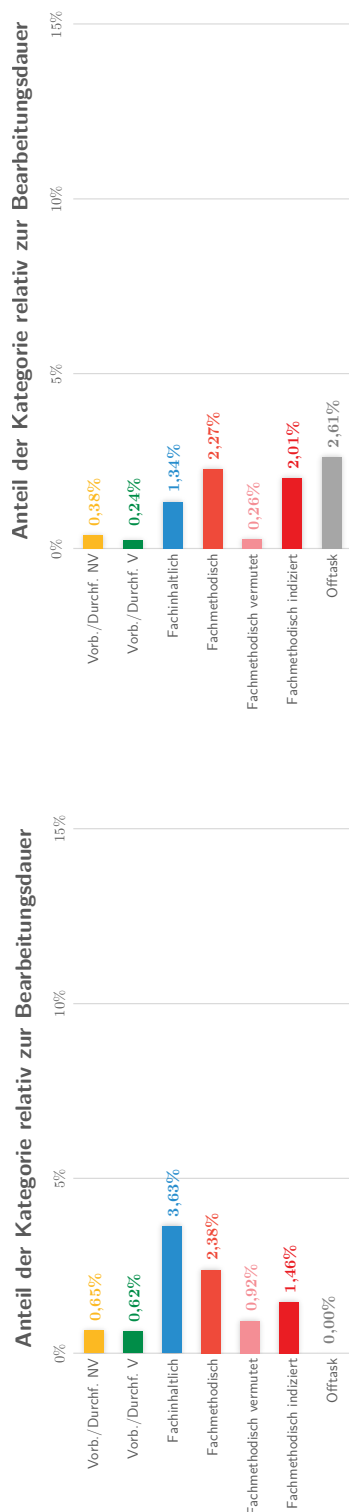
(c) E1-P-P1-C3 ($VDn \gg$ alles andere): I27

(b) E1-P-P1-C2 ($FMI > FI \approx VDN$): I25

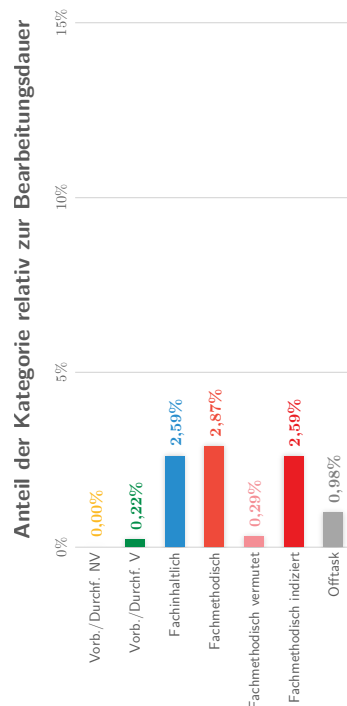


(d) E1-P-P1-C4 ($FI > FMI$): L34

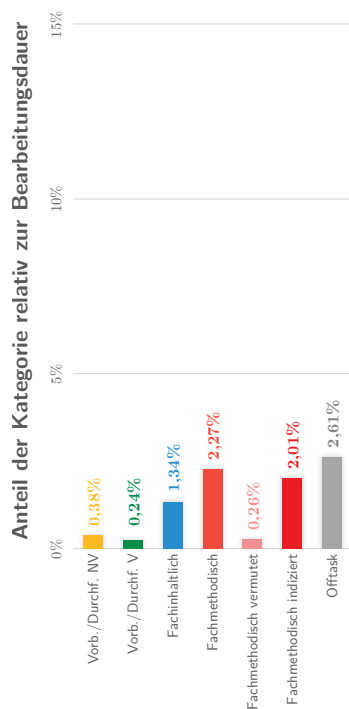
Abbildung 5.18: Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 1 der explizit-fachmethodischen Instruktionsvariante (Teil 1 der Abbildung: E1-P-P1-..., d. h. $VDn \gg 0$).



(e) E1-P-P2-C1 (FI $\gg$ FMI): A1



(g) E1-P-P2-C3 (FMI $\approx$ FI): F18



(f) E1-P-P2-C2 (FMI $\gg$ FI): B4

Abbildung 5.18: Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 1 der explizit-fachmethodischen Instruktionsvariante (Teil 2 der Abbildung: E1-P-P2-..., d. h. $VD_n \geq 0$).

E1-P-P2-C3 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *in etwa so groß* wie der zeitliche Anteil fachinhaltlicher Beiträge.

(Kurznotation: FMi $\approx$ FI.)

Personen: F18, H24, K33, M37.

Die inhaltlichen Beschreibungen der Aktivitätsprofiltypen stellen (hier und auch in späteren Unterunterabschnitten) Beschreibungen der aus Sicht des Autors zentralen Merkmale dar, anhand derer sich die Typen gut unterscheiden lassen. Teilweise liegen weitere Unterschiede zwischen den Aktivitätsprofilen vor, so dass die hier beschriebenen Unterschiede für ausgewählten Repräsentanten z. T. nicht so deutlich hervortreten wie im gemittelten Vergleich die Merkmale der Typen.

Diese Ergebnisse wurden entlang der Schritte generiert, die in Unterunterabschnitt 5.4.2.1 beschrieben sind. Nachfolgend sind einige Konkretisierungen des Vorgehens aufgeführt:

Schritt 1: Ausreißer identifizieren und Partionierungen erkennen. Alle Lernenden, die Einheit 1 bearbeitet haben (auch solche mit mittleren Kompetenzzuwächsen), wurden als Fälle der Clusteranalyse betrachtet. Die Variablen für die Clusteranalyse entsprechen den Kategorien *VorbereitenDurchführenNonverbal*, *VorbereitenDurchführenVerbal*, *Fachinhaltlich*, *Fachmethodisch-Indiziert* und *Fachmethodisch-Vermutet*; genau genommen wird der relative zeitliche Anteil der jeweilige Kategorie innerhalb der ersten Einheit der explizit-fachmethodischen Instruktionsvariante an der angesetzten Gesamtbearbeitungsdauer dieser Einheit als Variable gewählt. In Abbildung 5.19 findet sich das Dendrogramm, das den Prozess des Clusterings bei Anwendung des Algorithmus der Analyse nächstgelegener Nachbarn darstellt. Auf der horizontalen Achse ist ein auf 25 normiertes Maß für die Heterogenität der jeweiligen Clusterung dargestellt. Vertikale Linien stellen die Zusammenführung von Clustern (angefangen mit den Ein-Personen-Clustern ganz links) dar, während horizontale Linien bedeuten, dass die entsprechenden Cluster bestehen bleiben, während der Schwellwert für das Maß der Heterogenität erhöht wird und ggf. andere Cluster zusammengeführt werden.

Aus Abbildung 5.19 lässt sich entnehmen, dass nach und nach zwei große Cluster gebildet werden, die erst ganz am Ende zusammengeführt werden. Diese beiden Cluster sind im Dendrogramm durch eine händisch eingetragene gestrichelte horizontale Trennlinie (etwa in der Mitte) voneinander grafisch getrennt. Sie lassen sich inhaltlich mit dem Anteil der Kategorie *VorbereitenDurchführenNonverbal* in Verbindung bringen: Während sich Personen, die einen verglichen mit allen Personen großen Anteil der Gesamtbearbeitungsdauer (ca. 4–10 %) nonverbale Aktivitäten

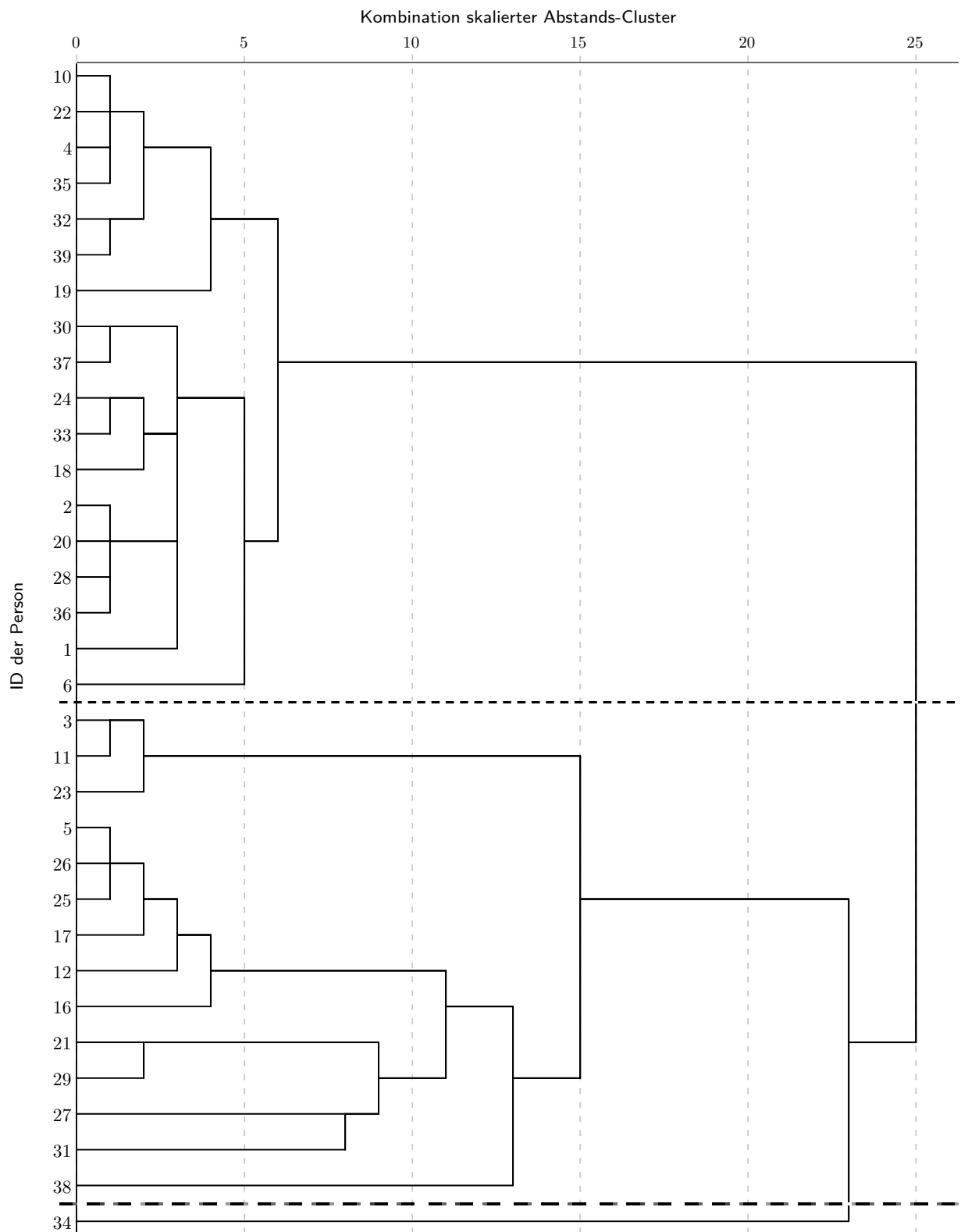


Abbildung 5.19: Dendrogramm zum Prozess des Clusterings der Personen, die Einheit 1 bearbeitet haben (Schritt 1: Pearson-Korrelation für nächstgelegene Nachbarn).

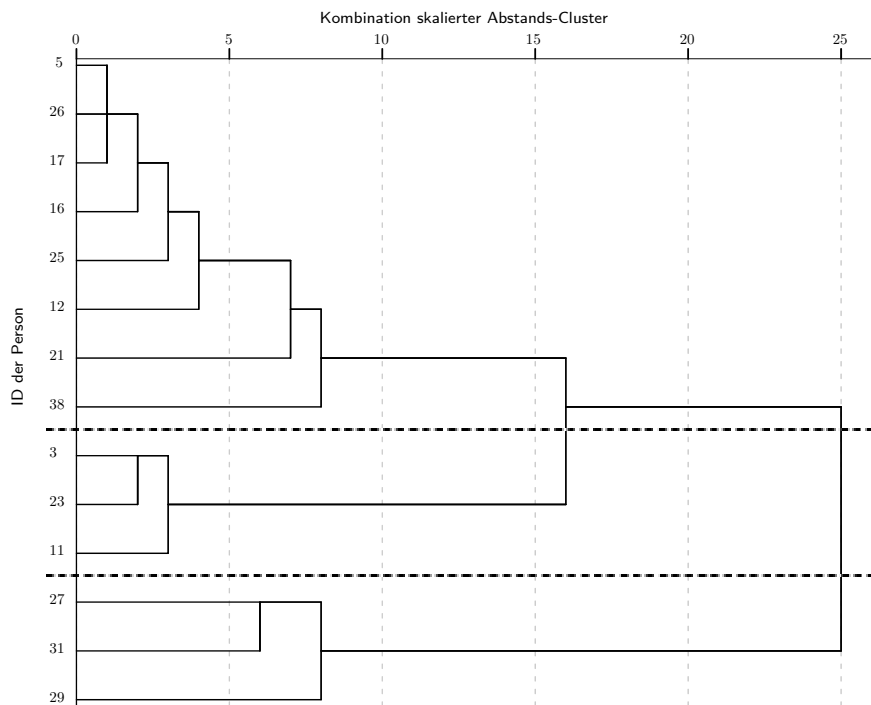
der Versuchsdurchführung vornehmen, im unteren Cluster befinden, finden sich Personen mit geringen Anteilen der Kategorie im Cluster oberhalb beider Trennlinien in Abbildung 5.19. Genauer betrachtet, kommt es nicht auf die absoluten Werte der Anteile an, sondern darauf, ob die Personen im Verhältnis zu ihren anderen Aktivitäten viel oder wenig Zeit mit **VorbereitenDurchführenNonverbal** verbringen. Die beiden Cluster von Personen wurden in nachfolgenden Schritten der Analyse als getrennte Partitionen betrachtet, um diesen grundsätzlichen Unterschied hinsichtlich der Aktivitäten zum nonverbalen Vorbereiten und Durchführen von Versuchen zu betonen.

Aus Abbildung 5.19 wird ferner deutlich, dass die Person L34 erst bei sehr hoher zugelassener Heterogenität dem unteren großen Cluster hinzugefügt wird. Sie ist grafisch ebenfalls durch eine Linie abgetrennt (ganz unten). Um Verzerrungen bei der Betrachtung des unteren großen Clusters zu vermeiden, wurde diese Person als Ausreißer gesondert betrachtet und wird weiter unten gesondert diskutiert. Auch wenn aus dem Dendrogramm noch weitere Hinweise auf spätere Cluster ersichtlich sind (etwa, dass die Personen A3, D11 und H23 recht sicher ein Cluster bilden), wurde dem Ziel des ersten Schrittes gemäß die nun zweigeteilte Stichprobe an den Algorithmus des zweiten Schritts übergeben.

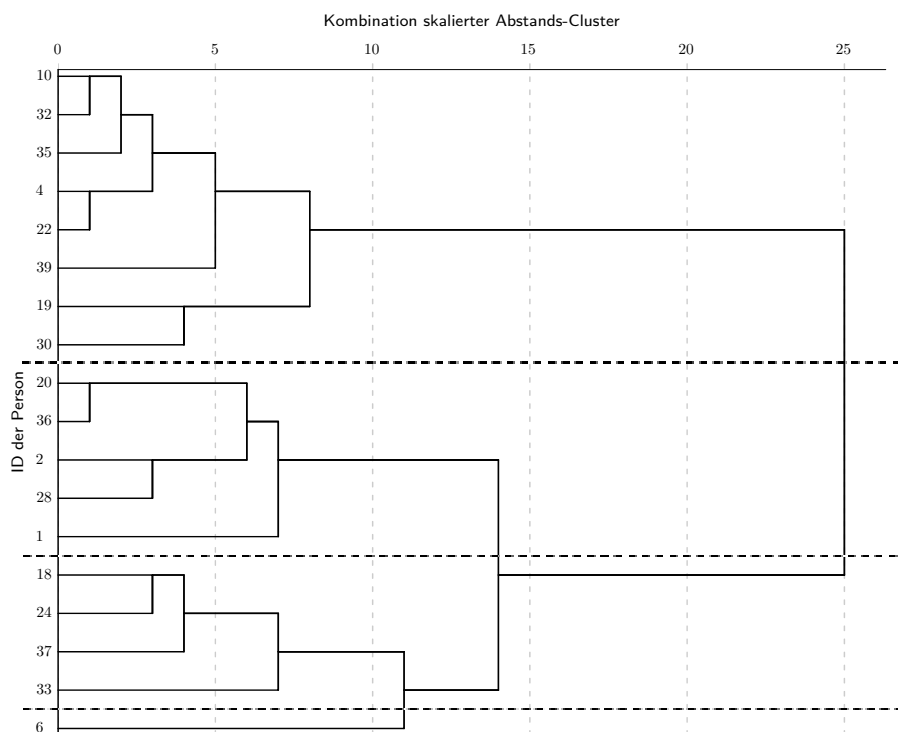
Schritt 2: Ähnlichkeiten der Aktivitätsprofile messen. Der zweite Schritt dient dazu, die Ähnlichkeiten der Aktivitätsprofile darzustellen. Erneut wird das Werkzeug des Dendrogramms verwendet. In diesem Fall fanden sich die beiden Dendrogramme in den Abbildungen 5.20a (Partition P1) und 5.20b (Partition P2).

In den Dendrogrammen sind bereits Trennlinien eingezeichnet, die sich für eine Clusterung aus der Betrachtung des Dendrogramms ableiten ließen. Dabei wurden vor allem die Längen der horizontalen Linien (Maß für die Zunahme der Heterogenität der Clusterung bei Zusammenführung) grafisch berücksichtigt. Das Hinzufügen einer einzelnen Person in einem späten Step zu einem deutlich früher gebildeten Cluster kann einen Hinweis darauf enthalten, dass diese Person deutlich von dem Cluster abweicht, dem sie hinzugefügt wird. Daher wurde Person B6 durch eine Trennlinie abgetrennt. Dennoch wurde im dritten Schritt inhaltlich überprüft, ob B6 dem Cluster nicht doch zugeordnet werden sollte.

Schritt 3: Entscheidung für Cluster und Beschreibung der Aktivitätsprofil-typen. Basierend auf dem Schritt 2 werden zur Entscheidung für die Cluster neben dem statistischen Maß der Heterogenität der Clusterung auch inhaltliche Überlungen hinzugezogen. Ferner wird, soweit dies möglich ist, davon abgesehen, allzu kleine



(a) Personen, die hohe Anteile der Aktivität VorbereitenDurchführenVerbal aufweisen



(b) Personen, die niedrige Anteile der Aktivität VorbereitenDurchführenVerbal aufweisen

Abbildung 5.20: Dendrogramme zu den Prozessen des Clusterings in Schritt 2 zur ersten Einheit der Instruktion für Personen (Pearson-Korrelation für Verlinkung innerhalb der Cluster).

(oder große – aufgrund der bereits in Schritt 1 erfolgten Partitionierung allerdings hier nicht zu beachten) Cluster zu bilden.

Für die Personen, die bei der ersten Einheit der explizit-fachmethodischen Instruktionsvariante wenig Zeit mit *VorbereitenDurchführenNonverbal* verbringen (Partition P2), fanden sich unter Berücksichtigung inhaltlicher Erwägungen die Cluster der Personen A1, A2, G20, J28 und L36 (Cluster E1-P-P2-C1), der Personen B4, D10, H22, K32, L35, M39, G19 und J30 (Cluster E1-P-P2-C2) sowie der Personen F18, H24, K33 und M37 (Cluster E1-P-P2-C3). Die inhaltlichen Erwägungen umfassten beispielsweise, dass bei die Person B6 *nicht* dem Cluster E1-P-P2-C3 zugeordnet wird, obwohl es vom Dendrogramm her ein Kandidat dafür wäre, *weil* der Anteil von *Fachinhaltlich* höher ist als der Anteil von *Fachmethodisch-Indiziert*, während es für alle Personen im Cluster E1-P-P2-C3 andersherum ist.

Bei den Aktivitätsprofiltypen aus der Partition P1 (mit großen zeitlichen Anteilen der Kategorie *VorbereitenDurchführenNonverbal*) wurde ähnlich vorgegangen. Anzu-merken ist Folgendes: Dass zwei Personen, die nicht zu Ende gearbeitet haben (I29 und K31) dem Cluster E1-P-P1-C3 zugeordnet wurden, scheint nicht vorrangig damit zusammenzuhängen, dass sie die letzten Karten nicht mehr bearbeitet haben. Eine Betrachtung der Aktivitäten für alle einzelnen Karten sowie Sichtungen der Videos zeigen vielmehr, dass die beiden Personen und I27 (ebenfalls im Cluster E1-P-P1-C3) tatsächlich den gesamten Prozess über immer wieder auf das Durchführungen von Versuchshandlungen fokussiert sind und sich verbal wenig betätigen. Ferner kann das Cluster E1-P-P1-C4 nicht im gleichen Sinne wie die anderen Cluster als Aktivitätsprofiltyp gedeutet werden, weil *nur eine* Person dem Cluster zugehörig ist. Zuletzt kann angemerkt werden, dass alternativ zur Beschreibung, dass der Anteil von *VorbereitenDurchführenNonverbal* für das Cluster E1-P-P1-C1 dem Anteil von *Fachinhaltlich* und für das Cluster E1-P-P1-C2 dem von *Fachmethodisch-Indiziert* entspricht, auch die Beschreibung adäquat ist, dass der Anteil von *VorbereitenDurchführenNonverbal* für E1-P-P1-C1 geringer ist als für das Cluster E1-P-P1-C2. Wenn also der genaue Anteil nonverbaler Aktivitäten zum Vorbereiten und Durchführen als eher irrelevant angesehen würde, hätten die beiden Cluster zusammengeführt werden können. Dies scheint allerdings inhaltlich nicht plausibel zu sein, weil diese spezifischen nonverbale Aktivitäten als zentral für das experimentbezogene Denken und Arbeiten erachtet werden. Die Trennung unterstützend zeigt sich, dass Personen im Cluster E1-P-P1-C2 fast alle niedrige Kompetenzzuwächse aufweisen, während die drei Personen im Cluster E1-P-P1-C1 alle hohe Kompetenzzuwächse haben. Dieser Befund wird in Unterabschnitt 5.4.3 genauer beschrieben und gedeutet.

5.4.2.3 Aktivitätsprofile für Teams bei Einheit 1: Ergebnisse und exemplarische ausführliche Darstellung des Vorgehens bei Teams

Auch eine Clusteranalyse zu den Aktivitätsprofilen von Teams wird ausführlicher als für weitere Einheiten dargestellt, weil aufgrund der geringen Anzahl zu vergleichender Aktivitätsprofile leicht von den Personenvergleichen abweichende Aspekte zu beachten sind. Für Einheit 1 liegen für 11 Teams Aktivitätsprofile vor (die Teams, die gemittelt mittlere Kompetenzzuwächse haben, eingeschlossen). Das mittlere Aktivitätsprofil über alle Teams ist in Abbildung 5.21 dargestellt.

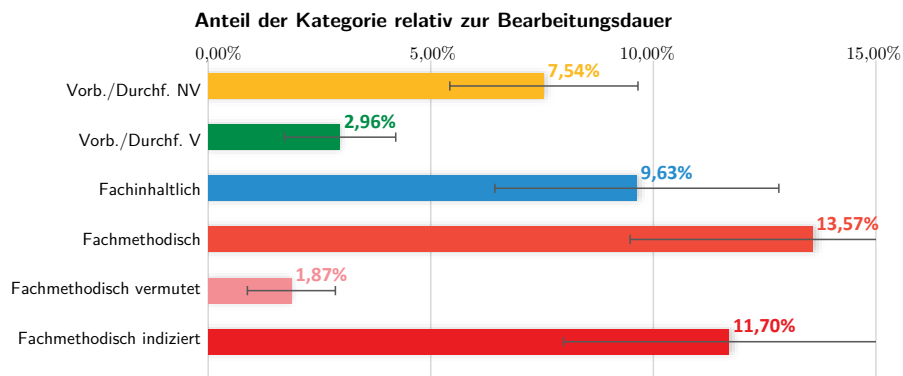


Abbildung 5.21: Mittlere Anteile der betrachteten Aktivitäten für 11 Teams bei der ersten Einheit der explizit-fachmethodischen Instruktionsvariante. Zusätzlich sind jeweils eine Standardabweichung des Mittelwerts nach oben und nach unten eingetragen.

Die Aktivitätsprofile aller Teams sind in Abbildung 5.22 dargestellt und bereits nach Aktivitätsprofiltypen sortiert. Es ergeben sich statistisch gesehen keine Partitionen, allerdings lassen sich die Aktivitätsprofiltypen inhaltlich in zwei Gruppen beschreiben:

- (1) Aktivitätsprofiltypen, bei denen der Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen vergleichsweise groß ist:

E1-T-C1 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *deutlich größer* als der zeitliche Anteil fachinhaltlicher Beiträge und nonverbaler Aktivitäten des Vorbereitens und Durchführens. Fachinhaltliche Beiträge nehmen einen kleineren zeitlichen Anteil ein als nonverbale experimentelle Aktivitäten.

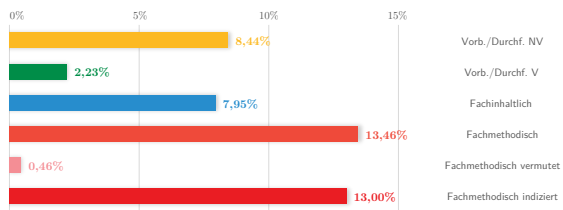
(Kurznotation: FMi $\gg$ VDn > FI.)

Teams: D, M.

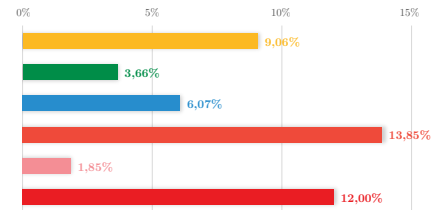
E1-T-C2 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *größer* als der zeitliche Anteil fachinhaltlicher Beiträge; beide

E1-T-C1

Anteil der Kategorie relativ zur Bearbeitungsdauer

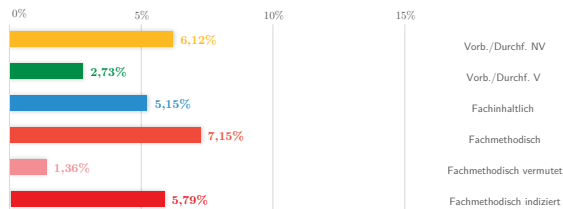


Anteil der Kategorie relativ zur Bearbeitungsdauer

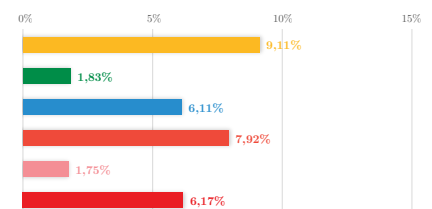


E1-T-C2

Anteil der Kategorie relativ zur Bearbeitungsdauer

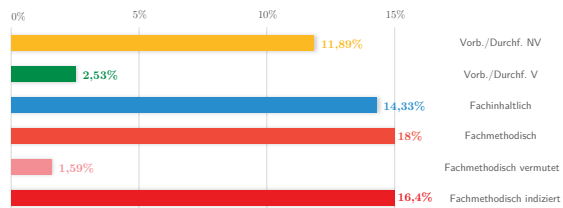


Anteil der Kategorie relativ zur Bearbeitungsdauer

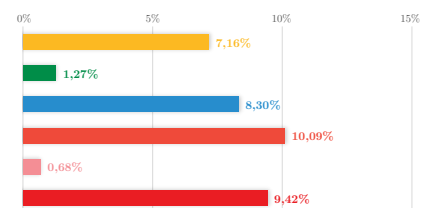


E1-T-C3

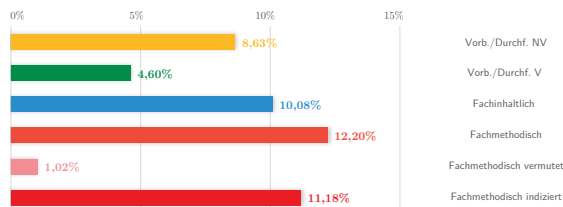
Anteil der Kategorie relativ zur Bearbeitungsdauer



Anteil der Kategorie relativ zur Bearbeitungsdauer

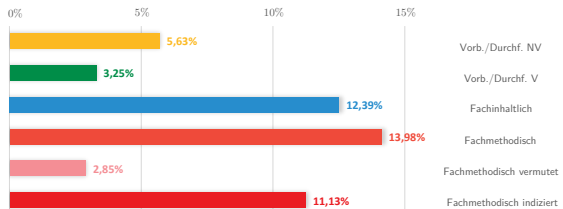


Anteil der Kategorie relativ zur Bearbeitungsdauer

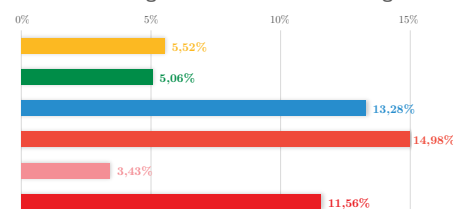


E1-T-C4

Anteil der Kategorie relativ zur Bearbeitungsdauer

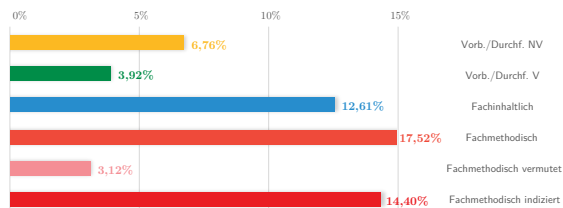


Anteil der Kategorie relativ zur Bearbeitungsdauer



E1-T-C5

Anteil der Kategorie relativ zur Bearbeitungsdauer



Anteil der Kategorie relativ zur Bearbeitungsdauer

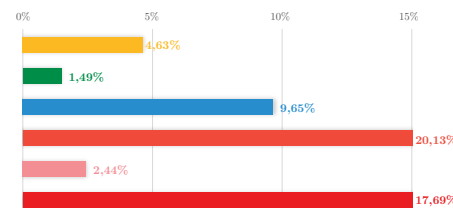


Abbildung 5.22: Aktivitätsprofile aller Teams sortiert nach Clustern (Einheit 1).

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

Anteile sind kleiner als der zeitliche Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen.

(Kurznotation: $VDn > FMi > FI$.)

Teams: G, J.

E1-T-C3 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist nur *wenig größer* als der zeitliche Anteil fachinhaltlicher Beiträge. Beide Anteile sind nur *wenig größer* als der zeitliche Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen.

(Kurznotation: $FMi \geq FI \geq VDn$.)

Teams: B, F, I.

- (2) Aktivitätsprofiltypen, bei denen der Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen vergleichsweise gering ist:

E1-T-C4 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist ein *wenig kleiner* als der zeitliche Anteil fachinhaltlicher Beiträge.

(Kurznotation: $FI \geq FMi$.)

Teams: A, L.

E1-T-C5 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *deutlich größer* als der zeitliche Anteil fachinhaltlicher Beiträge.

(Kurznotation: $FMi > FI$.)

Teams: K, H.

Aus Abbildung 5.22 geht über diese Beschreibungen hinaus hervor, dass für die Aktivitätsprofiltypen E1-T-C4 und E1-T-C5 die zeitlichen Anteile fachmethodisch-indizierter und fachinhaltlicher Beiträge deutlich größer sind als die zeitlichen Anteile aller anderen betrachteten Aktivitäten. Außerdem nimmt die Summe fachmethodisch-indizierter und -vermuteter Beiträge in beiden Typen einen größeren zeitlichen Anteil ein als fachinhaltliche Beiträge (in dieser Hinsicht weicht E1-T-C4 von allen anderen dargestellten Typen ab, weil sich die Summe beider fachmethodischen Kategorien anders verhält als fachmethodisch-indizierte Beiträge für sich genommen).

Diese Ergebnisse wurden entlang der Schritte generiert, die in Unterunterabschnitt 5.4.2.1 beschrieben sind. Nachfolgend sind bezogen auf die Team-Vergleiche einige Konkretisierungen des Vorgehens aufgeführt:

Schritt 1: Ausreißer identifizieren und Partionierungen erkennen. Aufgrund der geringeren Anzahl an Aktivitätsprofilen, die geclustert werden sollten, war das Identifizieren von Ausreißern schwieriger als bei den Analysen zu den Personen. Die Daten wurden trotzdem im ersten Schritt mithilfe des Algorithmus der Analyse nächstgelegener Nachbarn untersucht. Für die hier betrachteten Teams ergibt sich

5 Aktivitäten von Lernenden

das Dendrogramm in Abbildung 5.23. Die beiden Teams G und J wurden zunächst als Ausreißer herausgenommen und nicht im zweiten Schritt betrachtet, weil es nur zwei Teams sind und sie im Dendrogramm erst ganz am Ende mit den anderen Clustern zusammengeführt werden. Dieses Vorgehen verhinderte, dass der in Schritt 2 genutzte Algorithmus Unstimmigkeiten erzeugt.

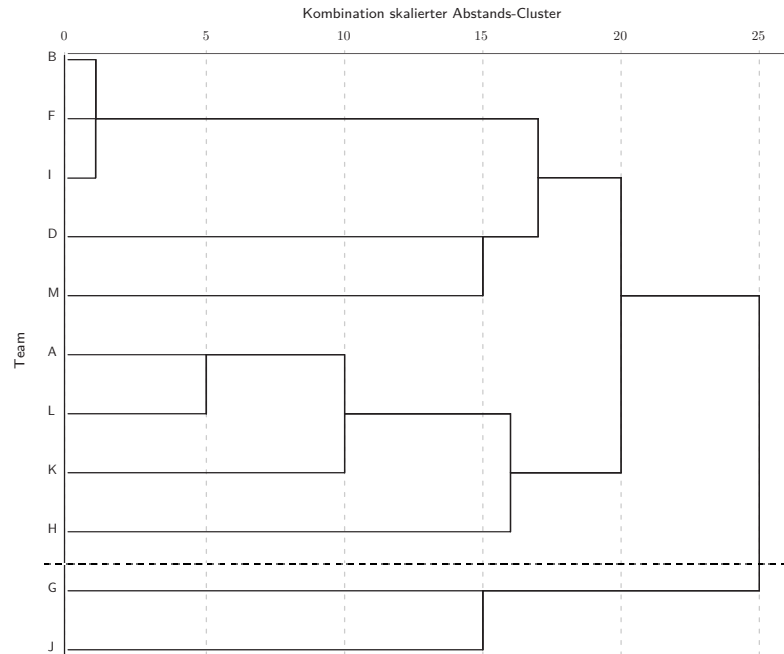


Abbildung 5.23: Dendrogramm zum Prozess des Clusterings der Teams, die Einheit 1 bearbeitet haben (Schritt 1: Pearson-Korrelation für nächstgelegene Nachbarn).

Schritt 2: Ähnlichkeiten der Aktivitätsprofile messen. Der zweite Schritt diente erneut dazu, die Ähnlichkeiten der Aktivitätsprofile darzustellen. In diesem Fall findet sich das Dendrogramm Abbildung 5.24. Die Teams G und J wurden als Ausreißer nicht bei der Anwendung des Algorithmus der Verlinkung innerhalb der Cluster berücksichtigt.

Schritt 3: Entscheidung für Cluster und Beschreibung der Aktivitätsprofil-typen. In Abbildung 5.24 sind bereits Trennlinien eingezeichnet. Die Teams D und M werden zeitgleich zusammengeführt wie A und L mit K, weshalb aus statistischer Sicht beide Cluster gebildet werden sollten, weil erst danach längere horizontale Linien (großer Heterogenitätszuwachs bei der darauffolgenden Zusammenführung von Clustern) folgen. Eine Trennlinie zwischen den Teams L und K ist dennoch

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

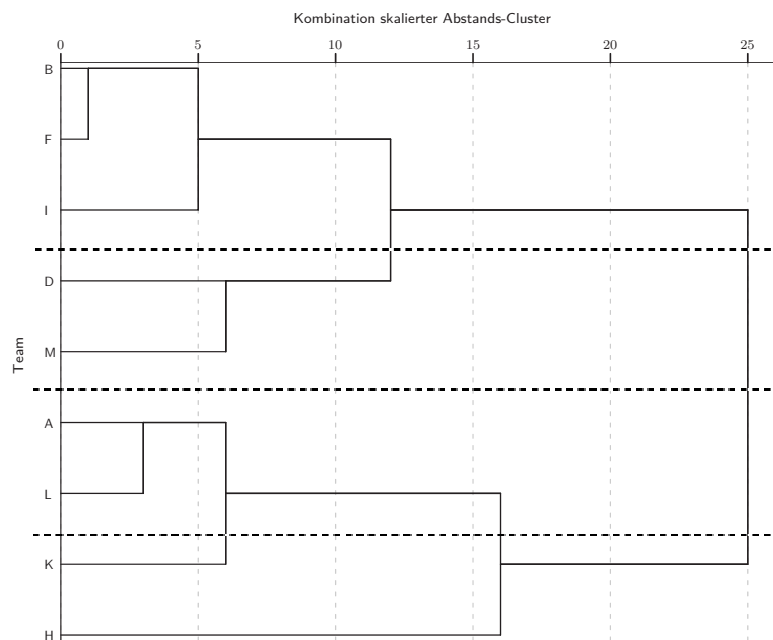


Abbildung 5.24: Dendrogramm zum Prozess des Clusterings der Teams, die Einheit 1 bearbeitet haben, außer Ausreißer (Schritt 2: Pearson-Korrelation für Verlinkung innerhalb der Cluster).

ingezeichnet, weil sie sich aus inhaltlichen Überlegungen ergab: Der Anteil fachinhaltlicher Beiträge ist im einen Fall größer, im anderen Fall kleiner als der Anteil fachmethodisch-indizierter Beiträge. Die weiteren inhaltlichen Überlegungen gehen aus den Beschreibungen der Aktivitätsprofiltypen hervor, die oben aufgeführt sind. Beispielsweise wird deutlich, dass die als Ausreißer in Schritt 1 aussortierten Teams einen eigenen Aktivitätsprofiltypen E1-T-C2 bilden, weil ihre Gemeinsamkeiten untereinander (u. a. $VDn > FMi$ anders als bei allen anderen Profilen) größer sind als die Gemeinsamkeiten mit anderen Typen (bspw. weisen vier von fünf Typen die Gemeinsamkeit auf, dass der Anteil von Fachmethodisch-Indiziert größer ist als der von Fachinhaltlich).

5.4.2.4 Aktivitätsprofiltypen für Personen bei Einheit 2

Die Betrachtung der mittleren Anteile aller betrachteten Aktivitäten für die zweite Einheit der explizit-fachmethodischen Instruktionsvariante liefert Abbildung 5.25.

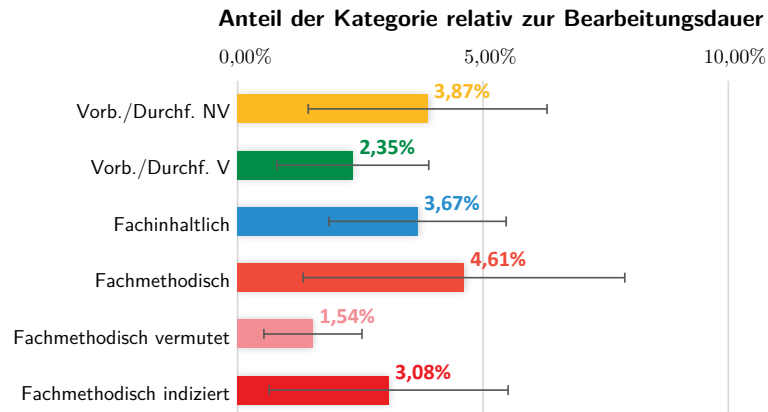


Abbildung 5.25: Mittlere Anteile der betrachteten Aktivitäten von 19 Personen bei der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante. Zusätzlich sind jeweils eine Standardabweichung des Mittelwerts nach oben und nach unten eingetragen.

Für die 19 Personen, deren Bearbeitungsprozesse für Einheit 2 kodiert wurden, ergeben sich 6 Aktivitätsprofiltypen in zwei Partitionen (siehe Abbildung 5.26). In der ersten Partition ist der Anteil der Aktivität des nonverbalen Vorbereitens und Durchführens von Versuchen (VorbereitenDurchführenNonverbal) in etwa von der Größenordnung der anderen Aktivitäten; dazu gehört ein einziger Aktivitätsprofiltyp:

E2-P-P1-C1 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist in etwa *gleich groß* wie der zeitliche Anteil fachinhaltlicher Beiträge. Der zeitliche Anteil fachmethodisch-vermuteter Beiträge ist in etwa *gleich groß* wie der zeitliche Anteil verbaler Beiträge zum Vorbereiten und Durchführen von Versuchen (und halb so groß wie der zeitliche Anteil fachmethodisch-indizierter Beiträge).⁵³ (Kurznotation: $FMi \approx FI \approx 2 \cdot FMv \approx 2 \cdot VDv$.)

Personen: B5, D11, D12, E14, F16 und F18.⁵⁴

⁵³Die große Ausführlichkeit der Beschreibung des Aktivitätsprofiltyps E2-P-P1-C1 dient der nachvollziehbaren Abgrenzung vom Aktivitätsprofiltyp E1-P-P2-C4, bei dem der Anteil fachmethodisch-vermuteter Beiträge anders als hier der letzte in der Reihenfolge und nur sehr gering ist.

⁵⁴Methodisch lässt sich anmerken, dass sich die Personen B5, D11, D12, E14, F16 und F18 aufgrund von Kettenbildung im Dendrogramm nicht in weitere Cluster aufteilen, sondern alle dem Cluster E2-P-P1-C1 zugehören. In diesem Fall ist die Kettenbildung ein angemessenes Phänomen, weil sie inhaltlich nachvollziehbar ist und mit den inhaltlichen Überlegungen übereinstimmt. In jedem Step

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

In der zweiten Partition ist der Anteil von **VorbereitenDurchführenNonverbal** (nicht immer deutlich, aber immer) größer als der Anteil aller anderen Kategorien. Es ergeben sich die Aktivitätsprofiltypen (siehe Abbildung 5.26):⁵⁵

E2-P-P2-C1 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *deutlich kleiner* als der zeitliche Anteil fachinhaltlicher Beiträge. Der zeitliche Anteil verbaler Beiträge zum Vorbereiten und Durchführen von Versuchen ist *kleiner* als der zeitliche Anteil fachmethodisch-indizierter Beiträge. (Kurznotation: $FI > FMi > VDv$.)

Personen: B4, B6, E15, F17.

E2-P-P2-C2 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *deutlich kleiner* als der zeitliche Anteil fachinhaltlicher Beiträge. Der zeitliche Anteil verbaler Beiträge zum Vorbereiten und Durchführen von Versuchen ist *in etwa so groß* wie der zeitliche Anteil fachmethodisch-indizierter Beiträge. (Kurznotation: $FI \approx VDv > FMi$.)

Personen: A2, E13.

E2-P-P2-C3 Der zeitliche Anteil der fachmethodisch-vermuteter Beiträge ist *kleiner* als der zeitliche Anteil fachinhaltlicher Beiträge und *etwas größer* als der zeitliche Anteil fachmethodisch-indizierter Beiträge. Der zeitliche Anteil verbaler Beiträge zum Vorbereiten und Durchführen von Versuchen ist *etwas größer* als der zeitliche Anteil fachinhaltlicher Beiträge. (Kurznotation: $VDv \geq FI > FMv \geq FMi$.)

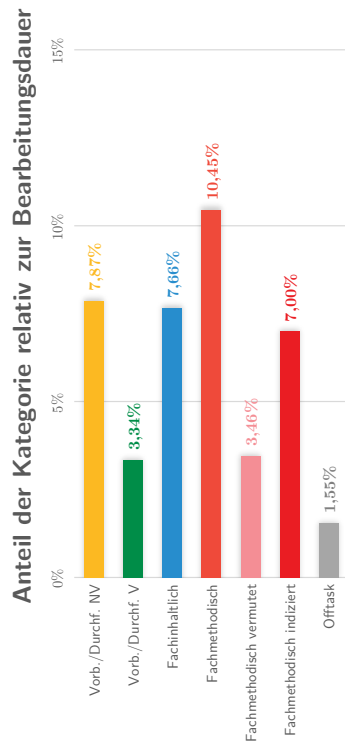
Personen: A1, G21.

Außerhalb der Partitionen finden sich die beiden Aktivitätsprofiltypen:

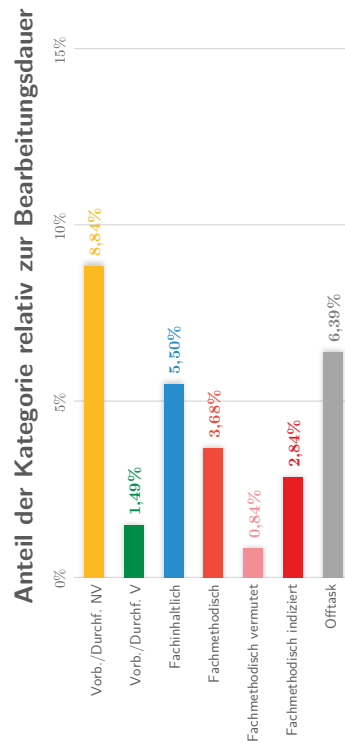
E2-P-Px-C1 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist in etwa *gleich groß* wie der zeitliche Anteil fachinhaltlicher Beiträge. Der

des Algorithmus wird der Schwellwert für die Heterogenität erhöht, was sich darin widerspiegelt, dass die in der inhaltlichen Beschreibung gegebene Faustformel zunehmend etwas (aber eben nicht deutlich) schlechter passt.

⁵⁵Da die Analysen aus den Schritten 2 und 3 für die zweite Partition Ergebnisse liefern, die sich inhaltlich nur schwer nachvollziehen lassen, werden die beiden Schritte für alle Lernenden, die nicht bereits E2-P-P1-C1 angehören unter Ausschluss der Variablen **VorbereitenDurchführenNonverbal** durchgeführt. Der Nachteil, der für den Vergleich mit den anderen Analysen durch das Ausklammern nonverbaler Beiträge zum Vorbereiten und Durchführen entsteht, wird gewissermaßen durch die Partitionierung beschränkt, da diese vorrangig gemäß dieser Aktivität zu deuten ist. Die nachfolgend gegebenen Beschreibungen der als Aktivitätsprofiltypen gedeuteten Cluster passen der Einschätzung des Autors nach immer noch schlechter als die für andere Einheiten bzw. für Teams gegebenen Beschreibungen. Vermutlich hängen die methodischen Schwierigkeiten mit der geringen Anzahl an Personen zusammen, die für Einheit 2 der explizit-fachmethodischen Instruktionsvariante analysiert wurden.



(a) E2-P-P1-C1 ($FMi \approx FI \approx 2 \cdot FMv \approx 2 \cdot VDv$): B5



(b) E2-P-P2-C1 ($FI > FMi > VDv$): B6

Abbildung 5.26: Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 2 der explizit-fachmethodischen Instruktionsvariante (Teil 1 der Abbildung).

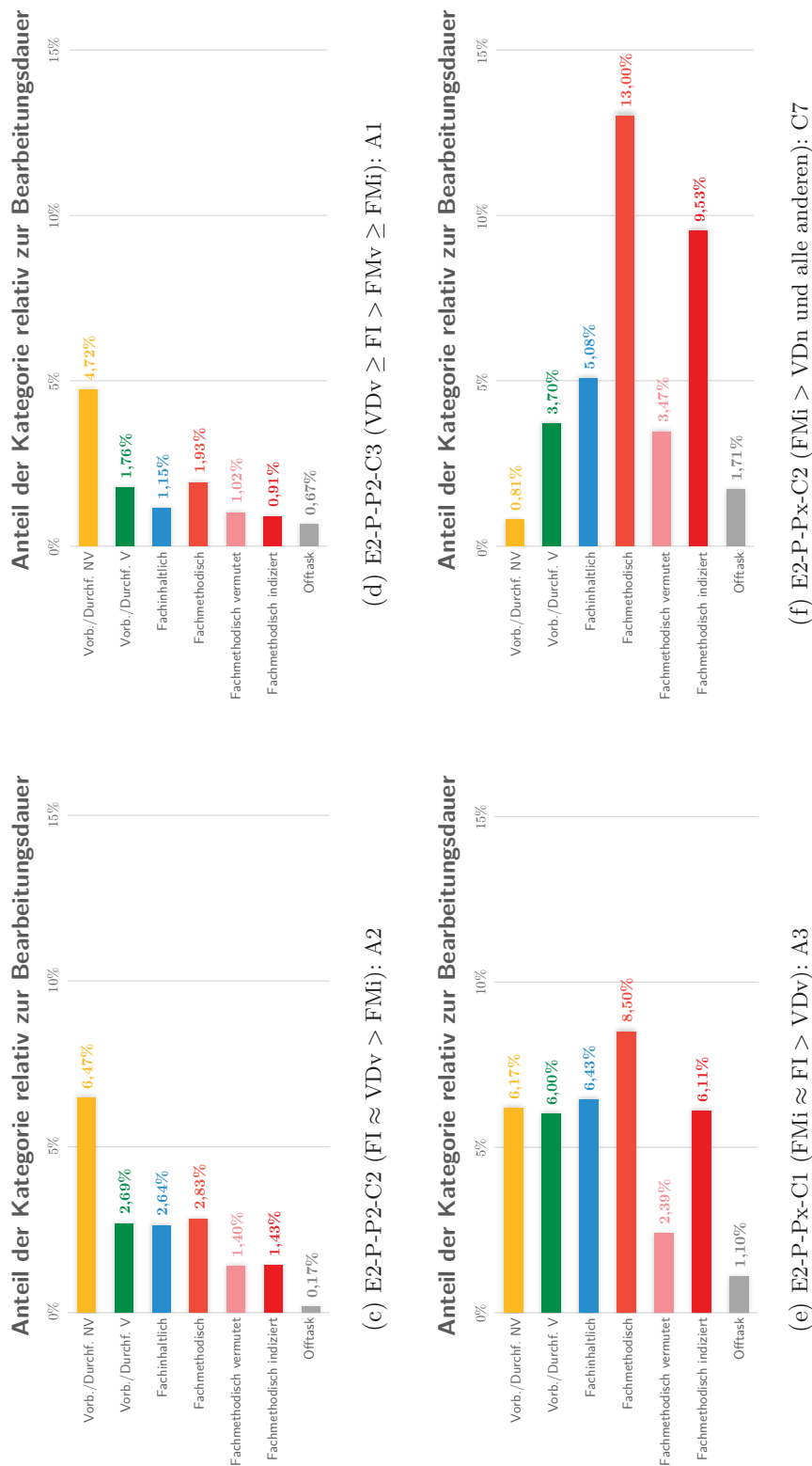


Abbildung 5.26: Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 2 der explizit-fachmethodischen Instruktionsvariante (Teil 2 der Abbildung).

zeitliche Anteil verbaler Beiträge zum Vorbereiten und Durchführen von Versuchen ist kleiner als der zeitliche Anteil dieser beiden Kategorien (und der zeitliche Anteil fachmethodisch-vermuteter Beiträge ist noch kleiner⁵⁶). (Kurznotation: $FMi \approx FI > VDv$.)

Personen: A3, G19, G20.

E2-P-Px-C2 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *größer* als der zeitliche Anteil aller anderen Aktivitäten (inklusive der nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen). (Kurznotation: $FMi > VDn$ und alle anderen.)

Personen: C7, C9.

Aufgrund der geringeren Personenzahl im Vergleich zu den beiden anderen Einheiten (und vermutlich auch der Vielfalt der Bearbeitungsprozesse) finden sich für einige der Aktivitätsprofiltypen nur zwei Personen. Dieser Einschränkung könnte durch die Bildung größerer Cluster begegnet werden – worauf allerdings verzichtet wurde, weil keine inhaltlichen Übereinstimmungen ausgemacht werden konnten. Besonders problematisch ist der Fall des Aktivitätsprofiltyps E2-P-Px-C2, weil ihm nur zwei Personen zugeordnet sind, die beide demselben Team entstammen. Vergleichbares gilt auch für E2-P-Px-C1. Einige Aktivitätsprofiltypen sind daher und aufgrund der kleinen Zahl zugehöriger Personen in folgenden Analysen mit besonderer Vorsicht zu diskutieren.

⁵⁶Der Klammereinschub zu E2-P-Px-C1 verdeutlicht den Unterschied dieses Typens zum Typ E2-P-P1-C1: Nicht nur der Anteil von *VorbereitenDurchführenNonverbal*, sondern auch der Anteil von *Fachmethodisch-Vermutet* ist unterschiedlich.

5.4.2.5 Aktivitätsprofiltypen für Teams bei Einheit 2

Das mittlere Aktivitätsprofil über alle Teams ist in Abbildung 5.27 dargestellt.

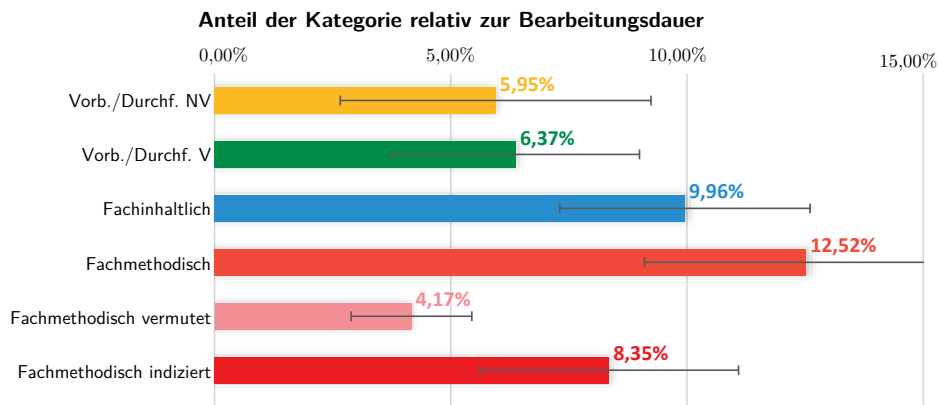


Abbildung 5.27: Mittlere Anteile der betrachteten Aktivitäten für 7 Teams bei der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante. Zusätzlich sind jeweils eine Standardabweichung des Mittelwerts nach oben und nach unten eingetragen.

Die 7 für Einheit 2 analysierten Teams lassen sich drei Aktivitätsprofiltypen zuordnen. Die Aktivitätsprofile der Repräsentanten finden sich in Abbildung 5.28. Für alle Teams bis auf das Team G ist der Anteil der Aktivität **Fachmethodisch-Indiziert** etwa doppelt so groß wie der von **Fachmethodisch-Vermutet**. Angesichts der großen Vielfalt der Aktivitätsprofile für die einzelnen Personen bei dieser Einheit sind diese Befunde bedeutsam und zeigen, dass grundsätzliche Ähnlichkeiten in den Bearbeitungen bestehen. Die Beschreibungen der Aktivitätsprofiltypen heben die Unterschiede hervor, die zwischen den Teams bestehen:

E2-T-C1 Der zeitliche Anteil der Aktivität **Fachinhaltlich** ist *größer* als der aller anderen Aktivitäten (einzeln betrachtet). Die am nächsthäufigsten auftretende Aktivität ist **Fachmethodisch-Indiziert**.⁵⁷

(Kurznotation: FI > FMi > alle anderen.)

Teams: B, E, F.

E2-T-C2 Der zeitliche Anteil der Aktivität **VorbereitenDurchführenVerbal** nimmt einen *größeren* Anteil ein als alle anderen Aktivitäten (einzeln betrachtet).⁵⁸

(Kurznotation: VDv > FMi > alle anderen.)

Teams: A, G.

⁵⁷Für die Teams B und E ist die Summe der indizierten und der vermuteten fachmethodischen Beiträge größer als der Anteil der fachinhaltlichen Beiträge.

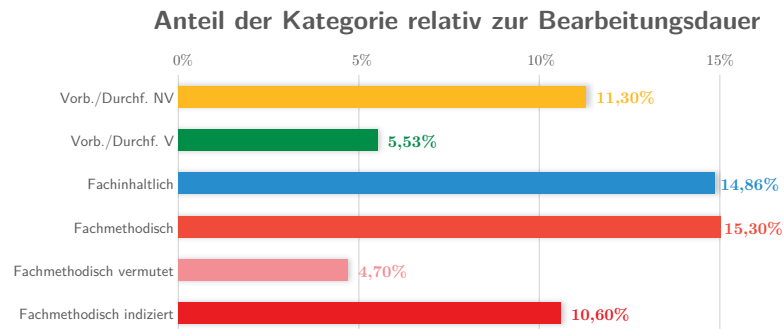
⁵⁸Für alle Teams dieses Aktivitätsprofiltyps ist die Summe der indizierten und der vermuteten fachmethodischen Beiträge größer als der Anteil von **VorbereitenDurchführenNonverbal**.

5 Aktivitäten von Lernenden

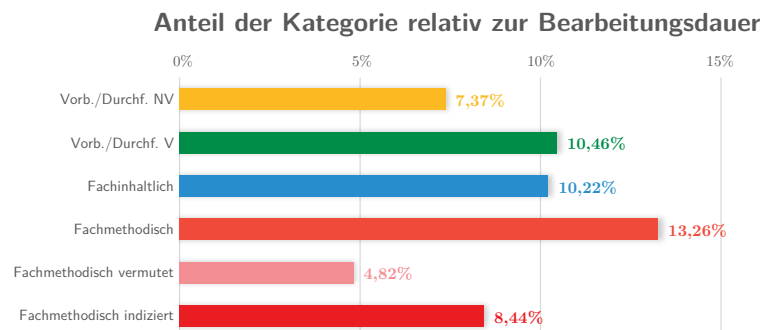
E2-T-C3 Der zeitliche Anteil der fachmethodisch-indizierter Beiträge ist *größer* als der zeitliche Anteil fachinhaltlicher Beiträge. Der zeitliche Anteil nonverbaler Beiträge zum Vorbereiten und Durchführen von Versuchen ist im Vergleich zu den anderen Aktivitätsprofiltypen *gering* und in der Größenordnung der anderen Aktivitäten.

(Kurznotation: FMi > FI; VDn $\approx$ alle.)

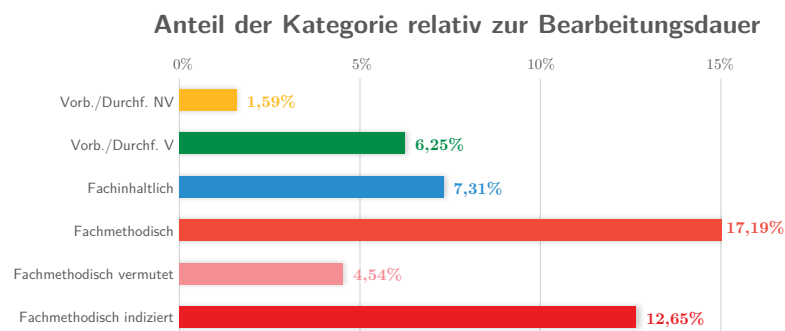
Teams: C, D.



(a) E2-T-C1 (FI > FMi > alle anderen): B



(b) E2-T-C2 (VDv > FMi > alle anderen): A



(c) E1-T-C3 (FMi > FI; VDn $\approx$ alle): C

Abbildung 5.28: Repräsentanten für die Aktivitätsprofiltypen für Teams bei Einheit 2 der explizit-fachmethodischen Instruktionsvariante.

5.4.2.6 Aktivitätsprofiltypen für Personen bei Einheit 3

Die Betrachtung der mittleren Anteile aller betrachteten Aktivitäten für die dritte Einheit der explizit-fachmethodischen Instruktionsvariante liefert Abbildung 5.29.

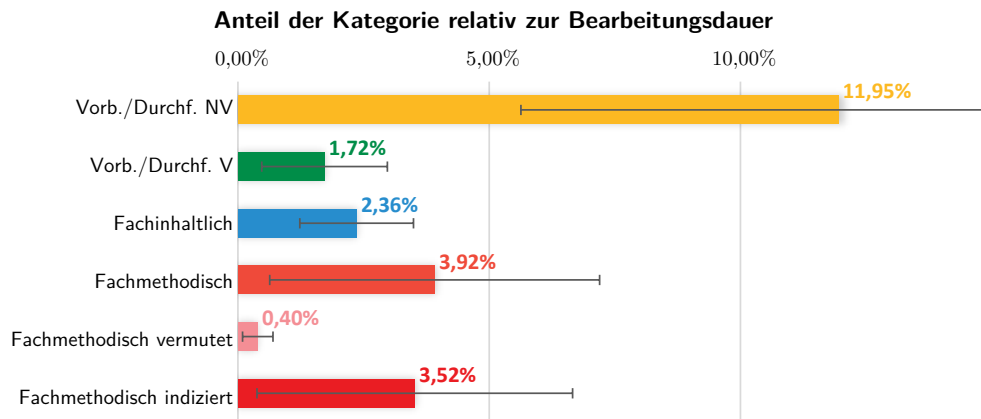


Abbildung 5.29: Mittlere Anteile der betrachteten Aktivitäten von 22 Personen bei der dritten Einheit der explizit-fachmethodischen Instruktionsvariante. Zusätzlich sind jeweils eine Standardabweichung des Mittelwerts nach oben und nach unten eingetragen.

Die 22 Personen, deren Bearbeitung von Einheit 3 kodiert wurde, lassen sich fünf Aktivitätsprofiltypen zuordnen. Repräsentanten für jeden Aktivitätsprofiltypen finden sich in Abbildung 5.30. In allen Aktivitätsprofilen ist der Anteil von **Fachmethodisch-Vermutet** deutlich kleiner als in den Aktivitätsprofilen von Personen bei Einheit 2. In den meisten Aktivitätsprofilen ist der Anteil von **VorbereitenDurchführenNonverbal** größer als der aller anderen Aktivitäten. Nur für die als *Ausreißer* identifizierten Personen J30 und M39 verhält es sich anders. Für J30 ist der Anteil von **Fachmethodisch-Indiziert** am größten, während für M39 der Anteil von **Fachinhaltlich** am größten ist.⁵⁹ Die Aktivitätsprofile der anderen Personen lassen sich den folgenden fünf Aktivitätsprofiltypen zuordnen (zur Darstellung inhaltlich gruppiert):

- (1) Aktivitätsprofiltypen, bei denen nur der Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen sehr viel größer ist als die (einzelnen) Anteile aller anderen Aktivitäten (zwei- bis viermal so groß):

E3-P-C1 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *kleiner* als der zeitliche Anteil fachinhaltlicher Beiträge. Der zeitli-

⁵⁹Ein weiterer Ausreißer ist G20; diese fällt allerdings nicht mit Blick auf **VorbereitenDurchführenNonverbal** heraus, sondern auf die anderen Aktivitäten.

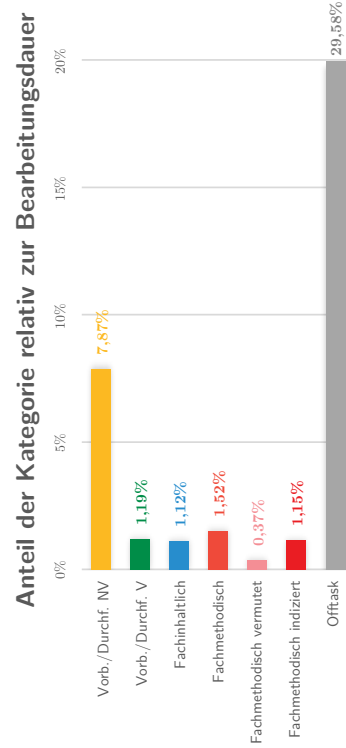
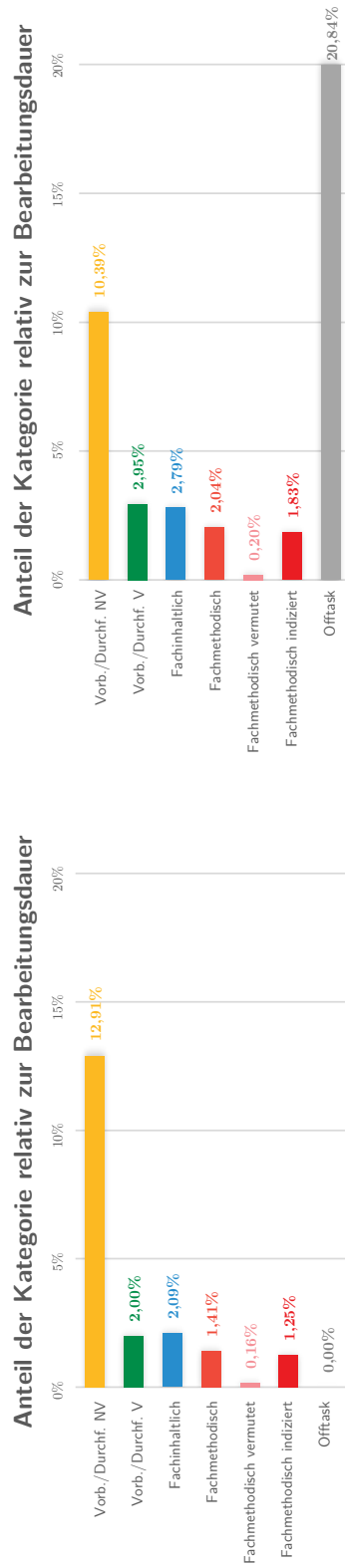
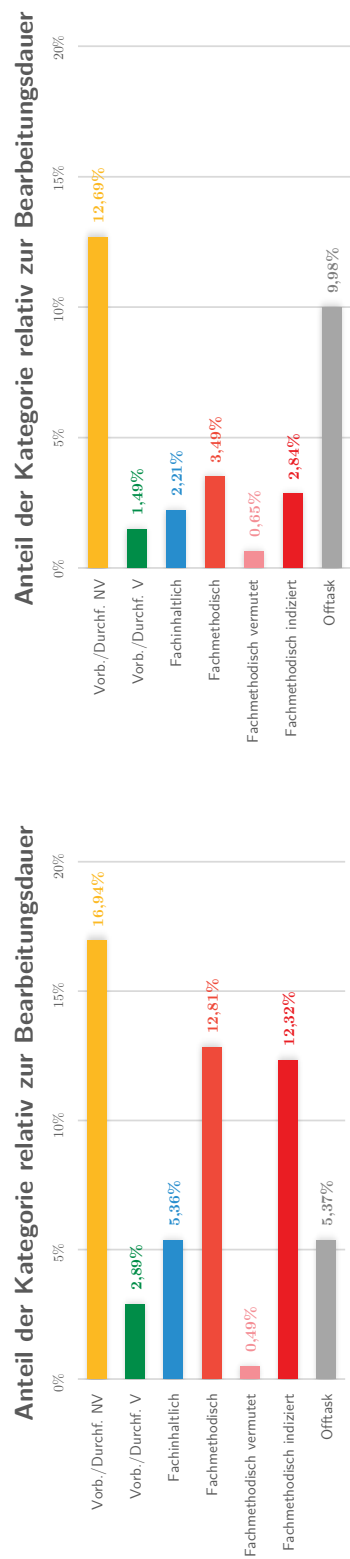


Abbildung 5.30: Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 3 der explizit-fachmethodischen Instruktionsvariante (Teil 1 der Abbildung: $VDn \gg$ alle anderen).



(d) E3-P-C4 (VDn ≥ FMi > FI): B5

(e) E3-P-C5 (VDn ≥ FMi > FI; FMv ≥ 0): B4

Abbildung 5.30: Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 3 der explizit-fachmethodischen Instruktionsvariante (Teil 2 der Abbildung: $VDn \geq FMi$ > alle anderen).

che Anteil fachmethodisch-vermuteter Beiträge ist verschwindend.
(Kurznotation: $FI > FMi$; $FMv \approx 0$.)

Personen: G19, G21, H24, J28, M37

E3-P-C2 Wie E3-P-C1, aber der Anteil verbaler Beiträge zum Vorbereiten und Durchführen von Versuchen ist *etwas größer* als der Anteil fachinhaltlicher Beiträge (statt wie in E3-P-C1 kleiner).

(Kurznotation: $VDv \geq FI > FMi$; $FMv \approx 0$.)

Personen: D10, F16.

E3-P-C3 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *größer* als der zeitliche Anteil fachinhaltlicher Beiträge. Zudem ist der zeitliche Anteil fachmethodisch-vermuteter Beiträge deutlich größer als in allen anderen Aktivitätsprofiltypen für Personen zu Einheit 3.⁶⁰

(Kurznotation: $FMi > FI$; $FMv \gg 0$.)

Personen: D12, F18.

- (2) Aktivitätsprofiltypen, bei denen der Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen und (nicht ganz so stark ausgeprägt) auch der Anteil fachmethodisch-indizierter Beiträge sehr viel größer ist als die (einzelnen) Anteile aller anderen Aktivitäten:

E2-P-C4 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *nur wenig kleiner* als der zeitliche Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen. Der nächstgrößere Anteil entfällt auf die Aktivität fachinhaltlicher Beiträge.

(Kurznotation: $VDn \geq FMi > FI$.)

Personen: B5, H23, J29, L34, L35, L36, M38.

E2-P-C5 Der zeitliche Anteil der fachmethodisch-indizierten Beiträge ist *viel kleiner* als der zeitliche Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen. Der nächstgrößere Anteil entfällt auf die Aktivität fachinhaltlicher Beiträge. Zudem ist (ähnlich wie in E2-P-C3) der zeitliche Anteil fachmethodisch-vermuteter Beiträge deutlich größer als in den anderen Aktivitätsprofiltypen für Personen zu Einheit 3.⁶¹ (Kurznotation: $VDn \gg FMi > FI$; $FMv \gg 0$.)

Personen: B4, D11, H22.

⁶⁰Der zweite Satz ist zur Beschreibung bzw. Abgrenzung des Aktivitätsprofiltyps nicht relevant, aber für spätere Deutungen.

⁶¹Der letzte Satz ist zur Beschreibung bzw. Abgrenzung des Aktivitätsprofiltyps nicht relevant, aber für spätere Deutungen. Er hebt die Parallelität zu E2-P-C3 hervor.

5.4.2.7 Aktivitätsprofiltypen für Teams bei Einheit 3

Für Einheit 3 wurden 8 Teams analysiert. Das mittlere Aktivitätsprofil über alle Teams ist in Abbildung 5.31 dargestellt.

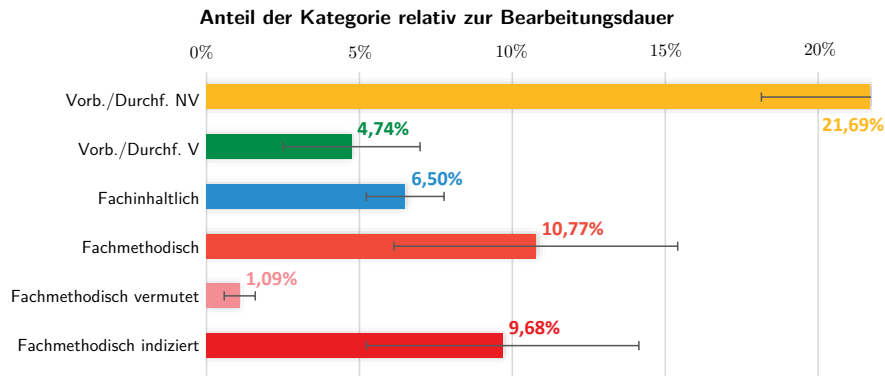


Abbildung 5.31: Mittlere Anteile der betrachteten Aktivitäten für 8 Teams bei der ersten Einheit der explizit-fachmethodischen Instruktionsvariante. Zusätzlich sind jeweils eine Standardabweichung des Mittelwerts nach oben und nach unten eingetragen.

Die sich ergebenden Cluster werden als zwei Aktivitätsprofiltypen interpretiert (siehe Abbildung 5.32), wobei die Aktivitätsprofile von zwei Teams (F und G) als Ausreißer außen vor bleiben. In allen Aktivitätsprofilen für Teams überwiegt bei Einheit 3 der Anteil der Aktivität *VorbereitenDurchführenNonverbal* deutlich; er liegt in etwa im Bereich 20–25 % (der Gesamtbearbeitungsdauer von Einheit 3; für Team J etwas darunter, für Team M etwas darüber), während alle anderen Aktivitäten außer *Fachmethodisch-Indiziert* maximal 10 % der Gesamtbearbeitungsdauer von Einheit 3 ausmachen.

E3-T-C1 Der Anteil von *Fachmethodisch-Indiziert* ist nur *minimal kleiner* als der von *VorbereitenDurchführenNonverbal*. Der nächstgrößte Anteil wird von der Aktivität *Fachinhaltlich* ausgemacht, die aber maximal halb so viel Zeit in Anspruch nimmt wie *Fachmethodisch-Indiziert*.⁶²

(Kurznotation: $VD_n \geq FMI > FI$.)

Teams: B, J, L.

E3-T-C2 Der Anteil von *VorbereitenDurchführenNonverbal* ist *deutlich größer* als der aller anderen Aktivitäten (einzeln betrachtet). Der nächstgrößte

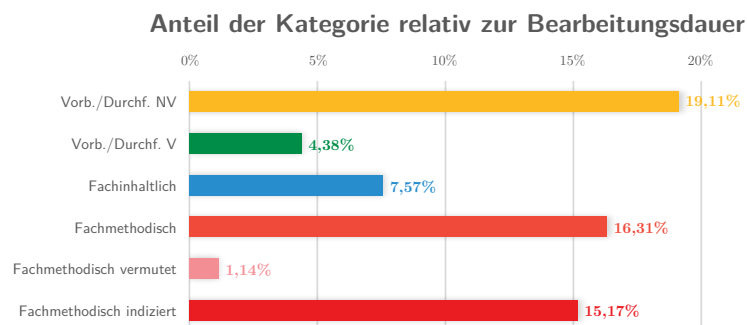
⁶²Schon an dieser Stelle fällt auf, dass die Profile E3-P-C4 (für Personen) und E3-T-C1 (für Teams) strukturell sehr ähnlich sind. Alle Personen aus dem Team L sowie jeweils eine Person aus den Teams B und J, aber auch zwei weitere Personen gehören E3-T-C1 an; was später diskutiert wird.

5 Aktivitäten von Lernenden

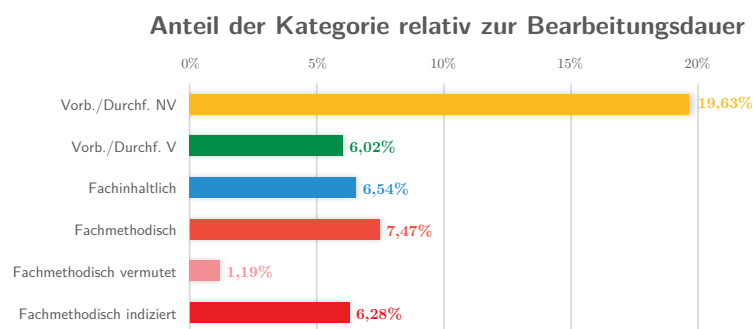
Anteil wird von der Aktivität Fachmethodisch-Indiziert ausgemacht, gefolgt von Fachinhaltlich. (Kurznotation: $VDn \gg FMi > FI$.)

Teams: D, H, M.

Das Team G stellt einen Ausreißer dar, weil der Anteil von Fachinhaltlich größer ist als der von Fachmethodisch-Indiziert (und sogar aller fachmethodischen Beiträge, d. h. auch inkl. Fachmethodisch-Vermutet). Ansonsten ähnelt es dem Aktivitätsprofiltypen E3-T-C1. Das Team F ist ebenfalls ein Ausreißer, weil hier der nach VorbereitenDurchführenNonverbal nächstgrößte Anteil von der Aktivität VorbereitenDurchführenNonverbal eingenommen wird.



(a) E3-T-C1 ($VDn \geq FMi > FI$): B



(b) E3-T-C2 ($VDn \gg FMi > FI$): D

Abbildung 5.32: Repräsentanten für die Aktivitätsprofiltypen für Teams bei Einheit 2 der explizit-fachmethodischen Instruktionsvariante.

5.4.2.8 Gesamtschau auf Aktivitätsprofile

In der Gesamtschau auf die Aktivitätsprofile stellt sich zunächst die Frage, wie vielfältig die Aktivitätsprofile innerhalb jeder der analysierten Einheiten für sich genommen sind, das heißt unter anderem, welche Unterschiede und Gemeinsamkeiten zwischen Aktivitätsprofilen innerhalb von einer einzelnen Einheit vorliegen. Für jede der drei Einheiten ergibt sich, dass die Aktivitätsprofile sich Aktivitätsprofiltypen zuordnen lassen, die ein plausibles Abbild von der Verteilung der Anteile der Aktivitäten an der Gesamtbearbeitungsdauer liefern. (Dies gilt für Einheit 3 mit leichten Einschränkungen.)

Sowohl Unterschiede und Gemeinsamkeiten bezüglich absoluter Prozentwerte als auch bezüglich relationaler Aspekte gehen aus den Analysen hervor. Beispielsweise betragen die absolute Werte für die Anteile nonverbaler Aktivitäten des Vorbereitens und Durchführens in Einheit 1 entweder ca. 5 % (Cluster E1-P-P1) oder sind sehr klein ($< 1\%$; Cluster E1-P-P2); relational betrachtet können große Anteile nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen allerdings sowohl mit kleinen Anteilen anderer Aktivitäten (Typ E1-P-P1-C3) einhergehen oder die Anteile der anderen Aktivitäten sind in einer vergleichbaren Größenordnung (die anderen drei Typen aus dem Cluster E1-P-P1).

Werden die Anteile der verbalen Aktivitäten in absteigender Reihenfolge sortiert, ergibt sich für viele der Aktivitätsprofiltypen (und jeweils für etwa zwei Drittel aller Personen) dieselbe Reihenfolge. Für Einheit 1 und Einheit 3 lautet diese Reihenfolge: $FM_i > FI > VD_v > FM_v$ (nur die 3 Personen mit dem Profiltyp E1-P1-C1 und die 6 Personen mit den Profiltypen E1-P-P1-C4 und E1-P-P2-C1 sowie die 7 Personen mit den Profiltypen E3-P-C1 und E3-P-C2 weichen davon ab). Für Einheit 2 lautet die Reihenfolge: $FI > FM_i > VD_v > FM_v$ (nur die 2 Personen mit dem Profiltyp E2-P-Px-C2 weichen davon ab). Die beiden Reihenfolgen unterscheiden sich nur darin, ob fachmethodisch-indizierte oder fachinhaltliche Beiträge einen größeren zeitlichen Anteil einnehmen (und auch die Personen, die von der für die jeweilige Einheit typischen Reihenfolge abweichen, tun dies in den allermeisten Fällen nur in dieser Hinsicht, eine Ausnahme stellt E1-P1-C1 dar, wo der Anteil fachmethodisch-vermuteter Beiträge größer ist als der von verbalen Aktivitäten der Vorbereitung und Durchführung von Versuchen). Daran werden zwei Ergebnisse deutlich:

- Die Aktivitätsprofile weisen in vielen Teilen Gemeinsamkeiten hinsichtlich der Reihenfolgen der Anteile auf, weichen allerdings hinsichtlich der Verhältnisse (d.h. auch der Abstände zwischen den Anteilen der Aktivitäten) so deutlich voneinander ab, dass sich verschiedene Aktivitätsprofiltypen bilden lassen.

- Eine zentrale Bedeutung für Gemeinsamkeiten und Unterschiede zwischen Aktivitätsprofilen (innerhalb jeweils einer Einheit und auch über die Einheiten hinweg) spielt der Anteil fachmethodisch-indizierter Beiträge (als absoluter Prozentwert und in Relation zu den Anteilen der anderen Aktivitäten).

Bezüglich der drei Einheiten stellt sich die Frage nach der (Un)Ähnlichkeit von Einheit zu Einheit. Als Gemeinsamkeit zwischen den Einheiten lässt sich die insgesamt ähnliche Reihenfolge von Anteilen der Aktivitäten nennen. Abweichungen zwischen den Einheiten liegen zwar in einem zentralen Punkt vor (siehe oben), aber eben nur in einem Punkt; ansonsten besteht große Ähnlichkeit über alle drei Einheiten hinweg. Eine weitere Gemeinsamkeit besteht darin, dass jeweils eine gute Hand voll Aktivitätsprofiltypen aufgefunden wird, die vergleichbare Ausprägungsunterschiede abbilden. Zwei Unterschiede fallen direkt auf:

- Der Anteil der nonverbalen Aktivitäten des Vorbereitens und Durchführens ist für Einheit 3 in allen Aktivitätsprofiltypen deutlich größer als in nahezu allen Aktivitätsprofilen der beiden anderen Einheiten (Ausnahme E3-P-C3 im Vergleich mit E2-P-P2-C1, weil für den Typ zu Einheit 2 ein sehr großer Anteil vorliegt). Es gibt in Einheit 3 keine Personen, die geringe zeitliche Anteile für diese Aktivität einsetzen; in den beiden anderen Einheiten finden sich allerdings sehr deutlich Personen, die kaum Zeit darauf verwenden.
- Die Reihenfolge der Anteile der Aktivitäten (siehe Aufzählungspunkte oben) weicht für Einheit 2 von der Reihenfolge in den anderen beiden Einheiten ab, weil der Anteil fachmethodisch-indizierter Beiträge kleiner ist als der Anteil fachinhaltlicher Beiträge.

Darüber hinaus liegt für Einheit 2 kein Aktivitätsprofiltyp vor, in dem Personen ihre Zeit vorrangig mit nonverbalen Aktivitäten des Vorbereitend und Durchführens von Versuchen zubringen und dabei die verbalen Aktivitäten vernachlässigen, während für Einheit 1 der Typ E1-P-P1-C3 und für Einheit 3 die Typen E3-P-C1 und -C3 existieren.

Ein genauerer Vergleich der Anteile der einzelnen Aktivitäten für jede der drei Einheiten zeigt, dass gleichwohl gewisse Aktivitäten in manchen Einheiten stärker ausgeprägt sind. Der deutlichste Unterschied findet sich für die Aktivität **Fachmethodisch-Vermutet**, die in Einheit 2 einen signifikant größeren Anteil einnimmt als in beiden anderen Einheiten (Vergleich mit E1: $p = .005$, $r = .508$; mit E3: $p < .001$, $r = .668$). Auch für **VorbereitenDurchführenVerbal** findet sich ein Unterschied, weil

diese Aktivität in Einheit 1 deutlich weniger Zeit einnimmt als in den anderen beiden Einheiten (E2: $p = .006$, $r = .521$; E3: $p = .049$, $r = .337$). Während der Unterschied für VorbereitenDurchführenVerbal sich plausibel auf die Anteile von VorbereitenDurchführenNonverbal zurückführen lässt (letztlich also auf die Zeit, die im Material für Versuchsvorbereitungen und -durchführungen eingeplant ist, was auch den Unterschieden in der Anlage der Einheiten der Instruktion entspricht), scheint dies für Fachmethodisch-Vermutet nicht der Fall zu sein. Vielmehr scheint der größere Anteil von Fachmethodisch-Vermutet ein Spezifikum der Einheit 2 darzustellen.

Zusammengenommen lässt sich somit festhalten, dass es trotz unterschiedlicher Anlage der Instruktionseinheiten (siehe Kapitel 4) große Gemeinsamkeiten zwischen den Einheiten gibt. Allerdings weicht Einheit 2 hinsichtlich mehrerer Aspekte von den anderen Einheiten ab. Es ist eine Einschränkung der Arbeit, dass für Einheit 2 weniger Lernende analysiert wurden als für die anderen Einheiten. Allerdings dürfte auch die Hinzunahme weiterer Lernender beispielsweise nichts an der sehr deutlich typischen Reihenfolge der Anteile der Aktivitäten ändern. Der eventuell weniger guten Auflösung von verschiedenen Aktivitätsprofiltypen, die durch eine geringe Stichprobengröße bei Einheit 2 bedingt ist, wurde dadurch begegnet, dass für diese Einheit auch Cluster mit zwei Personen erlaubt wurden und real vorliegen – der damit verbundene potentielle Nachteil, dass Aktivitätsprofiltypen sich nicht hinreichend gut voneinander differenzieren lassen, ist für die beschriebenen Unterschiede zwischen den Einheiten von nur geringer Relevanz.

Eine weitere potentielle Einschränkung der Arbeit besteht in der Auflösungsgrenze des eventbasierten Kodierens. Wenngleich keine feste Grenze wie beim intervallbasierten Kodieren vorliegt, ist durch die 3-Sekunden-Regel (siehe S. 75) eine ungefähre Grenze gegeben. Dass sich die Prozentwerte verschiedener Aktivitäten in einem grob ähnlichen Bereich finden, stellt allerdings keine Einschränkung hinsichtlich der für die Analysen betrachteten Reihenfolgen und Verhältnisse von Anteilen der Aktivitäten dar: die Auflösungsgrenze von drei Sekunden entspricht einem Anteil von weniger als 0.1 % (bei 45-minütiger Bearbeitungsdauer, für die Einheiten 2 und 3 also noch niedriger) und die Unterschiede der Anteile der Aktivitäten liegen im Bereich von mindestens 0.5 % (meist aber deutlich größer).

Abschließend sei noch eine Übersicht über die Aktivitätsprofiltypen der Personen gegeben, die für alle drei Einheiten analysiert wurden. In Tabelle 5.12 (auf S. 162) sind für diese 9 Personen die jeweils für die Einheiten zugewiesenen Aktivitätsprofiltypen vermerkt. Als Gemeinsamkeit lässt sich identifizieren, dass Personen, die große Anteile

5 Aktivitäten von Lernenden

nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen in einer der Einheiten aufweisen, in den allermeisten Fällen auch für die anderen Einheiten große Anteile derselben Aktivität aufweisen. Für die Stellung fachmethodisch-indizierter Beiträge, die exemplarisch betrachtet wurden, zeigt sich jedoch keine Konsistenz über die Einheiten hinweg. Diese Ergebnisse stützen unter anderem die Vermutung, dass es Personen gibt, die die Rolle des Experimentators einnehmen (siehe bereits weiter oben). Personen, die sich über die Einheiten hinweg zum (fachmethodischen) Sprecher machen, scheint es hingegen nicht zu geben; unter anderem sind daher Analysen zur spezifischen Bedeutung fachmethodischer Beiträge, die die Einheiten einzeln auflösen, nach wie vor von Relevanz. Zuletzt sei angemerkt, dass es vermutlich erst in einer größeren Stichprobe auszumachen sein dürfte, ob es Regelmäßigkeiten in den Mustern von Aktivitätsprofiltypen der Lernenden über die Einheiten hinweg gibt.

Tabelle 5.12: Aktivitätsprofile der Personen, die für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante analysiert wurden.

	Aktivitätsprofiltyp														
	Einheit 1: E1-P-...					Einheit 2: E2-P-				Einheit 3: E3-P-...					
	P1-C1	P1-C2	P2-C1	P2-C2	P2-C3	P1-C1	P2-C1	P2-C3	Px-C1	C1	C2	C3	C4	C5	Ausr
Person															
B4				X			X							X	
B5		X				X							X		
D11	X					X								X	
D12		X				X						X			
F16		X				X					X				
F18					X	X						X			
G19				X					X	X					
G20			X						X						X
G21		X						X		X					
Profiltyp- beschreibung															
VDn	sehr groß		mäßig groß			groß	mäßig groß		gering	groß			groß		klein
FMi	s. gr.	m. gr.	< FI	> FI	≈ FI	≈ FI	< FI	< FI	≈ FI	< FI	< FI	> FI	> alle ander.		

Anmerkungen. Ausr = Ausreißer (keine Systematik bzgl. FMi), der keinem Profil zugeordnet werden konnte. Die Profiltypenbeschreibungen geben nur Ausschnitte der vollständigen Beschreibungen aus den vorigen Abschnitten der Arbeit wieder. VDn = Kategorie Vorbereiten und Durchführen von Versuchen nonverbal; FMi = Kategorie Fachmethodisch-indiziert.

5.4.3 Aktivitätsprofile und Kompetenzzuwächse

Die bisher berichteten Analysen und Ergebnisse zu Aktivitätsprofilen beziehen sich auf eine Beschreibung der Video-Teilstichproben für die jeweiligen Einheiten der explizit-fachmethodischen Instruktionsvariante. Vor dem Hintergrund der Forschungsfrage F-Akt1 stellt sich anknüpfend an die Ergebnisse zur Vielfalt der Aktivitätsprofile die Frage danach, inwiefern Unterschiede zwischen Aktivitätsprofilen mit erreichten Kompetenzzuwächsen einhergehen.

Eine Betrachtung der mittleren Anteile der einzelnen Aktivitäten im Vergleich von Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen ist in Abbildung 5.33 dargestellt. Anhand der eingetragenen Standardfehler der Mittelwerte wird deutlich, dass sich aufgrund großer Streuungen signifikante Unterschiede nur für wenige Balken, vor allem bei Einheit 2, ergeben: Fachmethodisch-Indiziert für Einheiten 1 und 2, Fachmethodisch-Vermutet für Einheit 3, Fachinhaltlich für Einheit 2 sowie VorbereitenDurchführenNonverbal für Einheit 2. Für die sehr kleine Stichprobe von 8 Personen, die für alle drei Einheiten analysiert wurden, ergibt sich über den Verlauf der gesamten Instruktion hinweg nur ein signifikanter Unterschied für den zeitlichen Anteil fachmethodisch-vermuteter Beiträge.

Differenziertere Ergebnisse können Analysen zum Zusammenhang von Aktivitätsprofiltypen und Kompetenzzuwächsen liefern, welche durch den Rückgriff auf die Aktivitätsprofiltypen berücksichtigen, dass hinter den Mittelwertmustern aus Abbildung 5.33 vielfältige Muster auf Ebene der einzelnen Personen bzw. der jeweiligen Teams stehen.⁶³ Dafür wird nachfolgend für jeden Aktivitätsprofiltypen gezählt, wie viele Personen der Gruppe der Personen mit hohen, mittleren, niedrigen oder negativen Kompetenzzuwächsen (bzw. analog für Teams mit gemittelten Kompetenzzuwächsen) angehören.⁶⁴

⁶³Zudem erfolgt eine detaillierte isolierte Betrachtung der fachmethodischen Beiträge in den Abschnitten 5.4.4 — 5.4.6, ab S. 173.

⁶⁴Grundsätzlich wäre auch ein getrenntes Bilden von Aktivitätsprofiltypen für Personen (bzw. Teams) mit (gemittelt) hohen bzw. niedrigen Kompetenzzuwächsen denkbar, dem ein nachträglicher inhaltlicher Vergleich der Profile folgt. Gegen ein derartiges Vorgehen spricht im Rahmen der Arbeit allerdings, dass die Stichprobengrößen so klein sind, dass eine zusätzliche Aufteilung der Personen nach Kompetenzzuwächsen die Bildung von inhaltlich ausdeutbaren Clustern erschwert. Ferner kann eine getrennte Clusterung dazu führen, dass inhaltliche Unterschiede und Gemeinsamkeiten für Personen (bzw. Teams) mit (gemittelt) hohen bzw. niedrigen Kompetenzzuwächsen aus den Clustern nur schwer erkennbar sind. Ein Vorteil dieses Vorgehens wäre allerdings, dass die Gemeinsamkeiten innerhalb der nach Kompetenzzuwächsen gebildeten Gruppen klarer hervorstechen könnten. Für größere Stichproben scheint ein derartiges Vorgehen also mindestens ergänzend anzuraten. Für Einheit 1 (größte Stichprobe) wurde daher eine Clusterung einzeln für die nach Kompetenzzuwächsen gruppierten Personen vorgenommen (die Team-Stichprobe ist zu klein). Es ergeben sich ähnliche Ergebnisse wie nachfolgend dargestellt, wenn für die den Kompetenzzuwächsen zugehörigen Gruppen zugelassen wird, dass bereits ein einziges Aktivitätsprofil einen eigenen Aktivitätsprofiltypen bilden kann, dem dann ggf. nachträglich Ausreißer zugeordnet werden.

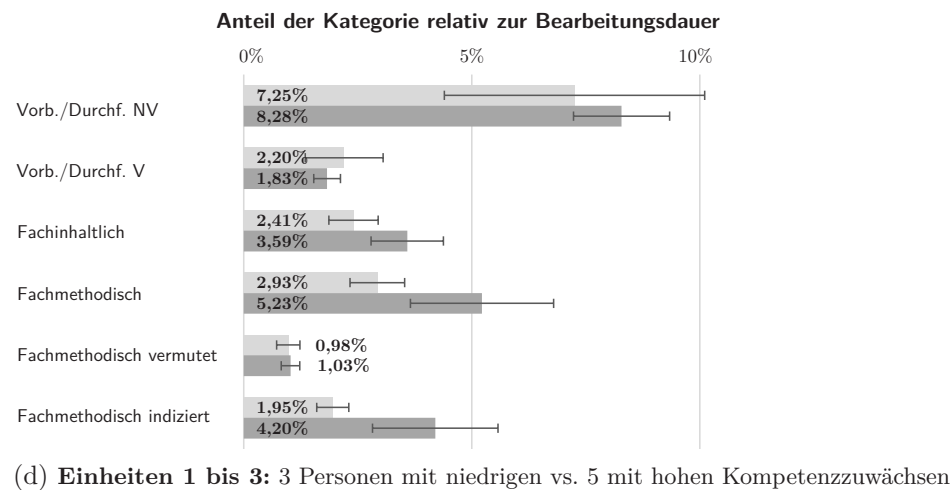
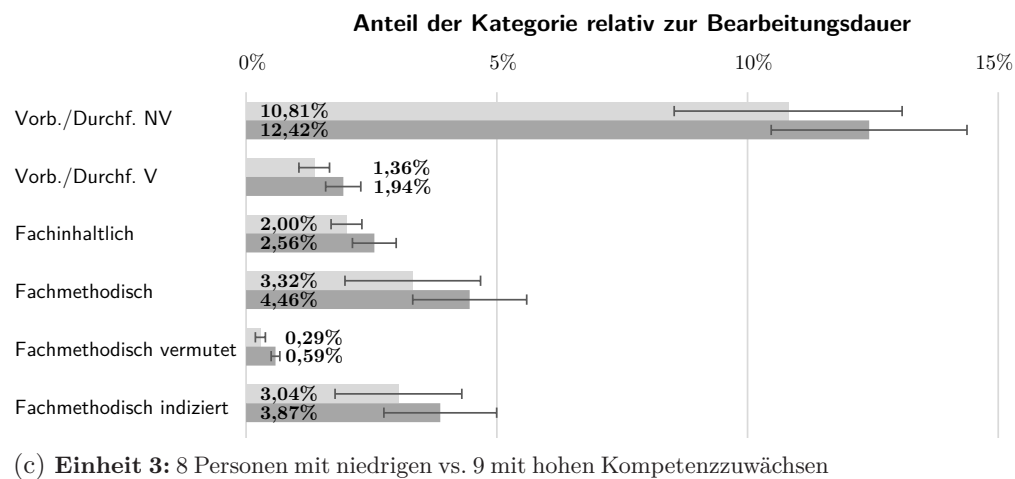
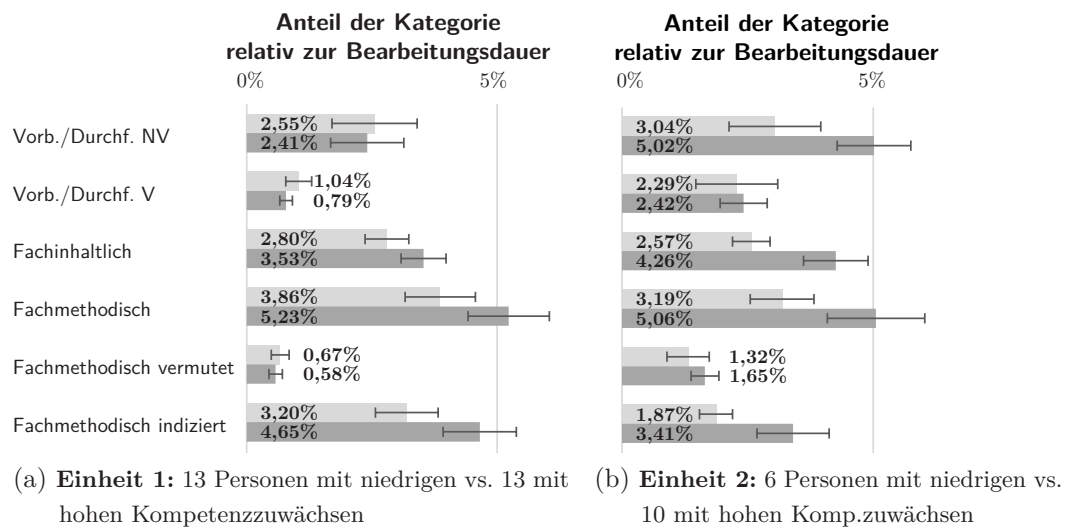


Abbildung 5.33: Mittlere Anteile der betrachteten Aktivitäten von Personen bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante (hellgraue Balken: niedrige, dunkelgraue: hohe Kompetenzzuwächse). Fehlerbalken entsprechen einem Standardfehler des Mittelwertes nach oben und nach unten.

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

In der Übersichtstabelle 5.13 ist für alle Aktivitätsprofiltypen die (mathematisierte) Kurzbeschreibung aus dem vorigen Unterabschnitt aufgeführt. Ferner ist aufgelistet, wie viele Personen (bzw. Teams) insgesamt, mit (gemittelt) hohen und mit (gemittelt) niedrigen Kompetenzzuwächsen einem jeweiligen Typ zugehörig sind. Die Anzahl der Personen (bzw. Teams) mit (gemittelt) mittleren Kompetenzzuwächsen ist nicht in der Tabelle dargestellt, ergibt sich aber rechnerisch aus den angegebenen Zahlen.

Der Vergleich der Personen- bzw. Teamanzahlen für (gemittelt) hohe und (gemittelt) niedrige Kompetenzzuwächse ergibt, dass zu den meisten Aktivitätsprofiltypen Personen bzw. Teams aus beiden Kompetenzzuwachs-Gruppen gehören. *Für weniger als die Hälfte aller Typen lassen sich klare Tendenzen bezüglich des Kompetenzzuwachses ausmachen.* Diese finden sich vor allem für die Profile auf Personen-Ebene bei den Einheiten 1 und 3. Dies scheint eher nicht damit zusammenzuhängen, dass für Einheit 2 weniger Personen als bei den anderen beiden Einheiten untersucht wurden. Vielmehr scheint es nahezuliegen, dass entweder die untersuchten Einheiten in unterschiedlicher Weise mit dem betrachteten Gesamtkompetenzzuwachs in Zusammenhang stehen oder dass die untersuchten Aktivitäten die Prozesse des Kompetenzaufbaus zu den verschiedenen Teilprozessen des experimentbezogenen Denkens und Arbeitens unterschiedlich gut abbilden (und insbesondere für Einheit 2 z. B. erst eine genauere Betrachtung der nonverbalen Aktivitäten des Vorbereitens und Durchführens oder der fachmethodischen Beiträge Aufschluss darüber geben würde, wo tatsächlich für Kompetenzaufbau zum Teilprozess Planung von Versuchen relevante Aktivitäten vollzogen werden).

Personen für Einheit 1 (Aktivitätsprofiltypen auf S. 135). Für den Aktivitätsprofiltyp E1-P-P1-C1 finden sich nur Personen, die hohe Kompetenzzuwächse erreichen. Abgrenzungsüberlegungen zu anderen Typen legen nahe, dass *sowohl der große Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen* (für Typen, die der Partition E1-P-P2 mit niedrigen Anteilen dieser Aktivität angehören, finden sich durchgängig ausgeglichene Verhältnisse bezüglich der erreichten Kompetenzzuwächse) *als auch der sehr große Anteil fachmethodisch-indizierter Beiträge von Bedeutung für die erreichten Kompetenzzuwächse sind* (nur große Anteile zum nonverbalen Vorbereiten und Durchführen reichen nicht aus, sondern sind im Gegenteil anscheinend eher kontraproduktiv wie sich bei E1-P-P1-C3 zeigt). Das könnte als ein Hinweis darauf gedeutet werden, dass für Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten insgesamt die *Verknüpfung von nonverbalem Vorbereiten und Durchführen von Versuchen mit Überlegungen bzw. Verbalisierungen zum Formulieren von Fragen und Hypothesen* (wofür der Anteil von Fachmetho-

Tabelle 5.13: Übersicht über alle Aktivitätsprofiltypen und die Anzahl der Personen (bzw. Teams) mit (gemittelt) hohen und niedrigen Kompetenzzuwächsen. Die Anzahl der Personen (bzw. Teams) mit (gemittelt) mittleren Kompetenzzuwächsen ist nicht in der Tabelle dargestellt.

Aktivitätsprofiltyp		Anzahl d. Personen / Teams		
Name	Kurzbeschreibung	Insg. im Profil	Zuwachs hoch	Zuwachs niedrig
<i>Einheit 1 – Personen</i>				
E1-P-P1-C1	FMi $\gg$ FI $\approx$ VDn $\gg$ 0	3	3	0
E1-P-P1-C2	FMi $>$ FI $\approx$ VDn $\gg$ 0	7	2	4*
E1-P-P1-C3	VDn $\gg$ alles andere	3	0	2
E1-P-P1-C4	FI $>$ FMi; VDn $\gg$ 0	1	0	1*
E1-P-P2-C1	FMi $\ll$ FI; VDn $\geq$ 0	5	2	2
E1-P-P2-C2	FMi $\gg$ FI; VDn $\geq$ 0	8	4	4
E1-P-P2-C3	FMi $\approx$ FI $>$ VDv $\approx$ FMv; VDn $\geq$ 0	4	1	2
<i>Einheit 1 – Teams</i>				
E1-T-C1	FMi $\gg$ VDn $>$ FI	2	1	0
E1-T-C2	VDn $>$ FMi $>$ FI	2	0	2
E1-T-C3	FMi $\geq$ FI $\geq$ VDn	3	1	0
E1-T-C4	FI $\geq$ FMi $\gg$ VDn	2	1	1
E1-T-C5	FMi $>$ FI $\gg$ VDn	2	1	1
<i>Einheit 2 – Personen</i>				
E2-P-P1-C1	FMi $\approx$ FI $\approx$ 2 $\cdot$ FMv $\approx$ 2 $\cdot$ VDv	6	5	1*
E2-P-P2-C1	FI $>$ FMi $>$ VDv	4	2	2
E2-P-P2-C2	FI $\approx$ VDv $>$ FMi	2	1	**
E2-P-P2-C3	VDv $\geq$ FI $>$ FMv $\geq$ FMi	2	1	1
E2-P-Px-C1	FMi $\approx$ FI $\geq$ VDv	3	1	2
E2-P-Px-C2	FMi $>$ VDn und alle anderen	2	0	1
<i>Einheit 2 – Teams</i>				
E2-T-C1	FI $>$ FMi $>$ alle anderen	3	1	0
E2-T-C2	VDv $>$ FMi $>$ alle anderen	2	1	1
E2-T-C3	FMi $>$ FI; VDv $\approx$ alle	2	1	1
<i>Einheit 3 – Personen</i>				
E3-P-C1	FI $>$ FMi; FMv $\approx$ 0	5	1	3
E3-P-C2	VDv $\geq$ FI $>$ FMi; FMv $\approx$ 0	2	1	1*
E3-P-C3	FMi $>$ FI; FMv $\gg$ 0	2	2	0
E3-P-C4	VDn $\geq$ FMi $>$ FI	7	2	3*
E3-P-C5	VDn $\gg$ FMi $>$ FI; FMv $\gg$ 0	3	3	0
<i>Einheit 3 – Teams</i>				
E3-T-C1	VDn $\geq$ FMi $>$ FI	3	1	2
E3-T-C2	VDn $\gg$ FMi $>$ FI	2	1	0

Anmerkungen. * = Davon hat eine Person stark negativen Kompetenzzuwachs (NEG); ** = Für eine Person liegt kein Wert für den Kompetenzzuwachs vor.

Zur Kurzbeschreibung: Verglichen werden Anteile von Aktivitäten an der Gesamtbearbeitungsdauer.

Aktivitäten: FMi = Fachmethodisch-Indiziert; FMv = Fachmethodisch-Vermutet; FI = Fachinhaltlich;

VDn = VorbereitenDurchführenNonverbal; VDv = VorbereitenDurchführenVerbal.

Verhältnisse der Anteile: $\gg$ = Viel größer als; $>$ = größer als; $\geq$ = nur minimal größer als; $\approx$ = etwa gleich groß.

disch-Indiziert in Einheit 1 vor allem steht) bedeutsam ist. Natürlich wurden die inhaltlichen Bezüge der jeweiligen Äußerungen nicht analysiert, so dass eine Aussage über real getätigte Verknüpfungen nicht möglich ist; dennoch scheint es zumindest nicht undenkbar, dass es hilfreich für den Kompetenzaufbau zum Formulieren von Fragen und Hypothesen ist, mit dem jeweiligen Versuchskontext durch ausführliche nonverbale Beschäftigung vertraut zu sein.⁶⁵

Ein Kontrast von E1-P-P1-C1 mit dem sehr ähnlichen E1-P-P1-C2 deutet darauf hin, dass ein *zeitlicher Anteil fachmethodisch-indizierter Beiträge*, der die Anteile aller anderen Aktivitäten deutlich übersteigt, mit höheren Kompetenzzuwächsen einhergeht (sofern ein gewisser Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen vorhanden ist, contra E1-P-P2-C2). Alternativ könnten die unterschiedlichen Verteilungen hinsichtlich der erreichten Kompetenzzuwächse für diese beiden Typen auch mit dem *Anteil fachmethodisch-vermuteter Beiträge* zusammenhängen, welcher nur bei E1-P-P2-C1 von ähnlicher (relativer) Größenordnung ist wie bei dem erfolgreichen Typ E1-P-P1-C1. Bei dem Kontrast der beiden Typen E1-P-P1-C1 und E1-P-P1-C2 ist ferner zu bedenken, dass eine Person aus dem Typen E1-P-P1-C2 einen stark negativen Kompetenzzuwachs aufweist. Eventuell hat diese Person den Test intendiert fehlerhaft bearbeitet oder sie lernt – wie andere Personen mit hohem Kompetenzzuwachs – viel dazu, allerdings aufgrund von Missverstehen der Instruktion etwas Kompetenzabträgliches. Wird diese Person ausgeklammert, wird deutlich, dass auch bei Zusammenfassung der beiden Typen E1-P-P1-C1 und E1-P-P1-C2 grundsätzlich ähnliche Aussagen darüber abgeleitet werden können, welche Relationen von zeitlichen Anteilen der Aktivitäten mit hohen Kompetenzzuwächsen einhergehen.

Für den Aktivitätsprofiltyp E1-P-P1-C3 finden sich *nur Personen mit niedrigen*

⁶⁵ *Methodische Randbemerkung.* Aufgrund des Vorgehens beim Training der Kodierer besteht die Möglichkeit, dass die Hilfskräfte 2 und 3 beim Kodieren der Einheit 1 nicht optimal für das Erkennen von Beiträgen der Kategorie *Fachmethodisch-Vermutet* geschult waren (insbesondere solche Beiträge, die zum Planen und Durchführen von Versuchen oder zum Auswerten und Interpretieren von Daten sind, da die Hilfskräfte zwar selbst die Einheiten der Instruktion zu diesen Teilprozessen des experimentbezogenen Denkens und Arbeitens bearbeitet hatten, aber bis dato nur in Ansätzen mit der umfangreichen Eindenkarbeit in Schüleräußerungen zu diesen Teilprozessen begonnen hatten). Der Autor der Arbeit hat daher nachträglich einzelne Abschnitte des Materials gesichtet, in denen *VorbereitenDurchführenNonverbal* kodiert ist, sowie eigene Kodiererproben aus einem sehr frühen Stadium hinzugezogen. Eine Prüfung dieser Daten ergibt, dass sich einige wenige fachmethodisch-vermutete Beiträge mit Bezug zum Planen und Durchführen von Versuchen finden. Da die Sichtung nur stellenweise stattfand, kann nur schwer geschätzt werden, wie stark sich die Anteile von *Fachmethodisch-Vermutet* für Einheit 1 bei einer nachträglichen Korrektur ändern würden. Allerdings legen sie Sichtungen nahe, dass sich die vermuteten fachmethodischen Beiträge insbesondere bei Personen mit hohen Zuwächsen finden lassen. Dieses nicht optimal abgesicherte Ergebnis passt gut zu den Ergebnissen für die anderen beiden Einheiten, die weiter unten berichtet werden.

Kompetenzzuwachsen. Personen dieses Typs scheinen verbale Aktivitäten zugunsten nonverbaler Aktivitäten zurückzustellen, was eventuell als ein Hinweis darauf gedeutet werden kann, dass diese Personen *stärker auf Handlungen als auf methodische Überlegungen fokussiert* sind. Diese Deutung liegt auch deshalb nahe, weil für den Typ E1-P-P1-C3 der Anteil aller fachmethodischen Beiträge zusammengenommen auch absolut niedriger ist als in allen anderen Profilen. Gleichzeitig ist allerdings anzumerken, dass auch Personen, für die die Anteile fachmethodischer Beiträge geringer sind als die Anteile fachinhaltlicher Beiträge (insbesondere E1-P-P2-C1), hohe Kompetenzzuwächse erreichen (allerdings nur etwa die Hälfte der zum Typ gehörigen Personen).

Personen für Einheit 2 (Aktivitätsprofiltypen auf S. 149). Der Aktivitätsprofiltyp E2-P-P1-C1 geht sehr deutlich mit hohen Kompetenzzuwachsen einher, während die anderen Typen keine Tendenzen erkennen lassen. Wird berücksichtigt, dass für die meisten Aktivitätsprofiltypen zu Einheit 2 der Anteil fachinhaltlicher Beiträge größer ist als der fachmethodisch-indizierter, wird deutlich, dass sich für das Profil E2-P-P1-C1 *vergleichsweise große Anteile fachmethodisch-indizierter Beiträge* finden, obwohl diese immer noch knapp unter den Anteilen fachinhaltlicher Beiträge liegen. Wie bereits für Einheit 1 findet sich also auch für Einheit 2 eine Relevanz von fachmethodisch-indizierten Beiträgen für den Kompetenzzuwachs; wie für Einheit 1 (siehe Vergleich der Partitionen E1-P-P1 und E1-P-P2) findet sich allerdings auch für Einheit 2 eine Einschränkung, weil der Typ E2-P-Px-C2 trotz sehr großer, die aller anderen Aktivitäten übersteigende Anteile nicht mit hohen Kompetenzzuwachsen einhergeht, sondern tendenziell eher mit niedrigen.

Eventuell stellt sich ein positiver Zusammenhang von Kompetenzzuwachs und fachmethodisch-indizierten Beiträgen für Einheit 2 *erst bei hinreichend großem zeitlichem Anteil fachinhaltlicher Beiträge* ein (relativ zu andere Aktivitäten, vgl. Typen E2-P-P1-C1 und E2-P-P2-C2). Dieses Ergebnis wäre spannend, da sich zwar auf Ebene der Gesamtkohorte kein Einfluss von vor der Instruktion vorliegenden fachinhaltlichen Kompetenzen auf den Kompetenzzuwachs finden lässt (Vorholzer, 2016), sich allerdings darauf keine Rückschlüsse darauf ableiten lassen, dass Rückgriffe auf Fachinhalte für den fachmethodischen Kompetenzaufbau irrelevant sind. Das mögliche Ergebnis aus den Analysen der Aktivitätsprofile zu Einheit 2 würde möglicherweise nun gerade zeigen, dass Rückgriffe auf Fachinhalte mindestens einen gewissen zeitlichen Anteil einnehmen sollten, damit das Profil der Aktivitäten in positivem Zusammenhang mit dem Kompetenzzuwachs steht. Die Einschränkungen dieser Deutung sind allerdings offensichtlich. Unter anderem sagt der zeitliche Umfang

fachinhaltlicher Beiträge nichts darüber, ob fachinhaltliche Kompetenzen genutzt wurden oder gar auf die zu lernenden fachmethodischen Kompetenzen Anwendung gefunden haben.

Auch wenn sich im Vergleich der Aktivitätsprofiltypen keine spezifische Aussage darüber treffen lässt, ob fachmethodisch-vermutete Beiträge eine Rolle in kompetenzförderlichen Gefügen von Aktivitäten spielen, weil die Ausgewogenheit bezüglich der Kompetenzzuwächse für nahezu alle Typen findet, lässt sich innerhalb des Typs E2-P-P1-C1 ein spannender Befund berichten, der zufällig entdeckt wurde, als die Profile zur Präsentation aufbereitet wurden: Die einzige Person mit sehr geringem Anteil von Fachmethodisch-Vermutet im Aktivitätsprofiltyp E2-P-P1-C1 ist F16. Sie ist die einzige Person, die keinen hohen Kompetenzzuwachs hat. Demnach könnte für Einheit 2 wie bereits für Einheit 1 die *Verknüpfung von Versuchsanweisungen und Versuchshandlungen mit fachmethodischen Überlegungen* von Bedeutung für die erreichten Kompetenzzuwächse sein (denn fachmethodisch-vermutete Beiträge betreffen im Kodiermanual insbesondere solche Fälle, siehe Abschnitt 5.1).⁶⁶

Vor dem Hintergrund der deutlichen Mittelwertunterschiede für sowohl fachinhaltliche sowie fachmethodisch-indizierte Beiträge als auch nonverbale Aktivitäten des Vorbereitens und Durchführens von Versuchen im Vergleich von Personen mit hohen vs. niedrigen Kompetenzzuwächsen bei Einheit 2 (siehe Abbildung 5.33) könnte *umfassendere kognitive Aktivität der Lernenden eine Erklärung für höhere Kompetenzzuwächse* darstellen; weil sich diese allerdings in unterschiedlichen Aktivitäten widerspiegelt, ist sie – so eine mögliche Deutung – nicht mit einem spezifischen Profil verbunden. Vielmehr müssten genauere Analysen der Inhalte der jeweiligen Aktivitäten zeigen, ob es Unterschiede zwischen Lernenden mit unterschiedlichen Kompetenzzuwächsen gibt, die neben umfangreicherer kognitiver Aktivität auf konkret mit den Lernzielen der Instruktion verbundene Spezifika der Aktivitäten der Lernenden mit höheren Kompetenzzuwächsen hinweisen (oder ob unspezifische umfangreichere kognitive Aktivität bereits zuträglich für den Kompetenzzuwachs ist).

⁶⁶Grundsätzlich lässt sich an dieser Stelle zum methodischen Vorgehen anmerken, dass die Ausführungen zum Zusammenhang von Kompetenzzuwächsen und Aktivitätsprofilen sich in allen anderen Absätzen des Unterabschnitts nur auf die Aktivitätsprofiltypen im Vergleich zueinander beziehen. Für eine zusätzliche Analyse *innerhalb* jedes einzelnen Aktivitätsprofiltyps spräche, dass innerhalb der Typen von einer größeren Vergleichbarkeit der Lernenden hinsichtlich verschiedener Variablen ausgegangen werden könnte. Gegebenenfalls zeigen sich nicht innerhalb der Gesamtheit aller Lernenden, die eine spezifische Einheit bearbeitet haben, sondern eben erst innerhalb der Lernenden, die eine spezifische Einheit mit einem vergleichbaren Aktivitätsprofiltyp bearbeitet haben, Zusammenhänge einzelner Aktivitäten mit den erreichten Kompetenzzuwächsen. Das Beispiel im obigen Text illustriert dies. Auf derartige Analysen wurde allerdings verzichtet, weil zum einen die Anzahl der einem jeweiligen Typ zugehörigen Lernenden häufig sehr gering ist und zum anderen der Umfang der Arbeit begrenzt werden musste.

Personen für Einheit 3 (Aktivitätsprofiltypen auf S.155). In Einheit 3 scheint erneut die Verknüpfung von nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen bedeutsam in erfolgreichen Prozessen des Kompetenzaufbaus zu sein. Die Einheit bietet sehr viel Gelegenheit zum nonverbalen Vorbereiten und Durchführen von Versuchen, was sich in den großen Anteilen der Kategorie *VorbereitenDurchführenNonverbal* widerspiegelt. Aktivitätsprofiltypen, in denen überwiegend Personen mit hohen Kompetenzzuwächsen sind (E3-P-C3, E3-P-C5) haben einen substantiell größeren Anteil von *Fachmethodisch-Vermutet* im Vergleich zu den anderen Aktivitäten als andere Aktivitätsprofiltypen. Diese Aktivität scheint tatsächlich eine große Bedeutung für den Kompetenzzuwachs zu haben, weil beim Typ E3-P-C3 alle anderen verbalen Aktivitäten im Vergleich zum nonverbalen Vorbereiten und Durchführen von Versuchen nur gering ausgeprägt sind⁶⁷ und trotzdem große Kompetenzzuwächse erreicht werden. Dieses Ergebnis ist im Einklang mit dem aus Abbildung 5.33 hervorgehenden signifikante Unterschied bezüglich fachmethodisch-vermuteter Beiträge zwischen Lernenden mit unterschiedlichen Kompetenzzuwächsen, bei dem Lernende mit hohen Kompetenzzuwächsen deutlich größere zeitliche Anteile dieser Aktivität aufweisen.

Die Deutung, dass fachmethodisch-vermutete Beiträge bei Einheit 3 einen relevanten Faktor für Kompetenzzuwächse der Lernenden darstellen, wird auch dadurch gestützt, dass sowohl im Typ E3-P-C1 als auch bei den zwei Ausreißern (die beide niedrige Kompetenzzuwächse haben) der Anteil fachmethodisch-vermuteter Beiträge sehr nah bei 0 % liegt. (Relativ) Viel über Fachmethoden zu sprechen, reicht hingegen nicht aus, wie an dem Typen E3-P-C4 deutlich wird, in dem der Anteil von *Fachmethodisch-Indiziert* relativ zu den anderen Aktivitäten sehr groß ist, die Kompetenzzuwächse allerdings durchmischt sind. Dass es nicht ausreicht, viel über Fachmethoden zu sprechen, könnte damit zusammenhängen, dass indizierte fachmethodische Beiträge nicht so große Verknüpfung von Fachmethoden erfordern wie vermutet fachmethodische Beiträge. Insbesondere in Einheit 3 bezieht sich ein großer Teil des Materials auf die Wiedergabe von vorher adressierten Konzepten, vermutlich rührt der sehr große Anteil fachmethodisch-indizierter Beiträge im Aktivitätsprofiltyp E3-P-C4 daher.

Teamebene. Die Betrachtungen der Aktivitätsprofiltypen auf Teamebene ergeben ein sehr durchmishtes Bild. Klare Hinweise auf Relevanz für Prozesse des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten lassen sich nicht

⁶⁷Beim in Abbildung 5.30 dargestellten Repräsentanten D12 im Bereich von ca. 1,%; im anderen zum Typ gehörigen Profil von F18 alle in der Größenordnung von ca. 3 %. Dabei nimmt *VorbereitenDurchführenNonverbal* einmal ca. 8 %, einmal ca. 20 % der Gesamtbearbeitungsdauer der Einheit ein. Es scheint also nicht nur um absolute Prozentwerte zu gehen.

ableiten. Interessanterweise finden sich im Typen E1-T-P1-C2, bei dem der Anteil von Fachmethodisch-Vermutet auf Teamebene relativ zu den anderen Aktivitäten höher als bei anderen Teams ist, nur zwei (von sieben) Teams mit gemittelt hohen Kompetenzzuwächsen. Aus Abbildung 5.22 geht allerdings hervor, dass der Wert des Anteils im Vergleich zum Wert des Anteils in anderen Profilen relativ gering ist (entspricht in etwa einem geringen absoluten Anteil dieser Aktivität). Damit spricht auch dieser Befund letztlich eher für die Bedeutung der fachmethodisch-vermuteten Beiträge. Insgesamt lässt sich aber folgern, dass es mit Blick auf die erreichten Kompetenzzuwächse deutlich stärker auf Aktivitätsprofile der einzelnen Personen ankommt als auf die der Teams.

Aktivitätsprofile für komplette Instruktion. Für die 9 Personen, die für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante analysiert wurden, können auch die Aktivitätsprofile betrachtet werden, die sich bei gewichteter Zusammenfassung über jeweils alle drei Einheiten ergeben. Ergänzend zu der in Abbildung 5.33(d) auf Seite 164 vorgenommenen Differenzierung von Lernenden mit hohen und niedrigen Kompetenzzuwächsen wird in Tabelle 5.14 für alle Lernenden das gesamte Aktivitätsprofil angegeben. Dabei wird jeweils das mit der Einheitsdauer gewichtete Mittel⁶⁸ der Anteile der Aktivitäten für die drei Einheiten gebildet. Um die Übersichtlichkeit zu wahren, wurde etwas von der ansonsten gewählten Darstellungsform abgewichen, die Reihenfolge der Aktivitäten allerdings beibehalten.

Es scheint so zu sein, dass Personen, für die der Anteil fachmethodisch-indizierter Beiträge höher ist als der Anteil fachinhaltlicher Beiträge hohe Kompetenzzuwächse erreichen (aber andersherum nicht unbedingt, vgl. F18). Eine Korrelationsanalyse nach Pearson zeigt, dass die Kompetenzzuwächse (bei Setzung von $H = 3$, $N = 1$, $NEG = -1$) nicht optimal mit dem Anteil fachmethodischer Beiträge korrelieren

Tabelle 5.14: Aktivitätsprofile über alle Einheiten hinweg gewichtet gemittelt für die 9 für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante untersuchten Personen.

Person	B4	B5	D11	D12	F16	F18	G19	G20	G21
VDn	 7,14%	 11,76%	 8,35%	 5,36%	 8,69%	 8,81%	 6,61%	 2,66%	 12,49%
VDv	 0,93%	 2,72%	 1,71%	 1,69%	 3,02%	 2,09%	 1,55%	 1,16%	 3,89%
FI	 1,83%	 6,47%	 3,56%	 2,49%	 3,64%	 3,59%	 1,68%	 2,08%	 3,46%
FM	 2,32%	 11,20%	 5,32%	 3,95%	 3,28%	 3,34%	 2,18%	 2,50%	 4,11%
FMv	 0,48%	 1,72%	 1,10%	 1,01%	 0,39%	 0,82%	 0,58%	 0,89%	 1,46%
FMI	 1,84%	 9,48%	 4,22%	 2,94%	 2,89%	 2,51%	 1,59%	 1,61%	 2,66%
Zuwachs	H	H	H	H	NEG	H	N	N	N

Anmerkungen. Dargestellt sind für jede der Aktivitäten die gemäß der Einheitsdauer gewichteten Mittelwerte. Aktivitäten sind: VBn = Vorbereiten und Durchführen von Versuchen (nonverbal), VDv = Vorbereiten und Durchführen von Versuchen (verbal), FI = Fachinhaltlich, FM = Summe aus FMv und FMI, FMv = Fachmethodisch-vermutet, FMI = Fachmethodisch-indiziert.

⁶⁸Für Einheit 1 ist mit 45 Minuten Bearbeitungsdauer halb so viel Zeit angesetzt. Demnach wird der Anteil der jeweiligen Kategorie in Einheit 1 mit 45, der Anteil in Einheit 2 und 3 jeweils mit 90 multipliziert. Die gebildete Summe der gewichteten zeitlichen Anteile wird dann durch $45 + 90 + 90 = 235$ geteilt.

($r = .51$, $p = .16$). Werden die angegebenen Anteile für jede Person allerdings ins Verhältnis zu dem von ihr insgesamt für alle verbalen Aktivitäten erreichten gewichteten Anteil gesetzt, ergibt sich eine Korrelation fachmethodischer Beiträge mit dem Kompetenzzuwachs von $r = .86$, $p = .003$. Fachmethodisch-vermutete Beiträge scheinen in der gegebenen Darstellung eher unbedeutend zu sein; relativ zu den verbalen Beiträgen korrelieren sie sogar sehr deutlich negativ mit den erreichten Kompetenzzuwächsen. Für die Teilstichprobe der Personen, die für alle drei Einheiten analysiert wurden, ergibt sich also – vermutlich auch, weil sie recht klein ist – ein Bild, das etwas von der Betrachtung der einzelnen Einheiten abweicht. Insbesondere scheinen die fachmethodisch-vermuteten Beiträge eine andere, negative, Rolle einzunehmen. Da auch Beiträge des verbalen Vorbereitens und Durchführens von Versuchen negativ mit dem Kompetenzzuwachs korrelieren (basierend auf den Werten aus Tabelle 5.14 zwar nur mit $r = -.23$, $p = .55$, bei Betrachtung relativ zu den verbalen Aktivitäten allerdings mit $r = -.78$, $p = .01$), liegt es nahe, dass es über alle drei Einheiten hinweg darauf ankommt, möglichst viel der insgesamt auf die Aktivitäten verwendeten Zeit mit fachmethodisch(-indiziert)en Beiträgen zugebracht zu haben.

Aktivitätsprofiltypen für komplette Instruktion. Werden die in Tabelle 5.12 (auf S. 162) dargestellten Muster der Aktivitätsprofile über alle drei Einheiten hinweg mit den erreichten Kompetenzzuwächsen abgeglichen, ergibt sich, dass die Personen mit hohen Kompetenzzuwächsen sich für alle Einheiten solchen Aktivitätsprofiltypen zuordnen lassen, die eher große Anteile mit nonverbalen Vorbereiten und Durchführen von Versuchen zubringen; außerdem finden sich für die Personen mit hohen Kompetenzzuwächsen eher vergleichsweise große Anteile fachmethodisch-indizierter Beiträge. Besonders eindrücklich ist das Ergebnis bezüglich der fachmethodisch-indizierten Beiträge allerdings für die dritte Einheit: drei der Personen weisen fachmethodische Beiträge auf, die größer als alle anderen Aktivitäten sind, zwei weitere weisen sehr große Anteile fachmethodischer Aktivitäten auf. Personen mit niedrigen Kompetenzzuwächsen weisen für die dritte Einheit jedoch, festgemacht an den ihnen zugeordneten Aktivitätsprofiltypen, relativ zu den anderen Aktivitäten geringe Anteile fachmethodisch-indizierter Beiträge auf. Hier stellt sich erneut heraus, dass die einheitsweise Betrachtung bedeutsame Ergebnisse liefern kann – und in diesem Fall insbesondere die dritte Einheit von den anderen Einheiten abweichen könnte (dies stellt u. a. eine Grundlage für die Diskussion in Unterabschnitt 7.3.3, ab S. 467, dar).

5.4.4 Fachmethodische Beiträge in der Gesamtschau (über alle Einheiten hinweg und für die Einheiten) und Kompetenzzuwächse

Zur spezifischen Untersuchung der Relevanz fachmethodischer Beiträge im Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten werden in diesem Unterabschnitt (sowie in nachfolgenden) Vergleiche von Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen vorgenommen. Fachmethodische Beiträge werden in der Arbeit grundsätzlich durch die beiden Kategorien **Fachmethodisch-Indiziert** und **Fachmethodisch-Vermutet** abgebildet, die zwischen Äußerungen mit deutlichem Bezug zu Fachmethoden (zu fachmethodischen Konzepten) und Äußerungen ohne klare Indizien auf Bezugnahme zu Fachmethoden, die allerdings im Kontext auf fachmethodische Überlegungen schließen lassen, unterscheiden. Im Kodiermanual wurde zusätzlich eine Unterscheidung der Bezüge zu den verschiedenen Teilprozessen vorgenommen. Für diese zusätzlichen Kategorien *FH* (fachmethodische Beiträge mit Bezug zum Formulieren von Fragen und Hypothesen), *PU* (Bezug zum Planen und Durchführen von Untersuchungen) und *AI* (Bezug zum Auswerten und Interpretieren von Daten) ist in Abbildung 5.34 (auf S. 174) dargestellt, wie die Anteile über die verschiedenen Einheiten der explizit-fachmethodischen Instruktionsvariante verteilt sind.

Es ist deutlich, dass der Großteil der fachmethodischen Beiträge bei Einheit 1, die den Teilprozess des Formulierens von Fragen und Hypothesen explizit adressiert, auf denselben Teilprozess bezogen ist; analog gilt dies für die beiden anderen Einheiten. Auffällig ist, dass einzig Äußerungen mit Bezug zum Planen und Durchführen von Untersuchungen in allen drei Einheiten vorkommen. Äußerungen mit Bezug zu den anderen Teilprozessen kommen nur in den jeweils direkt angrenzenden Einheiten der explizit-fachmethodischen Instruktionsvariante vor. In Tabelle 5.15 sind die absoluten Anteile der Aktivitäten dargestellt. Es wird unter anderem deutlich, dass insgesamt

Tabelle 5.15: Zeitliche Anteile fachmethodischer Beiträge (absolut) aufgelöst nach Zugehörigkeit zu den Teilprozessen des experimentbezogenen Denkens und Arbeitens (FH: Formulieren von Fragen und Hypothesen; PU: Planen und Durchführen von Untersuchungen; AI: Auswerten und Interpretieren von Daten) und nach den Einheiten der explizit-fachmethodischen Instruktionsvariante (E1, E2, E3).

Bezug zu Teilprozess	E1		E2		E3		Gesamt			
							Fall 1		Fall 2	
FH	104,1	(58,4)	22,6	(24,6)	0,1	(0,6)	100,3	(68,9)	113,8	(56,7)
PU	4,4	(7,7)	142,7	(117,2)	32,5	(34,2)	100,6	(121,5)	165,9	(138,7)
AI	0,2	(1)	2,4	(5,3)	105,9	(96,2)	66,6	(91,3)	83,8	(82,2)

Anmerkungen. Angaben in Sekunden; Standardabweichung in Klammern. Fall 1 = Mittelwert über alle 39 Personen, die mindestens eine Einheit bearbeitet haben. Fall 2 = Mittelwert für die 12 Personen, die bei allen Einheiten anwesend waren.

Mittlerer Anteil der Kategorie an der Bearbeitungsdauer der jeweiligen Einheit

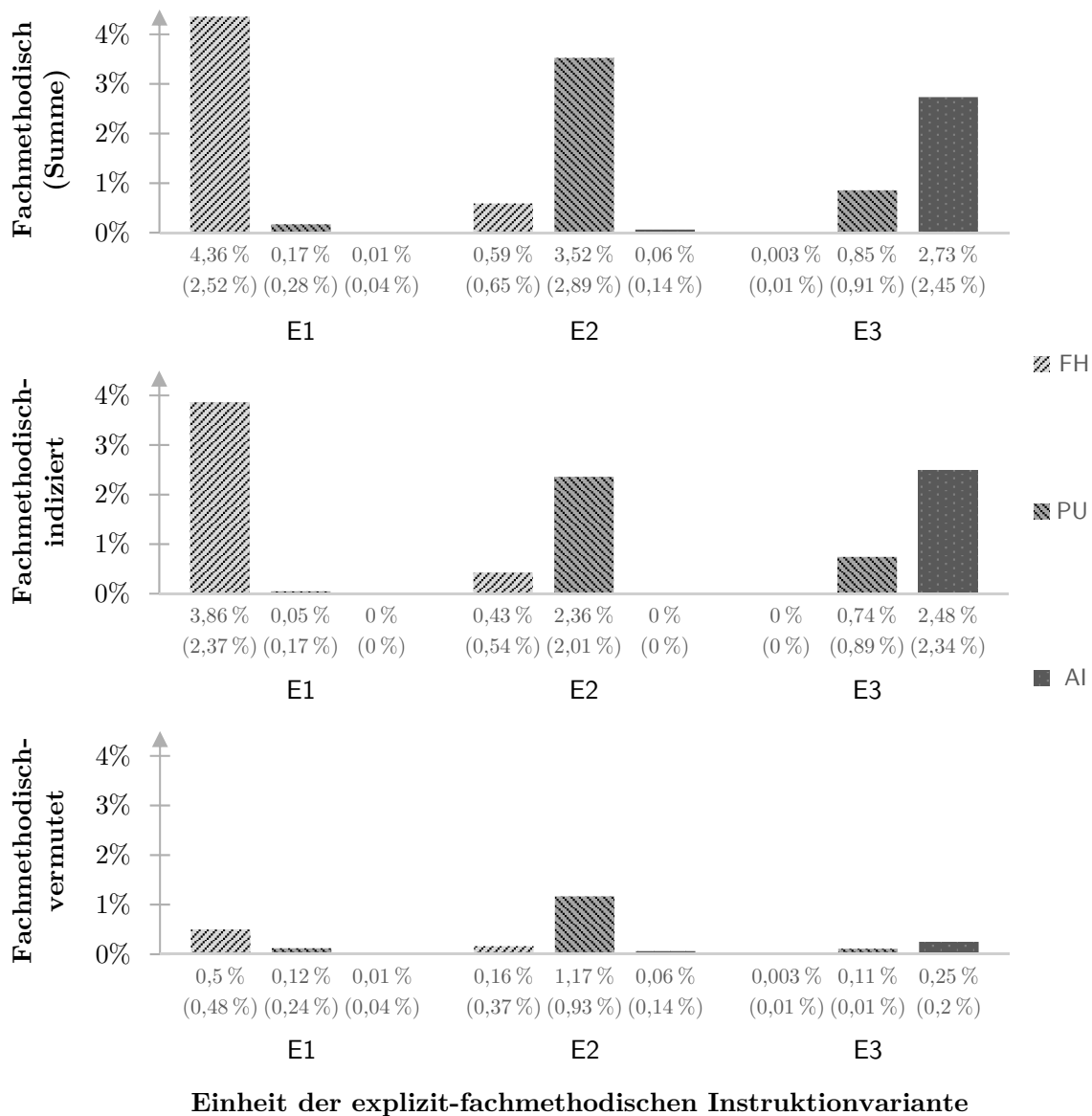


Abbildung 5.34: Relative zeitliche Anteile fachmethodischer Beiträge (oben: Summe; mitte: indizierte; unten: vermutete) aufgelöst nach Zugehörigkeit zu den Teilprozessen des experimentbezogenen Denkens und Arbeitens (verschieden schraffiert) und nach den Einheiten der explizit-fachmethodischen Instruktionsvariante (E1, E2, E3). Angegeben ist jeweils der Mittelwert über die Personen, die die jeweilige Einheit bearbeitet haben (siehe Tabelle A.2 im Anhang). In Klammern ist die jeweilige Standardabweichung angegeben.

dem Teilprozess des Planens und Durchführens von Untersuchungen tendenziell der größte Zeitanteil gewidmet wird; der nächstgrößte Anteil entfällt interessanterweise auf den Teilprozess des Formulierens von Fragen und Hypothesen, obwohl Einheit 1 mit 45 Minuten angesetzter Bearbeitungsdauer nur halb so umfangreich wie die anderen beiden Einheiten ist.

Wenngleich mit der Betrachtung der Bezüge zu den Teilprozessen spannende Fragen verbunden sind (diskutiert unter anderem in Unterabschnitt 7.3.3, ab S. 467), wird in der vorliegenden Arbeit zumeist keine Unterscheidung dieser Aspekte vorgenommen. Weil die Schwerpunkte der Einheiten sehr gut zu den gefundenen Schwerpunkten der Beiträge passen, werden vielmehr für alle Betrachtungen die fachmethodischen Beiträge unabhängig von ihren Bezügen zu Teilprozessen des experimentbezogenen Denkens und Arbeitens hinzugezogen und Deutungen anhand der groben Zuordnung zur Einheit vorgenommen. Dies scheint eine sinnvolle Reduktion der Komplexität zu sein. Nur im Vergleich der Instruktionsvarianten wird erneut auf die Teilprozesse eingegangen (begründet in Unterabschnitt 5.5.3, ab S. 228).

Die Ergebnisse für die einzelnen Kategorien (FMi = Fachmethodisch-Indiziert; FMv = Fachmethodisch-Vermutet) sowie für alle fachmethodischen Beiträge (FI = FMi + FMv) werden nachfolgend differenziert berichtet, weil a priori nicht vollständig klar ist, ob mit Unterschieden der Relevanz der jeweiligen Kategorien für den Kompetenzaufbau zu rechnen ist. Das Vorgehen besteht in einem Vergleich der zeitlichen Anteile fachmethodischer Beiträge für Lernende mit hohen vs. niedrigen Kompetenzzuwächsen, weil identifiziert werden soll, inwiefern zeitliche Anteile fachmethodischer Beiträge sich für Lernende mit unterschiedlichen Kompetenzzuwächsen unterscheiden. Dabei werden die zeitlichen Anteile als relative Anteile der jeweiligen Kategorie an der Bearbeitungsdauer des jeweils zugrundeliegenden Instruktionsschnittes (Einheit, Abschnitt, eine Karte oder mehrere Karten) angegeben. Diese Entscheidung beruht auf dem Ergebnis, dass Lernende mit hohen Kompetenzzuwächsen tendenziell größere Bearbeitungsdauern vorweisen als Lernende mit niedrigen Kompetenzzuwächsen (siehe Unterabschnitt 5.4.1). Um in den Analysen zu fachmethodischen Beiträgen nicht die Ergebnisse zu den Bearbeitungsdauern zu großen Teilen erneut abzubilden, werden die zeitlichen Anteile fachmethodischer Beiträge relativ zu den Bearbeitungsdauern betrachtet.

5.4.4.1 Vergleich der Gesamtdauer fachmethodischer Beiträge für Personen mit unterschiedlichen Kompetenzzuwächsen

Eine Näherung an fachmethodische Beiträge besteht darin, Personen mit hohen vs. niedrigen Kompetenzzuwächsen hinsichtlich der *Gesamtdauer* ihrer fachmethodischen

Beiträge zu vergleichen. Im Prinzip entpacken die berichteten Analysen die bereits in Abbildung 5.33 (auf S. 164) verarbeiteten Informationen zu den mittleren Aktivitätsprofilen von Personen mit hohen vs. niedrigen Kompetenzzuwächsen. Die Diagramme sind nicht erneut abgebildet, können aber zum Abgleich hinzugezogen werden.

Werden alle drei Einheiten zusammengefasst, finden sich nur für fachmethodisch-indizierte Beiträge Unterschiede zwischen Personen mit unterschiedlichen Kompetenzzuwächsen. Werden nur Personen betrachtet, die an allen drei Einheiten teilgenommen haben und wird das gewichtete Mittel der Anteile der Aktivitäten über *alle drei Einheiten*⁶⁹ gebildet, ergeben sich aus durchgeführten t-Tests die in Tabelle 5.16 dargestellten Werte. Die großen *p*-Werte dürften mit den sehr kleinen Stichproben zusammenhängen. Zusammengenommen deuten die Ergebnisse darauf hin, dass fachmethodisch-indizierte Beiträge nicht nur den Großteil aller fachmethodischen Beiträge ausmachen, sondern Unterschiede zwischen Personen mit unterschiedlichen Kompetenzzuwächsen sich auch in der Zusammenfassung aller drei Einheiten nur für fachmethodisch-indizierte Beiträge finden. Personen mit hohen Kompetenzzuwächsen sprechen im Mittel etwa 300 Sekunden (5 Minuten) länger über Fachmethoden als Personen mit niedrigen Kompetenzzuwächsen; dies erscheint vor dem Hintergrund von Gesamtbearbeitungsdauern von ca. 10000 Sekunden (ca. 170 Minuten)⁷⁰ zwar kurz, ist allerdings im Verhältnis zur Gesamtdauer der Kategorie fachmethodischer Beiträge groß (siehe letzte Spalte von Tabelle 5.16) und auch vor dem Hintergrund der zeitlichen Anteile der anderen Aktivitäten als nicht unbedeutend anzusehen (insbesondere, wenn berücksichtigt wird, dass eventbasiert kodiert wird und daher der zeitliche Umfang der Kodierungen recht gut den tatsächlichen Sprechdauern entspricht).

Tabelle 5.16: Students t-Tests für die Gesamtdauer fachmethodischer Beiträge von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante.

Vergleich	Zuwachs Niedrig			Zuwachs Hoch			df	<i>t</i> (df)	<i>p</i>	<i>r</i>	Unterschied	
	<i>N_N</i>	MW	SD	<i>N_H</i>	MW	SD					Zeit in s	Relativ
FMi	3	1,9524%	0,6093%	5	4,1995%	3,0765%	6	-1,213	0,271	0,444	303,35	72,2%
FMv		0,9769%	0,4427%		1,0256%	0,4544%	6	-0,148	0,887	0,060	6,58	6,4%
FM		2,9293%	1,0375%		5,2251%	3,5106%	6	-1,074	0,324	0,401	309,93	59,3%

Anmerkungen. Students t-Test; *N_{H/N}* = Stichprobengröße für Gruppe der Teams mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Zeit in s = Zeitliches Äquivalent der Mittelwertdifferenz (Anteil der Aktivität mal angesetzte Dauer der jeweiligen Einheit in Sekunden); Unterschied Relativ = Mittelwertdifferenz geteilt durch größeren der beiden Mittelwerte.

⁶⁹Siehe bereits Fußnote 68.

⁷⁰Mittelwert für E1: 2416 Sekunden mit Standardabweichung von 281 Sekunden; Mittelwert für E2: 4017 Sekunden mit Standardabweichung von 487 Sekunden; Mittelwert für E3: 3814 Sekunden mit Standardabweichung von 479 Sekunden.

Werden die drei Einheiten jeweils einzeln betrachtet, finden sich bei Einheit 1 Unterschiede für fachmethodisch-indizierte Beiträge und bei Einheit 3 Unterschiede für fachmethodisch-vermutete Beiträge zwischen Personen mit unterschiedlichen Kompetenzzuwächsen. Für eine *einheitweise Betrachtung* unter Einbezug aller jeweils analysierten Personen ist in Tabelle 5.17 dargestellt, welche Ergebnisse die durchgeführten Welch-Tests liefern. Während die zeitlichen Anteile fachmethodisch-indizierter Beiträge für Einheit 1 mit mittlerer Effektstärke vorsignifikant mit der Gruppierung nach Kompetenzzuwächsen zusammenhängen, ist dieser Zusammenhang für Einheit 3 nur bei fachmethodisch-vermuteten Beiträgen in ähnlicher Deutlichkeit vorhanden – dies sind die beiden deutlichsten Unterschiede. Insgesamt kann angemerkt werden, dass in allen Fällen der Zusammenhang tendenziell positiv ist (wenngleich sehr häufig nicht signifikant und auch nicht von bedeutender Effektstärke); das heißt, dass Personen mit hohen Kompetenzzuwächsen im Mittel mehr über Fachmethoden sprechen als Personen mit niedrigen Kompetenzzuwächsen. Für *fachmethodisch-indizierte* Beiträge liegt bei Einheit 1 ein deutlicher Zusammenhang vor; bei Einheit 2 könnten der hohe p -Wert sowie der nicht so deutlich wie bei Einheit 1 ausgeprägte Effekt durch den kleineren Stichprobenumfang bedingt sein; bei Einheit 3 scheint hingegen die Bedeutung fachmethodisch-indizierter Beiträge geringer zu sein als in den anderen beiden Einheiten, weil der Unterschied sowohl absolut (Zeit in Sekunden) als auch relativ (Verhältnis des Mittelwertunterschieds zum größeren der beiden Mittelwerte) deutlich kleiner als in den anderen beiden Einheiten ist (siehe auch kleine Effektstärke) und

Tabelle 5.17: Welch-Tests für die Gesamtdauer fachmethodischer Beiträge von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für jede der drei Einheiten der explizit-fachmethodischen Instruktionsvariante.

Vergleich		Zuwachs Niedrig			Zuwachs Hoch			Korr.		Unterschied			
Einheit	Aktivität	N_N	MW	SD	N_H	MW	SD	df	$t(df)$	p	r	Zeit in s	Relativ
1	FMi	8	0,028366	0,019088	12	0,048962	0,025936	17,737	-2,043	0,056	0,436	55,61	42,1%
	FMv		0,005077	0,005179		0,006061	0,004939	14,636	-0,424	0,678	0,110	2,66	16,2%
	FM		0,033443	0,023114		0,055022	0,028780	17,233	-1,852	0,081	0,407	58,26	39,2%
2	FMi	6	0,018717	0,008055	11	0,030963	0,023869	13,440	-1,548	0,145	0,389	66,13	39,6%
	FMv		0,013163	0,010313		0,015023	0,009681	9,826	-0,363	0,724	0,115	10,04	12,4%
	FM		0,031881	0,015546		0,045987	0,032862	14,880	-1,199	0,249	0,297	76,17	30,7%
3	FMi	9	0,027012	0,034941	10	0,034860	0,034084	16,689	-0,495	0,627	0,120	42,38	22,5%
	FMv		0,002538	0,002810		0,005283	0,003073	16,992	-2,034	0,058	0,442	14,82	52,0%
	FM		0,029550	0,037441		0,040143	0,035104	16,491	-0,634	0,535	0,154	57,20	26,4%

Anmerkungen. Welch-Test; $N_{H/N}$ = Stichprobengröße für Gruppe der Teams mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Zeit in s = Zeitliches Äquivalent der Mittelwertdifferenz (Mittelwertdifferenz der Anteile der Aktivität mal angesetzte Dauer der jeweiligen Einheit in Sekunden); Unterschied Relativ = Mittelwertdifferenz geteilt durch größeren der beiden Mittelwerte.

zudem bei weitem nicht signifikant wird. Während in den ersten beiden Einheiten *fachmethodisch-vermutete* Beiträge nur von sehr geringer Bedeutung für den Kompetenzzuwachs zu sein scheinen, tritt für Einheit 3 für diese Kategorie ein relativ und absolut bedeutender Unterschied auf, der zudem mit einem kleinen p -Wert und einer vergleichsweise großen Effektstärke vorliegt.

Personen mit großen zeitlichen Anteilen fachmethodischer Beiträge erreichen hohe Kompetenzzuwächse; unter den Personen mit geringen zeitlichen Anteilen fachmethodischer Beiträge gibt es sowohl welche mit hohen als auch welche mit niedrigen Kompetenzzuwächsen. Aus Streudiagrammen, in denen für jede Person der zeitliche Anteil fachmethodischer Beiträge gegen den Kompetenzzuwachs aufgetragen ist, geht für die einzelnen Einheiten hervor, dass die oben berichteten Zusammenhänge von Kompetenzzuwächsen und zeitlichen Anteilen fachmethodischer Beiträge etwas nuancierter zu beschreiben sind. In Abbildung 5.35 ist exemplarisch für Einheit 2 dargestellt, dass es sowohl Personen mit hohen als auch niedrigen Kompetenzzuwächsen gibt, die geringe zeitliche Anteile fachmethodisch(-indiziert)er Beiträge aufweisen.

Die beschriebene Art des Zusammenhangs, zusammengefasst in Tabelle 5.18, findet sich für fachmethodisch-indizierte Beiträge auch *in der Zusammenfassung aller drei Einheiten* für die acht Personen, die alle drei Einheiten und sowohl Prä- als auch Post-Test zu den Kompetenzen zum experimentbezogenen Denken und Arbeiten

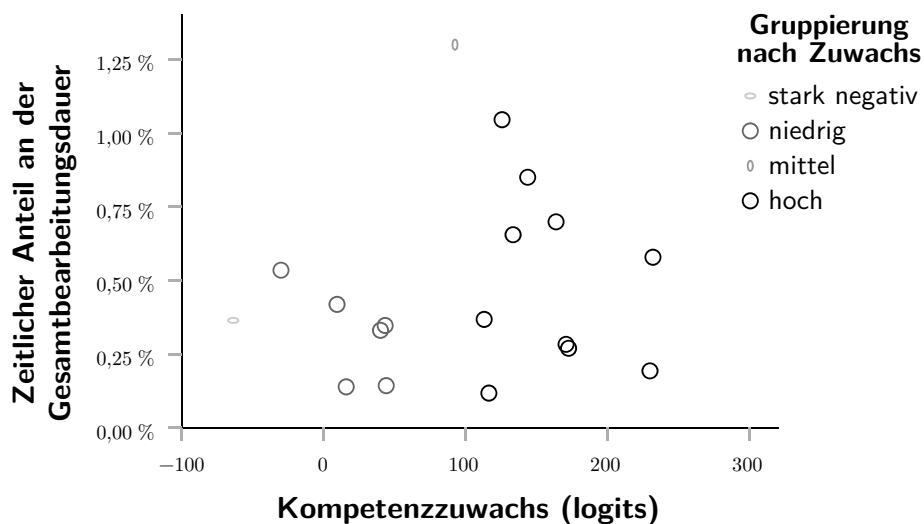


Abbildung 5.35: Streudiagramm für den zeitlichen Anteil fachmethodisch-indizierter Beiträge bei Einheit 2 und für Kompetenzzuwächse der Personen.

Tabelle 5.18: Übersicht über Ergebnisse zu ausgewählten Einheiten zum Zusammenhang von dem zeitlichen Anteil fachmethodischer Beiträge mit den Kompetenzzuwächsen für die einzelnen Personen, die die explizit-fachmethodische Instruktionsvariante bearbeitet haben.

Zeitlicher Anteil fachmethodischer Beiträge	Kompetenzzuwachs	
	Niedrig	Hoch
größer	Nicht vorhanden	vorhanden
geringer	vorhanden	vorhanden

Anmerkungen. Die Tabelle ist eine schematische Übersicht über die Streudiagramme. (Nicht) Vorhanden = Es gibt (keine bzw. nur sehr wenige) Personen, für die die Zeilen- und die Spaltenüberschrift zutreffen.

bearbeitet haben. Außerdem findet sie sich für fachmethodisch-*vermutete* Beiträge bei *Einheit 2* (nicht ganz so deutlich) und für fachmethodisch-*indizierte* Beiträge bei *Einheit 3*. Für *Einheit 1* lässt sich diese Art von Zusammenhang nur in Ansätzen ausmachen (siehe Abbildung 5.36).

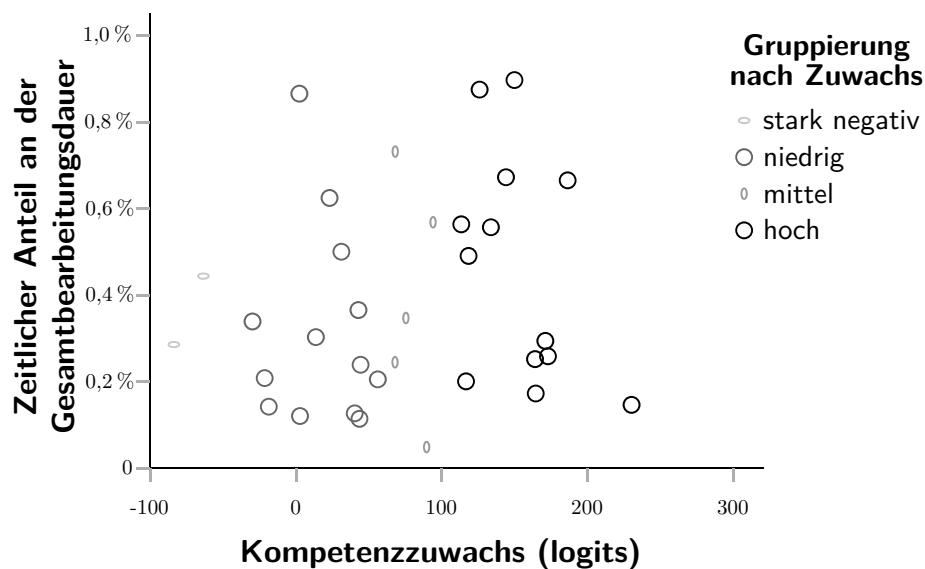


Abbildung 5.36: Streudiagramm für den zeitlichen Anteil fachmethodisch-indizierter Beiträge bei Einheit 1 und für Kompetenzzuwächse der Personen.

Der Zusammenhang fachmethodisch-vermuteter Beiträge mit den Kompetenzzuwächsen der einzelnen Personen stellt sich für Einheit 3 etwas anders dar, weil Personen mit hohen Kompetenzzuwächsen abweichend von dem in Tabelle 5.18 zusammengefassten Zusammenhang grundsätzlich hohe Anteile fachmethodisch-vermuteter Beiträge haben und somit quasi nur die Diagonale von Tabelle 5.18 gefüllt ist. Aus dem Streudiagramm in Abbildung 5.37 wird deutlich, dass diese Deutung nur näherungsweise gilt, weil es einzelne Personen gibt, die Ausreißer darstellen (niedrige Kompetenzzuwächse mit hohen zeitlichen Anteilen fachmethodischer Beiträge sowie hohe Kompetenzzuwächse mit niedrigen zeitlichen Anteilen fachmethodischer Beiträge). Dennoch lässt sich festhalten, dass der positive Zusammenhang fachmethodisch-vermuteter Beiträge mit den Kompetenzzuwächsen innerhalb von Einheit 3 *deutlicher* ist als der in Tabelle 5.18 (auf S. 179) dargestellte.

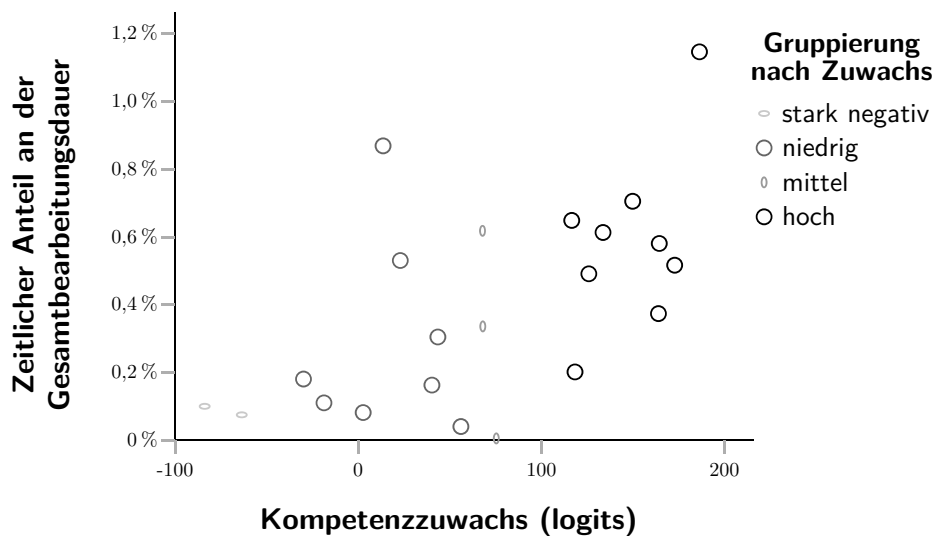


Abbildung 5.37: Streudiagramm für den zeitlichen Anteil fachmethodisch-vermuteter Beiträge bei Einheit 3 und für Kompetenzzuwächse der Personen.

5.4.4.2 Vergleich der Gesamtdauer fachmethodischer Beiträge des gesamten Teams von Personen mit unterschiedlichen Kompetenzzuwächsen

Wird danach gefragt, ob Personen mit hohen Zuwächsen eher in Teams arbeiten, die insgesamt viel über Fachmethoden sprechen, ergeben sich vergleichbare Ergebnisse zum Vergleich der Personenaktivitäten.

Der Vergleich für die Zusammenfassung der drei Einheiten ist stark eingeschränkt, unterstreicht aber zumindest die Bedeutung fachmethodisch-indizierter Beiträge. Da die drei Personen mit niedrigen Kompetenzzuwächsen aus einem Team stammen, haben sie denselben Wert für die Team-Gesamtdauern der Aktivitäten (jeweils die Summe der Aktivitäten-Dauern aller drei Personen). Dies schränkt den statistischen Vergleich deutlich ein. Die in Tabelle 5.19 berichteten Ergebnisse lassen sich daher nur als grober erster Hinweis auf die Bedeutung fachmethodisch-indizierter Beiträge auch auf Team-Ebene ansehen.

Die nach Einheiten aufgelöste Betrachtung der Aktivitäten der Teams von Personen mit unterschiedlichen Kompetenzzuwächsen liefert Ergebnisse, die denen aus dem Vergleich der Aktivitäten von Personen in ihrer Deutung recht ähnlich, aber weniger deutlich sind. Die zum Vergleich der zeitlichen Anteile fachmethodischer Beiträge der Teams, in denen Personen mit unterschiedlichen Kompetenzzuwächsen jeweils arbeiten, durchgeführten Welch-Tests sind in Tabelle 5.20 dokumentiert. Die jeweils größten Effekte, die die niedrigsten p -Werte aufweisen, finden sich bei den Einheiten 1 und 2 für fachmethodisch-indizierte Beiträge, bei Einheit 3 für fachmethodisch-vermutete Beiträge. Die Unterschiede sind weniger deutlich als im Vergleich der Personenaktivitäten (Effektstärken und relative Unterschiede sind kleiner; absolute Unterschiede können aufgrund der Addition

Tabelle 5.19: Students t-Tests für die Gesamtdauer fachmethodischer Beiträge im Team von Personen mit hohen vs. niedrigen Kompetenzzuwächsen in der Gesamtschau über alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante hinweg.

Vergleich	Zuwachs Niedrig			Zuwachs Hoch			df	$t(df)$	p	r	Unterschied	
	N_N	MW	SD	N_H	MW	SD					Zeit in s	Relativ
FMi	3	0,058573	0*	5	0,102347	0,032443	6,000	-2,263	0,064	0,679	590,95	42,8%
FMv		0,029306	0*		0,022341	0,005036	6,000	2,319	0,059	-0,688	-94,03	-23,8%
FM		0,087879	0*		0,124688	0,037199	6,000	-1,659	0,148	0,561	496,92	29,5%

Anmerkungen. Students t-Test; $N_{H/N}$ = Stichprobengröße für Gruppe der Personen mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Zeit in s = Zeitliches Äquivalent der Mittelwertdifferenz (Anteil der Aktivität mal angesetzte Dauer der jeweiligen Einheit in Sekunden); Unterschied Relativ = Mittelwertdifferenz geteilt durch größeren der beiden Mittelwerte; * = Nur ein Team, daher nur ein Wert.

Tabelle 5.20: Welch-Tests für die Gesamtdauer fachmethodischer Beiträge im Team von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante.

Vergleich		Zuwachs Niedrig			Zuwachs Hoch			Korr. df	<i>t</i> (df)	<i>p</i>	<i>r</i>	Unterschied	
Einheit	Aktivität	<i>N_N</i>	MW	SD	<i>N_H</i>	MW	SD					Zeit in s	Relativ
1	FMi	8	0,099018	0,041433	12	0,138656	0,029072	11,557	-2,348	0,038	0,568	107,02	28,6%
	FMv		0,016869	0,008793		0,016891	0,009907	16,391	-0,005	0,996	0,001	0,06	0,1%
	FM		0,115887	0,046465		0,155547	0,031641	11,309	-2,110	0,058	0,531	107,08	25,5%
2	FMi	6	0,070452	0,027733	11	0,087829	0,017868	7,332	-1,386	0,206	0,456	93,84	19,8%
	FMv		0,045877	0,017159		0,041762	0,008803	6,472	0,549	0,601	-0,211	-22,22	-9,0%
	FM		0,116329	0,033606		0,129591	0,025468	8,220	-0,844	0,423	0,282	71,62	10,2%
3	FMi	9	0,085319	0,038492	10	0,097281	0,039698	16,890	-0,666	0,514	0,160	64,59	12,3%
	FMv		0,010253	0,005430		0,012935	0,003924	14,449	-1,223	0,241	0,306	14,49	20,7%
	FM		0,095572	0,041316		0,110216	0,039953	16,650	-0,784	0,444	0,189	79,08	13,3%

Anmerkungen. Welch-Test; $N_{H/N}$ = Stichprobengröße für Gruppe der Teams mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Zeit in s = Zeitliches Äquivalent der Mittelwertdifferenz (Mittelwertdifferenz der Anteile der Aktivität mal angesetzte Dauer der jeweiligen Einheit in Sekunden); Unterschied Relativ = Mittelwertdifferenz geteilt durch größeren der beiden Mittelwerte.

der Werte dreier Personen für einen Vergleich durch drei geteilt werden und sind dann ebenfalls deutlich kleiner als im Vergleich der Personenaktivitäten). Dass sich im Gegensatz zum Vergleich der Personenaktivitäten für den Vergleich von Personen mit unterschiedlich hohen Kompetenzzuwächsen hinsichtlich der zeitlichen Anteile fachmethodisch-vermuteter Beiträge auf Ebene des gesamten Teams in der zweiten Einheit ein Zusammenhang einstellt, der in der Tendenz negativ ist, wird vor dem Hintergrund des geringen absoluten und relativen Effekts und des hohen *p*-Wertes als nicht-interpretierbares Ergebnis angesehen. Es bleibt also zunächst offen, ob ein derartiger negativer Zusammenhang nur ein Artefakt ist. Zu Einheit 3 ist anzumerken, dass die Ergebnisse für die Teamaktivitäten zwar in den Tendenzen den Ergebnissen aus den Personenvergleichen entsprechen, aber im Vergleich zu den anderen Einheiten am wenigsten stark ausgeprägt sind, obwohl dies auf Personenebene (hinsichtlich der Effektstärken und relativen Unterschiede) nicht der Fall war. Zeitliche Anteile fachmethodischer Beiträge des Teams, in dem eine Person arbeitet, stehen also grundsätzlich, und insbesondere bei Einheit 3, in einem nicht so deutlichen Zusammenhang mit dem Kompetenzzuwachs der Person wie der zeitliche Anteil der fachmethodischen Beiträge der Person selbst.

Die nach Einheiten aufgelöste Betrachtung der Aktivitäten der Teams von Personen mit unterschiedlichen Kompetenzzuwächsen liefert Ergebnisse, die denen aus dem Vergleich der Aktivitäten von Personen in ihrer Deutung recht ähnlich sind. Neben der in der Gesamtschau über alle drei Einheiten nicht zu vernachlässigenden Bedeutung fachmethodisch-indizierter Beiträge, wird erneut deutlich, dass in der zuletzt

bearbeiteten Einheit 3 fachmethodisch-vermutete Beiträge im Zusammenhan für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten haben. Dass die Effekte in den Analysen zu den Aktivitäten des gesamten Teams von Personen mit unterschiedlichen Kompetenzzuwächsen besonders für die Kategorie **Fachmethodisch-Indiziert** positiv sind, lässt sich eventuell so deuten, dass bei fachmethodisch-indizierten Beiträgen auch Äußerungen von anderen Lernenden, die mit der betrachteten Person zusammenarbeiten, einen Beitrag zum Kompetenzaufbau leisten. Dies scheint besonders in Einheit 1 der Fall zu sein, in der Vorbereitung und Durchführung von realen Experimenten eine geringere Rolle spielt. Fachmethodisch-vermutete Beiträge des gesamten Teams stehen hingegen für die Einheiten 1 und 2 in keinem Zusammenhang mit den Kompetenzzuwächsen. In Einheit 3 könnte hingegen – wie weiter oben schon erwähnt – die Verknüpfung der erlernten fachmethodischen Konzepte mit konkreten Experimentiersituationen bedeutsam sein und sich, sowohl auf Team- als auch auf Personen-Ebene, vorrangig in fachmethodisch-vermuteten Beiträgen widerspiegeln. Dieser Deutung entsprechend wäre es auch plausibel, dass der Effekt für die Einheiten auf Team-Ebene jeweils nicht so deutlich ist wie auf Personen-Ebene, weil zum einen die Teams hinsichtlich der Kompetenzzuwächse heterogen sind und zum anderen die Verknüpfung von Konzepten und konkreten Situationen eher eigenständig (als von irgendjemand im Team der betrachteten Person) geleistet werden muss.

5.4.5 Fachmethodische Beiträge für spezifische Ausschnitte der Instruktion (Abschnitte und Karten) und Kompetenzzuwächse

Eine weitere Näherung an die Bedeutung fachmethodischer Beiträge im Kompetenzaufbau wird so vorgenommen, dass das *Auftreten* fachmethodischer Beiträge *zeitlich eingegrenzt* wird. Die *Abschnitte* der Instruktion sind mit verschiedenen adressierten Konzepten assoziiert (z. B. ein Konzept dazu, was Beobachtungen von Deutungen unterscheidet in Einheit 3), die *Karten* der Instruktion mit konkreten Situationen des experimentbezogenen Denkens und Arbeitens (z. B. Karte 3-05 mit dem Fallenlassen eines Lineals und dem Notieren von Messwerten zur Reaktionszeit). Wie groß der zeitliche Anteil fachmethodischer Beiträge in Bezug auf Bearbeitungsprozesse zu spezifischen Konzepten bzw. zu konkreten Situationen ist, kann Einblicke in die Relevanz fachmethodischer Beiträge und Überlegungen liefern. Daher wird nachfolgend für die Abschnitte sowie für die Karten der Instruktion nach Unterschieden und Gemeinsamkeiten für Personen mit unterschiedlichen Kompetenzzuwächsen hinsichtlich der Kategorien Fachmethodisch-Indiziert und Fachmethodisch-Vermutet gesucht.

5.4.5.1 Fachmethodische Beiträge und inhaltliche Abschnitte der Instruktion

Um für Personen mit unterschiedlichen Kompetenzzuwächsen die zeitlichen Anteile fachmethodischer Beiträge zu vergleichen, wird für jede Person der absolute zeitliche Umfang der Abschnittsbearbeitungsdauer für jeden der Abschnitte der explizit-fachmethodischen Instruktionsvariante berechnet (als Summe der Kartenbearbeitungsdauern, siehe Unterabschnitt 5.4.1 ab S. 106) sowie der relative zeitliche Anteil fachmethodischer (bzw. weiter unten aufgeteilt: fachmethodisch-indizierter sowie = vermuteter) Beiträge, die innerhalb des entsprechenden Abschnittsintervalls auftauchen, an der jeweiligen Abschnittsbearbeitungsdauer berechnet. Um stabil interpretierbare Werte zu erhalten, ergeben sich die in den Tabellen 5.21, 5.22 und 5.23 (ab S. 186) angegebenen Absolutwerte (vorletzte Spalte) für die Unterschiede zwischen den Gruppen mit hohen vs. niedrigen Kompetenzzuwächsen jeweils als Produkt von der mittleren Abschnittsbearbeitungsdauer aller Personen mit entsprechenden Kompetenzzuwächsen und von dem mittleren relativen Anteil der entsprechenden Kategorie fachmethodischer Beiträge; sie stellen also zeitliche Äquivalente der relativen Anteile dar (und nicht reale zeitliche Mittelwerte). Basierend auf grafischen Darstellungen der zeitlichen Anteile (zusammenfassend Abbildung 5.38) und Welch-Tests, in denen die Gruppen mit hohen vs. niedrigen Kompetenzzuwächsen gegeneinander getestet werden, finden sich die folgenden Ergebnisse:

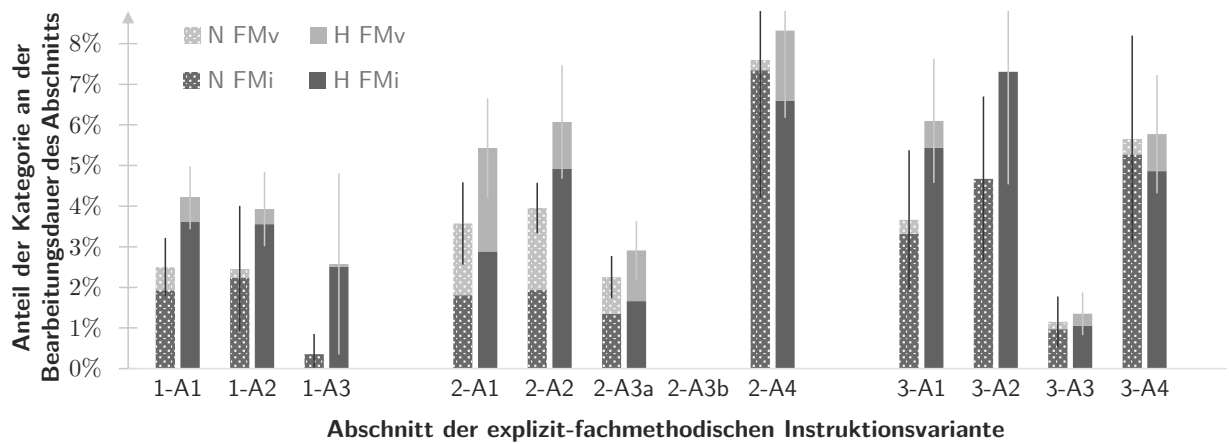


Abbildung 5.38: Mittelwerte der relativen Anteile der fachmethodisch-indizierten (FMi) und -vermuteten (FMv) Beiträge an den jeweiligen Abschnittsbearbeitungsdauern für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante für Personen mit hohen (H) vs. niedrigen (N) Kompetenzzuwächsen. Fehlerbalken entsprechen einem Standardfehler des Mittelwerts für fachmethodische Beiträge (also für den Mittelwert über die Summe der beiden dargestellten Kategorien).

In allen Abschnitten finden sich in der Tendenz für Personen mit hohen Kompetenzzuwächsen größere relative zeitliche Anteile fachmethodischer Beiträge als für Personen mit niedrigen Kompetenzzuwächsen. Dieses Ergebnis gilt zunächst bei Zusammenfassung der beiden Kategorien fachmethodisch-indizierter und fachmethodisch-vermuteter Beiträge (also bei Betrachtung der Gesamthöhe der Balken in Abbildung 5.38). Es gilt mit kleinen Einschränkungen auch bei der isolierten Betrachtung fachmethodisch-indizierter bzw. fachmethodisch-vermuteter Beiträge: Nur für die Abschnitte 2-A4 und 3-A4 finden sich tendenziell (aber nicht signifikant und auch nur mit sehr kleinen Effektstärken) größere zeitliche Anteile fachmethodisch-indizierter Beiträge für Personen mit niedrigen Kompetenzzuwächsen (Details in Tabelle 5.22). Nur für den Abschnitt 2-A2 finden sich tendenziell ($p = .192$; $r = .441$) größere zeitliche Anteile fachmethodisch-vermuteter Beiträge für Personen mit niedrigen Kompetenzzuwächsen (Details in Tabelle 5.23 auf S. 187).

Die Muster der relativen Anteile fachmethodischer Beiträge über die Abschnitte hinweg sind für Personen mit hohen und Personen mit niedrigen Kompetenzzuwächsen sehr ähnlich. Für Abschnitte, bei denen für Personen mit hohen Kompetenzzuwächsen fachmethodische Beiträge (insgesamt oder aus einer der beiden Kategorien) im Vergleich zu anderen Abschnitten größere relative zeitliche Anteile einnehmen, ist dies grundsätzlich auch für Personen mit niedrigen Kom-

5 Aktivitäten von Lernenden

petenzzuwächsen der Fall. Ausnahmen stellen 2-A2 und 2-A4 für fachmethodisch-indizierte sowie 2-A2 und 3-A4 für fachmethodisch-vermutete Beiträge dar.

Tabelle 5.21: Welch-Tests für die Dauer fachmethodischer Beiträge (insgesamt als Summe der beiden Subkategorien) von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für jeden Abschnitt der drei Einheiten der explizit-fachmethodischen Instruktionsvariante.

Vergleich		Zuwachs Niedrig			Zuwachs Hoch			korr.df	<i>t</i> (df)	<i>p</i>	<i>r</i>	Unterschied	
Einheit	Abschnitt	<i>N_N</i>	MW	SD	<i>N_H</i>	MW	SD					Absolut	Relativ
1	1	8	0,02500	0,02029	12	0,04207	0,02682	17,605	-1,616	0,1238	0,359	27,42	40,6%
	2	8	0,05837	0,04383	12	0,07678	0,03179	11,844	-1,022	0,3271	0,285	11,69	24,0%
	3	8	0,01181	0,01396	12	0,06639	0,07731	12,057	-2,388	0,0342	0,567	9,63	82,2%
2	1	6	0,03573	0,02483	11	0,05430	0,04055	14,638	-1,170	0,2608	0,292	16,89	34,2%
	2	6	0,03953	0,01520	11	0,06071	0,04631	13,304	-1,387	0,1884	0,355	17,02	34,9%
	3a	6	0,02253	0,01277	11	0,02912	0,02390	14,993	-0,741	0,47	0,188	12,38	22,6%
	3b	6	0	0	11	0	0	0	0	0	0	0	0,0%
3	4	6	0,07602	0,08318	11	0,08319	0,07117	9,068	-0,179	0,862	0,059	2,93	8,6%
	1	9	0,03661	0,05147	10	0,06099	0,04835	16,502	-1,061	0,304	0,253	30,69	40,0%
	2	9	0,04674	0,06075	10	0,07309	0,08774	16,027	-0,767	0,4543	0,188	7,61	36,0%
	3	9	0,01151	0,01872	10	0,01349	0,01662	16,151	-0,243	0,8111	0,060	3,31	14,7%
	4	9	0,05651	0,07649	10	0,05771	0,04602	12,849	-0,041	0,9681	0,011	0,72	2,1%

Anmerkungen. Welch-Test (korr.df = korrigierte Freiheitsgrade); $N_{H/N}$ =Stichprobengröße für Gruppe der Personen mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Absolutwertäquivalent zur Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Tabelle 5.22: Welch-Tests für die Dauer fachmethodisch-indizierter Beiträge von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für jeden Abschnitt der drei Einheiten der explizit-fachmethodischen Instruktionsvariante.

Vergleich		Zuwachs Niedrig			Zuwachs Hoch			korr.df	<i>t</i> (df)	<i>p</i>	<i>r</i>	Unterschied	
Einheit	Abschnitt	<i>N_N</i>	MW	SD	<i>N_H</i>	MW	SD					Absolut	Relativ
1	1	8	0,01926	0,01703	12	0,03612	0,02454	17,930	-1,814	0,0865	0,394	27,09	46,7%
	2	8	0,02233	0,01452	12	0,03554	0,01786	17,140	-1,816	0,087	0,402	8,39	37,2%
	3	8	0,00358	0,00477	12	0,02501	0,03053	11,794	-2,388	0,0346	0,571	3,78	85,7%
2	1	6	0,01805	0,006763	11	0,028854	0,026469	12,188	-1,279	0,2246	0,344	9,82	37,4%
	2	6	0,019364	0,0148	11	0,049244	0,042148	13,638	-2,123	0,0525	0,498	24,00	60,7%
	3a	6	0,013512	0,008006	11	0,01664	0,01562	14,995	-0,546	0,5933	0,140	5,87	18,8%
	3b	6	0	0	11	0	0	0	0	0	0	0	0,0%
3	4	6	0,073473	0,085507	11	0,066005	0,062927	8,040	0,188	0,8556	-0,066	-3,06	-10,2%
	1	9	0,033236	0,050583	10	0,054399	0,048009	16,557	-0,933	0,3644	0,223	26,64	38,9%
	2	9	0,046744	0,060755	10	0,073089	0,087736	16,027	-0,767	0,4543	0,188	7,61	36,0%
	3	9	0,009686	0,015049	10	0,010503	0,014834	16,734	-0,119	0,9067	0,029	1,37	7,8%
	4	9	0,052736	0,069338	10	0,048612	0,045431	13,575	0,152	0,8818	-0,041	-2,47	-7,8%

Anmerkungen. Welch-Test (korr.df = korrigierte Freiheitsgrade); $N_{H/N}$ =Stichprobengröße für Gruppe der Personen mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Absolutwertäquivalent zur Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Tabelle 5.23: Welch-Tests für die Dauer fachmethodisch-*vermuteter* Beiträge von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für jeden Abschnitt der drei Einheiten der explizit-fachmethodischen Instruktionsvariante.

Vergleich		Zuwachs Niedrig			Zuwachs Hoch			korr.df	$t(df)$	p	r	Unterschied	
Einheit	Abschnitt	N_N	MW	SD	N_H	MW	SD					Absolut	Relativ
1	1	8	0,00574	0,00596	12	0,00594	0,00773	17,500	-0,066	0,9482	0,016	0,32	3,4%
	2	8	0,00222	0,00276	12	0,00376	0,00374	17,729	-1,059	0,304	0,244	0,98	40,9%
	3	8	0,00000	0,00000	12	0,00075	0,00202	11,000	-1,293	0,2225	0,363	0,13	100,0%
2	1	6	0,01768	0,02456	11	0,02545	0,01997	8,691	-0,665	0,5236	0,220	7,07	30,5%
	2	6	0,02016	0,01305	11	0,01147	0,01005	8,319	1,418	0,1924	-0,441	-6,98	-43,1%
	3a	6	0,00902	0,00576	11	0,01248	0,00990	14,836	-0,912	0,3763	0,230	6,51	27,8%
	3b	6	0	0	11	0	0	0	0	0	0	0,00	0,0%
	4	6	0,00254	0,00419	11	0,01718	0,02420	11,064	-1,954	0,0765	0,506	5,99	85,2%
3	1	9	0,00337	0,00261	10	0,00659	0,00420	15,214	-2,027	0,0605	0,461	4,05	48,8%
	2	9	0	0	10	0	0	0	0	0	0	0,00	0,0%
	3	9	0,00183	0,00385	10	0,00299	0,00328	15,840	-0,706	0,4904	0,175	1,95	38,9%
	4	9	0,00377	0,00753	10	0,00910	0,01429	13,917	-1,029	0,3208	0,266	3,18	58,5%

Anmerkungen. Welch-Test (korr. df = korrigierte Freiheitsgrade); $N_{H/N}$ =Stichprobengröße für Gruppe der Personen mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Absolutwertäquivalent zur Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Unterschiede finden sich im Wesentlichen für einen Abschnitt: Zu kontrollierten und unkontrollierten Versuchen. Aus den beiden bisher dargestellten Ergebnissen geht unter anderem hervor, dass es nur wenige Abschnitte gibt, für die Personen mit hohen Kompetenzzuwächsen sich – vor dem Hintergrund des Gesamtbildes über alle Abschnitte – grundlegend von Personen mit niedrigen Kompetenzzuwächsen unterscheiden. Wird zusätzlich beachtet, dass es sich im Fall von 2-A4 und 3-A4 jeweils um Unterschiede handelt, die wenigen Sekunden mehr oder weniger Sprechen über Fachmethoden bedeuten und zudem statistisch unsicher sind (vgl. Signifikanz und Effektstärke), bleibt ein zentraler Unterschied übrig: für den Abschnitt 2-A2. Dieser Abschnitt betrifft Aufgaben zum Unterscheiden von kontrollierten und unkontrollierten (fairen und unfairen) Versuchsdesigns.

Die Unterschiede zwischen Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen scheinen innerhalb jeder Einheit für fachmethodisch-indizierte Beiträge mit fortschreitender Zeit abzunehmen und für fachmethodisch-vermutete Beiträge zuzunehmen. Eine statistisch aufgrund der Streuungen nicht optimal absicherbare Beobachtung zu Abbildung 5.38 lautet, dass sich die Personengruppen mit hohen vs. niedrigen Kompetenzzuwächsen im Mittel zu Beginn der Einheiten mehr und dann immer weniger bezüglich der relativen zeitlichen Anteile fachmethodisch-indizierter Beiträge unterscheiden. Für fachmethodisch-vermutete

Beiträge scheint dies umgekehrt zu sein. Diese gegenläufigen Trends heben sich in etwa auf, so dass das zweite (weiter oben berichtete) Ergebnis entsteht, dass das Muster über die Abschnitte hinweg recht unabhängig vom Kompetenzzuwachs ist.

5.4.5.2 Fachmethodische Beiträge und konkrete Situationen (Karten der Instruktion)

Der Frage nach dem Auftreten von fachmethodischen Beiträgen kann auch nachgegangen werden, indem es zeitlich noch spezifischer eingegrenzt wird und die *einzelnen Karten* betrachtet werden. Wie bereits für die Abschnitte werden dafür Welch-Tests für jede Karte durchgeführt, die Personen mit hohen Kompetenzzuwächsen und Personen mit niedrigen Kompetenzzuwächsen hinsichtlich der Kategorien zu den fachmethodischen Beiträgen kontrastieren. Signifikante Unterschiede mit großen Effektstärken finden sich bei 15 Karten, meist nur für fachmethodisch-indizierte oder für fachmethodisch-vermutete Beiträge, teilweise aber auch für beide Aktivitäten. Eine Liste der Karten mit einer kurzen inhaltlichen Beschreibung sowie der Tendenz im Unterschied findet sich in Tabelle 5.24.

Für die meisten Karten überwiegen die zeitlichen Anteile einer der fachmethodischen Kategorien für Personen mit hohen Kompetenzzuwächsen gegenüber denen der Personen mit niedrigen Kompetenzzuwächsen. Viele der Karten mit Unterschieden zwischen Personen mit unterschiedlichen Kompetenzzuwächsen finden sich in Einheit 2. Die meisten dieser Karte und die Karte in Einheit 3 sind mit der Einschätzung oder Planung von Versuchen verbunden. Eine Verbindung zur Variablenkontrollstrategie, insbesondere zur Zuweisung von unabhängigen und abhängigen Variablen sowie Kontrollvariablen, lässt sich vermuten. Interessanterweise finden sich für die Karte 2-07, die vor der Einführung dieser Begriffe liegt, mehr fachmethodische Beiträge (indiziert) für Personen mit niedrigen Kompetenzzuwächsen.

Tabelle 5.24: Übersicht über die Karten der explizit-fachmethodischen Instruktionsvariante, für die sich signifikante Unterschiede mit großer Effektstärke beim Vergleich von Personen mit unterschiedlichen Kompetenzzuwächsen hinsichtlich einer der beiden Kategorien zu fachmethodischen Beiträgen ergeben.

Karte	FMi	FMv	Kurzbeschreibung
1-14	H > N		Entscheiden, ob Frage präzise oder allgemein ist. Versuch kurz durchführen.
1-32	H > N		Aus Beispielen ableiten, was Vermutung von Hypothese unterscheidet.
2-07	N > H		Versuchsplanung notieren.
2-12	H > N		Kontrollvariablen diskutieren, Frage auf Naturwissenschaftlichkeit einschätzen.
2-13	H > N		Deckblatt: Faire und unfaire Versuche erkennen.
2-16	H > N		Versuchsaufbau variablenkontrolliert auswählen.
2-17	H > N		Diskutieren, warum andere Versuche nicht verglichen werden können.
2-19	H > N		Versuchsaufbau variablenkontrolliert zu einer Frage passend auswählen.
2-20	H > N		Diskutieren, warum andere Versuche nicht zur Frage passen.
2-21		N > H	Passende Frage zu Versuch formulieren.
2-27		H > N	Kurz diskutieren, wie einer Frage mit Versuch nachgegangen werden kann.
2-32		H > N	Unabhängige, abhängige und Kontrollvariablen für Versuch in Grafik eintragen.
2-37	H > N	H > N	Fiktivem Team erklären, warum ihr Versuch nicht fair ist.
2-51		H > N	Diskutieren, warum nicht alle Variablen Kontrollvariablen sind.
3-48	H > N		Diskutieren, ob der gegebene Versuch unter vollständiger Variablenkontrolle möglich ist.

Anmerkungen. H > N steht dafür, dass Personen mit hohen Kompetenzzuwächsen signifikant höhere zeitliche Anteile der jeweiligen Kategorie aufweisen als Personen mit niedrigen Kompetenzzuwächsen (N > H analog andersherum).

FMi = Fachmethodisch-Indiziert, FMv = Fachmethodisch-Vermutet.

5.4.6 Fachmethodische Beiträge im Kontext von Experimentiersituationen und Kompetenzzuwächse

Die in Unterabschnitt 5.4.3 (beginnend auf S. 163) formulierte Deutung, dass die Verknüpfung von fachmethodischen Beiträgen mit Versuchshandlungen als relevant für Prozesse des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten angesehen werden kann, legt nahe, diese Verknüpfung genauer zu betrachten. Die Verknüpfung wurde zwar insbesondere an der Kategorie **Fachmethodisch-Vermutet** festgemacht, dennoch liegt es – auch vor dem Hintergrund der in den vorherigen Unterabschnitten herausgearbeiteten Bedeutung fachmethodisch-indizierter Beiträge – für eine erste Annäherung an den Gegenstand nahe, auch Beiträge der Kategorie **Fachmethodisch-Indiziert** als mit Versuchshandlungen verknüpft anzusehen, wenn diese in demselben Kontext wie Versuchshandlungen auftreten.⁷¹ Nachfolgend werden daher Lernende mit hohen vs. niedrigen Kompetenzzuwächsen daraufhin verglichen, ob und inwiefern sich die Anteile fachmethodischer Beiträge der Kategorien **Fachmethodisch-Indiziert** und **Fachmethodisch-Vermutet** für Situationen der Vorbereitung und Durchführung von Versuchen unterscheiden.

5.4.6.1 Wahl der Situationen und Vorgehen beim Vergleich

In einer Annäherung an den Gegenstand lassen sich Situationen der Vorbereitung und Durchführung von Versuchen mit einigen der eingesetzten Karten der Instruktion in Verbindung bringen. Daher werden die Aktivitäten der Lernenden für die folgenden Karten genauer untersucht:

- 1-04 Die Karte enthält eine Versuchsanleitung zum Pusten in Trinkhalme, die zylinderförmige Becken mit unterschiedlicher Wasserhöhe, aber gleichen Durchmesser gehalten werden. Zu dieser Anleitung soll zunächst eine passende Fragestellung formuliert werden. Danach soll der Versuch durchgeführt werden und geprüft werden, ob die formulierte Frage sich beantworten lässt.
- 1-11 Die Karte enthält eine Frage dazu, was beim Pusten in Trinkhalme passiert, wenn diese in zylinderförmige Becken gleichen Durchmessers gehalten werden, die unterschiedlich hoch mit Wasser gefüllt sind. Dazu soll eine Vermutung ausgewählt, begründet und anschließend in einem Versuch überprüft werden.

⁷¹Eine Überprüfung der spezifischen Inhalte der einzelnen Beiträge wäre anhand der generierten Daten möglich, erfolgt aber in der vorliegenden Arbeit nicht.

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

- 1-14 In einem selbst zu planenden Versuch soll die Frage untersucht werden, welchen Einfluss der Durchmesser des Trinkhalms darauf hat, in welchem zylinderförmigen Becken zuerst Luftblasen aufsteigen. (Implizit wird durch eine Zeichnung Variablenkontrolle nahegelegt.) Vor der Durchführung ist zu entscheiden, ob die Frage präzise oder allgemein ist (Kategorien vorher im Material eingeführt).
- 2-33 Es soll ein Versuch geplant werden, mit dem ohne Aufnehmen von Messwerten untersucht werden kann, ob der Durchmesser eines Papierkegels (in der Materialsammlung für die Schülerinnen und Schüler vorhanden, vorher in der Instruktion bereits thematisiert) einen Einfluss auf die durchschnittliche Fallgeschwindigkeit in Luft hat. Danach soll der Versuch durchgeführt und die Frage mit ja oder nein beantwortet werden.
- 2-34 Die Karte enthält den Hinweis, dass für den Versuch von Karte 2-33 die Masse konstant gehalten werden muss. Es soll überprüft werden, ob das in der eigenen Versuchsdurchführung geschehen ist. Andernfalls soll der Versuch modifiziert und erneut durchgeführt werden.
- 2-39 Ein Versuch wird beschrieben, bei dem durch Ineinanderstecken mehrerer Papierkegel nach und nach die Masse des fallenden Objekts erhöht wird. eine passende Frage soll formuliert und nach Durchführung des Versuchs eine Antwort auf diese Frage gegeben werden.
- 3-05 Unabhängige und abhängige Variable sind gegeben. Als Frage ist formuliert, ob alle Teammitglieder ein Lineal, das fallen gelassen wird, nach gleicher Strecke fangen. Messwerte sollen notiert werden. Im Anschluss wird aufgefordert, zu prüfen, ob die Variablenkontrolle eingehalten wurde.
- 3-34f (Steht für die Karten 3-34 und 3-35.) Auf der Karte 3-34 wird dazu aufgefordert, einen auf den vorigen Karten angeleitet aufgebauten Versuch durchzuführen. Bei dem Versuch geht es um die Wurfweite eines Wasserstrahls in Abhängigkeit vom Abwurfwinkel (Frage auf Karte 3-32 in dieser Weise formuliert). Auf der Karte 3-35 ist Raum zum Notieren der Messwerte. Da dieses Notieren parallel zum Durchführen geschehen soll (und dabei Karte 3-35 kodiert wurde), werden die Karten zusammengefasst. Zu bedenken ist allerdings, dass sich auf der Karte 3-35 auch eine Aufforderung findet, eine Antwort auf die Forschungsfrage zu formulieren.

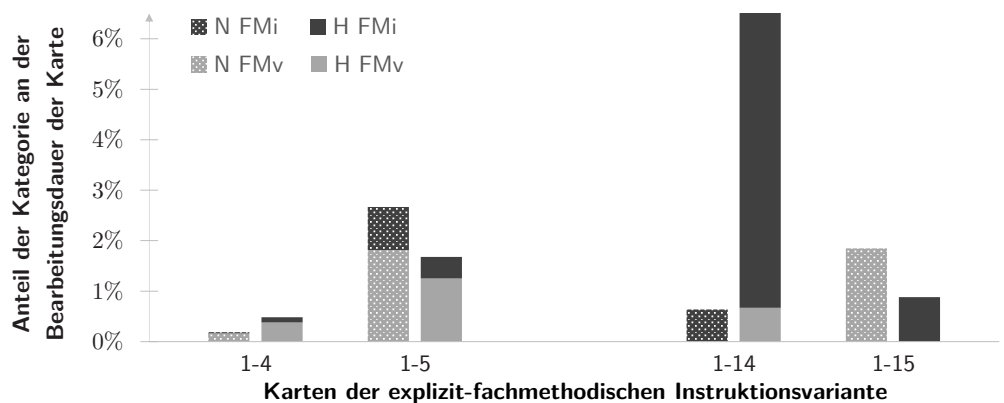
Für jede der genannten Karten wird jeweils eine Karte vorher und eine Karte danach mitbetrachtet, um mögliche Verschiebungen von Überlegungen der Vor- oder Nachbereitung auf andere Karten zu berücksichtigen. Für die Karten 1-10 (vor 1-11), 1-13 (vor 1-14) und 2-38 (vor 2-39) liegt beim Übergang von den hinzugenommenen Karten zu den Versuchssituationen ein eher deutlicher inhaltlicher Wechsel vor. (Beispielsweise wird auf 1-10 diskutiert, warum es nicht immer möglich ist, dass naturwissenschaftliche Fragen präzise formuliert werden, bevor auf 1-11 eine neue Situation vorgestellt wird, zu der eine Vermutung formuliert und ein Versuch durchgeführt werden soll.) Diese Karten werden nachfolgend nicht in den Diagrammen dargestellt, um keine Fehlinterpretationen nahezulegen; sie wurden allerdings zur Absicherung der Interpretationen mitbetrachtet und stehen diesen nicht entgegen.

Für jede der analysierten Karten werden die zeitlichen Anteile fachmethodischer Beiträge an der Gesamtbearbeitungsdauer der Karte dargestellt. Diese werden für die beiden Kategorien **Fachmethodisch-Indiziert** und **Fachmethodisch-Vermutet** getrennt berichtet, weil die jeweils spezifische Bedeutung dieser Kategorien untersucht werden soll. Da ein Vergleich von Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen erfolgen soll, werden die Anteile zudem einzeln für die Personen mit hohen sowie die Personen mit niedrigen Kompetenzzuwächsen dargestellt. Die Vergleiche von Lernenden mit hohen und niedrigen Kompetenzzuwächsen erfolgen dann basierend auf den Diagrammen.

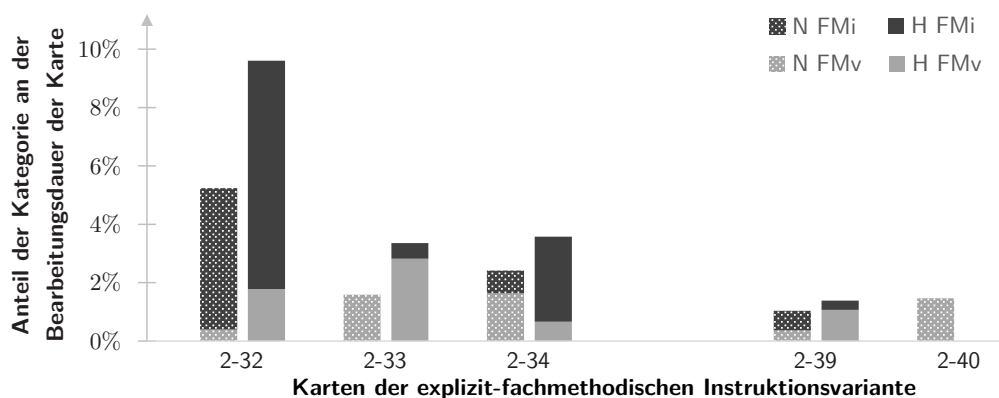
5.4.6.2 Ergebnisse der Vergleiche

In Abbildung 5.39 sind die Diagramme für die untersuchten Karten aus den drei Einheiten dargestellt. Für die Karte 1-11 finden sich keine fachmethodischen Beiträge; Karte 1-11 und die Karten in ihrem Umfeld werden daher nicht dargestellt. Es stellen sich in der Gesamtschau über alle Experimentiersituationen keine deutlichen systematischen Unterschiede in den Anteilen fachmethodischer Beiträge *in den Experimentiersituationen selbst* (d. h. für die oben aufgelisteten Karten) ein, wenn Personen mit hohen vs. niedrigen Kompetenzzuwächsen verglichen werden. Weder finden sich für überwiegend viele Situationen deutlich mehr Beiträge der Kategorie **Fachmethodisch-Indiziert** (wie durch 1-14 und 2-33 nahegelegt werden könnte), noch deutlich mehr Beiträge der Kategorie **Fachmethodisch-Vermutet** (vgl. 2-33, 2-39). Auch wenn die beiden Kategorien zusammengelegt werden (Addition der zeitlichen Anteile) ergibt sich kein über die Experimentiersituationen hinweg systematisch einheitlicher Unterschied (nur für 2-33 und 2-39 finden sich absolut und relativ als eventuell relevant anzusehende Unterschiede).

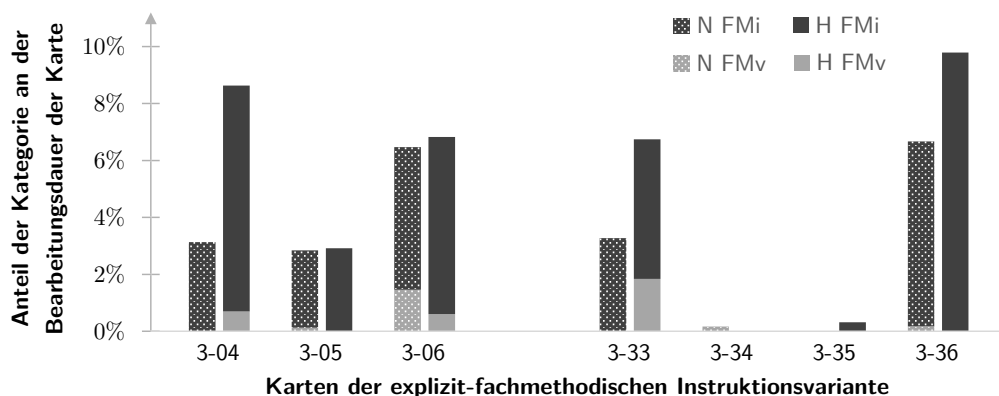
5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante



(a)



(b)



(c)

Abbildung 5.39: Fachmethodische Beiträge (vermutet: FMv; indiziert: FMi) bei Experimentiersituationen aus den drei Einheiten (Einheit 1: a, Einheit 2: b, Einheit 3: c) aufgelöst für Personen mit hohen vs. niedrigen Kompetenzzuwächsen (H bzw. N). Fehlerbalken sind zur Komplexitätsreduktion omitiert. Personenanzahlen können aus Tabelle A.2 entnommen werden.

Die beiden Karten 2-33 und 2-39, die beide der Einheit zum Planen und Durchführen von Untersuchungen entstammen, lassen sich als ggf. vom sonstigen Befund abweichende Experimentiersituationen identifizieren (was auch an den in den Klammern des vorigen Absatzes angeführten Beispiele für vorliegende Unterschiede deutlich wird). Zum einen könnte dies damit zusammenhängen, dass es in der besagten Einheit eben gerade um die Verbindung von fachmethodischen Äußerungen und Versuchshandlungen geht. Lernende mit hohen Kompetenzzuwächsen würden demnach die Instruktion eher so bearbeiten, wie es der Anlage der Instruktion entspricht. Zum anderen könnten die beiden Experimentiersituationen sich grundsätzlich von anderen Experimentiersituationen aus der Instruktion unterscheiden. Eine Vermutung dazu wäre, dass Versuche mit den eingesetzten Papierkegeln Überlegungen zur Kontrolle von Variablen besonders notwendig und zielführend machen, während andere Versuche im Material eher so angelegt sind, dass die Variablen bereits kontrolliert sind oder sehr naheliegend kontrolliert werden können, weshalb die Lernenden bei ihnen weniger fachmethodische Beiträge machen.

Dort, wo es inhaltlich sinnvoll ist, die Karten vor den Experimentiersituationen zu untersuchen, finden sich für Personen mit hohen Kompetenzzuwächsen größere Anteile fachmethodischer Beiträge (insgesamt, wenn also **Fachmethodisch-Indiziert** und **Fachmethodisch-Vermutet** zusammengenommen werden). Ferner nimmt der Anteil der Kategorie **Fachmethodisch-Indiziert** an allen fachmethodischen Beiträgen bei diesen Karten über die Situationen hinweg in der Tendenz zu. Zusammengenommen mit dem Ergebnis zu den Bearbeitungsdauern (Karten, die inhaltlich auf Versuche vorbereiten, werden von Lernenden mit hohen Kompetenzzuwächsen im Mittel länger bearbeitet als von Lernenden mit niedrigen Kompetenzzuwächsen) ergibt sich die Deutung, dass Lernende mit hohen Kompetenzzuwächsen sich intensiver und mit einer stärkeren Orientierung auf Fachmethoden auf Versuche vorbereiten. Die Zunahme des Anteils der Kategorie **Fachmethodisch-Indiziert** legt nahe, dass Lernende mit hohen Kompetenzzuwächsen im Verlauf der Bearbeitung der Instruktion bereits in der Vorbereitung auf Versuche dazu in der Lage sind, fachmethodische Überlegungen deutlich zu verbalisieren. Eine kurze inhaltliche Sichtung der Äußerungen scheint darauf hinzudeuten, dass die Lernenden sich an diesen Stellen zu größer werdenden Teilen verbal auf in der Instruktion früher adressierte Konzepte beziehen.

Der Unterschied zwischen Personen mit hohen vs. niedrigen Kompetenzzuwächsen scheint am deutlichsten darin zu liegen, dass Lernende mit niedrigen Kompetenzzuwächsen auf den Karten, die *nach* den oben gelisteten Experimentiersituationen betrachtet werden, einen höheren Anteil der Kategorie **Fachmethodisch-Vermutet**

aufweisen als Lernende mit hohen Kompetenzzuwächsen (siehe 1-05, 1-15, 2-40, 3-06, 3-36;⁷² vgl. auch die in Unterunterabschnitt 5.4.6.3 gesondert diskutierte Karte 2-34). Dies könnte so gedeutet werden, dass Lernende mit hohen Kompetenzzuwächsen nach dem Abschließen der Experimentiersituation dazu in der Lage sind, möglicherweise geforderte fachmethodische Äußerungen so zu formulieren, dass sie mit **Fachmethodisch-Indiziert** kodiert werden, während Lernende mit niedrigen Kompetenzzuwächsen die Experimentiersituation verbal nachbereiten und daher Verknüpfungen mit den konkreten Experimentiersituationen herstellen (die eher mit **Fachmethodisch-Vermutet** kodiert werden).

5.4.6.3 Detaillierte Betrachtung einer Experimentiersituation

Für die Karten 2-33 und 2-34 (siehe Abbildung 5.40, ab S. 196) ergibt sich aus Abbildung 5.39 (auf S. 193) ein erster Hinweis auf ein gegebenenfalls spannendes Ergebnis, wenn berücksichtigt wird, dass die Schülerinnen und Schüler auf der Karte 2-34 den von ihnen eigenständig zur Karte 2-33 geplanten und durchgeführten Versuch hinsichtlich der Variablenkontrolle prüfen sollen. Bei Nicht-Einhaltung der Variablenkontrolle sollen die Lernenden auf der Karte 2-34 den Versuch erneut durchführen. Das Ergebnis, dass sich im Vergleich von 6 Personen mit niedrigen und 11 Personen mit hohen Kompetenzzuwächsen findet, lässt sich detailliert folgendermaßen beschreiben (siehe insbesondere Abbildung 5.41 auf S. 197):

- Für beide Gruppen von Lernenden nimmt der mittlere Anteil der Kategorie **Fachmethodisch-Indiziert** (an der Gesamtbearbeitungsdauer der Einheit) zu, allerdings ist der Mittelwertunterschied für die Gruppe der Personen mit hohen Kompetenzzuwächsen etwas deutlicher als für die Gruppe der Personen mit niedrigen Kompetenzzuwächsen (visuell in Abbildung 5.41 auf S. 197; Details zu gepaarten t-Tests in Tabelle 5.25 auf S. 198). (Der Unterschied zwischen den Personengruppen lässt sich durch den Unterschied der Kartenmittelwertdifferenzen beschreiben: er ist nur tendenziell und von kleiner Effektstärke; siehe Tabelle 5.26 auf S. 198.)

⁷²Auch Karte 1-12 enthält für Personen mit niedrigen Kompetenzzuwächsen fachmethodisch-vermutete Beiträge, während für Personen mit hohen Kompetenzzuwächsen keine vermutet fachmethodischen Beiträge vorkommen. Es dürfte plausibel sein, dass Lernende mit hohen Kompetenzzuwächsen sich vollständig mit der Karte 1-12 beschäftigen (Kennzeichen von Vorhersagen und Aufforderung zur Diskussion über die (Nicht-)Notwendigkeit von Vorhersagen für die Durchführung von Versuchen), während Lernende mit niedrigen Kompetenzzuwächsen von illustrativen Bezügen zum Versuch auf Karte 1-11, die im Text der Karte 1-12 enthalten sind, immer wieder dahin gelangen, über den Versuch zu sprechen.

33 Versuch 1: Durchmesser des Kegels

AUFGABE

Überlegen Sie sich einen Versuch, mit dem Sie ohne Messwerte aufzunehmen prüfen können, ob der Durchmesser des Kegels einen Einfluss auf die durchschnittliche Fallgeschwindigkeit hat. Beschreiben/skizzieren Sie, wie Sie diesen Versuch durchführen wollen.

Führen Sie den Versuch durch und notieren Sie eine Antwort.

Hinweis: Lassen Sie die Kegel immer mit der Spitze nach unten fallen!

Antwort: Der Durchmesser des Kegels hat ☐ **einen** / ☐ **keinen** Einfluss auf die durchschnittliche Fallgeschwindigkeit.

34 Versuch 1: Durchmesser des Kegels

AUFGABE

Haben Sie bei der Planung des Versuchs bedacht, dass ein Kegel mit einem größeren Durchmesser auch eine größere Masse hat?

- 1) Diskutieren Sie kurz, warum es wichtig ist, die Masse des Kegels zu kontrollieren.
- 2) Sollten Sie die Masse bei Ihrem ersten Versuch nicht konstant gehalten haben, überlegen Sie sich einen zweiten Versuch, bei dem der Durchmesser verändert und die Masse konstant gehalten wird. Führen Sie den neuen Versuch durch.

Hinweis: Sollte Ihnen nach einer Minute kein Versuch eingefallen sein, nutzen Sie den Tipp auf der nächsten Karte!

Notieren Sie ggf. eine neue Antwort auf die Untersuchungsfrage.

Der Durchmesser des Kegels hat ☐ **einen** / ☐ **keinen** Einfluss auf die durchschnittliche Fallgeschwindigkeit.

Abbildung 5.40: Darstellung der detailliert betrachteten Karten 2-33 und 2-34 aus der explizit-fachmethodischen Instruktionsvariante.

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

- Für Personen mit hohen Kompetenzzuwächsen nimmt der mittlere Anteil der Kategorie **Fachmethodisch-Vermutet** (an der Gesamtbearbeitungsdauer der Einheit) deutlich ab, während er für Personen mit niedrigen Kompetenzzuwächsen nahezu gleich bleibt (visuell in Abbildung 5.41 auf S. 197; Details zu gepaarten t-Tests in Tabelle 5.25 auf S. 198). (Der Unterschied zwischen den Personengruppen lässt sich durch den Unterschied der Kartenmittelwertdifferenzen beschreiben: er ist nur tendenziell und von kleiner Effektstärke; siehe Tabelle 5.26 auf S. 198.)
- Während der mittlere Anteil aller fachmethodischen Beiträge (FMi + FMv; Anteil an der Gesamtbearbeitungsdauer der Einheit) für Personen mit niedrigen Kompetenzzuwächsen zunimmt, bleibt er (aufgrund der Abnahme bei der Kategorie **Fachmethodisch-Vermutet**) für Personen mit hohen Kompetenzzuwächsen nahezu gleich (visuell in Abbildung 5.42 auf S. 198; Details zu gepaarten t-Tests in Tabelle 5.25 auf S. 198). Der Unterschied zwischen den beiden Gruppen ist allerdings nur tendenziell und von kleiner Effektstärke (siehe Tabelle 5.26 auf S. 198).

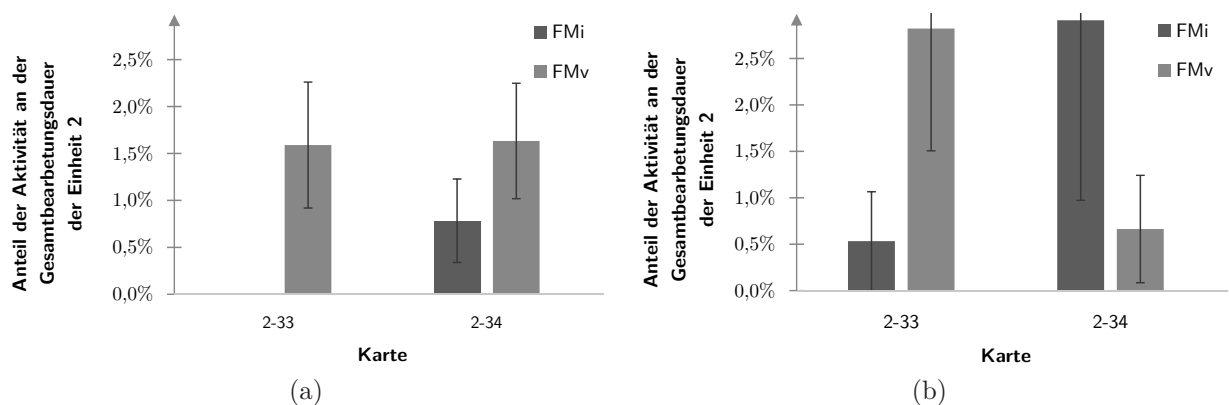


Abbildung 5.41: Mittelwerte der relativen Anteile der fachmethodisch-indizierten (FMi) und -vermuteten (FMv) Beiträge an der Gesamtbearbeitungsdauer von Einheit 2 der explizit-fachmethodischen Instruktionsvariante für 6 Personen mit niedrigen (a) und 11 Personen mit hohen (b) Kompetenzzuwächsen. Fehlerbalken entsprechen einem Standardfehler des Mittelwerts.

5 Aktivitäten von Lernenden

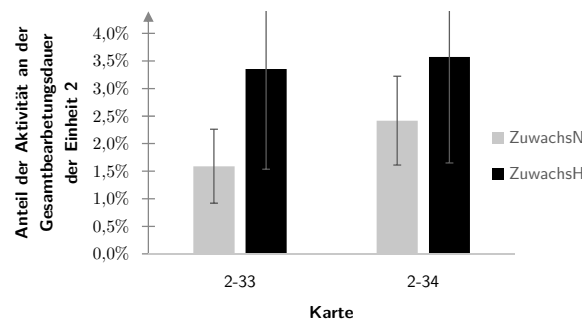


Abbildung 5.42: Mittelwerte der relativen Anteile der fachmethodischen Beiträge an der Gesamtbearbeitungsdauer von Einheit 2 der explizit-fachmethodischen Instruktionsvariante für 11 Personen mit hohen (ZuwachsH) vs. 6 mit niedrigen (ZuwachsN) Kompetenzzuwächsen. Fehlerbalken entsprechen einem Standardfehler des Mittelwerts.

Tabelle 5.25: Personenweise gepaarte t-Tests für die zeitlichen Anteile (der verschiedenen Kategorien) fachmethodischer Beiträge bei den Karten 33 und 34 der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante einzeln für die Gruppen von Personen mit niedrigen Kompetenzzuwächsen (ZuwachsN/ZuwachsH).

Vergleich	Karte 2-33			Karte 2-34			df	t (df)	p	r	Unterschied	
	N ₃₃	MW	SD	N ₃₄	MW	SD					Absolut	Relativ
FMi (ZuwachsN)	6	0,00%	0,00%	6	0,78%	1,09%	5	-1,761	0,139	0,619	0,78%	100,00%
FMi (ZuwachsH)	11	0,53%	1,09%	11	2,91%	6,43%	10	-1,632	0,134	0,459	2,38%	81,70%
FMv (ZuwachsN)	6	1,59%	1,64%	6	1,63%	1,51%	5	-0,056	0,957	0,025	0,04%	2,66%
FMv (ZuwachsH)	11	2,82%	1,51%	11	0,66%	1,92%	10	1,437	0,181	0,414	-2,16%	-76,49%
FM (ZuwachsN)	6	1,59%	1,64%	6	2,42%	1,97%	5	-1,291	0,253	0,500	0,83%	34,21%
FM (ZuwachsH)	11	3,35%	6,03%	11	3,57%	6,38%	10	-0,212	0,836	0,067	0,22%	6,13%

Anmerkungen. Gepaarte t-Test; $N_{33/34}$ =Stichprobengröße für die entsprechenden Gruppe der Personen bei der untersuchten Karte der explizit-fachmethodischen Instruktionsvariante; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Tabelle 5.26: Students t-Tests zum Vergleich der Personen mit niedrigen und hohen Kompetenzzuwächsen bezüglich der Differenzen der zeitlichen Anteile der jeweiligen fachmethodischen Beiträge für die Karten 33 und 34 der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante.

Vergleich	Zuwachs Niedrig			Zuwachs Hoch			df	t (df)	p	r	Unterschied	
	N _N	MW	SD	N _H	MW	SD					Absolut	Relativ
FMi	6	0,78%	1,09%	11	2,38%	4,83%	15	-0,786	0,444	0,199	1,59%	67,06%
FMv	6	0,04%	1,89%	11	-2,16%	4,98%	15	1,031	0,319	0,257	-2,20%	-102,01%
FM	6	0,83%	1,57%	11	0,22%	3,43%	15	0,407	0,690	0,104	-0,61%	-73,49%

Anmerkungen. Students t-Test; $N_{N/H}$ =Stichprobengröße für Personen mit hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Für die Deutung des Ergebnisses zu den verschiedenen Kategorien fachmethodischer Beiträge werden zudem folgende Ergebnisse hinzugezogen (siehe insbesondere die Abbildungen 5.43 und 5.44 auf S. 200):

- Die Lernenden mit hohen Kompetenzzuwächsen weisen (dem Ergebnis für die Gesamtbearbeitungsdauer der Einheit und den Ergebnissen zu den Abschnittsbearbeitungsdauern entsprechend, siehe Unterabschnitt 5.4.1, ab S. 106) größere mittlere *Bearbeitungsdauern für beide Karten* auf als die Lernenden mit niedrigen Kompetenzzuwächsen (siehe Abbildung 5.43 auf S. 200; Details in Tabelle 5.27 auf S. 201).
- Die mittleren Bearbeitungsdauern sind für beide Gruppen von Lernenden bei Karte 2-34 tendenziell geringer als bei Karte 2-33 (siehe Abbildung 5.44 auf S. 200; Details in Tabelle 5.28 auf S. 201). Der Unterschied zwischen den beiden Gruppen ist allerdings verschwindend gering (Statistik in Tabelle 5.29 auf S. 201).
- Die Lernenden mit hohen Kompetenzzuwächsen weisen bei jeder der beiden Karten etwas größere mittlere *Anteile verbaler und nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen* auf als die Lernenden mit niedrigen Kompetenzzuwächsen (siehe Abbildung 5.44 auf S. 200; Details in Tabelle 5.27 auf S. 201).
- Die mittleren Anteile verbaler und nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen sind für beide Gruppen von Lernenden bei Karte 2-34 tendenziell geringer als bei Karte 2-33 (siehe Abbildung 5.44 auf S. 200; Details in Tabelle 5.28 auf S. 201). Die Unterschiede zwischen den beiden Gruppen von Lernenden mit unterschiedlichen Kompetenzzuwächsen sind nur von sehr geringer Größe (siehe Tabelle 5.29 auf S. 201).

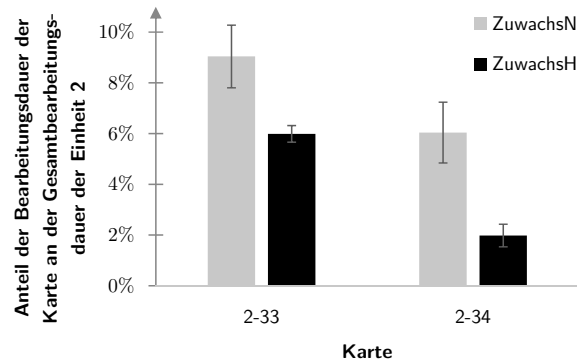


Abbildung 5.43: Mittelwerte der relativen Anteile der Bearbeitungsdauer der untersuchten Karten an der Gesamtbearbeitungsdauer von Einheit 2 der explizit-fachmethodischen Instruktionsvariante für 11 Personen mit hohen (ZuwachsH) vs. 6 Personen mit niedrigen (ZuwachsN) Kompetenzzuwächsen. Fehlerbalken entsprechen einem Standardfehler des Mittelwerts.

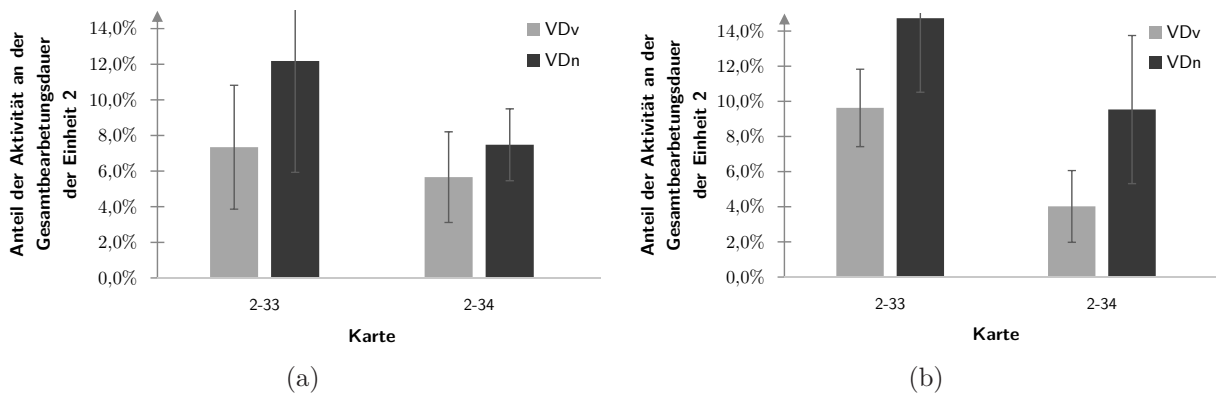


Abbildung 5.44: Mittelwerte der relativen Anteile der Aktivitäten des verbalen (VDv) und des non-verbalen (VDn) Vorbereitens und Durchführens von Versuchen an der Gesamtbearbeitungsdauer von Einheit 2 der explizit-fachmethodischen Instruktionsvariante für 6 Personen mit niedrigen (a) und 11 mit hohen (b) Kompetenzzuwächsen. Fehlerbalken entsprechen einem Standardfehler des Mittelwerts.

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

Tabelle 5.27: Students t-Tests zum Vergleich von Lernenden mit niedrigen und hohen Kompetenzzuwächsen bezüglich der verbalen und nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen jeweils bei Karte 33 und Karte 34 der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante sowie bezüglich der Bearbeitungsdauern der Karten (BearbD).

Vergleich	Zuwachs Niedrig			Zuwachs Hoch			df	t (df)	p	r	Unterschied	
	N_N	MW	SD	N_H	MW	SD					Absolut	Relativ
VDv bei Karte 2-33	6	7,34%	8,52%	11	9,62%	7,31%	15	-0,582	0,569	0,149	2,29%	23,75%
VDv bei Karte 2-34	6	5,66%	6,24%	11	4,02%	6,77%	15	0,490	0,631	0,126	-1,64%	-28,99%
VDn bei Karte 2-33	6	12,18%	15,30%	11	14,73%	13,97%	15	-0,349	0,732	0,090	2,56%	17,35%
VDn bei Karte 2-34	6	7,48%	4,94%	11	9,53%	13,98%	15	-0,344	0,735	0,089	2,06%	21,57%
BearbD bei 2-33	6	9,04%	3,02%	11	5,99%	1,08%	15	3,072	0,008	0,621	-3,05%	-33,75%
BearbD bei 2-34	6	6,04%	2,94%	11	1,98%	1,48%	15	3,845	0,002	0,704	-4,06%	-67,21%

Anmerkungen. Students t-Test; $N_{N/H}$ =Stichprobengröße für Personen mit hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Tabelle 5.28: Personenweise gepaarte t-Tests für die zeitlichen Anteile verbaler und nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen (VDv/VDn) bei den Karten 33 und 34 der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante sowie die Bearbeitungsdauern der Karten (BearbD) einzeln betrachtet für die Gruppen von Personen mit niedrigen Kompetenzzuwächsen (ZuwachsN/ZuwachsH).

Vergleich	Karte 2-33			Karte 2-34			df	t (df)	p	r	Unterschied	
	N_{33}	MW	SD	N_{34}	MW	SD					Absolut	Relativ
VDv (ZuwachsN)	6	7,34%	8,52%	6	5,66%	6,24%	5	0,899	0,410	0,373	-1,68%	-22,90%
VDv (ZuwachsH)	11	9,62%	6,24%	11	4,02%	6,77%	10	3,172	0,010	0,708	-5,61%	-58,25%
VDn (ZuwachsN)	6	12,18%	15,30%	6	7,48%	4,94%	5	0,793	0,464	0,334	-4,70%	-38,60%
VDn (ZuwachsH)	11	14,73%	4,94%	11	9,53%	13,98%	10	1,383	0,197	0,401	-5,20%	-35,29%
BearbD (ZuwachsN)	6	9,04%	3,02%	6	6,04%	2,94%	5	34,450	0,000	0,998	-3,00%	-33,19%
BearbD (ZuwachsH)	11	5,99%	1,08%	11	1,98%	1,48%	10	8,647	0,000	0,939	-4,01%	-66,93%

Anmerkungen. Gepaarte t-Test; $N_{33/34}$ =Stichprobengröße für die entsprechenden Gruppe der Personen bei der untersuchten Karte der explizit-fachmethodischen Instruktionsvariante; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Tabelle 5.29: Students t-Tests zum Vergleich der Personen mit niedrigen und hohen Kompetenzzuwächsen bezüglich der Differenzen der zeitlichen Anteile der verbalen und nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen (VDv/VDn) für die Karten 33 und 34 der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante sowie für deren Bearbeitungsdauern (BearbD).

Vergleich	Zuwachs Niedrig			Zuwachs Hoch			df	t (df)	p	r	Unterschied	
	N_N	MW	SD	N_H	MW	SD					Absolut	Relativ
VDv	6	-1,68%	4,58%	11	-5,61%	5,86%	15	1,415	0,178	0,028	-3,93%	-70,03%
VDn	6	-3,00%	0,21%	11	-4,01%	1,54%	15	1,574	0,136	0,006	-1,01%	-25,14%
BEarbD	6	-4,70%	14,52%	11	-5,20%	12,46%	15	0,075	0,942	0,067	-0,50%	-9,60%

Anmerkungen. Students t-Test; $N_{N/H}$ =Stichprobengröße für Personen mit hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

5 Aktivitäten von Lernenden

Ferner wurde eine Einschätzung vorgenommen, welche Teams auf Karte 2-34 (siehe Abbildung 5.40, ab S. 196) den Versuch (erneut⁷³) durchführen, zu dem auf Karte 2-33 aufgefordert wurde. Die Ergebnisse der Einschätzung sind in Tabelle 5.30 überblicksartig dokumentiert. Es ergab sich, dass 5 von den insgesamt 9 für die zweite Einheit der explizit-fachmethodischen Instruktionsvariante analysierten Teams den Versuch bei Karte 2-34 durchführten. Davon enthielten je ein Team nur Personen mit niedrigen und mit hohen Kompetenzzuwächsen; unter den Teams, die den Versuch nicht (erneut) bei Karte 2-34 durchführten, finden sich zwei Teams, die nur Lernende

Tabelle 5.30: Ergebnisse des Ratings dazu, ob die Lernenden bei Karte 2-34 den auf Karte 2-33 durchzuführenden Versuch (erneut) durchgeführt haben sowie auf der Videodurchsicht basierende Anmerkungen zu den Bearbeitungen der beiden Karten.

Team	Versuch bei 2-34 durchgeführt	Anmerkungen
A	ja	Lösung zur variablenkontrollierten Durchführung wird schnell gefunden (kleiner Hinweis des betreuenden Forschers auf Karte 2-31).
B	nein	Bereits bei Karte 2-33 variablenkontrolliert durchgeführt (es scheint, dass B6 mitbekommt, wie der betreuende Forscher mit einem Team am anderen Ende des Klassenraums über die Notwendigkeit der Konstanthaltung der Masse spricht).
C	ja	Fehler bei Karte 2-33 nicht bemerkt; unklar, ob Fehler von Karte 2-33 bei der Bearbeitung von Karte 2-34 bemerkt wird. Es wird eine fachlich korrekte Einschätzung zum Einfluss der Masse vorgenommen (nicht von Aufgabe gefordert); danach eigenständig Versuch variablenkontrolliert geplant und durchgeführt.
D	nein	Bereits bei Karte 2-32 (unaufgefordert) einen Versuch geplant und diesen (mit Hilfe des betreuenden Forschers) variablenkontrolliert angelegt.
E	ja	Lösung schnell gefunden, weil die Schülerin E14 bei einem anderen Team sieht, dass die Kegel ineinandergesteckt werden und sich darüber wundert: „Achso, das war Absicht“ und daraufhin auf das Konstanthalten der Masse kommt.
F	nein	Unklar, wie der Versuch bei Karte 2-33 durchgeführt wird, da die Schülerinnen das Bild verlassen; es wirkt allerdings so, dass durch zu schnelles Umblättern zu Karte 2-34 zwei Teammitglieder die Frage nach der Beachtung der Variablenkontrolle auf Karte 2-34 gar nicht lesen.
G	ja	Große Bearbeitungsdauer bei Karte 2-34; die Schülerinnen führen erst nach Hilfe durch den betreuenden Forscher den Versuch variablenkontrolliert durch.

Anmerkungen. Team A besteht aus 3 Personen mit hohen Kompetenzzuwächsen, Team B ebenfalls, Team C besteht aus einer Person mit niedrigem Kompetenzzuwachs und einer mit mittlerem (die dritte Person ist nicht anwesend), Team D besteht aus drei Personen mit hohen Kompetenzzuwächsen, Team E besteht aus einer Person mit niedrigem und einer mit hohem Kompetenzzuwachs und einer Person, für die kein Wert vorliegt, Team F besteht aus einer Person mit niedrigem und einer mit hohem Kompetenzzuwachs und einer (für die Zählung nicht berücksichtigen) Person mit stark negativem Kompetenzzuwachs, Team G besteht aus drei Personen mit niedrigen Kompetenzzuwächsen.

⁷³Auf Karte 2-34 wird nur unter der Bedingung einer vorherigen fehlerhaften Planung oder Durchführung zum erneuten Durchführen aufgefordert. Es könnte jedoch sein, dass die Lernenden den Versuch trotz korrekt variablenkontrollierter Durchführung bei Karte 2-33 erneut bei Karte 2-34 planen und durchführen. Es könnte auch sein, dass Lernende bei Karte 2-33 den Versuch nicht durchführen und ihn daher bei Karte 2-34 zum ersten Mal (und eben nicht erneut) durchführen. Daher wird das Wort *erneut* im Text eingeklammert.

mit hohen Kompetenzzuwächsen enthalten, und keine Teams, die nur aus Lernenden mit niedrigen Kompetenzzuwächsen bestehen. Für die in den auf Personen-Ebene vorgenommenen Vergleich einbezogenen 6 Personen mit niedrigen Kompetenzzuwächsen fanden sich 5 Personen, die den Versuch erneut durchführten; für die einbezogenen 11 Personen mit hohen Kompetenzzuwächsen fanden sich 4 Personen, die den Versuch erneut durchführten.

Als Deutung für die Ergebnisse wird Folgendes vorgeschlagen:⁷⁴

- Für den hier betrachteten Versuch, bei dem Kegel unterschiedlicher Durchmesser variablenkontrolliert miteinander verglichen werden sollen, wobei insbesondere die Kontrolle der Masse von Bedeutung ist, und für die auf den Karten der Instruktion gestellten Aufgaben könnte der *zeitliche Anteil fachmethodisch-vermuteter Beiträge* stark damit verbunden sein, ob der Versuch bei Karte 2-34 (erneut) durchgeführt wird. Der Anteil fachmethodisch-vermuteter Beiträge könnte für die Gruppe von Personen mit hohen Kompetenzzuwächsen aus dem Grund geringer als für die Gruppe der Personen mit niedrigen Kompetenzzuwächsen sein, dass weniger Personen den Versuch bei Karte 2-34 (erneut) durchführen. Dass der Anteil fachmethodisch-vermuteter Beiträge für Personen mit hohen Kompetenzzuwächsen für Karte 2-33 etwa doppelt so groß ist wie für Personen mit niedrigen Kompetenzzuwächsen und sehr viel größer ist als bei Karte 2-34, könnte daran liegen, dass Lernende mit hohen Kompetenzzuwächsen eigenständig bei dem Versuch stehen bleiben, bis die Problematik der Variablenkontrolle gelöst ist. Die Videoeinsicht (siehe u. a. die Anmerkungsspalte in Tabelle 5.30 auf S. 202) hat gezeigt, dass einige Lernenden zwar sehr wohl die Problematik erkennen, diese aber nicht selbstständig lösen. Es bleibt dabei allerdings offen, ob die Lernenden die *Problematik aufgrund fachinhaltlicher Kompetenzen bemerken* (weil das *Ergebnis* ihrer Intuition widerspricht), wenngleich die oberflächlichen Videoeinsichten dies nahelegen. Dies wäre spannend, weil dann fachinhaltliche Kompetenzen förderlich für den Aufbau fachmethodischer Kompetenzen wären, da vorrangig die Lernenden mit hohen Kompetenzen bei Karte 2-33 verweilen bis die Problematik geklärt ist. Außerdem bleibt offen,

⁷⁴Die formulierte Deutung geht vorrangig auf die Ergebnisse zu fachmethodischen Beiträgen (siehe S. 195 und anschließende) und das Rating (siehe S. 202 und anschließende) ein; die ab S. 199 dargestellten Hintergrundinformationen zu den zeitlichen Anteilen verbaler und nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen sowie zu den Bearbeitungsdauern werden implizit zur Deutung hinzugezogen, um die Ausführungen nicht unnötig komplex werden zu lassen. In den Diskussionen der Deutungen (ab S. 204) wird, wenn notwendig, auf die entsprechenden Ergebnisse verwiesen.

ob die Lernenden *die entstehende Problematik selbstständig lösen könnten* – denn bei den analysierten Teams war häufig ein Hinweis von außen oder eine direkte äußerliche Hilfe zeitnah verfügbar.⁷⁵

- Für den hier betrachteten Versuch und die hier betrachteten Aufgaben ist der *zeitliche Anteil fachmethodisch-indizierter Beiträge* zum Teil mit den konkreten Aufforderungen zum Diskutieren fachmethodischer Inhalte (Teil 1 auf Karte 2-34, siehe Abbildung 5.40 auf S. 196) verbunden, zum anderen mit dem bewussten bzw. fachmethoden-fokussierten Durchführen des Versuchs. Diese Deutung wird insbesondere basierend auf den Ergebnissen formuliert, dass Lernende mit niedrigen Kompetenzzuwächsen bei Karte 2-33 keine fachmethodisch-indizierten Beiträge machen und Lernende mit hohen Kompetenzzuwächsen bei Karte 2-34 sehr viel größere zeitliche Anteile fachmethodisch-indizierter Beiträge aufweisen als sie es bei Karte 2-33 tun. Für Lernende mit niedrigen Kompetenzzuwächsen liegt es daher nahe, die fachmethodisch-indizierten Beiträge bei Karte 2-34 als von der Aufgabe angeregt anzusehen,⁷⁶ während es für Lernende mit hohen Kompetenzzuwächsen neben dem zeitlichen Anteil, der durch das Antworten auf diese instruktionalen Anregungen zu fachmethodisch-indizierten Beiträgen zustande kommen dürfte, weitere zeitliche Anteile fachmethodisch-indizierter Beiträge gibt, die dem Verbinden von fachmethodischen Konzepten und dem Durchführen von Versuchen zugeordnet sein könnten.⁷⁷

Die skizzierte Deutung ist in mehrfacher Hinsicht spannend und lässt sich unter verschiedenen Gesichtspunkten diskutieren und weiterführen. Neben der bereits im ersten Aufzählungspunkt der Deutung diskutieren Verbindung zu fachinhaltlichen Kompetenzen der Lernenden gehören unter anderem folgende Gesichtspunkte dazu:

- Es ist möglich, dass die etwas längeren Zeiten des Experimentierens (Kategorie **VorbereitenDurchführenNonverbal, VDn**), die sich für Lernende mit hohen

⁷⁵Eventuell könnte die Analyse weiterer Teams aus dem gleichen Datensatz Aufschluss darüber geben, inwiefern externe Hinweise und Hilfen notwendig sind. Andernfalls wären andere Settings zur Bearbeitung der Instruktion denkbar: Team alleine ohne Betreuung, Team alleine ohne andere Teams im Klassenraum, eine Person alleine (um Einfluss anderer Teammitglieder zu vermeiden), ...

⁷⁶Es wurde allerdings nicht im Detail nachgeprüft, inwieweit die vorliegenden fachmethodisch-indizierten Beiträge auf die Aufforderungen der Instruktion bezogen sind.

⁷⁷Das Mehr an zeitlichem Anteil fachmethodisch-indizierter Beiträge für Lernende mit hohen Kompetenzzuwächsen könnte allerdings auch anders plausibilisiert werden; beispielsweise ist es möglich, dass Lernende mit hohen Kompetenzzuwächsen eine stärkere Tendenz dazu haben, verbal zu wiederholen oder zu begründen, was sie in vorherigen Handlungen bereits korrekt vorgenommen haben.

5.4 Auswertungen und Ergebnisse zur explizit-fachmethod. Instruktionsvariante

Kompetenzzuwächsen im Vergleich zu Lernenden mit niedrigen Kompetenzzuwächsen (für die beiden hier untersuchten Karten 2-33 und 2-34, siehe Abbildung 5.44 auf S. 200, und auch generell bei den Einheiten 2 und 3 der explizit-fachmethodischen Instruktionsvariante) finden, damit zusammenhängen, dass *Lernende mit hohen Kompetenzzuwächsen eher ihre experimentellen Handlungen (verbal) mit fachmethodischen Konzepten verbinden* als Lernende mit niedrigen Kompetenzzuwächsen und sie daher etwas mehr Zeit benötigen.

- Die für die Lernenden mit hohen Kompetenzzuwächsen bedeutsam deutlich stärker als für Lernende mit niedrigen Kompetenzzuwächsen abnehmenden Bearbeitungsdauern (siehe Abbildung 5.43 auf S. 200) dürften vorrangig damit zusammenhängen, dass in der analysierten Stichprobe mehr Lernende mit niedrigen Kompetenzzuwächsen bei Karte 2-34 den Versuch (erneut) durchführen. Die Videosichtungen legen allerdings auch nahe, dass die Frage für den Kompetenzzuwachs weniger ist, ob der Versuch bei Karte 2-34 wiederholt/durchgeführt wird, sondern ob *intensiv über die korrekte variablenkontrollierte Durchführung nachgedacht und gesprochen* wird: Einige Teams mit hohen Kompetenzzuwächsen wiederholen den Versuch bereits bei Karte 2-33 (ohne die instruktionale Anregung zur Überprüfung, aber z. T. aufgrund externer Anregung durch andere Teams oder Betreuende), einige Teams müssen den Versuch nicht wiederholen, weil sie vorher ausführlich über Variablenkontrolle nachgedacht haben (z. T. aufgrund externer Anregung durch andere Teams oder Betreuende; siehe z. B. Team D in Tabelle 5.30 auf S. 202).
- Instruktionale Hilfen, Prompts und Erklärungen, die schriftlich und zu festgelegten Zeitpunkten der Bearbeitung hereingegeben werden (wie bei der vorliegenden Instruktion der Fall) oder von den Lernenden selbstständig bei Problemen eingeholt werden können (wie zum Teil bei der vorliegenden Instruktion der Fall, vgl. Karte 2-34: Abbildung 5.40 auf S. 196), könnten verbalen Hilfen, Prompts und Erklärungen geschulter Lehrkräfte, die variabel gegeben oder ausgelassen werden können und zu verschiedensten Zeitpunkten eingesetzt werden können, unterlegen sein, wenn es um die Unterstützung komplexer Lernprozesse geht. Dafür spricht, dass bei den Personen mit hohen Kompetenzzuwächsen zeitnah Hilfestellungen verfügbar waren und diese von den Personen auch (aus der Lautstärke des gesamten jeweils in Kleingruppen arbeitenden Klassenverbands herausgefiltert) wahrgenommen oder sogar aktiv gesucht wurden. Kommt die Hilfestellung jedoch zu spät, könnte es sein, dass Lernende diese nicht mehr

annehmen, wie zum Teil für Lernende mit niedrigen Kompetenzzuwächsen aus den Videos ableitbar sein könnte. Diese Vermutung ist allerdings aufgrund der kleinen Stichprobe und der durch vorrangige Fokussierung auf das (erneute) Durchführen des Versuchs ansonsten oberflächlichen Sichtung der Videos mit deutlichen Einschränkungen behaftet.

Detaillierte Studien zum Umgang der Lernenden mit Feedback, Prompts und Hilfestellungen könnten genauere Aufschlüsse dazu geben, inwiefern Zusammenhänge mit erreichten Kompetenzzuwächsen vorliegen. In Bezug auf die in der vorliegenden Arbeit untersuchte explizit-fachmethodische Instruktionsvariante könnte dazu auf eine Wissenschaftliche Hausarbeit von Schmandt (2017) aufgebaut werden, der für diese Instruktionsvariante verschiedene Arten von Feedback und Typen des Umgangs mit diesem beschreibt. Zu den Ergebnissen zählt beispielsweise, dass wiederholtes Feedback häufig übersprungen wird; das würde dazu passen, dass zu spät kommendes Feedback ebenfalls übersprungen wird. Allerdings stellt Schmandt (2017) keine Verknüpfungen mit Kompetenzzuwächsen her.

Anhand der aufgeführten Diskussionsschlagrichtungen wird mitunter deutlich, dass detailliertere Analysen inhaltlicher Verbindungen von Experimentierhandlungen und fachmethodischen Beiträgen notwendig sind, um die aufgeworfenen Fragen zu untersuchen. In Kapitel 6 werden derartige Analysen vorgenommen – allerdings (aufgrund der parallelen Arbeit an verschiedenen Teilen des Forschungsprojekts) nicht mit Bezug zur Variablenkontrolle, sondern zu anderen Konzepten des experimentbezogenen Denkens und Arbeitens.

5.5 Auswertungen und Ergebnisse zur implizit-fachmethodischen Instruktionsvariante sowie zum Vergleich der beiden Instruktionsvarianten

Ergänzend zu der schwerpunktmäßig betrachteten explizit-fachmethodischen Instruktionsvariante wird in der Arbeit auch eine implizit-fachmethodische Instruktionsvariante untersucht. Das vorrangig damit verfolgte Ziel besteht darin, die zur explizit-fachmethodischen Instruktionsvariante berichteten Ergebnisse einzubetten und zu kontextualisieren. Dabei stellen sich Fragen zum einen zum Vergleich der grundsätzlichen Bearbeitungen beider Instruktionsvarianten und zum anderen dazu, wie Lernende mit unterschiedlichen Kompetenzzuwächsen in jeder der beiden Instruktionsvarianten einzeln sowie im Vergleich beider die Instruktion gearbeitet haben. Weil die Analysen vor allem dazu beitragen sollen, die Ergebnisse zur explizit-fachmethodischen Instruktionsvariante besser zu verstehen, werden die beiden Forschungsfragen F-Akt2 und F-Akt3 im vorliegenden Abschnitt gemeinsam bearbeitet:

F-Akt2 Welche Unterschiede und Gemeinsamkeiten zeigen sich in den Aktivitäten von Lernenden – insbesondere mit hohen vs. niedrigen Kompetenzzuwächsen – bei der Bearbeitung der implizit-fachmethodischen Instruktionsvariante?

F-Akt3 Welche Unterschiede und Gemeinsamkeiten zeigen sich in den Aktivitäten von Lernenden, die unterschiedliche Instruktionsvarianten (explizit-fachmethodisch vs. implizit-fachmethodisch) bearbeiten?

In ähnlicher Struktur wie in Abschnitt 5.4 werden nachfolgend sowohl die Bearbeitungsdauern zur Instruktion, die Relationen von Aktivitäten in der Gesamtschau (Aktivitätsprofile) sowie detaillierter die fachmethodischen Beiträge der Lernenden in den Blick genommen. Vergleiche der beiden Instruktionsvarianten werden, sofern sinnvoll und möglich, zuerst berichtet. Danach wird eine Auflösung der Lernenden anhand ihrer erreichten Kompetenzzuwächse vorgenommen, wobei die Ergebnisse zur implizit-fachmethodischen Instruktionsvariante kurz und knapp berichtet werden und anschließend im Abgleich mit (Ergebnissen zu) der explizit-fachmethodischen Instruktionsvariante diskutiert werden.

Die Analysen zur implizit-fachmethodischen Instruktionsvariante werden grundsätzlich über alle drei Einheiten hinweg zusammenfassend vorgenommen. Einzelne Einheiten werden nur in wenigen Fällen aufgelöst. Dieses Vorgehen ist weniger umfangreich als die detailliert in Abschnitt 5.4 vorgenommene Auflösung zur explizit-

fachmethodischen Instruktionsvariante, was dem Ziel der Analysen zur implizit-fachmethodischen Instruktionsvariante entspricht. Inhaltlich ist diese Ebene der Auflösung im wesentlichen dadurch begründet, dass die implizit-fachmethodische Instruktionsvariante aufgrund ihrer impliziten Anlage nur bedingt nach fachmethodischen Konzepten gegliedert werden kann. Insbesondere sind die den Lernenden transparent gemachten Abschnitte nicht entlang fachmethodischer Konzepte formuliert. Weil die beiden Instruktionsvarianten parallelisiert sind, könnten bspw. eigene Abschnitte gebildet werden (z. B. zu jeweils denselben Versuchsschritten mit denselben Experimentiermaterialien); darauf wird allerdings in der vorliegenden Arbeit verzichtet, weil derartige Analysen unverhältnismäßig viel zusätzlichen Platz benötigen dürften und der Fokus der Arbeit auf den fachmethodischen Konzepten liegt. Dadurch, dass die Analysen zur implizit-fachmethodischen Instruktionsvariante über alle drei Einheiten hinweg zusammenfassend vorgenommen werden, wird also die Gesamtheit dieser Instruktionsvariante in den Blick genommen und kann mit der der explizit-fachmethodischen Instruktionsvariante abgeglichen werden, um so die Interpretationen zur explizit-fachmethodischen Instruktionsvariante zu stützen.

5.5.1 Bearbeitungsdauern

Eine Möglichkeit zur Annäherung an den Vergleich der Bearbeitungsprozesse der beiden Instruktionsvariante besteht in der Analyse der Bearbeitungsdauern, die sich für die Lernenden ergeben. Aus zugehörigen Ergebnissen können Hinweise dazu abgeleitet werden, wie aussagekräftig die Ergebnisse zu den im Detail untersuchten Aktivitäten sind.

5.5.1.1 Vergleich der Instruktionsvarianten

Erstens werden die Bearbeitungsdauern der 4 für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante kodierten Teams (umfassen 9 Personen) und der 6 für alle drei Einheiten der implizit-fachmethodischen Instruktionsvariante kodierten Teams (umfassen 15 Personen) mittels Students t-Tests miteinander verglichen. Da Personen desselben Teams dieselben Bearbeitungsdauern aufweisen, wird auf Ebene der Teams verglichen. Die Ergebnisse finden sich in Tabelle 5.31. Zweitens wird für jede der drei Einheiten ein Vergleich aller Teams vorgenommen, die (mindestens) für diese Einheit bei einer der beiden Instruktionsvarianten kodiert wurde. Die Ergebnisse sind in Tabelle 5.32 dargestellt.

5.5 Auswert. u. Ergebnisse zur impl.-fachmeth. IV sowie zum Vergl. d. beiden IVs

Tabelle 5.31: Students t-Tests zu den Gesamtbearbeitungsdauern der Einheiten für beide Instruktionsvarianten im Vergleich von Teams, die alle Einheiten bearbeitet haben.

Einheit	Implizit			Explizit			df	t (df)	p	r	Unterschied	
	N_I	MW	SD	N_E	MW	SD					Absolut	Relativ
1	6	2615,1	106,6	4	2357,0	283,7	8	2,072	0,072	0,591	-258,2	-9,87%
2		4180,2	346,8		3870,5	453,1		1,23	0,254	0,399	-309,7	-7,41%
3		4498,6	546,3		3539,2	509,2		2,79	0,024	0,702	-959,4	-21,33%
Ges	6	11294,0	870,8	4	9766,7	773,4	8	2,831	0,022	0,707	-1527,3	-13,52%

Anmerkungen. Students t-Test; $N_{E/I}$ =Stichprobengröße für Gruppe der Teams, die die explizit-/implizit-fachmethodische Instruktionsvariante bearbeitet haben; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Tabelle 5.32: Students t-Tests zu den Gesamtbearbeitungsdauern der Einheiten für beide Instruktionsvarianten im Vergleich von Teams, die die jeweilige Einheit bearbeitet haben.

Einheit	Implizit			Explizit			df	t (df)	p	r	Unterschied	
	N_I	MW	SD	N_E	MW	SD					Absolut	Relativ
1	6	2615,1	106,6	11	2411,8	297,5	15	1,599	0,131	0,382	-203,4	-7,78%
2	6	4180,2	346,8	7	4017,6	526,8	11	0,644	0,533	0,191	-162,6	-3,89%
3	6	4498,6	546,3	8	3814,6	512,7	12	2,404	0,033	0,57	-684,0	-15,21%
Ges	4	11294,0	870,8	6	9766,7	773,4	8	2,831	0,022	0,707	-1527,3	-13,52%

Anmerkungen. Students t-Test; $N_{E/I}$ =Stichprobengröße für Gruppe der Teams, die die explizit-/implizit-fachmethodische Instruktionsvariante bearbeitet haben; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

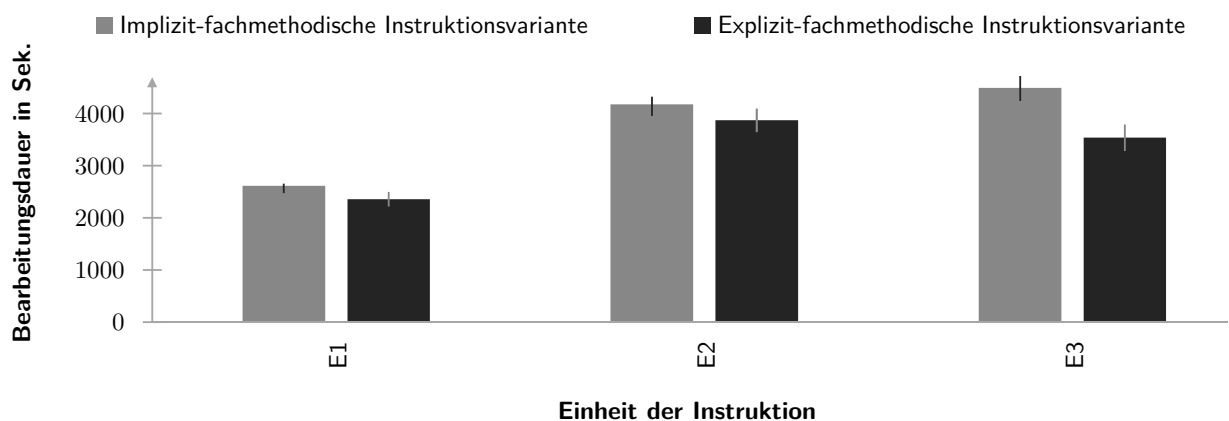


Abbildung 5.45: Mittelwerte der Gesamtbearbeitungsdauern zu beiden Instruktionsvarianten (Fehlerbalken entsprechen einer Standardabweichung des jeweiligen Mittelwerts).

Die Gesamtbearbeitungsdauern zur implizit-fachmethodischen Instruktionsvariante sind grundsätzlich größer als die zur explizit-fachmethodischen Instruktionsvariante. Wie in Tabelle 5.31 visualisiert, beträgt die Zeit, die Lernende für die Bearbeitung der implizit-fachmethodischen Instruktionsvariante mehr einsetzen als für die explizit-fachmethodische Instruktionsvariante im Mittel zwischen 4 und 15 Minuten (je nach Einheit). Die Effekte sind größtenteils groß, was darauf schließen lässt, dass die Ergebnisse sich im Sinne systematischer Unterschiede zwischen den beiden Instruktionsvarianten deuten lassen. Für die Einheiten 1 und 3 sind die Effekte bei den im Rahmen der Arbeit größtmöglichen Stichproben deutlich kleiner als beim Vergleich der Teams, die alle Einheiten bearbeitet haben. Gemäß der Anlage der Instruktion dürften daher insbesondere die Unterschiede für die dritte Einheit eher als stichprobenbedingt zu groß einzuschätzen sein, was durch die geringe Anzahl analysierter Teams zustande kommen dürfte. Für Einheit 2 kann die Stichprobe nicht deutlich vergrößert werden, allerdings finden sich in beiden Analysen die geringsten Effektstärken und relativen Unterschiede (absolut beträgt der Unterschied der Mittelwerte bei Einheit 2, die mit 90 Minuten angesetzt ist, ca. 5 Minuten, bei Einheit 1, die mit 45 Minuten angesetzt ist, ca. 4 Minuten).

Für die explizit-fachmethodische Instruktionsvariante finden sich höhere mittlere Kompetenzzuwächse als für die implizit-fachmethodische, obwohl die Bearbeitungsdauern im Mittel kürzer sind. Da die mittleren Kompetenzzuwächse der Lernenden bei der explizit-fachmethodischen Instruktionsvariante höher sind als die bei der implizit-fachmethodischen (siehe Kapitel 4), wäre es bezüglich der Bedeutung von time on task problematisch für den Vergleich, wenn die Bearbeitung der explizit-fachmethodischen Instruktionsvariante deutlich höhere Bearbeitungsdauern aufweisen würde. Dann könnten die größeren Kompetenzzuwächse aufgrund von vermehrtem Zeiteinsatz entstanden sein. Die Ergebnisse zeigen aber das Gegenteil und stellen somit die Effektivität der explizit-fachmethodischen Instruktionsvariante im Vergleich zur implizit-fachmethodischen nur noch deutlicher heraus.

5.5.1.2 Lernende mit unterschiedlichen Kompetenzzuwächsen

Innerhalb der implizit-fachmethodischen Instruktionsvariante stellt sich die Frage, ob Lernende mit unterschiedlichen Kompetenzzuwächsen unterschiedliche Bearbeitungsdauern aufweisen. Es ergibt sich:

Auf Ebene der Teams steht die Bearbeitungsdauer in keinem Zusammenhang mit den erreichten gemittelten Kompetenzzuwächsen. In Tabelle 5.33 sind die Bearbeitungsdauern der vier Teams (zwei mit gemittelt niedrigen, zwei mit gemittelt hohen Kompetenzzuwächsen) verzeichnet. Es wird deutlich, dass die

Tabelle 5.33: Übersicht über die Bearbeitungsdauern zur implizit-fachmethodischen Instruktionsvariante.

Team	U	R	S	T
Zuwachs	niedrig	niedrig	hoch	hoch
<i>Bearbeitungsdauer in Sekunden</i>				
Einheit 1	2529.1	2706.2	2541.6	2495.0
Einheit 2	3660.1	4437.0	4013.6	4638.6
Einheit 3	3983.6	4992.0	3719.2	4922.5
Gesamt	10172.8	12135.2	10274.4	12056.1

Teams U und S in den meisten Fällen niedrigere (nah beieinander liegende), die Teams R und T größere (ebenfalls nah beieinander liegende) Bearbeitungsdauern aufweisen. Für Einheit 1 weicht Team T von diesem Muster ab und es findet sich eine Bearbeitungsdauer, die deutlich geringer als die von Team R und sogar geringer als die der beiden anderen Teams ist. Werden die Teams U und S sowie die Teams R und T gepaart betrachtet, findet sich, dass Teams mit hohen gemittelten Zuwächsen für Einheit 2 jeweils tendenziell minimal länger, für Einheit 3 jeweils tendenziell minimal kürzer arbeiten – bei Zusammenfassung aller drei Einheiten ergibt sich keine solche Tendenz. Diese Ergebnisse spiegeln sich auch in t-Tests nach Student wider, die in Tabelle 5.34 berichtet sind.

Aufgrund der paarweise kleinen absoluten (Tabelle 5.33) und kleinen relativen (Tabelle 5.34) Unterschiede sowie der großen p -Werte (Tabelle 5.34) werden die Ergebnisse so interpretiert, dass die Bearbeitungsdauern in keinem Zusammenhang mit den Kompetenzzuwächsen stehen. Gegebenenfalls könnte für Einheit 1 (kürzere Bearbeitungsdauern bei höheren Kompetenzzuwächsen) und für Einheit 2 (längere Bearbeitungsdauern bei höheren Kompetenzzuwächsen) von Zusammenhängen gesprochen werden, sofern weitere Hinweise darauf gefunden werden oder sich bei größeren Stichproben dieselben Ergebnisse finden.

Tabelle 5.34: Students t-Tests zu den Gesamtbearbeitungsdauern im Vergleich von Teams mit unterschiedlichen Kompetenzzuwächsen, die die jeweilige Einheit der implizit-fachmethodischen Instruktionsvariante bearbeitet haben.

Einheit	Zuwachs Niedrig			Zuwachs Hoch			df	$t(df)$	p	r	Unterschied	
	N_N	MW	SD	N_H	MW	SD					Absolut	Relativ
1	2	2617,7	125,2	2	2518,3	33,0	2	1,085	0,391	0,609	-99,3	-3,80%
2		4048,6	549,4		4326,1	441,9		-0,557	0,634	0,366	277,6	6,42%
3		4487,8	713,0		4320,9	850,9		0,213	0,851	0,149	-167,0	-3,72%
Ges	2	11154,0	1387,6	2	11165,3	1259,9	2	-0,008	0,994	0,006	11,3	0,10%

Anmerkungen. Students t-Test; $N_{H/N}$ =Stichprobengröße für Gruppe der Teams mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Auf Ebene der Personen steht die Bearbeitungsdauer in keinem Zusammenhang mit den erreichten gemittelten Kompetenzzuwächsen. Auch wenn danach gefragt wird, ob Personen, die in Teams mit größeren Bearbeitungsdauern arbeiten, höhere Kompetenzzuwächse erreichen, ergibt sich ein ähnliches Bild wie auf Ebene der Teams. Verglichen werden dann 8 Personen mit niedrigen und 6 Personen mit hohen Kompetenzzuwächsen. Dass die Ergebnisse ähnlich zu denen im Vergleich von Teams sind, ist insbesondere deshalb relevant, weil die Teams jeweils recht heterogen bezüglich der Kompetenzzuwächse der Personen sind (das heißt, einige der Bearbeitungsdauern-Werte, die ja für alle Personen eines Teams dieselben sind, tauchen sowohl bei hohen als auch bei niedrigen Kompetenzzuwächsen auf).

Wie in Tabelle 5.35 zusammengefasst, finden sich für die einzelnen Einheiten tendenzielle Unterschiede von etwa 2–3 %, die zudem nicht signifikant sind und mit kleinen Effektstärken vorliegen. Ferner sind die Unterschiede über die Einheiten hinweg unsystematisch (anders als bspw. beim Vergleich der beiden Instruktionsvarianten, bei dem sich die Unterschiede der einzelnen Einheiten aufaddieren, siehe Tabelle 5.31). Der Unterschied beim Zusammenfassen der Einheiten (Zeile: Ges) ist von verschwindender Effektstärke. Es liegt nahe, dass Zeiteinsatz bei der implizit-fachmethodischen Instruktionsvariante in keinem Zusammenhang mit den erreichten Kompetenzzuwächsen steht.

Tabelle 5.35: Students t-Tests zu den Gesamtbearbeitungsdauern im Vergleich von Personen mit unterschiedlichen Kompetenzzuwächsen, die die jeweilige Einheit der implizit-fachmethodischen Instruktionsvariante bearbeitet haben.

Einheit	Zuwachs Niedrig			Zuwachs Hoch			df	t (df)	p	r	Unterschied	
	N _N	MW	SD	N _H	MW	SD					Absolut	Relativ
1	8	2631,6	98,3	6	2582,0	104,3	12	0,909	0,381	0,254	-49,5	-1,88%
2		4180,8	362,6		4272,7	300,5		-0,504	0,624	0,144	92,0	2,15%
3		4585,6	437,0		4443,0	592,0		0,521	0,612	0,149	-142,6	-3,11%
Ges	8	11397,9	814,1	6	11297,7	830,4	12	0,226	0,825	0,065	-100,2	-0,88%

Anmerkungen. Students t-Test; $N_{H/N}$ =Stichprobengröße für Gruppe der Teams mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Größere Bearbeitungsdauern gehen für die explizit-fachmethodische Instruktionsvariante tendenziell mit höheren Kompetenzzuwächsen einher, während es für die implizit-fachmethodische Instruktionsvariante (abgesehen von Einheit 2) andersherum und weniger deutlich ist. In Unterabschnitt 5.4.1 ist berichtet, dass für die explizit-fachmethodische Instruktionsvariante höhere Bearbeitungsdauern bei Einheit 2 sowie bei Einheit 3 tendenziell mit höheren Kompetenzzuwächsen einhergehen. Für Einheit 1 ist es andersherum. In dem vorlie-

genden Unterabschnitt 5.5.1 ist berichtet, dass sich für die implizit-fachmethodische Instruktionsvariante keine Zusammenhänge von Bearbeitungsdauern und Kompetenzzuwächsen finden; tendenziell gehen sogar kürzere Bearbeitungsdauern mit höheren Kompetenzzuwächsen einher (Einheiten 1 und 3 sowie bei Zusammenfassen aller drei Einheiten), nur für Einheit 2 ist es andersherum. Die (Beträge der) absoluten Differenzen sind für die explizit-fachmethodische Instruktionsvariante deutlich größer als für die implizit-fachmethodische (E1: ca. –170 s vs. –50 s; E2: 609 s vs. 92 s; E3: 506 s vs. –142 s), obwohl dieselben Dauern für die Einheiten angesetzt sind (E1: 45 Minuten; E2 sowie E3: 90 Minuten).

Die Ergebnisse zur implizit-fachmethodischen Instruktionsvariante basieren auf sehr kleinen Stichproben, liefern aber trotzdem Hinweise darauf, wie die Ergebnisse zur explizit-fachmethodischen Instruktionsvariante zu interpretieren sind. Auch wenn für die explizit-fachmethodische Instruktionsvariante zunächst davon ausgegangen wurde, dass die gefundenen Unterschiede aufgrund fehlender Signifikanz und fehlender Deutlichkeit in den Effekten so zu interpretieren sind, dass größere Bearbeitungsdauern nicht mit größeren erreichten Kompetenzzuwächsen zusammenhängen, so scheint dies für die explizit-fachmethodische Instruktionsvariante eher der Fall zu sein als für die implizit-fachmethodische Instruktionsvariante: Die schwachen, tendenziell eher negativen Zusammenhänge bei der implizit-fachmethodischen Instruktionsvariante scheinen darauf hinzuweisen, dass der für die explizit-fachmethodische Instruktionsvariante beschriebene Zusammenhang kein instruktionsunabhängiges Ergebnis ist, sondern spezifisch für die explizit-fachmethodische Instruktionsvariante ist. Weil für die explizit-fachmethodische Instruktionsvariante vorrangig tendenziell positive Zusammenhänge von Bearbeitungsdauer und Kompetenzzuwachs vorliegen, liegt es nahe, höhere Kompetenzzuwächse mit intensiverer Beschäftigung mit den Inhalten und Aufgaben in Verbindung zu bringen. Dies wird dadurch gestützt, dass Lernende mit hohen Kompetenzzuwächsen für nahezu alle Abschnitte der explizit-fachmethodischen Instruktionsvariante längere Bearbeitungsdauern aufweisen als Lernende mit niedrigen Kompetenzzuwächsen). An diese Deutung anschließend lässt sich – basierend auf der Tatsache, dass Lernende bei größerem Zeiteinsatz für die Bearbeitung der implizit-fachmethodischen Instruktionsvariante nicht zu im Mittel höheren Kompetenzzuwächsen gelangen – auch vermuten, dass bei der implizit-fachmethodischen Instruktionsvariante die Größe der Kompetenzzuwächse stärker von den Lernenden (Personenmerkmale etc.) als von der Instruktion abhängig ist.

5.5.2 Aktivitätsprofile

Die Analysen zu den Aktivitätsprofilen der Lernenden haben zum Ziel, die Gesamtheit der ausgewählten Aktivitäten im Vergleich der beiden Instruktionsvarianten sowie im Vergleich von Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen in den Blick zu nehmen. Dazu wird in ähnlicher Weise vorgegangen wie in Unterabschnitt 5.4.2.

Die in den Aktivitätsprofilen betrachteten Aktivitäten umfassen die verbalen Aktivitäten der Kategorien Fachinhaltlich, Fachmethodisch (mit Subkategorien Fachmethodisch-Indiziert und Fachmethodisch-Vermutet) und VorbereitenDurchführenVerbal sowie die nonverbalen Aktivitäten der Kategorie VorbereitenDurchführenNonverbal. Eine Illustration eines Aktivitätsprofils findet sich in Abbildung 5.15 auf Seite 123.

5.5.2.1 Vergleich der Instruktionsvarianten

Lernende setzen bei der implizit-fachmethodischen Instruktionsvariante in Einheit 3 größere relative Zeitanteile für verbale Aktivitäten ein als bei Einheit 2; die geringsten relativen Zeitanteile für verbale Aktivitäten finden sich für Einheit 1. Dies weicht vom Muster für die explizit-fachmethodische Instruktionsvariante ab. Analog zur Betrachtung der Anteile der verbalen Aktivitäten für die explizit-fachmethodische Instruktionsvariante in Tabelle 5.10 auf S. 124 werden die relativen Anteile der in den Aktivitätsprofilen betrachteten Aktivitäten auch für die implizit-fachmethodische Instruktionsvariante in Tabelle 5.36 aufgelöst.⁷⁸ Daraus geht hervor, dass der Anteil verbaler Aktivitäten an der tatsächlichen Gesamtbearbeitungsdauer der jeweiligen Person im Mittel für

Tabelle 5.36: Mittlerer Anteil der in den Aktivitätsprofilen betrachteten verbalen Aktivitäten an der Gesamtbearbeitungsdauer bei den einzelnen Einheiten aufgelöst auf Ebene der Personen und auf Ebene der Teams.

Einheit	Personen				Teams			
	MW	SD	Min	Max	MW	SD	Min	Max
1	6,94%	4,07%	1,14%	16,65%	19,25%	4,46%	15,18%	26,04%
2	7,67%	3,87%	2,32%	19,72%	23,07%	2,87%	18,77%	25,98%
3	9,51%	4,99%	1,72%	17,67%	24,71%	8,44%	13,57%	33,95%
Gesamt	7,95%	4,42%	1,14%	19,72%	22,53%	5,29%	13,57%	33,95%

Anmerkung. In der Gesamt-Zeile ist der Mittelwert über alle existierenden Paarungen von Einheiten mit Personen bzw. Teams aufgeführt.

⁷⁸Diese Werte weichen von den weiter unten in Abbildung 5.49a angegebenen Werten ab, weil sie anhand aller einzelnen Werte aller Personen berechnet sind, während die Werte in Abbildung 5.49a als Mittelwerte über alle personenspezifischen gewichteten Einheitenmittelwerte berechnet sind.

Einheit 3 am größten ist. Für die explizit-fachmethodische Instruktionsvariante ist der Anteil bei Einheit 2 größer als bei den anderen beiden Einheiten. Gemeinsam haben beiden Instruktionsvarianten, dass der Anteil verbaler Aktivitäten innerhalb von Einheit 1 jeweils am geringsten ist.

Lernende setzen bei der implizit-fachmethodischen Instruktionsvariante absolut tendenziell ähnlich viel Zeit für die in den Aktivitätsprofilen betrachteten verbalen Aktivitäten ein wie bei der explizit-fachmethodischen Instruktionsvariante; sie setzen absolut deutlich mehr Zeit für die nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen ein als Lernende, die die explizit-fachmethodische Instruktionsvariante bearbeiten. Aus Abbildung 5.46 auf S. 216 (Details im Anhang in den Tabellen B.3 auf S. 535 und B.4 auf S. 535) geht hervor, dass sich im Vergleich der Instruktionsvarianten nur geringe Unterschiede für die betrachteten verbalen und nonverbalen Aktivitäten auf Personenebene ergeben. Auf Teamebene finden sich allerdings sehr deutliche Unterschiede, die die nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen betreffen. Da einige der untersuchten Teams (zeitweise) nur aus zwei, statt aus drei Personen bestehen, könnte es sein, dass durch Umverteilung der notwendigen nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen die mittleren Personenanteile dieser Aktivität höher sind als in Dreier-Teams. Dieser Effekt könnte für die explizit-fachmethodische Instruktionsvariante stärker sein als für die implizit-fachmethodische (weil für erstgenannte Variante häufiger Teams nur mit zwei Personen besetzt sind), so dass sich auf Personenebene kein Unterschied bezüglich der nonverbalen Aktivitäten des Vorbereitens und Durchführens findet, obwohl er auf Teamebene sehr deutlich ist.

Dass die Unterschiede für die betrachteten verbalen Aktivitäten bei Zusammenfassung aller drei Einheiten gering und für die nonverbalen Aktivitäten des Vorbereitens und Durchführens groß sind, ergibt sich auch aus einer je spezifischen Betrachtung der Einheiten (in der jeweils mehr Teams/Personen berücksichtigt werden als in den im vorigen Absatz berichteten Analysen). In Abbildung 5.47 auf S. 216 ist aufgelöst, welche Unterschiede sich für die einzelnen Einheiten zwischen den Instruktionsvarianten ergeben (Details im Anhang in den Tabellen B.3 und B.4). Es wird deutlich, dass es sich für die hier betrachteten verbalen Aktivitäten bei Einheit 3 anders verhält als für die beiden anderen Einheiten und so vermutlich bei Zusammenfassung aller Einheiten ein Ausgleich entsteht. Für nonverbale Aktivitäten des Vorbereitens und Durchführens ergibt sich für alle Einheiten, dass Lernende in der implizit-fachmethodischen Instruktionsvariante mehr Zeit damit verbringen (für E2 und E3 sehr deutlich).

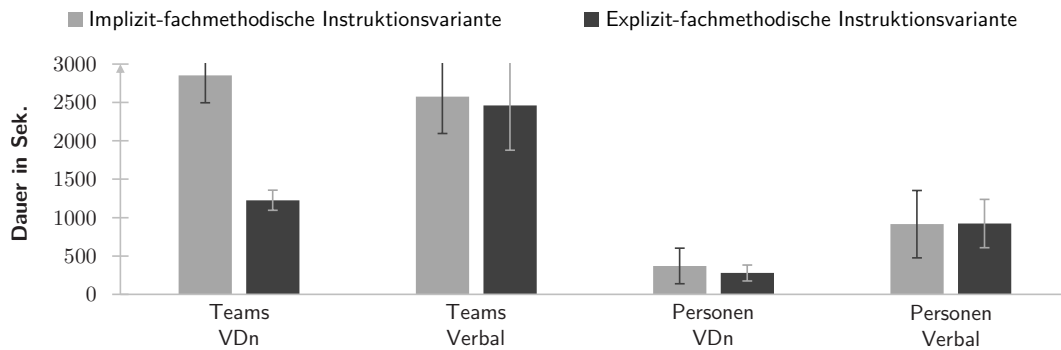


Abbildung 5.46: Mittelwerte der nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen (VDn) und aller betrachteten verbalen Aktivitäten (Fehlerbalken geben die 95 %-Konfidenzintervalle der Mittelwerte an) teamweise und personenweise im Vergleich der Instruktionsvarianten (9 Personen für die explizit-, 15 Personen für implizit-fachmethodische Variante).

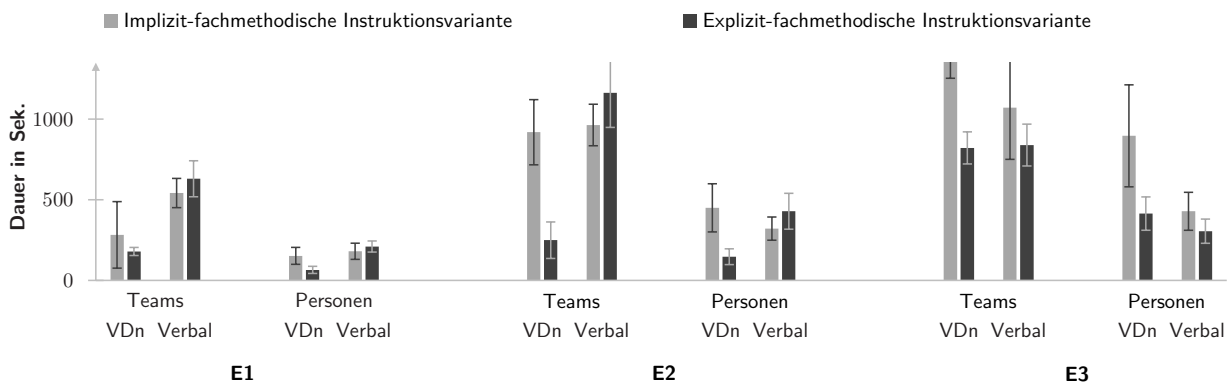


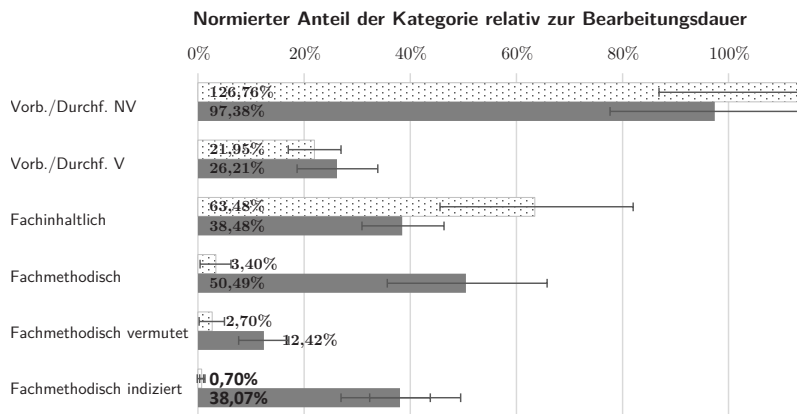
Abbildung 5.47: Mittelwerte der nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen (VDn) und aller betrachteten verbalen Aktivitäten (Fehlerbalken geben die 95 %-Konfidenzintervalle der Mittelwerte an) teamweise und personenweise im Vergleich der Instruktionsvarianten aufgeteilt nach den jeweils drei (zueinander parallelen) Instruktionseinheiten (E1, E2, E3). Personenanzahlen können aus Tabelle A.2 im Anhang entnommen werden.

In Abbildung 5.48 (auf S. 218) sind die mittleren Aktivitätsprofile normiert (durch eine einfache Division durch die je Person, je Einheit und je Instruktionsvariante vorliegenden Anteile der betrachteten verbalen Aktivitäten an der Gesamtbearbeitungsdauer) und für die beiden untersuchten Instruktionsvarianten in einem gemeinsamen Diagramm gegenübergestellt (jeweils pro Betrachtungszeitraum: über alle drei Einheiten hinweg bzw. für je eine der drei Einheiten).

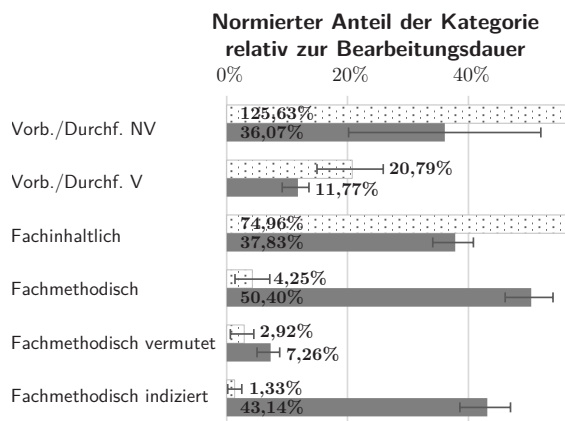
Die Normierung dient der besseren Vergleichbarkeit der beiden Instruktionsvarianten und ist als Korrektur der sich zwischen den Varianten systematisch unterscheidenden Anteile verbaler Aktivitäten anzusehen. Sie wird nur für den direkten Vergleich der Instruktionsvarianten genutzt. Für spezifisch auf die implizit-fachmethodische Instruktionsvariante bezogene Analysen, bspw. zu unterschiedlichen Kompetenzzuwächsen, werden nachfolgend die Anteile der Aktivitäten-Kategorien an der Gesamtbearbeitungsdauer genutzt (analog zu den Betrachtungen zur explizit-fachmethodischen Instruktionsvariante), zu denen sich ein Überblick in Abbildung 5.49 (auf S. 219) findet. Zum Vergleich der Instruktionsvarianten hinsichtlich der in den Aktivitätsprofilen betrachteten Kategorien lassen sich mehrere Ergebnisse benennen:

Die zeitlichen Anteile fachmethodischer Beiträge sind für die implizit-fachmethodische Instruktionsvariante deutlich geringer als für die explizit-fachmethodische – die zeitlichen Anteile der anderen Aktivitäten sind deutlich größer. Wie Balkenlängen und Fehlerbalken in Abbildung 5.49 (auf S. 219) visualisieren, finden sich für nahezu alle Aktivitäten zeitlich relevante (Balkenlängen unterscheiden sich deutlich) und häufig auch signifikante (Fehlerbalken überlappen einander nicht) Unterschiede. Anhand der normierten Abbildung 5.48 (auf S. 218) wird deutlich, dass das Ergebnis, dass bei der Zusammenfassung aller drei Einheiten die Aktivität des verbalen Vorbereitens und Durchführens von diesem Befund abweicht, sehr wahrscheinlich auf die stark verkleinerte Stichprobe zurückzuführen ist – denn für die drei Einheiten ergibt sich konsistent ein davon abweichendes Ergebnis (tendenzielle, wenngleich z. T. nicht 5 %-signifikante Unterschiede), wenn jede dieser einzeln betrachtet wird. Vermutlich sind genau die Personen, die große Anteile dieser Aktivität ausführen, nicht bei allen drei Einheiten anwesend. Es ist aber unklar, ob dies Zufall ist oder tatsächlich Zusammenhänge existent sind, auch wenn letzteres fachdidaktisch unplausibel sein dürfte.

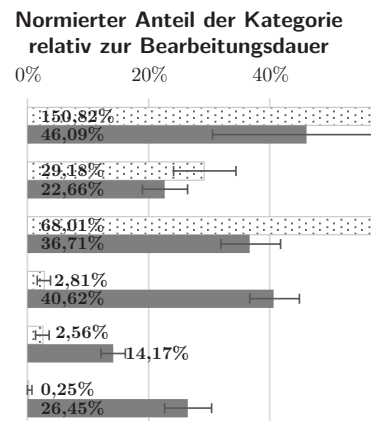
Die Unterschiede in Abbildung 5.49 (auf S. 219) sind insbesondere für die dritte Einheit sehr deutlich. Eine Betrachtung von in Abbildung 5.48 (auf S. 218) zeigt aber, dass dies vermutlich mit den dort zwischen den Instruktionsvarianten stärker ausgeprägt vorliegenden Unterschieden der Anteile von den betrachteten verbalen und nonverbalen Aktivitäten zusammenhängt (vgl. Abbildung 5.47 auf S. 216).



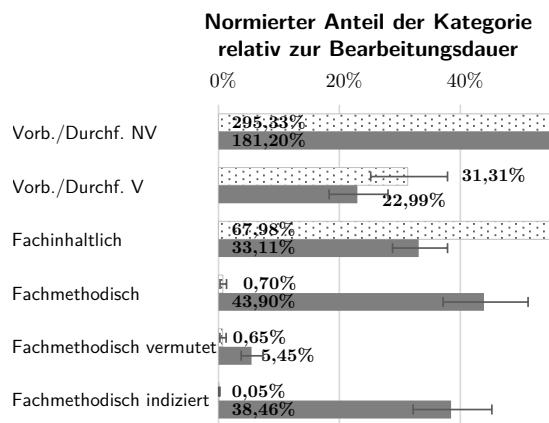
(a) **Einheiten 1 bis 3:** 9 Personen für die explizit-fachmethodische vs. 15 für die implizit-fachmethodische Instruktionsvariante



(b) **Einheit 1:** 33 Pers. für die explizit- vs. 18 für die implizit-fachmethodische IV

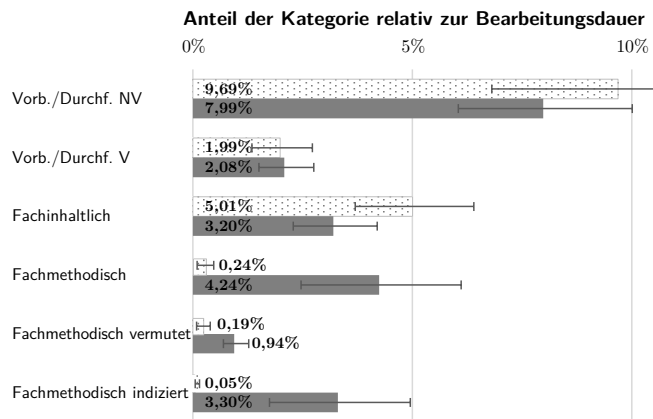


(c) **Einheit 2:** 19 Pers. für die explizit- vs. 18 für die implizit-fachmethodische IV

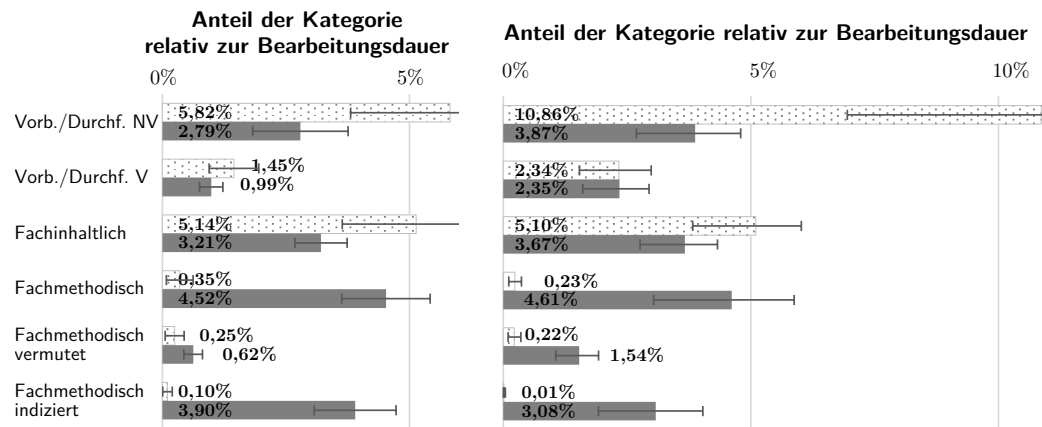


(d) **Einheit 3:** 22 Personen für die explizit-fachmethodische vs. 15 für die implizit-fachmethodische Instruktionsvariante

Abbildung 5.48: Auf den Anteil der betrachteten verbalen Aktivitäten normierte mittlere Anteile der für die Aktivitätsprofile betrachteten Aktivitäten von Personen bei der Bearbeitung der implizit-fachmethodischen (hellgrau gepunktete Balken) bzw. der explizit-fachmethodischen (dunkelgraue Balken) Instruktionsvariante. Fehlerbalken entsprechen den um die Normierungsfaktoren korrigierten 95 %-Konfidenzintervallen der Mittelwerte.

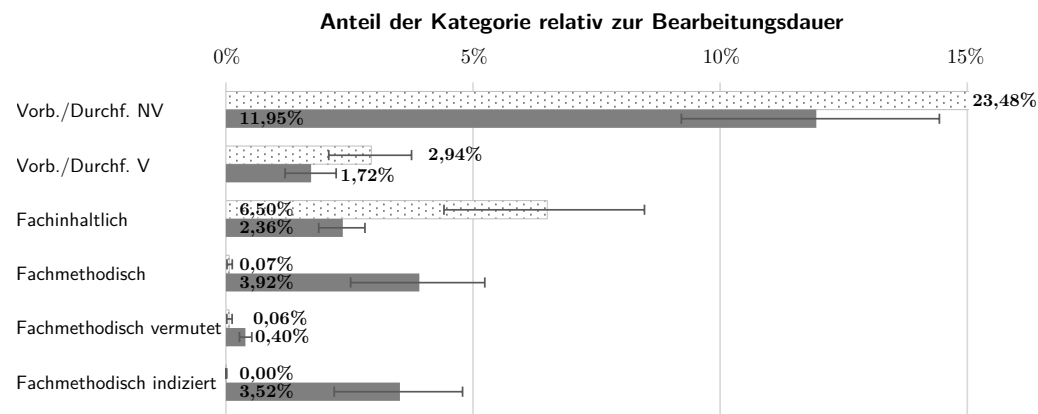


(a) **Einheiten 1 bis 3:** 9 Personen für die explizit-fachmethodische vs. 15 für die implizit-fachmethodische Instruktionsvariante



(b) **Einheit 1:** 33 Pers. für die explizit- vs. 18 für die implizit-fachmethodische IV

(c) **Einheit 2:** 19 Personen für die explizit-fachmethodische vs. 18 für die implizit-fachmethodische IV



(d) **Einheit 3:** 22 Personen für die explizit-fachmethodische vs. 15 für die implizit-fachmethodische Instruktionsvariante

Abbildung 5.49: Mittlere Anteile der betrachteten Aktivitäten von Personen bei der Bearbeitung der implizit-fachmethodischen (hellgrau gepunktete Balken) bzw. der explizit-fachmethodischen (dunkelgraue Balken) Instruktionsvariante. Fehlerbalken geben die 95 %-Konfidenzintervalle der Mittelwerte an.

Insgesamt ist davon auszugehen, dass die in den jeweils die drei Einheiten zusammenfassenden Diagrammen dargestellten Tendenzen die mittleren Aktivitätsprofile für beide Instruktionsvarianten und die somit vorliegenden Unterschiede angemessen darstellen. Die eingetragenen Konfidenzintervalle dürften für die Population allerdings zu groß sein und damit dürfte die Deutlichkeit der Unterschiede als zu gering eingeschätzt werden; dafür spricht, dass sich für die einzelnen Einheiten (und somit für die jeweils größeren Stichproben) deutlichere Unterschiede ergeben.

Es ist besonders anzumerken, dass der Anteil fachinhaltlicher Beiträge an den betrachteten verbalen Beiträgen bei der implizit-fachmethodischen Instruktionsvariante über 60 % (bei der explizit-fachmethodischen unter 40 %) und der Anteil fachmethodischer Beiträge bei der explizit-fachmethodischen Instruktionsvariante über 40 % (bei der implizit-fachmethodischen unter 5 %) beträgt. Vor dem Hintergrund der breit formulierten Kategoriendefinition zu **Fachinhaltlich** ist es angemessen, dass sich für beide Instruktionsvarianten große Anteile fachinhaltlicher Beiträge finden; es ist jedoch deutlich, dass sich fachmethodische Beiträge in beachtenswertem Umfang nur für die explizit-fachmethodische Instruktionsvariante finden, während die Anteile fachmethodischer Beiträge für die implizit-fachmethodische Instruktionsvariante relativ zu den betrachteten verbalen Beiträgen (unter 5 %) sowie relativ zur Gesamtbearbeitungsdauer (unter 0.5 %) sehr gering ausfallen.

Anders als bei der explizit-fachmethodischen Instruktionsvariante sind die Anteile fachmethodischer-vermuteter Beiträge bei der implizit-fachmethodischen Instruktionsvariante höher als die Anteile fachmethodisch-indizierter Beiträge. Auf Fachmethoden bezogene Äußerungen von Lernenden, die die implizit-fachmethodische Instruktionsvariante bearbeiten, nehmen also vorwiegend indirekt und unklar Bezug zu den Fachmethoden und sind selten auf spezifische fachmethodische Konzepte, Begriffe oder Regeln bezogen. Es liegt nahe, die Formulierung fachmethodisch-indizierter Beiträge bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante als Ergebnis der spezifischen Instruktionsanlage anzusehen. Da die Unterschiede zwischen den Instruktionsvarianten auch mit Blick auf fachmethodisch-vermutete Beiträge groß (wenngleich nicht so groß wie für fachmethodisch-indizierte) sind, liegt es auch hier nahe, diese mit der Anlage der Instruktionsvarianten in Zusammenhang zu bringen. Eine ausführlichere Diskussion erfolgt in Kapitel 7.

Anders als für die explizit-fachmethodische Instruktionsvariante unterscheiden sich die mittleren Aktivitätsprofile zu den drei Einheiten der implizit-fachmethodischen Instruktionsvariante nur geringfügig. Aus den Abbildungen 5.48 und 5.49 (auf den Seiten 218 und 219) geht hervor, dass die Anteile der einzelnen Kategorien zu den verbalen Beiträgen der Lernenden bei der

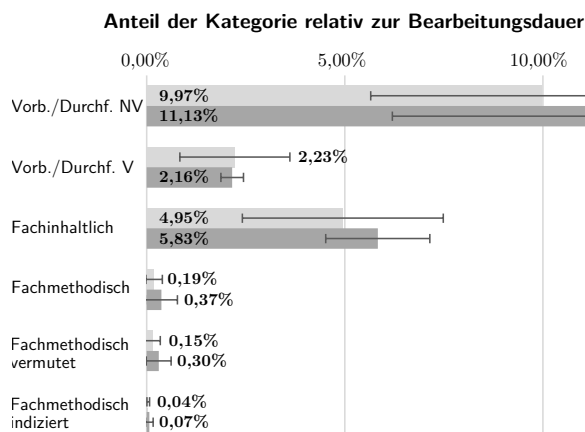
Bearbeitung der implizit-fachmethodischen Instruktionsvariante sich nur geringfügig unterscheiden. Die deutlichsten Unterschiede zwischen den Einheiten finden sich für die Kategorie der verbalen Beiträge des Vorbereitens und Durchführens von Versuchen, die auf eher technische Anweisungen u. ä. bezogen sind. Diese gehen (leicht sublinear) mit den Werten für die Aktivität der nonverbalen Aktivitäten zum Vorbereiten und Durchführen von Versuchen einher. Dass sich für die nonverbalen Aktivitäten der Kategorie `VorbereitenDurchführenNonverbal` Unterschiede zwischen den Einheiten finden, entspricht zum einen den Ergebnissen zur explizit-fachmethodischen Instruktionsvariante und ist zum anderen anhand der im Instruktionsmaterial parallelisiert eingesetzten Versuche und deren Umfang nachvollziehbar.

5.5.2.2 Lernende mit unterschiedlichen Kompetenzzuwächsen

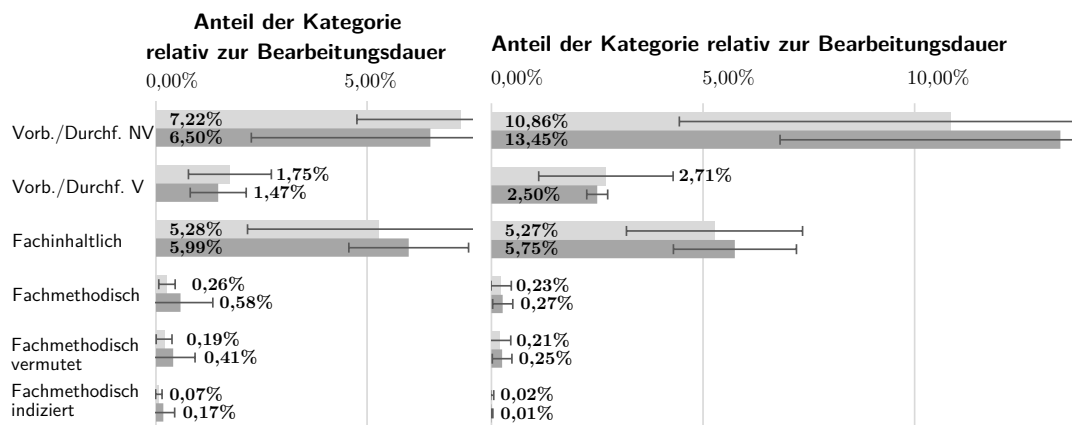
Der zeitliche Anteil fachlicher – aber nicht spezifisch fachmethodischer – verbaler und kognitiver Aktivität könnte in Zusammenhang mit erreichten Kompetenzzuwächsen stehen. In Abbildung 5.50 (auf S. 222) sind die mittleren Aktivitätsprofile für die implizit-fachmethodische Instruktionsvariante gruppiert nach Personen mit hohen vs. niedrigen Kompetenzzuwächsen gegenübergestellt.⁷⁹ Der Anteil fachmethodischer sowie fachinhaltlicher Beiträge ist für Personen mit hohen Kompetenzzuwächsen in allen vier Diagrammen Abbildung 5.50 tendenziell höher als für Personen mit niedrigen Kompetenzzuwächsen. Der Anteil verbaler Aktivitäten zum Vorbereiten und Durchführen von Versuchen schwankt für Personen mit niedrigen Kompetenzzuwächsen deutlich stärker als für Personen mit hohen Kompetenzzuwächsen und er ist in etwa so groß wie der Anteil für Personen mit hohen Kompetenzzuwächsen (außer bei Einheit 3).

Zusammengenommen könnte diese Werte darauf hinweisen, dass Lernende mit hohen Kompetenzzuwächsen durch insgesamt größere Anteile von solchen verbalen Aktivitäten gekennzeichnet sind, die auf kognitive Aktivität der Lernenden schließen lassen. Das Ergebnis bei Einheit 3 könnte dann so plausibilisiert werden, dass die fachmethodischen Beiträge dort nur einen sehr geringen Anteil einnehmen, aber stattdessen ein tendenzieller Unterschied für die Kategorie `VorbereitenDurchführenVerbal` auftritt (etwa, weil die Kategorie `Fachmethodisch-Vermutet` für die dritte Einheit nur unzureichend kodiert wurde, was den identifizierten Problemen im Kodiermanual entspräche, oder weil es für das Auswerten und Interpretieren von Versuchen

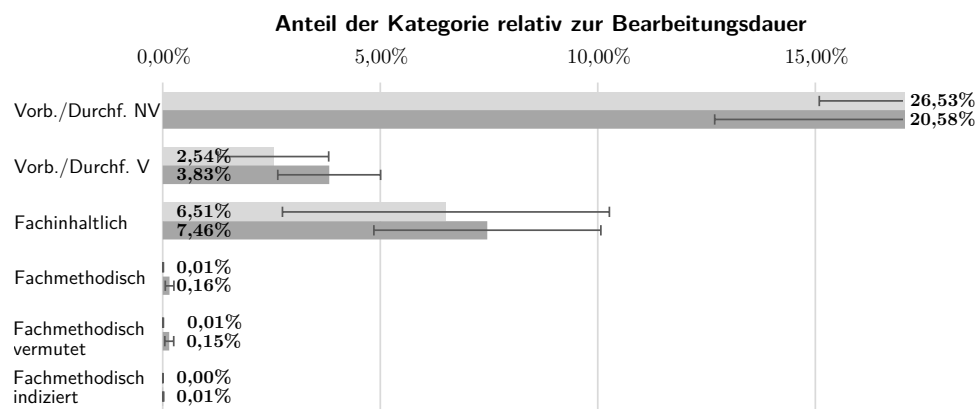
⁷⁹Die Werte erscheinen auf den ersten Blick zum Teil inkompatibel mit den in Abbildung 5.49 vermerkten Werten zur implizit-fachmethodischen Instruktionsvariante, weil beispielsweise beide vorliegenden Werte niedriger sind als der dort angegebene Mittelwert. Dies erklärt sich allerdings dadurch, dass die Gruppen der Personen mit mittleren sowie negativen Kompetenzzuwächsen nicht in die hier vorgenommene Gegenüberstellung einbezogen werden.



(a) **Einheiten 1 bis 3:** 7 Personen mit niedrigen vs. 6 mit hohen Kompetenzzuwächsen



(b) **Einheit 1:** 8 Personen mit niedrigen vs. 6 mit hohen Kompetenzzuwächsen (c) **Einheit 2:** 8 Personen mit niedrigen vs. 6 mit hohen Kompetenzzuwächsen



(d) **Einheit 3:** 7 Personen mit niedrigen vs. 6 mit hohen Kompetenzzuwächsen

Abbildung 5.50: Mittlere Anteile der betrachteten Aktivitäten von Personen bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante (hellgraue Balken gehören zu niedrigen, dunkelgraue zu hohen Kompetenzzuwächsen). Fehlerbalken geben die 95 %-Konfidenzintervalle der Mittelwerte an.

bereits ausreichend ist, die Tätigkeiten zu beschreiben und vermutete oder indizierte Hinweise auf Fachmethoden nicht notwendig sind).

Lernende mit hohen Kompetenzzuwächsen gehören zu Aktivitätsprofiltypen, die größere zeitliche Anteile fachmethodisch-vermuteter Beiträge aufweisen als andere Aktivitätsprofiltypen. Aufgrund der Ähnlichkeit der Einheiten wird eine Analyse der Aktivitätsprofile nur für die Zusammenfassung aller drei Einheiten vorgenommen. Dabei wird der absolute Anteil der jeweiligen Aktivität über alle drei Einheiten der implizit-fachmethodischen Instruktionsvariante hinweg durch die Dauer der bearbeiteten Einheiten geteilt; für die drei Personen, die die dritte Einheit nicht bearbeitet haben (Q51, R53, S56) wird also der Absolutwert der jeweiligen Aktivität durch die reale Bearbeitungsdauer der ersten beiden Einheiten geteilt, für alle anderen Personen durch die reale Bearbeitungsdauer aller drei Einheiten. Dieses Vorgehen erschien plausibel, weil die mittleren Anteile der Aktivitäten sich zwischen den Einheiten nur geringfügig unterscheiden, wenngleich es natürlich mit Einschränkungen verbunden ist, die die drei genannten Personen betreffen. Es kann jedoch angemerkt werden, dass die Aktivitätsprofile der drei Personen den anderen Aktivitätsprofilen stark ähneln und sich ohne Weiteres den Aktivitätsprofiltypen zuordnen lassen. Methodisch ist anzumerken, dass im zweiten Schritt des Vorgehens zum Clustern (siehe Unterunterabschnitt 5.4.2.1) anstelle der Pearson-Korrelation als Maß für die Ähnlichkeit die quadrierte euklidische Distanz gewählt wird, weil dadurch den für die gesamte implizit-fachmethodische Instruktionsvariante sehr kleinen Anteile fachmethodischer Beiträge besser Rechnung getragen werden kann. Zu den Ergebnissen ist anzumerken, dass sich aus dem ersten Schritt keine Partitionen ergeben, die im zweiten Schritt vorgeschlagenen Cluster sich allerdings gut anhand der Anteile nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen unterteilen lassen.

Für die implizit-fachmethodische Instruktionsvariante ergeben sich fünf Aktivitätsprofiltypen (Kurzübersicht in Tabelle 5.37 auf S. 227). In Abbildung 5.51 ist für jeden der nachfolgend beschriebenen Aktivitätsprofiltypen ein Repräsentant dargestellt:

- (1) Aktivitätsprofiltypen, bei denen der Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen etwa 20 % beträgt (groß ist):

E1b3-P-C1 Der zeitliche Anteil der fachinhaltlichen Beiträge ist groß (etwa 7 %). Der zeitliche Anteil fachmethodischer Beiträge ist sehr gering. (Kurznotation: $VD_n \approx 20\%$; $FI \approx 7\%$; $FM \approx 0.$)

Personen: Q50, R54, R55, T60, U62.

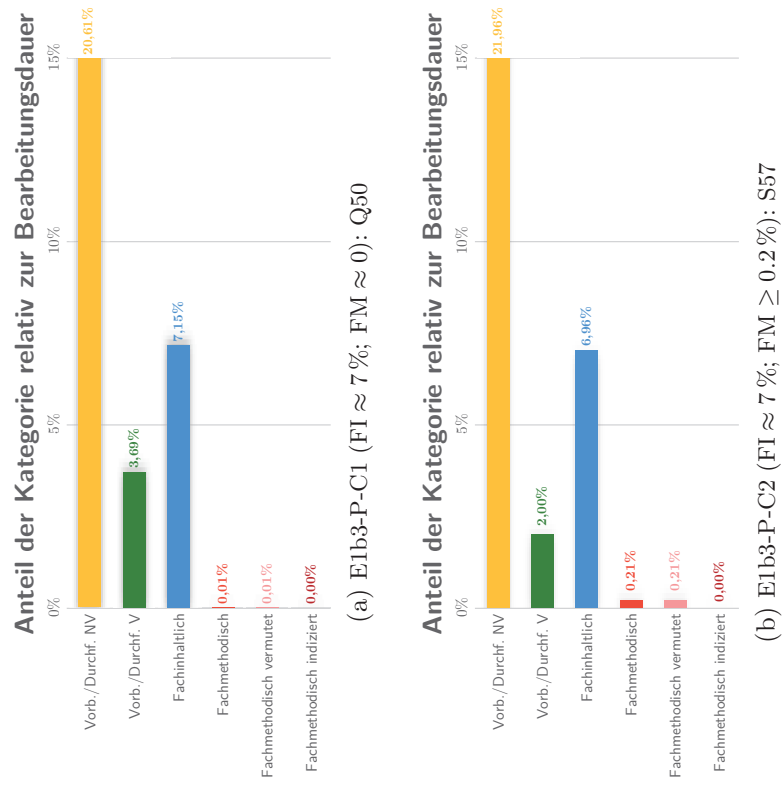


Abbildung 5.51: Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Zusammenfassung der Einheiten der implizit-fachmethodischen Instruktionsvariante (Teil 1 der Abbildung: VDN $\approx$ 20 %).

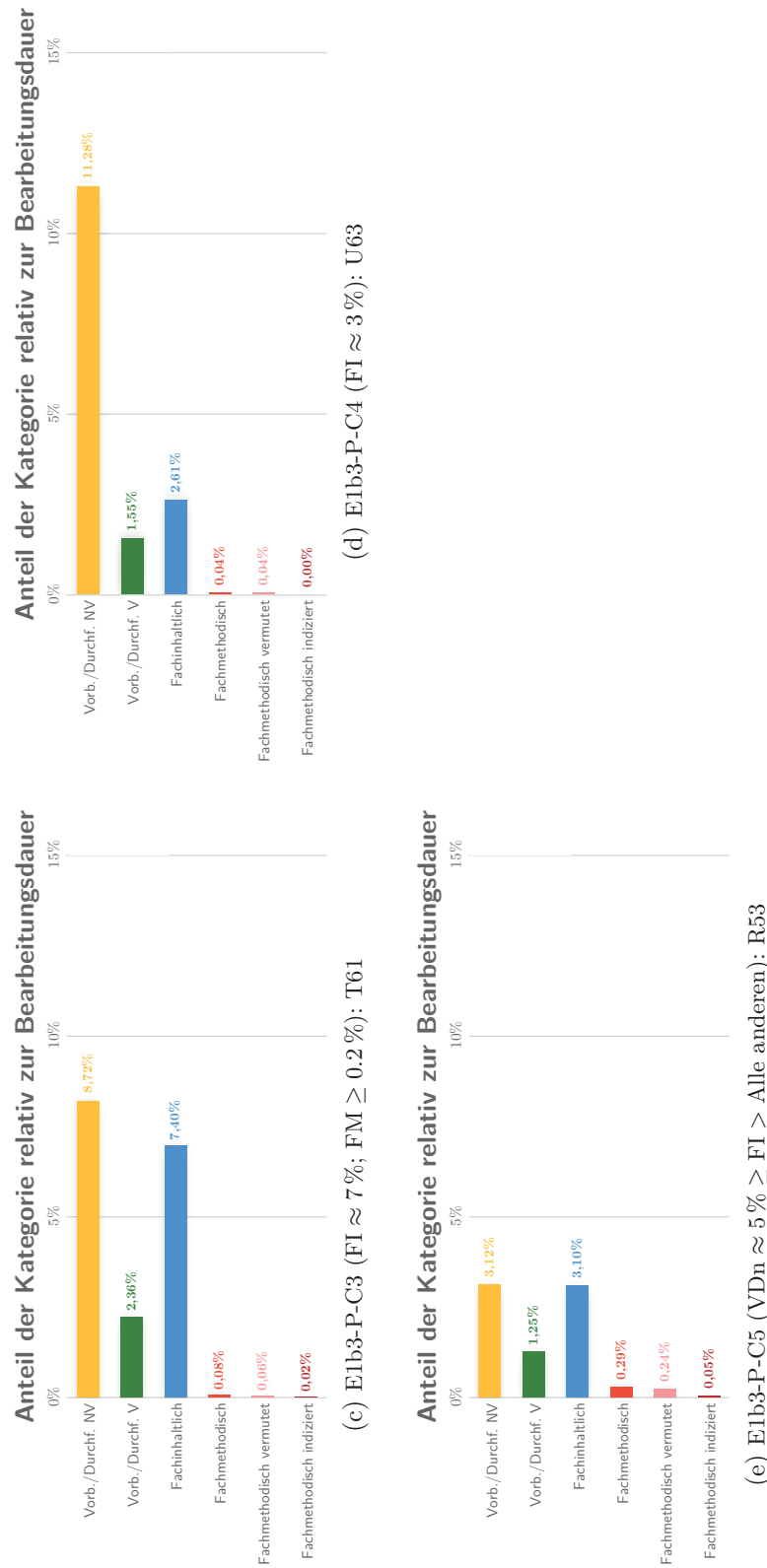


Abbildung 5.51: Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Zusammenfassung der Einheiten der implizit-fachmethodischen Instruktionsvariante (Teil 2 der Abbildung: c & d mit VDn $\approx 10\%$, e mit VDn $\approx 5\%$).

E1b3-P-C2 Wie E3-P-C1, aber der Anteil fachmethodischer Beiträge ist relativ zu anderen Aktivitätsprofiltypen stärker ausgeprägt (größer als 0.2 %). (Kurznotation: $VDn \approx 20\%$; $FI \approx 7\%$; $FM \geq 0.2\%$.)
Personen: S57, S58, V65.

- (2) Aktivitätsprofiltypen, bei denen der Anteil nonverbaler Aktivitäten des Vorbereitens u. Durchführens von Versuchen ca. 10 % beträgt (mittelgroß ist):

E1b3-P-C3 Bis auf den Anteil nonverbaler Aktivitäten wie E1b3-P-C2. (Kurznotation: $VDn \approx 10\%$; $FI \approx 7\%$; $FM \geq 0.2\%$.)
Personen: Q52, T59, T61.

E1b3-P-C4 Der zeitliche Anteil der fachinhaltlichen Beiträge ist gering (etwa 3 %). (Der zeitliche Anteil fachmethodischer Beiträge ist für die zugehörigen Aktivitätsprofile verschieden.)
(Kurznotation: $VDn \approx 10\%$; $FI \approx 3\%$.)
Personen: U63, V67.

- (3) Aktivitätsprofiltypen, bei denen der Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen kleiner als 5 % ist (klein ist):

E3-P-C5 Der zeitliche Anteil fachinhaltlicher Beiträge ist kleiner als der zeitliche Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen (und damit klein, ähnlich wie E3-P-C4). Der nächstgrößere Anteil entfällt auf die Aktivität verbaler Beiträge zum Vorbereiten und Durchführen. Zudem ist (ähnlich wie in E3-P-C4) der zeitliche Anteil fachmethodischer Beiträge für die zugehörigen Aktivitätsprofile verschieden.
(Kurznotation: $VDn \approx 5\% \geq FI > \text{Alle anderen.}$)
Personen: Q51, R53, S56, U64, V66.

Aus Tabelle 5.37 geht hervor, dass die Aktivitätsprofiltypen sehr klar nach Personen mit hohen vs. niedrigen Kompetenzzuwächsen differenzieren (obwohl die Typenbildung unabhängig davon stattgefunden hat). Einschränkend ist allerdings anzumerken, dass die Unterschiede zwischen den Profilen nicht so deutlich sind wie beispielsweise in den Clusteranalysen zur explizit-fachmethodischen Instruktionsvariante; insbesondere wurden drei der 18 untersuchten Personen nachträglich Aktivitätsprofiltypen zugewiesen, denen sie gemäß der statistisch vorgeschlagenen Cluster nicht zugehörig sind (bspw. ist der Unterschied zwischen E1b3-P-1 und E1b3-P-2 nicht ganz eindeutig und eine Person wurde vor dem Abgleich mit den Kompetenzzuwächsen händisch umsortiert).

Tabelle 5.37: Übersicht über alle Aktivitätsprofiltypen und die Anzahl der Personen mit hohen und niedrigen Kompetenzzuwächsen. Die Anzahl der Personen mit mittleren Kompetenzzuwächsen ist nicht in der Tabelle aufgeführt.

Aktivitätsprofiltyp		Anzahl d. Personen		
Name	Kurzbeschreibung	Insg. im Profil	Zuwachs hoch	Zuwachs niedrig
E1b3-P-C1	VDn $\approx$ 20 %; FI $\approx$ 7 %; FM $\approx$ 0	5	0	5
E1b3-P-C2	VDn $\approx$ 20 %; FI $\approx$ 7 %; FMv $\geq$ 0.2 %	3	3	0
E1b3-P-C3	VDn $\approx$ 10 %; FI $\approx$ 7 %; FMv $\geq$ 0.2 %	3	3	0
E1b3-P-C4	VDn $\approx$ 10 %; FI $\leq$ 3 %	2	0	0
E1b3-P-C5	VDn $\approx$ 5 % $\geq$ FI > Alle anderen	5	0	3

Anmerkungen. Zur Kurzbeschreibung: Verglichen werden Anteile von Aktivitäten an der Gesamtbearbeitungsdauer. Aktivitäten: FMI = Fachmethodisch-Indiziert; FMv = Fachmethodisch-Vermutet; FI = Fachinhaltlich; VDn = VorbereitenDurchführenNonverbal; VDv = VorbereitenDurchführenVerbal. Verhältnisse der Anteile: $\gg$ = Viel größer als; $>$ = größer als; $\geq$ = nur minimal größer als; $\approx$ = etwa gleich groß.

Für die beiden Aktivitätsprofiltypen, die durch relativ⁸⁰ hohe Anteile fachmethodisch(-vermutet)er Beiträge gekennzeichnet sind (E1b3-P-C1, E1b3-P-C2) finden sich fast nur nur Lernende mit hohen Kompetenzzuwächsen. Auch andersherum finden sich Lernende mit hohen Kompetenzzuwächsen meist in den Aktivitätsprofiltypen mit relativ hohen Anteilen fachmethodischer Beiträge. Eine Ausnahme stellt R53 dar, der relativ hohe Anteile fachmethodischer Beiträge aufweist und dem diesbezüglich durchmischten Aktivitätsprofiltyp E1b3-P-5 zugehörig ist. Er erreicht niedrige Kompetenzzuwächse; allerdings ist bei ihm der Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens recht niedrig – und auch die Anteile aller anderen Aktivitäten sind niedrig, was sich eventuell im Sinne insgesamt geringer kognitiver Aktivität deuten lässt (vgl. voriges Ergebnis zu Aktivitätsprofilen im Mittel).

Hohe Anteile nonverbaler Aktivitäten des Vorbereitens und Durchführens allein hängen nicht mit hohen Kompetenzzuwächsen zusammen (siehe Aktivitätsprofiltypen E1b3-P-1 und E1b3-P-2), sehr niedrige Anteile könnten allerdings mit niedrigen Kompetenzzuwächsen zusammenhängen. Vergleichsweise große Anteile fachinhaltlicher Beiträge (E1b3-P-1, E1b3-P-2 und E1b3-P-3) treten tendenziell eher mit hohen Kompetenzzuwächsen gemeinsam auf (aber E1b3-P-1). Finden sich sowohl relativ große zeitliche Anteile fachmethodisch(-vermutet)er Beiträge als auch relativ große

⁸⁰Relativ bezieht sich dabei auf die implizit-fachmethodische Instruktionsvariante; im Vergleich zur explizit-fachmethodischen Instruktionsvariante sind die Anteile fachmethodischer Aktivitäten sehr gering.

zeitliche Anteile fachinhaltlicher Beiträge und wird zudem hinreichend viel Zeit für nonverbale Aktivitäten des Vorbereitens und Durchführens von Versuchen eingesetzt (Typen E1b3-P-2 und E1b3-P-3), finden sich hohe Kompetenzzuwächse. Insgesamt lässt sich dies so deuten, dass neben dem positiven Zusammenhang von Kompetenzzuwächsen und fachmethodischen Beiträgen also ein positiver Zusammenhang von Kompetenzzuwachs und kognitiver Aktivität (im Kontext ausreichend gegebenen nonverbalen Experimentierens) vorliegt – zumindest dann, wenn davon ausgegangen wird, dass fachlichen (fachinhaltlichen und fachmethodischen) Aktivitäten Hinweise auf die kognitive Aktivität der Lernenden geben können.

5.5.3 Fachmethodische Beiträge

Ein Ziel der Arbeit besteht darin, die Bedeutung von Äußerungen mit fachmethodischem Gehalt zu untersuchen. In ähnlicher Weise wie in Unterabschnitt 5.4.4 wird nachfolgend detailliert auf die fachmethodischen Beiträge der Lernenden eingegangen, die die implizit-fachmethodische Instruktionsvariante bearbeitet haben. Wie bereits in den vorigen Abschnitten wird zunächst ein grober Abgleich mit der explizit-fachmethodischen Instruktionsvariante vorgenommen (Unterabschnitt 5.5.3.1), bevor Details zur implizit-fachmethodischen Instruktionsvariante berichtet werden (Unterabschnitt 5.5.3.2); zuletzt schließt sich ein auf die einzelnen Personen bezogener Vergleich der Instruktionsvarianten an, der differenzierte Ergebnisse zum Vergleich beiträgt (Unterabschnitt 5.5.3.3).

5.5.3.1 Vergleich der Instruktionsvarianten

Im Unterschied zu den Aktivitätsprofilen für die explizit-fachmethodische Instruktionsvariante finden sich, wie in den vorigen Unterabschnitten berichtet, in den Aktivitätsprofilen für die Bearbeitung der implizit-fachmethodischen Instruktionsvariante große zeitliche Anteile fachinhaltlicher Beiträge, während die zeitlichen Anteile fachmethodischer Beiträge sehr gering sind. Da der Fokus der implizit-fachmethodischen Instruktionsvariante nicht auf den Fachmethoden liegt, ist dieses Ergebnis plausibel. Auch dass sich für die implizit-fachmethodische Instruktionsvariante ein größerer zeitlicher Anteil für fachmethodisch-vermutete Beiträge als für fachmethodisch-indizierte Beiträge findet, ist erwartbar, weil direkte, klar auf adressierte Konzepte bezogene oder regelhafte Äußerungen zu Fachmethoden innerhalb der implizit-fachmethodischen Instruktionsvariante, welche die Konzepte nur implizit adressiert, kaum zu erwarten sind. Umso relevanter ist das Ergebnis, dass sich in Aktivitätsprofiltypen mit vergleichsweise großen zeitlichen Anteilen fachmethodisch-vermuteter Beiträge nur Lernende mit hohen Kompetenzzuwächsen finden (siehe Unterunterab-

schnitt 5.5.2.2, ab S. 221). Auch andersherum finden sich alle Lernenden mit hohen Kompetenzzuwächsen in solchen Typen (siehe ebenda). Obwohl der zeitliche Anteil fachmethodisch-vermuteter Beiträge also relativ zur Gesamtbearbeitungsdauer gering ausfällt (deutlich geringer als bei der explizit-fachmethodischen Instruktionsvariante) und noch deutlich weniger Zeit auf fachmethodisch-indizierte Beiträge entfällt, scheinen diese Beiträge in einem Zusammenhang mit dem Kompetenzzuwachs der Lernenden zu stehen.

Im Unterschied zu der explizit-fachmethodischen Instruktionsvariante finden sich für die Einheiten der implizit-fachmethodischen Instruktionsvariante keine klaren Schwerpunkte in den zu Teilprozessen des experimentbezogenen Denkens und Arbeitens existierenden Bezügen der fachmethodischen Beiträge. Aus Abbildung 5.52 (auf S. 230) und Tabelle 5.38 (auf S. 229) geht stattdessen hervor, dass fachmethodische Beiträge mit Bezügen zum Planen und Durchführen von Untersuchungen insgesamt und in allen Einheiten deutlich überwiegen. Interessanterweise finden sich Beiträge mit Bezügen zum Formulieren von Fragen und Hypothesen nur bei der ersten Einheit der implizit-fachmethodischen Instruktionsvariante, während sich Beiträge mit Bezügen zum Teilprozess des Auswertens und Interpretierens von Daten in allen drei Einheiten finden, sogar in der dritten Einheit, in der insgesamt nur sehr geringe zeitliche Anteile fachmethodischer Beiträge vorgefunden werden.

Dass sich anders als für die explizit-fachmethodische Instruktionsvariante bei der implizit-fachmethodischen Instruktionsvariante keine Schwerpunkte in den Bezügen zu gewissen Teilprozessen des experimentbezogenen Denkens und Arbeitens zeigen, ist vor dem Hintergrund der Anlage der Instruktionsvariante plausibel. Die implizit-fachmethodische Instruktionsvariante ist zwar implizit entlang der Konzepte des

Tabelle 5.38: Zeitliche Anteile fachmethodischer Beiträge (absolut) aufgelöst nach Zugehörigkeit zu den Teilprozessen des experimentbezogenen Denkens und Arbeitens (FH: Formulieren von Fragen und Hypothesen; PU: Planen und Durchführen von Untersuchungen; AI: Auswerten und Interpretieren von Daten) und nach den Einheiten der explizit-fachmethodischen Instruktionsvariante (E1, E2, E3).

Bezug zu Teilprozess	E1		E2		E3		Gesamt			
							Fall 1		Fall 2	
FH	1,1	(2)	0	(0)	0	(0)	1,1	(2)	0,7	(1,5)
PU	6,7	(10,4)	8,2	(11,1)	1,5	(2,4)	16,2	(20,5)	14,9	(21,3)
AI	1,6	(4,9)	1,3	(2,7)	1,4	(2,5)	4,1	(6,1)	3,9	(5,8)

Anmerkungen. Angaben in Sekunden; Standardabweichung in Klammern. Fall 1 = Mittelwert über alle 12 Personen, die mindestens eine Einheit bearbeitet haben. Fall 2 = Mittelwert für die 9 Personen, die bei allen Einheiten anwesend waren.

Mittlerer Anteil der Kategorie an der Bearbeitungsdauer der jeweiligen Einheit

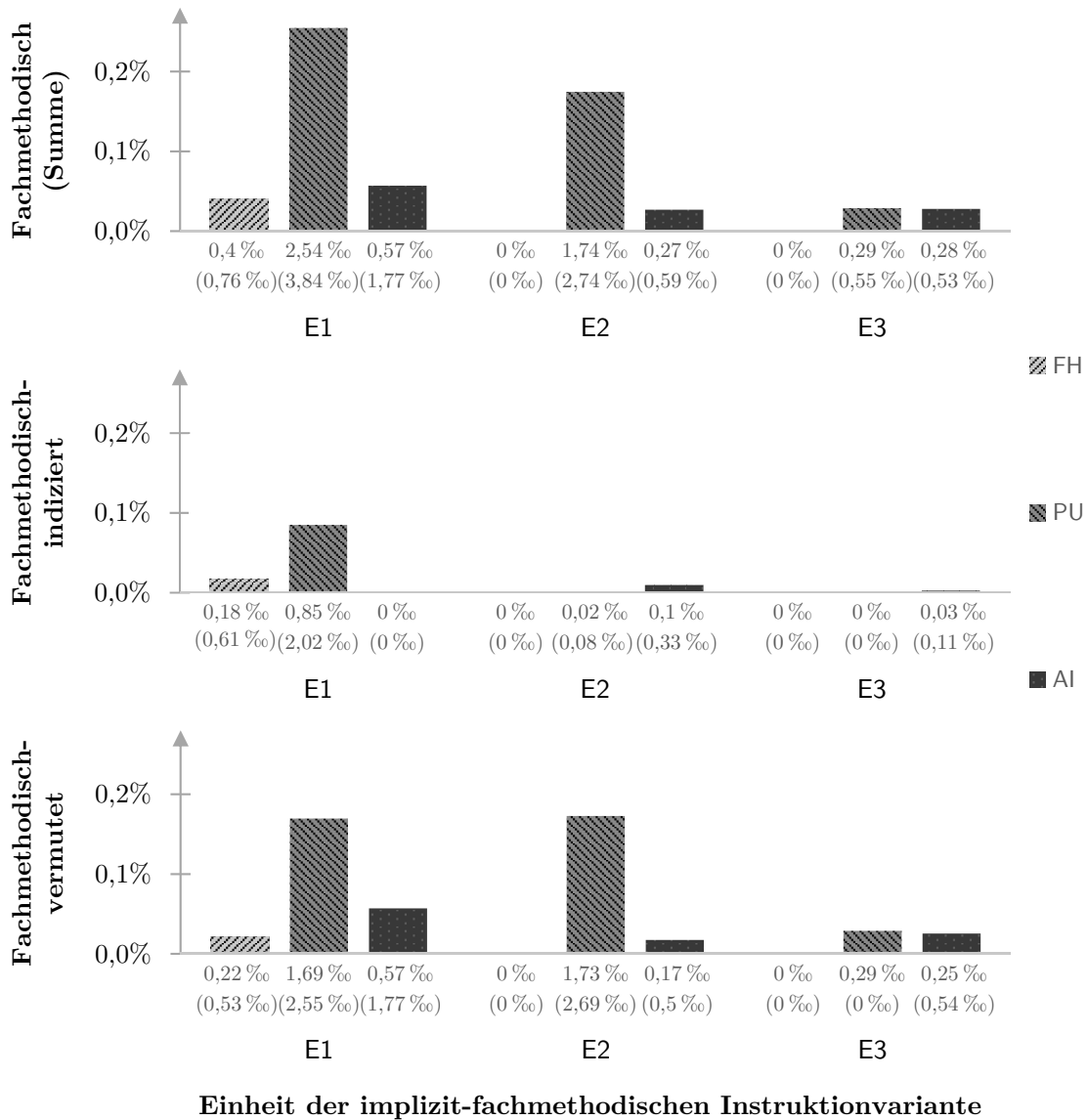


Abbildung 5.52: Zeitliche Anteile fachmethodischer Beiträge (oben: Summe; mitte: indizierte; unten: vermutete) aufgelöst nach Zugehörigkeit zu den Teilprozessen des experimentbezogenen Denkens und Arbeitens (verschieden schraffiert) und nach den Einheiten der explizit-fachmethodischen Instruktionsvariante (E1, E2, E3). Angegeben ist jeweils der Mittelwert über die Personen, die die jeweilige Einheit bearbeitet haben (siehe Tabelle A.2 im Anhang). In Klammern ist die jeweilige Standardabweichung angegeben.

experimentbezogenen Denkens und Arbeitens strukturiert; es werden allerdings keine explizit auf Fachmethoden gerichteten Aufgaben gestellt, die zur Fokussierung spezifischer Fachmethoden und damit auch spezifischer Teilprozesse anregen würden. Dass sich insgesamt vor allem Beiträge mit Bezügen zum Planen und Durchführen von Untersuchungen sowie zum Auswerten und Interpretieren von Daten finden, könnte damit zusammenhängen, dass in allen drei Einheiten der implizit-fachmethodischen Instruktionsvariante Untersuchungen durchgeführt werden müssen, auf die zum Teil auch instruktional angeregte Tätigkeiten des Auswertens und Interpretierens folgen. Da allerdings nicht grundsätzlich davon auszugehen ist, dass Lernende begleitend zu den Tätigkeiten auch vermutete oder gar indizierte fachmethodische Beiträge machen, dürfte die implizite Adressierung der Konzepte von Bedeutung für entsprechende fachmethodische Beiträge sein. Als Frage schließt sich unter anderem an, ob bei (intendiert oder unabsichtlich) fehlerhaft angelegten Experimentieraufträgen nicht sogar mehr fachmethodische Beiträge vorkommen würden als bei der hier eingesetzten impliziten Adressierung.

5.5.3.2 Lernende mit unterschiedlichen Kompetenzzuwächsen

Die Bedeutung der fachmethodisch-vermuteten Beiträge innerhalb der implizit-fachmethodischen Instruktionsvariante wird durch einen Befund aus einer einheitenspezifischen Analyse unterstrichen. Einheit 3 der implizit-fachmethodischen Instruktionsvariante wird nur von Lernenden bearbeitet, die an allen drei Einheiten teilgenommen haben. Werden die Personen mit hohen Kompetenzzuwächsen den Personen mit niedrigen Kompetenzzuwächsen gegenübergestellt, findet sich folgendes Ergebnis:

- Für alle 6 Personen mit hohen Kompetenzzuwächsen (bis auf eine) liegt innerhalb von Einheit 3 ein Anteil der Kategorie **Fachmethodisch-Vermutet** vor, der deutlich größer als Null ist.
- Für alle 8 Personen mit niedrigen Kompetenzzuwächsen (bis auf eine) liegt innerhalb von Einheit 3 ein Anteil der Kategorie **Fachmethodisch-Vermutet** vor, der Null ist.

Bereits für Einheit 2 zeigt sich ein ähnlicher Befund (der erste der beiden Aufzählungspunkte trifft analog zu, der zweite nicht). Allerdings scheint der Befund – insbesondere in dieser Deutlichkeit – besonders für Einheit 3 von großer Relevanz zu sein, weil sie als letztes bearbeitet wird und daher eventuell auch kumulierte Hinweise auf Prozesse des Kompetenzaufbaus gibt. So könnte es beispielsweise sein, dass zum Formulieren mancher fachmethodischer Beiträge vorlaufende Erfahrungen mit (spezifischen oder mehreren) Experimentiersituationen notwendig sind.

Interessanterweise findet sich keine Systematik für fachmethodisch-indizierte Beiträge innerhalb von Einheit 3 der implizit-fachmethodischen Instruktionsvariante, was so gedeutet werden könnte, dass diese für Prozesse des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten nur eine untergeordnete Rolle spielen. Eine Möglichkeit der Interpretation wäre, dass fachmethodische-indizierte Beiträge innerhalb der explizit-fachmethodischen Instruktionsvariante als Ausgangspunkt dafür dienen, fachmethodisch-vermutete Beiträge zu formulieren. Dieser Ausgangspunkt würde dann allerdings nur von einigen Personen genutzt (und auch nicht in Abhängigkeit der Dauer ihrer fachmethodisch-indizierten Beiträge), weil sich sonst statistische Zusammenhänge einstellen würde. Eine andere Möglichkeit der Deutung wäre, dass Lernende bei Bearbeitung der implizit-fachmethodischen Instruktionsvariante nicht die Notwendigkeit sehen, fachmethodisch-indizierte Beiträge zu formulieren (wie beispielsweise Regeln zum naturwissenschaftlichen Arbeiten) und daher die Prozesse ihres Kompetenzaufbaus nur an fachmethodisch-vermuteten Beiträgen festgemacht werden können. Zuletzt sei noch genannt, dass es auch die Möglichkeit gibt, dass Lernende, die die implizit-fachmethodische Instruktionsvariante bearbeiten, nicht dazu in der Lage sind, ihre fachmethodischen Überlegungen so zu formulieren, dass die Kategorie **Fachmethodisch-Indiziert** kodiert werden kann. Hohe Kompetenzzuwächse könnten dann deshalb auftreten, weil für die richtige Bearbeitung der Prä- und Post-Erhebungen nicht die Formulierung der Konzepte, sondern die Anwendung dieser erforderlich ist, welche die Lernenden ggf. auch ohne ein explizit-regelbasiertes Verständnis leisten könnten (vgl. z. B. von Aufschnaiter & Rogge, 2010b).

5.5.3.3 Lernende in unterschiedlichen Instruktionsvarianten und mit unterschiedlichen Kompetenzzuwächsen

Aus den (Unter-)Abschnitten 5.4.3, 5.4.4, 5.5.2.1, 5.5.3.1 und 5.5.3.2 geht u. a. hervor, dass größere zeitliche Anteile fachmethodischer Beiträge für beide Instruktionsvarianten mit höheren Kompetenzzuwächsen einhergehen; andersherum gehen höhere Kompetenzzuwächse allerdings zwar für die implizit-fachmethodische Instruktionsvariante, aber für die explizit-fachmethodische Instruktionsvariante im Allgemeinen nicht mit größeren zeitlichen Anteilen fachmethodischer Beiträge einher. Ferner ist es sehr deutlich, dass die zeitlichen Anteile fachmethodischer Beiträge der Lernenden bei der explizit-fachmethodischen Instruktionsvariante bedeutsam größer sind als für die Lernenden, die die implizit-fachmethodische Instruktionsvariante bearbeitet haben. Wird berücksichtigt, dass die mittleren Kompetenzzuwächse der Lernenden bei der explizit-fachmethodischen Instruktionsvariante bedeutsam höher sind als für

die implizit-fachmethodische Instruktionsvariante (Abschnitt 4.5, ab S. 63), könnte dies zunächst so verstanden werden, dass Unterschiede in den zeitlichen Anteilen der fachmethodischen Beiträge sowohl die Unterschiede zwischen den Instruktionsvarianten als auch zwischen den Lernenden innerhalb der Instruktionsvarianten erklären können. Nachfolgend werden Ergebnisse berichtet, die dafür sprechen, dass dies – zumindest in einer direkten, unmittelbaren Weise – *nicht* der Fall ist (anschließende Diskussion in Abschnitt 7.2, ab S. 431):⁸¹

Für Lernende mit vergleichbaren Kompetenzzuwächsen, die unterschiedliche Instruktionsvarianten bearbeitet haben, finden sich deutlich unterschiedliche zeitliche Anteile fachmethodischer Beiträge. Anhand einer Kovarianzanalyse lässt sich das Ergebnis finden, dass unter Kontrolle der Kompetenzzuwächse ein signifikanter Unterschied sehr großer Effektstärke zwischen den beiden Instruktionsvarianten findet (siehe Tabelle 5.39). Da die Voraussetzung einer über die beiden Gruppen (Instruktionsvarianten) hinweg homogenen Kovariate (Kompetenzzuwachs) nicht gegeben ist (vgl. die in Abschnitt 4.5, ab S. 63, berichteten Ergebnisse; siehe auch Vorholzer, 2016) und auch die anderen Voraussetzungen nur zum Teil erfüllt sind, wurden trotz der großen Effekte und guten p -Werte zusätzlich händische Vergleiche vorgenommen: Die Lernenden wurden gemäß ihrer Kompetenzzuwächse (und unabhängig von der Instruktionsvariante) gruppiert, indem in fünfzigersritten vorgegangen wurde⁸² und nur der Bereich zwischen –50 und 200 aufgelöst wurde. Die

Tabelle 5.39: Zusammenfassung der Ergebnisse dreier Kovarianzanalysen zum Zusammenhang des zeitlichen Anteils fachmethodischer Beiträge mit den Instruktionsvarianten unter Kontrolle der von den Lernenden jeweils erreichten Kompetenzzuwächse.

Vergleich	Unabh. Variable: Instruktionsvariante					Kovariate: Kompetenzzuw.			
	Freiheitsgr.	F -Wert	p -Wert	part. η^2		Freiheitsgr.	F -Wert	p -Wert	
Fachmethodisch	1 , 42	29.772	< .001	.415		1 , 42	0.346	.560	
Fachmethodisch-vermutet	1 , 42	16.254	< .001	.279		1 , 42	4.939	.032	
Fachmethodisch-indiziert	1 , 42	28.217	< .001	.402		1 , 42	.043	.837	

⁸¹Genau genommen sprechen die Ergebnisse dafür, dass die beiden Instruktionsvarianten hinsichtlich der zeitlichen Anteile fachmethodischer Beiträge nicht auf einer (qualitativen) Skala verortet werden sollten: Auch trotz der Ergebnisse könnte es sehr wohl sein, dass die zeitlichen Anteile fachmethodischer Beiträge sowohl die mittleren Unterschiede zwischen den Instruktionsvarianten als auch die mittleren oder individuellen Unterschiede innerhalb der Instruktionsvarianten gut erklären; allerdings kann den Ergebnissen folgend die allgemein Aussage, dass die zeitlichen Anteile fachmethodischer Beiträge die Kompetenzzuwächse erklären, nicht über die Instruktionsvarianten hinweg aufrechterhalten werden.

⁸²Der Mittelwert der Kompetenzzuwächse über beide Instruktionsvariante beträgt 51.25 mit einer Standardabweichung von 69.71.

zugehörigen Ergebnisse stützen die Deutung; für das am wenigsten deutliche Ergebnis zu den fachmethodisch-vermuteten Beiträgen findet sich in Abbildung 5.53 die entsprechende Darstellung; für die Zusammenfassung fachmethodischer Beiträge und für fachmethodisch-indizierte Beiträge sind die Balken zur implizit-fachmethodischen Instruktionsvariante verschwindend klein im Vergleich zu den Balken zur explizit-fachmethodischen Instruktionsvariante.

Dass sich unter den Lernenden, die *innerhalb jeweils einer der Instruktionsvarianten* geringere zeitliche Anteile fachmethodischer Beiträge aufweisen, sowohl Lernende mit hohen als auch niedrigen Kompetenzzuwächsen finden, könnte damit zusammenhängen, dass manche Personen (grundsätzlich oder in den jeweiligen Team-Konstellationen) dazu neigen, weniger zu reden als andere. Daher war es geplant, neben dem zeitlichen Anteil fachmethodischer Beiträge an der Gesamtbearbeitungsdauer der Instruktion auch den zeitlichen Anteil fachmethodischer Beiträge an der Gesamtdauer *verbaler* Beiträge zu betrachten. Aus Zeitgründen ist diese Analyse nicht mehr umfänglich erfolgt; erste Sichtungen der Daten liefern aber Hinweise, dass selbst bei Betrachtung fachmethodischer Beiträge relativ zum Redeanteil der jeweiligen Person noch Personen mit geringeren zeitlichen Anteilen fachmethodischer Beiträge existieren, die hohe Kompetenzzuwächse erreichen. Insbesondere die Betrachtung

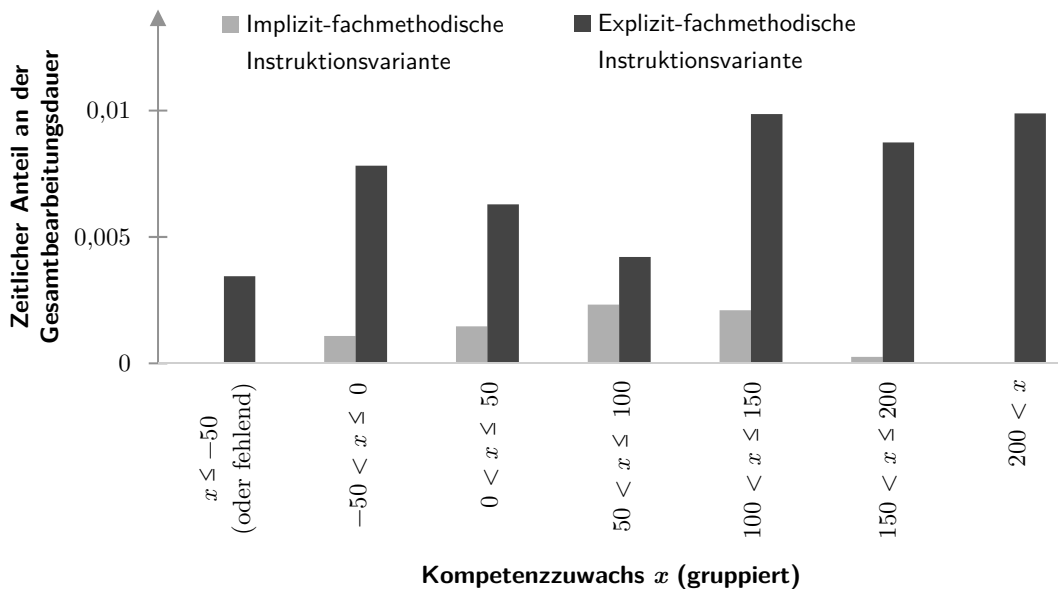


Abbildung 5.53: Mittelwerte der zeitlichen Anteile fachmethodisch-vermuteter Beiträge an der Gesamtbearbeitungsdauer aller drei Einheiten verglichen für Lernende aus beiden Instruktionsvarianten (27 Personen für explizit-, 18 Personen für implizit-fachmethodische Variante) gruppiert nach den erreichten Kompetenzzuwächsen.

tung der implizit-fachmethodischen Instruktionsvariante liefert somit als relevantes Teilergebnis, dass auch bei sehr geringem zeitlichen Umfang fachmethodischer Beiträge erfolgreich Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten stattfinden kann.

Für die meisten Lernenden mit niedrigen Kompetenzzuwächsen, die die explizit-fachmethodische Instruktionsvariante bearbeitet haben, finden sich deutlich größere zeitliche Anteile fachmethodischer Beiträge als für alle Lernenden, die die implizit-fachmethodische Instruktionsvariante bearbeitet haben. Dieses Ergebnis geht aus den Histogrammen in Abbildung 5.54 (auf die S. 235) und Abbildung 5.55 (auf S. 236) hervor. Erneut findet sich das am wenigsten eindeutige Ergebnis für die Kategorie der fachmethodisch-vermuteten Beiträge. Allerdings ist zu bedenken, dass die Lernenden, die bei der explizit-fachmethodischen Instruktionsvariante geringe zeitliche Anteile fachmethodisch-vermuteter Beiträge aufweisen, zumeist mittlere bis hohe Anteile fachmethodisch-indizierter Beiträge aufweisen, so dass sind die geringen Anteile nicht im Sinne fachmethodischer Inaktivität deuten lassen.

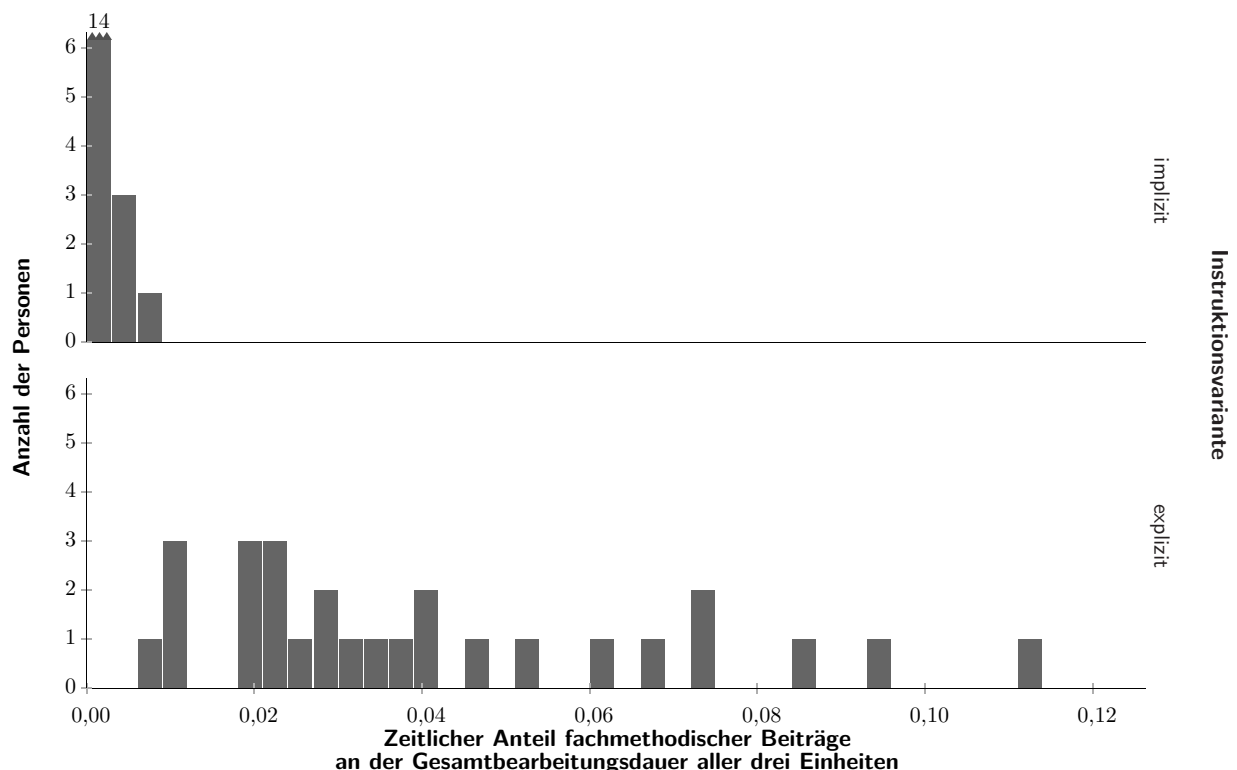


Abbildung 5.54: Histogramm zum zeitlichen Anteil der fachmethodischen Beiträge an der Gesamtbearbeitungsdauer aller drei Einheiten aufgelöst nach Instruktionsvarianten (27 Personen für explizit-, 18 Personen für implizit-fachmethodische Variante).

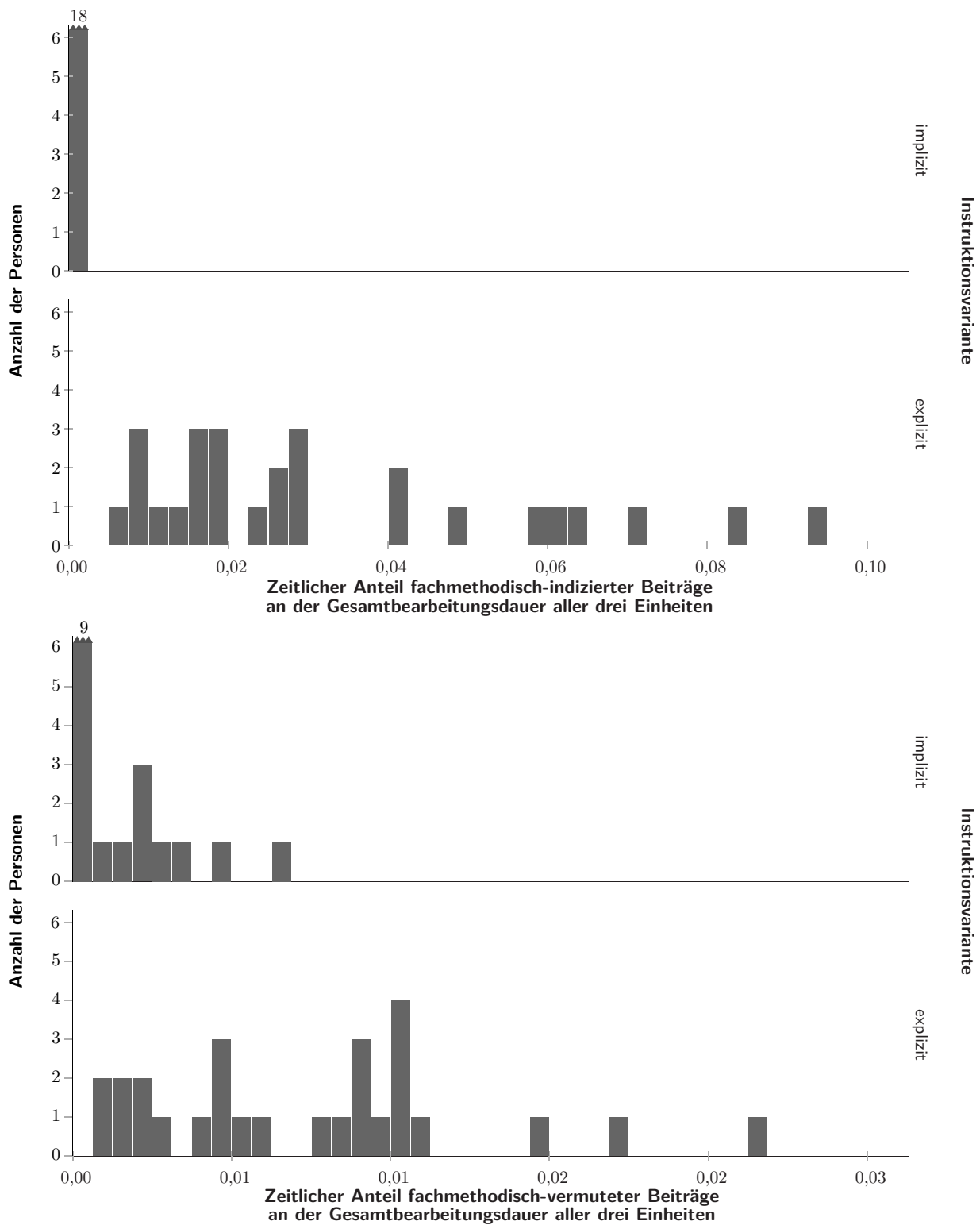


Abbildung 5.55: Histogramme zum zeitlichen Anteil der fachmethodisch-indizierten (oben) und -vermuteten (unten) Beiträge an der Gesamtbearbeitungsdauer aller drei Einheiten aufgelöst nach Instruktionsvarianten (27 Personen für explizit-, 18 Personen für implizit-fachmethodische Variante).

5.5.4 Inhalte der fachmethodischen Beiträge

Da die zeitlichen Anteile fachmethodischer Beiträge sehr gering sind, wird für die Betrachtung von Details zu fachmethodischen Beiträgen bei der Bearbeitung der implizit-fachmethodischen Instruktionsvariante die Anzahl fachmethodischer Beiträge betrachtet. Dieses Vorgehen ist angemessen, weil die fachmethodischen Beiträge einer Person meist mit großen zeitlichen Abständen auftreten, wie grafische Analysen von Zeitstrahlen für die Kodierungen zeigen. Der Einfachheit halber wird nicht zwischen fachmethodisch-indizierten und -vermuteten Beiträgen unterschieden. Es finden sich bei einem Vergleich von Personen mit hohen Kompetenzzuwächsen ($N = 6$ Personen, $MW = 6.83$ fachmethodische Beiträge, $SD = 5.382$, $SE = 2.197$, $Median = 6$) vs. Personen mit niedrigen Kompetenzzuwächsen ($N = 8$ Personen, $MW = 3.25$ fachmethodische Beiträge, $SD = 2.816$, $SE = 0.996$, $Median = 2.5$) sowohl mit einem t-Test nach Student als auch mit einem Mann-Whitney-U-Test nahezu vorsignifikante Unterschiede von fast großer Effektstärke (t-Test: $t(12) = -1.624$, $p = .181$, $r = .424$; Mann-Whitney-U-Test: $U(14) = 36.0$, $z = 1.561$, $p = .142$, $r = .417$).⁸³

Die Ergebnisse zum Vergleich der Anzahlen fachmethodischer Beiträge von Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen bei der implizit-fachmethodischen Instruktionsvariante deuten ähnliche wie die Ergebnisse zu den zeitlichen Anteilen darauf hin, dass fachmethodische Beiträge mit den erreichten Kompetenzzuwächsen zusammenhängen. Mithilfe des für die implizit-fachmethodische Instruktionsvariante ergänzten Teils des Kodiersystems, der sich auf die Inhalte der fachmethodischen Beiträge bezieht (siehe Abschnitt 5.1) und parallel zu den Abschnitten der explizit-fachmethodischen Instruktionsvariante formuliert ist (die jeweils auf einen der Inhalte bezogen sind, siehe Abschnitt 4.3), lässt sich nachfolgend auf die Frage eingehen, welche Unterschiede und Gemeinsamkeiten es innerhalb der implizit-fachmethodischen Instruktionsvariante für Lernende mit unterschiedlichen Kompetenzzuwächsen hinsichtlich der inhaltlichen Bezüge ihrer fachmethodischen Beiträge gibt. Im Ansatz ist aufgrund der Parallelität der Kategorien zu den Abschnitten der explizit-fachmethodischen Instruktionsvariante auch ein Vergleich der beiden Instruktionsvarianten möglich.

⁸³Für die Gruppe der Personen mit mittleren Kompetenzzuwächsen ($N = 4$ Personen), die nicht in den statistischen Vergleich einbezogen wird, ergeben sich folgende Werte: $MW = 3.25$ fachmethodische Beiträge, $SD = 2.872$, $SE = 1.436$, $Median = 2.5$. Anhand der Quartile für die drei Gruppen wird zudem deutlich, dass sich ein plausibler Trend vorfindet (angegeben ist Obergrenze):

Perzentil	Kompetenzzuwachs		
	Niedrig	Mittel	Hoch
25	1.25	1	2.75
50	2.5	2.5	6
75	4.75	6.25	9.5

In Tabelle 5.40 (auf S. 239) ist dargestellt, wie häufig jede Kategorie zum Inhalt von fachmethodischen Beiträgen bei der Bearbeitung der implizit-fachmethodischen Instruktionsvariante kodiert wurde. Aufgelöst wird sowohl nach den Einheiten der Instruktion als auch nach Kompetenzzuwachs der Lernenden.

Variablenkontrolle dominiert die fachmethodischen Beiträge bei der implizit-fachmethodischen Instruktionsvariante inhaltlich. Insgesamt ist für die implizit-fachmethodische Instruktionsvariante ersichtlich, dass der allergrößte Teil der fachmethodischen Äußerungen auf Bemerkungen zur Variablenkontrolle entfällt. Danach folgen Äußerungen zum Messbereich sowie verschiedene Aspekte zur Passung von Fragestellungen, Versuchsanlagen und Deutungen. Beiträge zur Formulierung und zu Unterscheidungen von (verschiedenen Arten von) Fragestellungen, Hypothesen/Vermutungen und Beobachtungen/Deutungen finden sich bei der Bearbeitung der implizit-fachmethodischen Instruktionsvariante nur sehr vereinzelt.

Die Gesamtheit der Lernende mit hohen Kompetenzzuwächsen weist zu nahezu allen Inhalten mehr Äußerungen auf als die Gesamtheit der Lernenden mit niedrigen Kompetenzzuwächsen. Sie deckt zudem ein größeres Spektrum an Inhalten ab. Im Vergleich von 7 Lernenden mit hohen vs. 8 Lernenden mit niedrigen Kompetenzzuwächsen zeigt sich für die Inhalte der fachmethodischen Beiträge bei der implizit-fachmethodischen Instruktionsvariante, dass die Anzahl fachmethodischer Beiträge mit einem spezifischen Inhalt (also ein Tabellenfeld) grundsätzlich größer für die Gruppe der Lernenden mit hohen Kompetenzzuwächsen ist. Ausnahmen stellen nur die Kategorien **PassungFrageVersuch** bei Einheit 1 und **VariablenKontrolle** bei Einheit 2 dar. Anhand der mehrfachen Nulleinträge für Lernende mit niedrigen Kompetenzzuwächsen, die mit Nichtnnulleinträgen für Lernende mit hohen Kompetenzzuwächsen einhergehen, wird zudem deutlich, dass die Gesamtheit der Lernenden mit hohen Kompetenzzuwächsen in jeder Einheit und auch insgesamt mehr fachmethodische Inhalte abdeckt.

Fachmethodische Beiträge von Lernende mit niedrigen Kompetenzzuwächsen beziehen sich in den meisten Einheiten auf Variablenkontrolle; fachmethodische Beiträge von Lernenden mit hohen Kompetenzzuwächsen beziehen sich ebenfalls in jeder Einheit schwerpunktmäßig auf Variablenkontrolle und zusätzlich auf das inhaltliche Pendant zur jeweiligen Einheit der explizit-fachmethodischen Instruktionsvariante. Sowohl für Lernende mit hohen als auch für Lernende mit niedrigen Kompetenzzuwächsen macht die Kategorie zur Variablenkontrolle insgesamt und in jeder Einheit jeweils den größten

5.5 Auswert. u. Ergebnisse zur impl.-fachmeth. IV sowie zum Vergl. d. beiden IVs

Tabelle 5.40: Übersicht über die Inhalte fachmethodischer Beiträge bei der Bearbeitung der implizit-fachmethodischen Instruktionsvariante.

Kategorie	Beschreibung: Hinweis auf Überlegung zu/zur/zum ...	Auftreten p. Einheit			Auftreten insges.	
		E1	E2	E3	Absolut	Relativ
FragenNawi	Kennzeichen von naturwiss. Fragestellungen etc.	N: 0	N: 0	N: 0	N: 0	N: 0 %
		H: 1	H: 0	H: 0	H: 1	H: 2,22 %
FragenPrAllg	präzisen oder allgemeinen Formulierung von Fragestellungen	N: 0	N: 0	N: 0	N: 0	N: 0 %
		H: 0	H: 0	H: 0	H: 0	H: 0 %
VermutungHypothese	Formulierung von Vermutungen/Hypothesen	N: 0	N: 0	N: 0	N: 0	N: 0 %
		H: 1	H: 1	H: 0	H: 2	H: 4,44 %
PassungFrageVorhersage	Passung von Fragestellung und Vermutung/Hypothese	N: 0	N: 0	N: 0	N: 0	N: 0 %
		H: 1	H: 0	H: 0	H: 1	H: 2,22 %
VariablenKontrolle	(un)abhängigen Var., Kontrollvar., Einhaltung der Var.-Kontrolle etc.	N: 8	N: 9	N: 1	N: 18	N: 69,23 %
		H: 11	H: 9	H: 2	H: 22	H: 48,89 %
PassungVersuchFrage	Passung von Versuch und Fragestellung	N: 3	N: 0	N: 0	N: 3	N: 11,54 %
		H: 1	H: 1	H: 1	H: 3	H: 6,67 %
PassungVersuchVorhersage	Passung von Versuch und Vermutung/Hypothese	N: 0	N: 0	N: 0	N: 0	N: 0 %
		H: 0	H: 0	H: 0	H: 0	H: 0 %
Messbereich	Nutzung des gesamten oder eines bestimmten Messbereichs	N: 0	N: 1	N: 0	N: 1	N: 3,85 %
		H: 0	H: 3	H: 2	H: 5	H: 11,11 %
TabellenDiagramme	Anlegen von Tabellen oder Diagrammen	N: 0	N: 1	N: 0	N: 1	N: 3,85 %
		H: 0	H: 0	H: 5	H: 5	H: 11,11 %
BeobachtungDeutung	Formulierung oder Unterscheidung von Beobachtungen und Deutungen	N: 0	N: 0	N: 0	N: 0	N: 0 %
		H: 0	H: 0	H: 1	H: 1	H: 2,22 %
PassungDeutung	Passung von Deutungen zu Messwerten und Fragestellungen	N: 0	N: 0	N: 0	N: 0	N: 0 %
		H: 0	H: 2	H: 0	H: 2	H: 4,44 %
ErfolgSinnhaftigkeit	Zielstellung von Versuchen (insb.: nicht das Erzeugen von Erfolg)	N: 0	N: 0	N: 0	N: 0	N: 0 %
		H: 2	H: 0	H: 0	H: 2	H: 4,44 %
SonstigesUnklar	Fachmethoden, aber thematisch nicht zuzuordnen	N: 1	N: 2	N: 0	N: 3	N: 11,54 %
		H: 0	H: 1	H: 0	H: 1	H: 2,22 %
Summen nach Kompetenzzuwachs	Niedrig	12	13	1	26	32,50%
	Hoch	17	17	11	45	56,25%
	Gesamt	32	35	13	80	100,00%

Anmerkungen. Die Anzahlen der Kategorien für Personen mit mittleren Kompetenzzuwächsen sind nicht aufgeführt und werden nur in die Gesamt-Zeile einbezogen. Die relativen Angaben im oberen Teil der Tabelle beziehen sich auf den Anteil der Häufigkeit der jeweiligen Kategorie an der Häufigkeit aller Kategorien bei demselben Kompetenzzuwachs. Im Summen-Teil der Tabelle sind Anteile an der Anzahl aller vergebenen Kategorien notiert.

Anteil aus. Ein Unterschied besteht in der Größe des Anteils, was mit dem breiteren Spektrum für Lernende mit hohen Kompetenzzuwächsen zusammenhängt.

Die Gesamtheit der Lernenden mit niedrigen Kompetenzzuwächsen weist vor allem Beiträge zur Variablenkontrolle auf, während sich für die Gesamtheit der Lernenden mit hohen Kompetenzzuwächsen zusätzlich zum Variablenkontrollschwerpunkt je nach Einheit auch stärker Beiträge mit Inhalten zum inhaltlichen Schwerpunkt der dazu parallelen Einheit der explizit-fachmethodischen Instruktionsvariante finden.

In Einheit 1 finden sich für Lernende mit hohen Kompetenzzuwächsen finden sich Beiträge zu drei der ersten vier Kategorien aus Tabelle 5.40; für Lernende mit niedrigen Kompetenzzuwächsen finden sich nur Beiträge zur Kategorie **PassungVersuchFrage**, die sich ebenfalls in Teilen dem Schwerpunkt des Formulierens von Fragen und Hypothesen zuordnen lässt.

In Einheit 2 ist der Unterschied am wenigsten deutlich, weil der inhaltliche Schwerpunkt des explizit-fachmethodischen Pendants die Planung von Versuchen ist, zu der zentral die Variablenkontrolle gehört. Allerdings findet sich auch hier für Beiträge zum Messbereich (ebenfalls zur Planung von Versuchen zugehörig) vor allem für Lernende mit hohen Kompetenzzuwächsen ein weiterer Schwerpunkt.

In Einheit 3 finden sich für Lernende mit niedrigen Kompetenzzuwächsen nur ein Beitrag zur Variablenkontrolle, während für die Gesamtheit der Lernenden mit hohen Kompetenzzuwächsen schwerpunktmäßig die Variablenkontrolle sowie Tabellen und Diagramme (zugehörig zum inhaltlichen Fokus Auswerten und Interpretieren von Daten der dritten Einheit der explizit-fachmethodischen Instruktionsvariante) finden und darüber hinaus viele der anderen Inhalte.

Der Unterschied in der Anzahl fachmethodischer Beiträge zwischen Lernenden mit hohen und niedrigen Kompetenzzuwächsen ist besonders deutlich bei Einheit 3. Dies geht aus den vorigen Ergebnissen hervor.

6 Vorstellungen von Lernenden während der Bearbeitung der explizit-fachmethodischen Instruktionsvariante

Die Analysen, anhand derer auf die Vorstellungen von Lernenden bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante geschlossen wird, sind darauf gerichtet, über welche fachmethodischen Inhalte die Lernenden sprechen und in welcher Weise sie dies tun. Ziele für die Analysen bestehen darin, auf Vorstellungen orientierte Beschreibungen und Deutungen der Bearbeitungsprozesse zu liefern, auf denen Hypothesen zu Prozessen des Kompetenzaufbaus basieren können (siehe Kapitel 7 für einige Hypothesen), sowie darin, auf Prozessen basierende Ergebnisse zu inhaltlichen und anderen Ausprägungen von Vorstellungen zum experimentbezogenen Denken und Arbeiten zu generieren.

Das eingesetzte methodische Vorgehen ist in Abschnitt 6.1 ausführlich verschriftlicht. Es lässt sich ob Einfachheit der Kommunikation im methodologischen Paradigma der qualitativen Inhaltsanalyse verorten und anhand dort genutzter Begrifflichkeiten beschreiben. Aufgrund der rekonstruierenden und detaillierten Vorgehensweise sind die Analysen zu den Vorstellungen – anders als die in Kapitel 5 dargestellten Analysen zu den Aktivitäten der Lernenden – nicht auf das gesamte bearbeitete Instruktionsmaterial und nicht auf die vollständigen Videoaufzeichnungen bezogen, sondern betreffen nur die Videosequenzen zu ausgewählten Instruktionsextrakten, also zu Auszügen aus dem Instruktionsmaterial.⁸⁴ Die Wahl der analysierten Instruktionsextrakte ist in Abschnitt 6.2 ab S. 269 dargestellt und begründet. Ferner werden nur für ausgewählte Lernende die Videosequenzen zu den Instruktionsextrakten analysiert. Gründe dafür sowie das Vorgehen bei der Auswahl der analysierten Personen sind in

⁸⁴Der Begriff *Auszug* wurde zur besseren Abgrenzung von bereits genutzten Begrifflichkeiten (insbesondere *Ausschnitt*) nicht dauerhaft gewählt, an die an dieser Stelle erinnert wird: (Video-) Ausschnitt (für einen zeitlich zusammenhängenden, je nach individuellem Bezug sehr kurzen bis längeren Schnitt aus einem Video), (*Instruktions-*)*Einheit* (für eine der drei zu unterschiedlichen Zeitpunkten eingesetzten Einheiten, siehe Abbildung 4.1 auf S. 51), (*Instruktions-*)*Abschnitt* (für mehrere zusammenhängende Karten des Instruktionsmaterials, die den Lernenden durch Abschnittsüberschriften transparent gemacht sind, siehe Tabellen 4.4, S. 60, und 4.5, S. 60), (*instruktionale*) *Serie* (von Aufgaben, die nicht notwendigerweise lückenlos aufeinander folgen, siehe Abbildung 4.3 auf S. 57) uvm. Außerdem wird der unspezifische Begriff *Teil* für verschiedene Teilmengenrelationen genutzt und eignet sich daher nicht. Der Begriff *Extrakt* verdeutlicht die bewusste Auswahl spezifischer Instruktionsteile, die grundsätzlich inhaltlich und zumeist auch zeitlich zusammenhängender Natur sind.

Abschnitt 6.3 ab S. 278 dargestellt. Diskussionen zur Angemessenheit und Validität des Vorgehens finden sich an den spezifischen Stellen in der Darstellung der Methoden sowie in einem eigenen Unterabschnitt zur Prüfung der Beurteilerübereinstimmung für die qualitative Inhaltsanalyse (Unterabschnitt 6.1.6 ab S. 261). Die Darstellung der Ergebnisse erfolgt für die einzelnen Forschungsfragen in den Abschnitten 6.4 bis 6.10 und beginnt auf Seite 283.

6.1 Qualitative Inhaltsanalyse zur Rekonstruktion von Vorstellungen der Lernenden

Mit der qualitative Inhaltsanalyse zu den ausgewählten Sequenzen wird fokussiert, über welche fachmethodischen Inhalte die Lernenden bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante sprechen und in welcher Weise sie dies tun. Die qualitative Inhaltsanalyse besteht aus induktiven und deduktiven Bestandteilen, deren Bezüge in Abbildung 6.1 schematisch dargestellt sind.

Die Teilanalysen lassen sich grob den bezüglich der Vorstellungen formulierten Forschungsfragen zuordnen: Zur Unterscheidung inhaltlich verschiedener Vorstellungen (insbesondere F-Vor1, formuliert auf S. 48) wird im Ergebnisteil auf induktiv generierte Ideen-Kategorien zurückgegriffen (Ebene 1a des Kodiermanuals); zur Betrachtung von Verallgemeinerungen (insb. F-Vor2) werden auf Ausschnitte aus den deduktiv-induktiv generierten Kategorien zum Modus der Äußerung genutzt (Ebene 2 des Kodiermanuals); eine Auflösung nach Zeit und Situationen (insb. F-Vor5, F-Vor6) wird durch das Kodieren der Aufgabe, auf die die Äußerung bezogen ist, ermöglicht

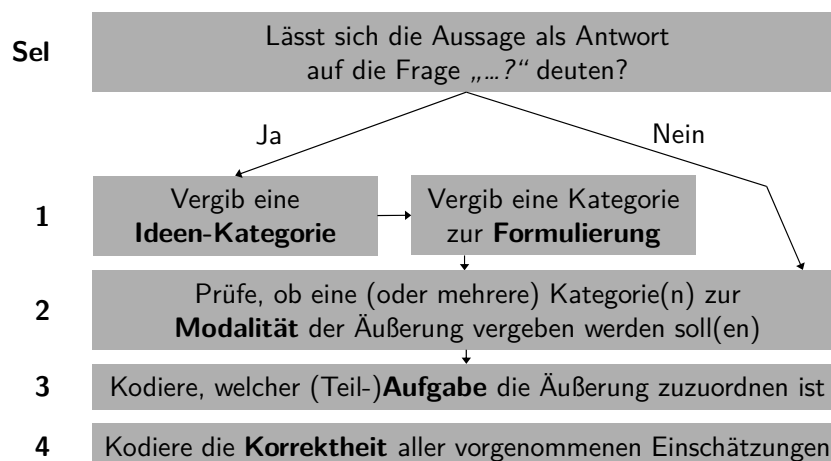


Abbildung 6.1: Übersicht über das Vorgehen bei der qualitativen Inhaltsanalyse.
(Sel = Selektionskriterium)

(Ebene 3 des Kodiermanuals); zuletzt wird direkt auf verbale Bekundungen von Entscheidungen zum experimentbezogenen Denken und Arbeiten (insb. F-Vor4) eingegangen (Ebene 4 des Kodiermanuals).

Der Kodiervorgang fand Äußerung für Äußerung statt, indem durch das Transkript oder das Video für ein Team hindurchgegangen wird. Beim Kodieren am Video wurde mit der Aussage begonnen, die im zeitlichen Verlauf früher beginnt. Die verschiedenen Bestandteile der qualitativen Inhaltsanalyse wurden zeitlich aufeinanderfolgend bearbeitet und sind im Detail in den nachfolgenden Unterabschnitten erläutert. Vor der detaillierten Beschreibung der Bestandteile wird ein Abriss des Vorgehens gegeben (siehe Abbildung 6.1 auf S. 242):

- Für jede Äußerung wird zunächst das Selektionskriterium (Ebene 0)

»Lässt sich die Äußerung als Antwort auf die Frage ›...?‹ deuten?«

angewendet.⁸⁵ Die Frage »...?« aus dem Selektionskriterium könnte beispielsweise für Analysen zu auf Kennzeichen von naturwissenschaftlichen Fragen bezogenen Vorstellungen »Was kennzeichnet eine naturwissenschaftliche Fragestellung?« lauten, für Analysen zu Vorstellungen bezüglich fairer Versuche »Was macht einen fairen Versuch aus?« und für Analysen zu Vorstellungen zur Unterscheidung von Beobachtung und Deutung »Was kennzeichnet eine Beobachtung oder Deutung, ggf. im Unterschied zu einer Deutung/Beobachtung?«⁸⁶

Beispielsweise⁸⁷ enthält die Äußerung »Das hier ist eine naturwissenschaftliche Fragestellung, kannst du ja messen« einen Hinweis auf mögliche Ideen zur Frage aus dem Selektionskriterium: Ob eine Fragestellung naturwissenschaftlich ist, hängt laut der Äußerung zumindest irgendwie damit zusammen, ob etwas gemessen werden kann. (Unklar bleibt hier allerdings, was gemessen werden kann, wie es gemessen werden müsste usw. Weitere Positiv-Beispiele sind in Unterabschnitt 6.1.1 ab S. 245 zu finden.)

Die Äußerung »Nein, das ist keine naturwissenschaftliche Fragestellung« lässt sich allerdings nicht unmittelbar als Antwort auf die Frage, was eine naturwissenschaftliche Fragestellung kennzeichnet, deuten. (Allerdings

⁸⁵Genau genommen ist das *Kriterium* zur Selektion, dass vom Kodierer *ja* geantwortet wird. Da allerdings auch bei einem Nein als Antwort weitere Kategorien vergeben werden, liegt eine Selektion streng genommen nur in Bezug darauf vor, ob für die Äußerung Ideen-Kategorien vergeben werden.

⁸⁶Weitere Inhaltsbereiche zu Konzepten des experimentbezogenen Denkens und Arbeitens sind denkbar. Aufgrund der Wahl in Abschnitt 6.2 werden allerdings nur diese Beispiele gegeben.

⁸⁷Beispiele beziehen sich der Einfachheit halber in diesem Abschnitt vorrangig auf Kennzeichen von naturwissenschaftlichen Fragen, zu denen für die vorliegende Arbeit ein Instruktionsextrakt erstellt wurde (basierend vor allem auf der ersten Einheit der explizit-fachmethodischen Instruktionsvariante, siehe Unterabschnitt 6.2.1), zu dem Videosequenzen analysiert werden.

steht sie in mittelbarem Zusammenhang mit der Forschungsfrage, weil aus der Äußerung zwar keine Idee hervorgeht, aber eine Handlung stattfindet, die mit Vorstellungen abgeglichen werden kann. Daher wird die Aussage auf Ebene 2 berücksichtigt; siehe Unterabschnitt 6.1.3 ab S. 254.)

Entsprechend der explizit als deutungsbehaftet ausgewiesenen Entscheidung anhand des Selektionskriteriums wird dann anschließend unterschiedlich verfahren.

- Äußerungen, für die das Selektionskriterium erfüllt ist, werden nachfolgend mit einer bereits vorhandenen oder neu generierten Ideen-Kategorie kodiert (Ebene 1a) und daran anschließend auch mit einer Kategorie zur Formulierung der Äußerung (Ebene 1b). Erst danach wird (mindestens) eine Kategorie zum Modus vergeben (Ebene 2).
- Äußerungen, für die das Selektionskriterium nicht erfüllt ist, werden keiner rekonstruktiv angelegten Analyse bezüglich der Vorstellungen unterzogen (d. h. Ebene 1 wird übersprungen), sondern direkt daraufhin geprüft, ob eine Kategorie zum Modus zuzuweisen ist (Ebene 2).
- Alle Äußerungen, denen mindestens eine Kategorie auf den ersten beiden Ebenen zugewiesen wurde, werden mit einer Kategorie zur Aufgabe, in deren Kontext die Äußerung formuliert wurde, versehen (Ebene 3). Für andere Äußerungen werden keine Aufgaben-Kategorien vergeben.
- Die Kodierung der Korrektheit der Einschätzungen werden in einem separaten Durchgang durch das Material vorgenommen (Ebene 4). Dazu werden alle Äußerungen, denen gewisse Kategorien zum Modus zugewiesen wurden, daraufhin beurteilt, ob die mit ihnen verbundene (ggf. implizite) Einschätzung aus fachmethodischer Sicht als korrekt anzusehen ist.

Die dargestellten Schritte wurden vom Autor der Arbeit vorgenommen und durch Nutzung von MAXQDA dokumentiert. Vorläufige Versionen des Vorgehens wurden anhand von 5 Teams vom Autor der Arbeit geprüft und ggf. weiterentwickelt. In Unterabschnitt 6.1.6 (ab S. 261) wird berichtet, dass auch andere Personen das so entwickelte Vorgehen anwenden (können) und zu vergleichbaren Ergebnissen kommen. Daher wurden anschließend alle ausgewählten Teams (siehe Abschnitt 6.3 beginnend auf S. 278) vom Autor der Arbeit analysiert. Dieses Vorgehen ermöglichte es, dass der Autor für alle Kodierungen einschätzen kann, welche Äußerungen tatsächlich damit verbunden sind, was für die angestrebten Interpretationen zielführend erschien.

6.1.1 Induktives Generieren von Ideen-Kategorien (Ebene 1a)

Ein zentraler Bestandteil der durchgeführten qualitativen Inhaltsanalyse ist das induktive Generieren von Ideen-Kategorien, die zugleich den Äußerungen zugewiesen werden. Die Ideen-Kategorien dienen dazu, die Äußerungen der Lernenden in ihren inhaltlichen Ausprägungen abzubilden, um verschiedene Vorstellungen voneinander unterscheiden zu können. Als inhaltliche Ausprägung oder Dimension wird dabei angesehen, was die Lernenden durch ihre Äußerung inhaltlich (in Bezug auf die Frage »...?« im Selektionskriterium) beitragen. Beispielsweise enthält die oben angeführte Äußerung »Das hier ist eine naturwissenschaftliche Fragestellung, kannst du ja messen« den (oben bereits erläuterten) Hinweis auf die Idee, dass naturwissenschaftliche Fragen dadurch gekennzeichnet sind, dass sie zumindest irgendwie etwas enthalten, was gemessen werden kann. Dieser Aspekt wird als inhaltlicher Aspekt aufgefasst und mit der Ideen-Kategorie abgebildet. Allerdings lässt sich die Äußerung anhand von weiteren Aspekten beschreiben, etwa dem Bezug zu einer konkreten Beispielfrage (»das hier«) oder der personenbezogenen Formulierung (»kannst du«), von denen manche Aspekte in weiteren Kategorien aufgegriffen werden.

Die Kategorien werden als Ideen-Kategorien bezeichnet und nicht als Vorstellungskategorien, weil – wie später anhand der Ergebnisse ausführlicher diskutiert (siehe ad F-Vor1, ab S. 285) – gegebenenfalls mehrere Ideen-Kategorien als eine Vorstellung interpretiert werden könnten oder eine Ideen-Kategorie mehrere Vorstellungen umfassen könnte. Die Ideen-Kategorien existieren durchaus zu dem Zweck, die Ideen, die Lernende situativ äußern, mit ihren Vorstellungen in Verbindung zu bringen. Die unterschiedlichen Begriffe *Ideen* und *Vorstellungen* stellen zugleich aber heraus, dass dabei zwei verschieden stark deutungsbehaftete Ebenen miteinander in Verbindung gebracht werden. Anders formuliert ist die begriffliche Unterscheidung ein Hilfsmittel, um zu verdeutlichen, dass nicht davon ausgegangen wird, dass die aus einzelnen Äußerungen rekonstruierten Ideen-Kategorien (oder deren Beschreibungen) bereits den Vorstellungen der Lernenden entsprechen.

Die Ideen-Kategorien werden induktiv generiert und im gleichen Zuge den Äußerungen zugewiesen. Das heißt: Für jede Äußerung wird geprüft, ob bereits eine passende Ideen-Kategorie vorhanden ist und dann diese zugewiesen; andernfalls wird eine Ideen-Kategorie (gemäß unten geschilderter Regeln) angelegt und der geprüften Äußerung zugewiesen. Eine Modifikation einer bereits bestehenden Ideen-Kategorie ist nur denkbar, wenn alle zugehörigen Äußerungen erneut kodiert werden. Das eingesetzte Verfahren basiert zwar grundlegend auf der bei Mayring (2015) geschilderten induktiven Kategorienbildung, weicht allerdings an einigen Stellen von

dieser ab und lässt sich eventuell in mancherlei Hinsicht eher der *codebook analysis* (Neuendorf, 2017) zuordnen. Eine genauere Diskussion der Zuordnung des Verfahrens ist für die Ergebnisse der Arbeit nicht ausschlaggebend; eine genaue Darstellung des Vorgehens reicht dafür aus. In Anlehnung an die von Mayring (2015) vorgeschlagene Definition wird das induktive Generieren der Ideen-Kategorien nachfolgend durch fünf Bereiche dargestellt:

- Kategoriendefinition: Wonach wird gesucht?
- Angestrebtes Abstraktionsniveau: Wie sollen die Kategorien aussehen?
- Kodiereinheit: Wie groß ist das Textstück, das kodiert wird?
- Kontexteinheit: Was wird zur Deutung einer Äußerung mitgenutzt?
- Auswertungseinheit: Auf welchem (Video-)Material basieren die entwickelten Kategorien?

Kategoriendefinition (Wonach wird gesucht?). Dieser Aspekt ist in das Selektionskriterium ausgelagert. Es wird nach Aussagen von Schülerinnen und Schülern gesucht, die als Antworten auf die Frage »...« (bspw. »Was kennzeichnet eine naturwissenschaftliche Fragestellung?« oder »Was kennzeichnet eine Beobachtung oder Deutung, ggf. im Unterschied zu einer Deutung/Beobachtung?«) gedeutet werden können. Ergänzende Beispiele, auf die weiter unten erneut eingegangen wird, lauten:

- (1) »Das ist keine naturwissenschaftliche Frage, weil es mit den Meinungen von den Personen zu tun hat.« Die Äußerung kann als Antwort auf die Frage »Was kennzeichnet diese spezielle nicht-naturwissenschaftliche Fragestellung?« gedeutet werden: Sie hat mit Meinungen von Personen zu tun. Die hier formulierte Frage lässt sich als Unterfrage der Frage aus dem Selektionskriterium für Kennzeichen von naturwissenschaftlichen Fragen auffassen, so dass sich die Äußerung als Antwort auf Frage von oben deuten lässt.
- (2) »Bei einer naturwissenschaftlichen Frage gibt es doch eine eindeutige Antwort, oder?« Auch wenn die Äußerung als Frage formuliert ist, enthält sie einen Hinweis auf eine Idee dazu, was naturwissenschaftliche Fragen kennzeichnet.
- (3) »Ähm man kann es durch eine äh naturwissenschaftliche Methode nachweisen?« Wenn »es« aus der Äußerung eine einzelne Frage oder eine Klasse von Fragen meint, dann ist die Äußerung als Antwort auf die Frage, was naturwissenschaftliche Fragen kennzeichnet, deutbar.

Die Formulierung des Selektionskriteriums, aus welchem die Kategoriendefinition abgeleitet ist, kann kritisch hinterfragt werden, weil für die Entscheidung, ob eine Äußerung kodiert werden soll, Deutungsprozesse vonnöten sind, die als relativ stark eingeschätzt werden könnten. Für die Kategoriendefinition auf alle Äußerungen zurückzugreifen, die spezifische Worte (etwa »(naturwissenschaftliche) Frage« bzw. »Beobachtung« oder »Deutung«) beinhalten, stellt allerdings keine Alternative dar, weil die Schülerinnen und Schüler häufig von diesen Dingen sprechen, ohne die entsprechenden Worte zu verwenden. Ferner sind die Äußerungen ohnehin für die Rekonstruktion der Ideen zu deuten, wenn Kategorien generiert und zugewiesen werden sollen, die die *Inhalte* der Äußerungen angemessen abbilden. Das Selektionskriterium ist so gewählt, dass dessen Anwendung möglichst wenig Interpretation erfordert, die für die Generierung und Zuweisung der Ideen-Kategorien erforderlichen Interpretationen übersteigt.

Angestrebtes Abstraktionsniveau der Kategorien-Formulierung (Wie sollen die Kategorien aussehen?). Jeder zentrale Aspekt der Idee wird in je ein Schlagwort gefasst (wobei ein Schlagwort auch aus mehreren Wörtern bestehen darf). Das beinhaltet, dass gegebenenfalls mehrere Schlagwörter gleichzeitig an einem Textsegment zugewiesen werden können. Das Schlagwort wird in der Arbeit durch serifenlose Schrift und ein vorangestelltes + dargestellt (und Leerzeichen werden aus technischen Gründen weggelassen) und repräsentiert die Ideen-Kategorie. Beispiele für Kategorie-Formulierungen, die die Nummerierung von oben (Punkt zur *Kategoriendefinition*, S. 246) aufgreifen:

- (1) +Meinung
- (2) +eindeutige Antwort (in der Arbeit letztlich: +AntwortEindeutig)
- (3) Zwei Ideen-Kategorien: +naturwissenschaftliche Methode (kurz: +NawiMethode) und +Nachweisen

Die Benennung der Ideen-Kategorien durch Schlagworte ist mit der Überlegung verbunden, dass die Ideen-Kategorien nicht die Vorstellungen der Lernenden abbilden sollen und können, sondern das inhaltliche Element beschreiben, das aus Beobachtersicht im Kern der Äußerung liegt. Damit stellen Ideen-Kategorien einen Zwischenschritt auf dem Weg der Rekonstruktion der Vorstellungen dar, der Situatивität der Äußerungen Rechnung trägt, u. a. durch den bewussten Gebrauch von Schlagworten.

Später in der Arbeit werden zu jeder der im Kodierprozess mit Schlagworten versehenen Ideen-Kategorien (oder Ideen-Oberkategorie, siehe dafür Unterabschnitt 6.1.7) Beschreibungen gegeben (bspw. zu +Meinung: »Eine nicht-naturwissenschaftliche

Frage fragt nach der Meinungen einer Person«). Diese wurden nach dem Kodieren aller für die Arbeit betrachteten Äußerungen formuliert (und finden sich daher im Ergebnisteil, z. B. in Tabelle 6.10). Sie dienen dazu, die zu den Schlagworten zugehörigen Äußerungen zugänglich zu machen, ohne all diese darstellen zu müssen. Die Beschreibungen der Ideen-Kategorien basieren jeweils auf der Durchsicht aller Äußerungen, die mit dem als Kategorien-Name fungierende Schlagwort kodiert sind. Daher sind die Beschreibungen nicht im strengen Sinne als Namen der Ideen-Kategorien anzusehen – das heißt, dass nur für die Schlagworte als Kategorie-Namen die Ausführungen zur Kategorien-Formulierung gelten: Während mit den Schlagworten (zumindest vom Autor der Arbeit) nahtlos weitergearbeitet werden könnte, wenn die Stichprobe vergrößert würde, wären die Beschreibungen am Ende eines solchen Erweiterungsprozesses erneut zu erstellen.

Kodiereinheit (Wie groß ist das Textstück, das kodiert wird?). Eine Äußerung ist begrenzt durch *Sprecherwechsel*, d. h. Ablösung durch anderen Sprecher. Bei Überlappungen wird vorsichtig vorgegangen, weil ggf. – trotz neuer Absätze im Transkript – kein Sprecherwechsel vorliegt, sondern sich zwei Personen parallel äußern. Ferner gilt beim gemeinsamen Formulieren von Sätzen durch mehrere Personen (Person vervollständigt Satz einer anderen Person) die Begrenzung durch Sprecherwechsel nicht. Eine Äußerung ist außerdem durch *Themenwechsel* begrenzt. Antworten auf die Frage »...?« aus dem Selektionskriterium sind nur dann zwei verschiedene Themen, wenn erkennbar eine gesonderte Aussage getroffen wird, ansonsten sind zwei Codes gleichzeitig zuzuweisen. Ein Hinweis auf einen Themenwechsel kann eine längere Pause (auch innerhalb eines Transkriptabsatzes) sein.

In dem Beispiel in Abbildung 6.2 werden zwei Ideen-Kategorien zugewiesen. Die Ideen-Kategorie +Objektiv (Unterstreichung) lässt sich eindeutig S2 zuordnen. Die Ideen-Kategorien +Fachdisziplin Physik und +Fachdisziplin Chemie (beide für die gezackte Markierung) geht über die beiden Personen S2 und S1 hinweg, weil der Satzanfang »handelt es sich« eine intendierte Zuordnung zur Fachdisziplin auch für S2 plausibel macht, auch wenn sie nachfolgend von S1 vorgenommen wird, nur S1 auf Chemie verweist, und unklar bleibt, warum S2 an dieser Stelle stockt.

S2: Weil es eine objektive Frage ist (5) und
 (3) es handelt sich um Phy-
 S1: (2) -sik, (zeigt mit der Hand einen Schrägstrich) Chemie.
 (Lacht)

Abbildung 6.2: Transkript mit zwei Markierungen für Ideen-Kategorien
 (die gezackte Markierung betrifft zwei Personen gleichzeitig).

6.1 Qualitative Inhaltsanalyse zur Rekonstruktion von Vorstellungen der Lernenden

Eine Alternative für die Festlegung der Kodiereinheit bestünde beispielsweise darin, die Vorstellungen einer Person aus allen ihren (untersuchten, siehe dazu *Auswertungseinheit* auf S. 249) Äußerungen zu rekonstruieren, die sich auf eine spezifische Teilaufgabe oder sogar auf eine ausgewählte Karte beziehen. Diese Äußerungen könnten zusammengefasst mit einer Ideen-Kategorie kodiert werden. Denkbar wäre auch, thematische Abschnitte aus dem Diskurs abzuleiten und zugehörige Vorstellungen für diese zu kodieren. Da allerdings für die einzelnen Lernenden und für die einzelnen Situationen untersucht werden soll, welche Vorstellungen vorliegen, ist das geschilderte Vorgehen gegenüber diesen Alternativen zu bevorzugen.

Kontexteinheit (Was wird zur Deutung einer Äußerung mitgenutzt?). In Fällen, in denen Äußerungen nicht sinnvoll alleinstehend gedeutet werden können, wird der unmittelbare (vor allem verbale) Kontext der Äußerung hinzugezogen. Dazu gehören zeitlich direkt vorausgehende Äußerungen (bei mehreren Schülern muss sich das nicht unbedingt auf den vorigen Transkriptabsatz beschränken), aber auch Aufgaben aus dem Lernmaterial (insbesondere wenn Äußerungen Bezug darauf nehmen, bspw. durch Pronomina, die anderweitig nicht gedeutet werden können, oder durch Kausalkonjunktionen, die auf eine Konklusion außerhalb der Äußerung verweisen).

Zur Deutung einer Äußerung werden keine Inhalte von Äußerungen derselben Person hinzugezogen, die in einem anderen Kontext gemacht wurden. Auch Äußerungen, die scheinbar mit geringer Variation wiederholt werden, werden eigenständig kodiert und nicht vor dem Hintergrund des Inhalts einer vorherigen Äußerung kodiert. Diese Beschränkung auf den unmittelbaren Kontext einer Äußerung, dient dazu, dass die mit dieser Äußerung verbundene Vorstellung rekonstruiert wird und dass die Ideen-Kategorie auch für zeit- und aufgabenbezogene Analysen sinnvoll genutzt werden kann. Stellen frühere (oder gar spätere) Äußerungen eine Deutungsfolie für die Kodierung der aktuellen Äußerung dar, kann nicht davon ausgegangen werden, dass Kodierungen sich tatsächlich auf Zeitpunkte bzw. einzelne (Teil-)Aufgaben beziehen. Diese Zuordnung stellt aber einen wichtigen Aspekt für die Forschungsfragen (insbesondere F-Vor6, S. 50) dar.

Auswertungseinheit (Auf welchem Material basieren die entwickelten Kategorien?). Die induktiv generierten Ideen-Kategorien basieren auf den Äußerungen der Lernenden aus den ausgewählten Sequenzen (siehe Abschnitt 6.2 ab S. 269) für alle analysierten Gruppen (Abschnitt 6.3 ab S. 278). Das heißt, dass für jedes weitere hinzugenommene Team auch weitere Ideen-Kategorien entstehen können.

Ein typisches Vorgehen wäre, eine Sättigung der Kategorien zu postulieren und anhand weiterer Videos zu überprüfen (Schreier, 2013). Nach einer festgelegten Anzahl von Teams oder nach einem vorher festgelegten Kriterium zur Sättigung würden dann für weitere Teams keine Ideen-Kategorien generiert werden. Äußerungen, die nicht zugeordnet werden könnten, würden allesamt mit einer einzigen Reste-Kategorie versehen. Davon wird in der Arbeit aber abgesehen, weil die Passung der generierten Ideen-Kategorien möglichst gut sein soll, die Arbeit zudem explorativ angelegt ist und es nicht so sehr auf die Absicherung der Vollständigkeit des Kodiersystems ankommt. Dennoch kann an dieser Stelle angemerkt werden, dass sich in den Analysen selbst herausstellte, dass bei den letzten analysierten Teams nur noch sehr wenige Ideen-Kategorien hinzukamen. Demnach scheint eine Sättigung plausibel zu sein, auch wenn sie nicht überprüft wurde. Im Kontext des bearbeiteten Lernmaterials und der Stichprobe von Elftklässlern eines Gymnasiums scheinen also die generierten Ideen-Kategorien als recht vollständig angesehen werden zu können.

6.1.2 Zuweisen von Kategorien der Formulierung (Ebene 1b)

Neben dem rekonstruierten inhaltlichen Aspekt einer Äußerung, der in der Ideen-Kategorie abgebildet ist, werden Kategorien zur Formulierung der Äußerung zugewiesen, die dazu dienen, negativ formulierte von positiv formulierten Äußerungen zu unterscheiden. Diese Unterscheidung basiert im weiteren Sinne auf der Überlegung, dass es Unterschiede zwischen dem Wissen, dass etwas stimmt oder funktioniert, und dem Wissen, dass etwas nicht stimmt/funktioniert geben könnte (vgl. Lange, 2012; Oser & Spychiger, 2005). Die Kategorienbildung erfolgte vollständig deduktiv.

Als Ausgangspunkt wurde die in Tabelle 6.1 (auf S. 251) schematisch dargestellte Unterscheidung genutzt, dass eine Äußerung, die (mindestens implizit) eine Einschätzung bezüglich eines fachmethodischen Merkmals enthält (bspw., ob eine Frage naturwissenschaftlich ist, ob ein Versuch fair geplant ist, ob eine Versuchsplanung zur Fragestellung passt, ...) und zudem einen Hinweis auf damit verbundene Vorstellungen zu einem fachmethodischen Konzept oder Bereich enthält (bspw., was eine naturwissenschaftliche Frage kennzeichnet, was einen fairen Versuch von einem unfairen unterscheidet, woran sich Passung von Frage und Versuchsplanung festmachen lässt, ...) hinsichtlich zweier Dimensionen der Formulierung eingeschätzt werden kann (sofern sie hinreichend einfach formuliert ist – was für verbalen Diskurs, der in der vorliegenden Arbeit untersucht wird, zumeist der Fall ist):

- (1) Fachmethodisches Merkmal laut (implizitem) Gehalt der Äußerung gegeben?
(p für plus, m für minus)

Tabelle 6.1: Allgemeines Schema für Kategorien der Formulierung von Äußerungen.

Fachmethodisches Merkmal ...	Ideen-Aspekt trifft zu (+)	Ideen-Aspekt trifft nicht zu (–)
gegeben / erfüllt (p)	p+	p–
nicht gegeben / erfüllt (m)	m+	m–
Unklarer Formulierungstyp	nn	

- (2) Äußerung positiv oder negativ formuliert?
(+ für positiv, – für negativ)

Als positive oder negative Formulierung wird dabei angesehen, ob darauf verwiesen wird, dass ein benanntes Kriterium erfüllt (positiv) oder nicht erfüllt (negativ) ist. Beispielsweise könnte das Kriterium, dass *Kontrollvariablen verändert* wurden, *negativ* genutzt werden, um zu begründen, dass der Versuch fair ist: »Der Versuch ist fair, weil die Kontrollvariablen *nicht* verändert wurden« (zugehörig zur Tabellenzelle oben rechts: p–). Anhand des Beispiels wird deutlich, dass sich die Einträge auf der Hauptdiagonalen (bzw. auf der Gegendiagonalen) von Tabelle 6.1 als doppelte Negationen voneinander begreifen lassen. Beispielsweise ist das Kriterium, dass die *Kontrollvariablen verändert* wurden eine Negation des Kriteriums im obigen Beispiel; der Satz »Der Versuch ist *nicht* fair, weil die Kontrollvariablen verändert wurden« ist sehr ähnlich zum obigen Beispiel (aber in der Logik nicht gleich!, daher die Kategorie m+). Das für verschiedene fachmethodische Merkmale verwendete Vorgehen wird nachfolgend anhand zweier Beispiele erläutert; diese beiden Bereiche werden im weiteren Verlauf der Arbeit für die weiteren Analysen benötigt.

Die Unterscheidung bezieht sich darauf, in welcher Weise die Frage »...« aus dem Selektionskriterium beantwortet wird. Für Kennzeichen von naturwissenschaftlichen Fragen lautet diese Frage »Was kennzeichnet eine naturwissenschaftliche Fragestellung?«. Lässt sich eine Äußerung als eine Antwort auf den *genauen* Wortlaut »Was kennzeichnet eine *naturwissenschaftliche* Fragestellung?« deuten, so wird eine Kategorie beginnend mit n zugewiesen (n steht für: Formulierung für *naturwissenschaftlich*). Eine Kategorie beginnend mit u wird hingegen vergeben, wenn es darum geht, was eine *nicht*-naturwissenschaftliche Frage kennzeichnet (u für: Formulierung für *unnaturwissenschaftlich*).

Wie in Tabelle 6.2 dargestellt, wird das allgemeine Vorgehen direkt auf Kennzeichen von naturwissenschaftlichen Fragen übertragen. Wird erwähnt, dass eine naturwissenschaftliche (bzw. nicht-naturwissenschaftliche) Frage dadurch gekennzeichnet ist, dass ein gewisser inhaltlicher Aspekt (durch das entsprechende Schlagwort im Kategorien-Namen abgebildet) erfüllt ist, so wird ein + angehängt (bspw. »Ist eine naturwissenschaftliche Frage $[\rightarrow n]$, denn du kannst es mit einem Versuch untersuchen $[\rightarrow +]$ « $\rightarrow n+$; »Ist keine $[\rightarrow u]$, weil es hier um Meinung geht $[\rightarrow +]$ « $\rightarrow u+$). Ein – wird angehängt, wenn die Nicht-Erfüllung des inhaltlichen Aspekts (abgebildet im Schlagwort) zu der Einschätzung führt (bspw. »Ist keine, Du kannst es ja nicht mit einem Versuch untersuchen« $\rightarrow u-$; »Das ist eine naturwissenschaftliche Frage, weil es hier nicht um Meinung geht« $\rightarrow n-$).

Die in den beiden vorigen Absätzen dargestellten Bestandteile der Kodierung lassen sich durch die Vierfeldermatrix in Tabelle 6.2 (auf S. 253) veranschaulichen. Die Zuweisungsentscheidungen werden sowohl für Zeilen als auch Spalten der Tabelle, soweit möglich, am Wortlaut der Äußerung festgemacht (wie für die Beispiele oben dargelegt). Es ist danach u. U. auch erlaubt, aus dem unmittelbaren Kontext zu schließen, wie die Aussage gemeint ist. Trifft aus Beobachtersicht keine der vier Varianten am plausibelsten zu, so wird der Code *nn* vergeben (und in einem Kommentarfeld vermerkt, ob und ggf. inwiefern die Zeilen- oder die Spalten-Entscheidung problematisch war).

Für die Unterscheidung von Beobachtung und Deutung wird davon ausgegangen, dass Beobachtungen und Deutungen als Gegensätze aufgefasst werden können. Sie stellen keinesfalls Negationen voneinander dar, wie dies naturwissenschaftliche und nicht-naturwissenschaftliche Fragen tun. Dennoch ist das Lernmaterial – mindestens in der ausgewählten Sequenz (siehe später in Unterabschnitt 6.2.2 ab S. 272) – so zu bearbeiten, dass entweder Beobachtungen oder Deutungen zuzuweisen sind. In diesem Sinne spannen sowohl *n* und *u* als auch *b* und *d* ein gesamtes Universum (einmal aller Fragen, einmal aller Aussagen) auf, in dem es nur jeweils zwei Optionen und keine Restmenge gibt. Im Kontext der Ergebnisdarstellung wird diese Trennung kritisch anhand der Ergebnisse diskutiert.

Basierend auf Tabelle 6.1 (auf S. 251) ergibt sich für die Unterscheidung von Beobachtung und Deutung vor dem Hintergrund der Ausführungen im letzten Absatz die in Tabelle 6.3 (auf S. 253) dargestellte Unterscheidung von Kategorien zur Formulierung. Dabei lässt sich die Kategorie *b–* beispielsweise (im Sinne eines orientierenden Merksatzes) für den Kodiervorgang derart umreißen: Es ist eine Beobachtung $[\rightarrow b]$, weil es nicht $[\rightarrow -]$... Dabei steht ... für den durch das Schlagwort jeweils abgebildeten inhaltlichen Aspekt der Äußerung steht, beispielsweise: Es ist eine Beobachtung, weil es nicht mit der Meinung von Personen zu tun hat.

Tabelle 6.2: Kategorien der Formulierung von Äußerungen für Kennzeichen von naturwissenschaftlichen Fragen.

	Ideen-Aspekt trifft zu (+)	Ideen-Aspekt trifft nicht zu (–)
Naturwissenschaftliche Frage- stellung (n)	n+	n–
Nicht-naturwissenschaftliche Fragestellung (u)	u+	u–
Unklarer Formulierungstyp	nn	

Tabelle 6.3: Kategorien der Formulierung von Äußerungen für die Unterscheidung von Beobachtung und Deutung.

	Ideen-Aspekt trifft zu (+)	Ideen-Aspekt trifft nicht zu (–)
Beobachtung (b)	b+	b–
Deutung (d)	d+	d–
Unklarer Formulierungstyp	nn	

6.1.3 Zuweisen von Kategorien des Modus (Ebene 2)

In der qualitativen Inhaltsanalyse wird der Modus der Äußerung durch in einem deduktiv-induktiven Verfahren generierte Kategorien abgebildet. Diese Kategorien dienen dazu, verschiedene Strukturmerkmale der Äußerungen abzubilden, die für die Einschätzung der Qualität der Vorstellungen der Lernenden relevant sein könnten. Beispielsweise wird kodiert, ob eine Äußerung sich auf einen allgemeinen Sachverhalt bezieht oder auf ein konkretes Beispiel.

Wie in Abbildung 6.3 (als Erweiterung von Abbildung 6.1 auf S. 242) dargestellt, werden abhängig von der Erfüllung des Selektionskriteriums und der Zuweisung einer Ideen-Kategorie auf der zweiten Ebene Kategorien zum Modus vergeben. Bei zugewiesener Ideen-Kategorie wird mindestens eine Modus-Kategorie aus 2a oder 2b vergeben. Die Modus-Kategorien aus 2a lassen sich nur sinnvoll zuweisen, wenn sich aus der Äußerung ein inhaltlicher Beitrag zum fachmethodischen Diskurs über Kennzeichen von naturwissenschaftlichen Fragen bzw. die Unterscheidung von Beobachtung und Deutung ableiten lässt, der die Kodierung mit einer Ideen-Kategorie möglich macht. Äußerungen, die nicht mit einer Ideen-Kategorie kodiert

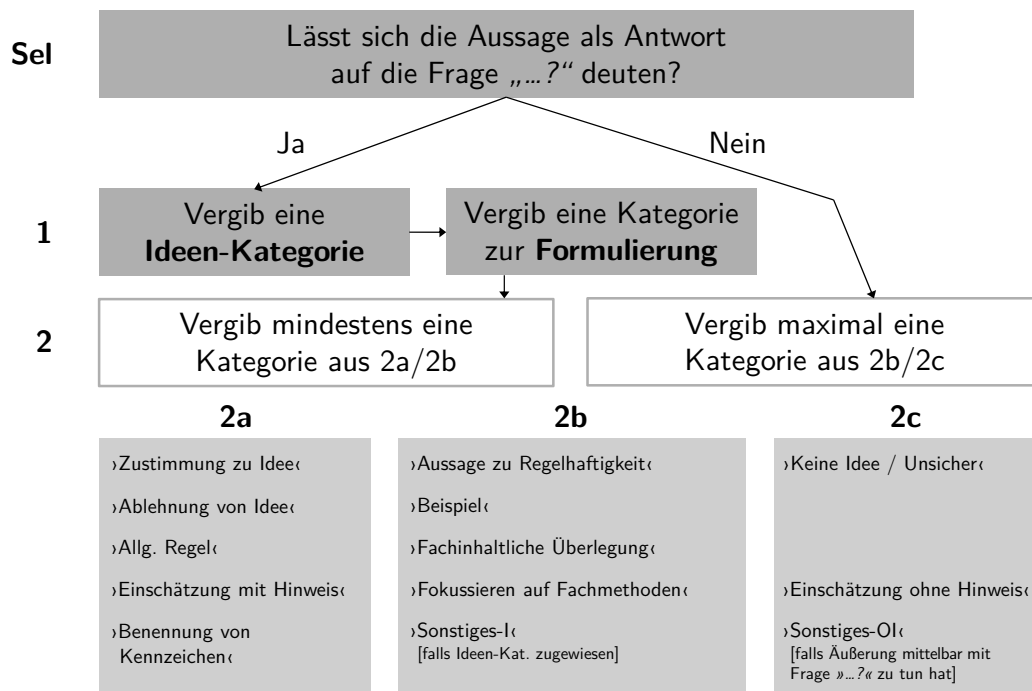


Abbildung 6.3: Vorgehen auf Ebene 2 der qualitativen Inhaltsanalyse zum Modus der Äußerungen. (Sel = Selektionskriterium)

6.1 Qualitative Inhaltsanalyse zur Rekonstruktion von Vorstellungen der Lernenden

sind, weil sie das Selektionskriterium nicht erfüllen, *können* mit maximal einer Kategorie aus 2b oder 2c versehen werden, wenn die Äußerung eine Stellung in der Aufgabenbearbeitung oder im Diskurs einnimmt, die mittelbar Bezug zur Frage »...?« aus dem Selektionskriterium hat. Beispiele dafür finden sich in den Tabellen 6.4 bis 6.6 in den Spalten der Ankerbeispiele.

Die Kategorien basieren auf Kategoriensystemen aus anderen Studien (u. a. Rogge, 2010; Steckenmesser-Sander, 2015) und leiten sich auch aus den Überlegungen zu den Forschungsfragen (siehe Abschnitt 3.2, ab S. 48) ab. Insofern, als dass während der Kodierung für die ersten 5 Teams noch Anpassungen an den Kategorien vorgenommen

Tabelle 6.4: Kategorien zum Modus mit Beschreibungen und Ankerbeispielen – zu Teil 2a aus Abbildung 6.3.

Code	Beschreibung	Beispiel
Allgemeine Regel	Allgemeine Regel zur Identifikation [z. B. von naturwissenschaftlichen Fragestellungen] oder Klassifikation [z. B. von Fragestellungen].	• »[Naturw. Fragen] Kannst du immer mit Versuchen beantworten.«
Einschätzung mit Hinweis	Einschätzung [z. B. einer (Beispiel-) Fragestellung] mit Hinweis auf Idee. [Beachte 1. $\exists$ Ohne Hinweis $\rightarrow$ siehe Tabelle für 2c. Beachte 2. Auch kodieren, wenn Einschätzung von Material oder anderer Person getroffen wird und von kodierter Person begründet wird.]	• »Das ist eine naturw. Frage, kannst du ja messen.« • »Das ist keine naturw. Frage, weil das subjektiv ist.«
Benennung von Kennzeichen	Benennung von identifizierten Objekten/Vorgängen [z. B. aus einer (Beispiel-)Fragestellung]. [Immer mit Hinweis verbunden, aber nicht notwendigerweise mit <i>Einschätzung</i> mit Hinweis verbunden, daher wird Einschätzung mit Hinweis ggf. zusätzlich kodiert.]	• »Du kannst messen, <i>wie lange das zum Runterfallen braucht</i>.« • »Das ist keine naturw. Frage, weil <i>der Geschmack von Wasser</i> subjektiv ist.«
Zustimmung zu Idee	Zustimmung zu einer zum Selektionskriterium passenden Äußerung [z. B. eine als Antwort auf die Frage nach Kennzeichen naturw. Fragestellungen deutbare Aussage]. [Kodierbar mit keinem anderen Code von Ebene 2.]	• [S2 liest Erläuterung vor, was eine naturw. Frage kennzeichnet] S1: »Ja, das stimmt« [Abhängig von vorgelesenem Text kann Ideen-Kat. evtl. Idee unklar sein.]
Ablehnung von Idee	Ablehnung von einer als Antwort auf die Frage nach Kennzeichen naturwissenschaftlicher Fragestellungen deutbaren Aussage. [Kodierbar mit keinem anderen Code von Ebene 2.]	• [S1: „Ist nicht-naturw., weil es nur ums Rechnen geht.“] S2: »Doch, Rechnen ist auch naturwissenschaftlich«

Anmerkungen. Für die Beispiele gelten die folgenden Markierungen: (1) Hinweis auf Idee: gesperrt gesetzt. (2) Aus einer Frage benannte Objekte/Vorgänge: *kursiv* gesetzt.

Tabelle 6.5: Kategorien zum Modus mit Beschreibungen und Ankerbeispielen – zu Teil 2b aus Abbildung 6.3.

Code	Beschreibung	Beispiel
Aussage zu Regelhaftigkeit	Regelhaftigkeit beim Einschätzen oder Begründen wird sprachlich benannt oder angedeutet. [Nicht notwendigerweise mit Einschätzung o. Ä. verbunden, daher wird Einschätzung o. Ä. ggf. zusätzlich kodiert; nicht notwendigerweise, aber oft, mit Idee verbunden.]	• »Wir können einfach immer sagen: weil es beweisbar ist.« • »Willst du wirklich immer <i>subjektiv</i> schreiben?« • »Kannst du immer die gleiche Begründung hinschreiben.«
Beispiel	Formulierung einer Beispielfragestellung. [Nicht notwendigerweise mit Einschätzung o. Ä. verbunden, daher wird Einschätzung o. Ä. ggf. zusätzlich kodiert.]	• »Hat X einen Einfluss auf Y, das wäre eine naturw. Frage.«
Fachinhaltliche Überlegung	Fachinhaltliche Überlegungen zur Beantwortung einer Beispielfragestellung.	• [Liest: Hat X einen Einfluss auf Y?] »Nein, das hat keinen Einfluss.«
Fokussieren auf Fachmethoden	Hinweis, dass es um Fachmethodisches und nicht um Fachinhaltliches geht. [Kodierbar mit keinem anderen Code von Ebene 2.]	• »Wir sollen die Frage nicht beantworten, sondern sagen, ob sie naturw. ist.«
Sonstiges-I	Falls Ideen-Code zugewiesen wird, aber kein anderer Code aus 2b oder 2c passt. [Kodierbar mit keinem anderen Code von Ebene 2.]	

Tabelle 6.6: Kategorien zum Modus mit Beschreibungen und Ankerbeispielen – zu Teil 2c aus Abbildung 6.3.

Code	Beschreibung	Beispiel
KeinIdee/ Unsicher	Person äußert, dass sie keine Idee hat oder keine Einschätzung treffen kann oder unsicher ist, ob andere Person recht hat. [Kodierbar mit keinem anderen Code von Ebene 2.]	• [S1 liest: Ist diese Frage eine naturw. Frage?] S2: »Ich weiß es nicht.« • »Bist du dir da wirklich sicher, dass es immer eine eindeutige Antwort geben muss?«
Einschätzung ohne Hinweis	Einschätzung einer Beispielfragestellung ohne Hinweis auf Idee oder Fachinhalt. [Kodierbar mit keinem anderen Code von Ebene 2.]	• »Das ist keine naturw. Frage.« • »Ja«/»Nein« [falls nicht eindeutig Fachinhaltliche Überlegung]
Sonstiges-OI	Falls kein Ideen-Code zugewiesen werden kann, aber die Aussage mittelbar mit der Frage »Was kennzeichnet eine naturw. Frage?« zu tun hat. [Kodierbar mit keinem anderen Code von Ebene 2.]	

6.1 Qualitative Inhaltsanalyse zur Rekonstruktion von Vorstellungen der Lernenden

wurden, die auf dem konkreten Material basieren und nicht nur Ausschärfungen der Kategorien-Definitionen waren, kann die Kategorienbildung als deduktiv-induktiv bezeichnet werden. Beispielsweise wird die Unterscheidung von Fall und Konzept (basierend auf z. B. von Aufschnaiter & von Aufschnaiter, 2003) auf das experimentbezogene Denken und Arbeiten übertragen – allerdings nicht direkt, sondern vermittelt der Modus-Kategorien *Allgemeine Regel* und *Einschätzung einer Frage*. Anhand der vorliegenden Daten hat sich eine derartige Unterscheidung als tragfähig erwiesen, wobei induktiv zusätzlich die Differenzierung von *Einschätzung einer Frage mit Hinweis* (auf eine Idee) und *Einschätzung einer Frage ohne Hinweis*⁸⁸ ergänzt wurde, um eine Trennung verschiedener Arten von Einschätzungen von Situationen/Fällen zu ermöglichen. Als weitere Beispiele für das deduktiv-induktive Vorgehen können die Kategorien *Fachinhaltliche Überlegung* und *Aussage zu Regelhaftigkeit* genannt werden. Die erste der beiden Kategorien ist klar mit den Überlegungen zur Zielstellung des Projekts und den Forschungsfragen verbunden, weil die Differenzierung von fachmethodischen Konzepten und Überlegungen vs. fachinhaltlichen Konzepten und Überlegungen eine zentrale Rolle spielt. Sie ist deduktiv erzeugt und hat das Ziel, fachinhaltliche Äußerungen im Kontext fachmethodischer Fokussierungsphasen zu identifizieren. Die zweite der beiden Kategorien wurde induktiv hinzugefügt, weil Lernende Regelmäßigkeiten benennen, die sie im Prozess erkennen. Es gibt einen Unterschied zwischen der Äußerung »Alle naturwissenschaftlichen Fragen beziehen sich nur auf Objektives« (*Allgemeine Regel*) und der Äußerung »Hier kannst Du schon wieder schreiben, dass diese Frage naturwissenschaftlich ist, weil sie sich nur auf Objektives bezieht« (*Aussage zu Regelhaftigkeit* sowie *Einschätzung einer Frage mit Hinweis*), der durch diese Kategorie abgebildet wird.

6.1.4 Kodieren der zugehörigen (Teil-)Aufgabe (Ebene 3)

Die qualitative Inhaltsanalyse umfasst auch das Kodieren der (Teil-)Aufgabe, auf die sich die Äußerung bezieht. Diese Kategorien dienen der Zuordnung von Äußerungen, ihren Strukturmerkmale sowie den zugehörigen Vorstellungen zu den Elementen der Instruktion und können genutzt werden, um Personen in Bezug auf die Elemente der Instruktion zu vergleichen.

Das Lernmaterial besteht wie bereits dargestellt aus nummerierten Karten, die teilweise nummerierte Aufgaben beinhalten. Damit jedoch identifiziert werden kann,

⁸⁸Die Kategorie-Namen sind in den nächsten (Unter-)Abschnitten zumeist ohne Leerzeichen dargestellt (teilweise bei Auslassung mancher Worte), damit erkennbar ist, dass jeweils genau eine Kategorie gemeint ist; beispielsweise: *AllgemeineRegel*, *EinschätzungMitHinweis* oder *EinschätzungOhneHinweis*.

auf einzelne Bestandteile der Aufgaben sich die Äußerungen der Schülerinnen und Schüler beziehen, werden auch die als Teilaufgaben bezeichneten Elemente (siehe Abschnitt 6.2 beginnend auf S.269) einzeln kodiert. Typischerweise werden die Teilaufgaben in den Aufgaben-Kategorien durch kleine römische Zahlen abgebildet. Dabei wird mit der Zählung links oben auf der Aufgabenkarte begonnen, nach unten und dann nach rechts gezählt. Ein Beispiel findet sich in Abbildung 6.4.

Auf Ebene 1 oder 2 kodierte Äußerungen werden mit einer Aufgaben-Kategorie versehen. Dazu wird rekonstruiert, auf welche Aufgabe sich die Äußerung bezieht bzw. im Kontext welcher Aufgabe sie gemacht wird (bspw. FH25-A; Aufgaben-Kategorien werden nicht serifenlos gesetzt). Bezieht sich eine Äußerung auf eine Teilaufgabe, so wird die Kategorie der Teilaufgabe (z. B. FH21-ii, FH25-A-i) vergeben; wird eine Äußerung getroffen, die mehrere Teilaufgaben umfasst (z. B. 1-21-ii und 1-21-iii) oder allgemein mit naturwissenschaftlichen Fragen zu tun hat, so wird die nächsthöhere Kategorie (also Aufgabe oder Karte) zugewiesen (im Beispiel: 1-21).

FH21
21

Eine naturwissenschaftliche Fragestellung?

AUFGABE

Nicht jede Fragestellung, die etwas mit Physik, Chemie oder Biologie zu tun hat, ist automatisch eine *naturwissenschaftliche* Fragestellung. Anders herum können auch alltägliche Fragen naturwissenschaftliche Fragen sein.

Bei welchen der folgenden Fragestellungen handelt es sich Ihrer Einschätzung nach um naturwissenschaftliche Fragestellungen?

FH21-i
☐

Ist naturwissenschaftliches Arbeiten ein spannendes Thema für Schüler/innen?

FH21-ii
☐

Wird Kaffee mit Milch schneller kalt als Kaffee ohne Milch?

FH21-iii
☐

Hat der Luftdruck einen Einfluss darauf, bei welcher Temperatur Wasser kocht?

FH21-iv
☐

Beeinflusst die Größe eines Hundes, wie oft er im Durchschnitt pro Stunde bellt?

FH21-v
☐

Schmeckt Mineralwasser mit Kohlensäure besser als Leitungswasser?

FH21-vi
☐

Von welchen Faktoren hängt die Endgeschwindigkeit eines Fallschirmspringes ab?

FH21-vii
☐

Sollten Schüler/innen etwas darüber lernen, welche Faktoren die Periodendauer eines Fadenpendels beeinflussen?

Abbildung 6.4: Nummerierung der (Teil-)Aufgaben auf Karte 1-21. Die Karte enthält nur eine Aufgabe mit mehreren Teilaufgaben, ansonsten ließen sich 1-21-1-i, 1-21-1-ii, ... und 1-21-2-i unterscheiden, wobei 1-21-1 und 1-21-2 die Aufgaben auf der Karte wären.

6.1.5 Kodieren der Korrektheit der Einschätzungen (Ebene 4)

Bei der Bearbeitung des Lernmaterials nehmen die Schülerinnen und Schüler Einschätzungen vor, deren Korrektheit ebenfalls im Rahmen der qualitativen Inhaltsanalyse kodiert wird. Das Kodieren der Korrektheit der Einschätzungen dient nicht dazu, direkte Aussagen über die Vorstellungen der Lernenden zu machen. Vielmehr ist das Kodieren der Korrektheit als Komponente der Analyse anzusehen, bei der die Fähigkeiten der Lernenden in den Blick genommen werden, u. a. um diese mit den Vorstellungen in Verbindung zu bringen (siehe F-Vor4, S. 50). Fähigkeiten zum experimentbezogenen Denken und Arbeiten umfassen unter anderem auch Entscheidungen im Prozess des experimentbezogenen Denkens und Arbeitens. Zu diesen Entscheidungen gehören beispielsweise die Einschätzungen von Fragen (bspw. im Kontext von Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen), Versuchsplanungen (bspw. im Kontext von Vorstellungen zu fairen Versuchen) und Aussagen (bspw. im Kontext von Vorstellungen zur Unterscheidung von Beobachtung und Deutung). Die Fähigkeiten der Lernenden können dann in den Auswertungen auf die Vorstellungen der Lernenden bezogen werden.

Für jede Äußerung, die mit *EinschätzungOhneHinweis* oder *EinschätzungMitHinweis* kodiert wurde, wird kodiert, ob die Person in Übereinstimmung mit der im Rahmen des Aufgabenmaterials als angemessen anzusehende Einschätzung entschieden hat (Kategorien: *korrekt*, *inkorrekt* und *unklar*). Fast immer geht die im Material intendierte Einschätzung aus später folgenden Lösungen (z. B. Lösungskarten) hervor. In den anderen Fällen wurde sie aus dem Gesamtkontext deduziert. Für naturwissenschaftliche Fragen sollen beispielsweise auf 1-21 (der einundzwanzigsten Karte der ersten Einheit der explizit-fachmethodischen Instruktionsvariante) sechs Beispielfragen eingeschätzt werden (siehe Abbildung 6.4 auf S. 258), auf 1-22 finden sich dann die intendierten Lösungen (siehe Abbildung 6.5). Die Einschätzungen der

22	Eine naturwissenschaftliche Fragestellung?	KONTROLLE
Überprüfen Sie Ihre Einschätzungen!		
☐	Ist naturwissenschaftliches Arbeiten ein spannendes Thema für Schüler/innen?	
☒	Wird Kaffee mit Milch schneller kalt als Kaffee ohne Milch?	
☒	Hat der Luftdruck einen Einfluss darauf, bei welcher Temperatur Wasser kocht?	
☒	Beeinflusst die Größe eines Hundes, wie oft er im Durchschnitt pro Stunde bellt?	
☐	Schmeckt Mineralwasser mit Kohlensäure besser als Leitungswasser?	
☒	Von welchen Faktoren hängt die Endgeschwindigkeit eines Fallschirmspringes ab?	
☐	Sollten Schüler/innen etwas darüber lernen, welche Faktoren die Periodendauer eines Fadenpendels beeinflussen?	

Abbildung 6.5: Lösungen auf Karte 1-22, die genutzt werden können, um Einschätzungen der Lernenden auf ihre Korrektheit zu überprüfen (aus der Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen).

Lernenden bei 1-21 und 1-22 werden als korrekt kodiert, wenn sie mit den auf 1-22 aufgeführten Lösungen inhaltlich übereinstimmen.

Wenn eine Entscheidung durch direkt darauf folgende Worte revidiert wird, wird nur die Entscheidung aus der Revision kodiert. Aus dem Vorgehen beim Transkribieren und beim Prüfen des Selektionskriteriums ergibt sich, dass es typischerweise innerhalb eines Absatzes nur eine Einschätzung gibt. Daher kann der Beginn eines neuen Absatzes als Orientierung dafür genutzt werden, welche Worte noch als direkt folgende Revision angesehen werden. Wenn im nächsten Absatz derselben Person eine andere Entscheidung getroffen wird, werden auf jeden Fall zwei verschiedene Entscheidungen kodiert, weil dann auch zwei Instanzen mit *EinschätzungOhneHinweis* bzw. *EinschätzungMitHinweis* kodiert wurden.

Im Kodierprozess wurde eine weitere Kategorie eingeführt, um abzubilden, dass es Äußerungen gibt, die sich nicht klar als korrekt oder inkorrekt kodieren lassen, weil sie nicht den aus dem Material (direkt oder deduktiv) entnommenen Einschätzungen entsprechen, aber aus fachmethodischer Sicht angemessen erscheinen. Die zugehörige Kategorie heißt *über dem Erwartbaren* und wurde nur für die Sequenz zur Unterscheidung von Beobachtung und Deutung vergeben. Dort haben Schülerinnen und Schüler einzuschätzende Aussagen (die zumeist vollständige Sätze waren) in Teilaussagen (bspw. Haupt- und Nebensatz) zerlegt und die jeweiligen Teilaussagen eingeschätzt, während später im Material eine Einschätzung der gesamten Aussage (d. h. des gesamten Satzes) vorgenommen ist. Ausführlicher wird dies im Ergebnisteil (insbesondere auf S. 570 und folgenden) behandelt.

6.1.6 Beurteilerübereinstimmung

Zur Überprüfung der Anwendbarkeit und Güte des geschilderten Vorgehens für die qualitative Inhaltsanalyse wurden zwei weitere Kodierer (P1 und P2) hinzugezogen. Diese Kodierer waren im selben Institut wie der Autor der Arbeit tätig und mit Physikdidaktik und naturwissenschaftlichen Methoden grundsätzlich vertraut. Ihnen wurde eine vorläufige Fassung der Beschreibung der qualitativen Inhaltsanalyse vorgelegt (die weniger Beispiele als die Fassung in dieser Arbeit umfasste und nur wenige Begründungen für das Vorgehen und keine Ausführungen zu Einschränkungen der Methode enthielt, aber ansonsten sehr ähnlich war). Ziel der Überprüfung war im Sinne einer formativen Reliabilitätsprüfung (vgl. Mayring, 2013), die Beschreibung des Vorgehens für die qualitative Inhaltsanalyse daraufhin zu prüfen, ob sie bei verschiedenen Personen zu ähnlichen Ergebnissen führt. Außerdem wurden Rückmeldungen zu Schwierigkeiten bei der Umsetzung der qualitativen Inhaltsanalyse eingeholt, anhand derer das Verfahren und dessen Beschreibung überarbeitet werden sollte.

Die Überprüfung basierte auf dem Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen (Details in Unterabschnitt 6.2.1, ab S. 269). Zu diesem Instruktionsextrakt wurde die Videosequenz für das Team der Personen C7m, C8m, C9w ausgewählt, weil die Qualität des zugehörigen Transkriptes und Audio- und Video-Qualität des Videos von den bisher vom Autor der Arbeit gesichteten Teams am besten war. Für die Überprüfung erhielten die beiden Kodierer eine MAXQDA-Datei, die das Video, ein zugehöriges Transkript sowie das vorher angelegte Kategoriensystem (freilich ohne die induktiv zu generierenden Ideen-Kategorien) enthielt.

Die Überprüfung der Übereinstimmung erfolgte mithilfe der Ausgabe von MAXQDA für jedes einzelne kodierte Segment. Für einen zusammenfassenden Bericht wird eine inhaltliche Gegenüberstellung der Ideen-Kategorien (siehe Tabelle 6.7 auf S. 263) sowie eine Berechnung der κ -Werte in den jeweiligen Ebenen des Kategoriensystems (Ideen-Kategorien, Formulierung, Modus, Aufgaben) vorgenommen (siehe Tabelle 6.8 auf S. 263; Berechnungen gemäß Unterabschnitt 5.1.2 beginnend auf S. 76, bei einer Minimalüberlappung der kodierten Textsegmente von 50 %).

Die inhaltliche Gegenüberstellung (Tabelle 6.7 auf S. 263) zeigt, dass von den drei Personen zwar meist unterschiedliche Schlagworte gewählt wurden, diese aber einander klar zugeordnet werden können. Ein großer Teil der Ideen-Kategorien wurde in sehr ähnlicher Weise von den drei Personen generiert. Auf dieser Ebene wurden nur die Schlagworte miteinander abgeglichen, nicht jedoch untersucht, ob die drei Personen jeweils auch die gleichen Äußerungen mit den generierten Ideen-Kategorien kodiert haben. Aus Tabelle 6.8 (S. 263) lässt sich allerdings ableiten, dass die Personen

die Ideen-Kategorien auch ähnlichen Instanzen zugewiesen haben. Zur Berechnung der Übereinstimmung wurden für jede Zeile aus Tabelle 6.7 (abwechselnd weiß oder grau hinterlegt) eine neue Kategorie erstellt und den Äußerungen zugewiesen, die mit einer der Kategorien aus der jeweiligen Zeile kodiert waren. Die κ -Werte sind für die Ideen-Kategorien mit der geschilderten Anpassung und auch für die anderen Ebenen des Kodiersystems überall gut (teilweise sehr gut; nach Wirtz & Caspar, 2002). Da ein Großteil der Kategorien als hochinferent einzuschätzen ist, sind die Übereinstimmungen vermutlich sogar eher als etwas besser einzuschätzen als in typischen Vergleichstabellen suggeriert. Auch detaillierte Betrachtungen der einzelnen kodierten Segmente legen nahe, dass die Ideen-Kategorien mit guter bis sehr guter Übereinstimmung generiert und zugewiesen wurden. Im Gespräch über die Äußerungen, zu denen abweichende Kodierungen vorlagen, stellte sich zudem häufig heraus, dass die Option der anderen Personen ebenfalls angedacht worden war. Ferner lässt sich der vergleichsweise niedrige κ -Wert von .66 bei den Aufgaben gemäß dieser Gespräche vermutlich dadurch erklären, dass sich P2 – anders als P1 und der Autor der Arbeit – häufiger dazu entschieden hat, einzelne Äußerungen einer höheren Ebene (bspw. der Karte statt der Aufgabe) zuzuordnen, sobald dazu in Ansätzen die Möglichkeit bestand. Minimale Anpassungen im Kodiermanual wurden im Anschluss an die Gespräche vorgenommen, um einen ähnlicheren Umgang mit den identifizierten Unsicherheiten herbeizuführen.

Aus den vertieften Diskussionen des Autors der Arbeit mit den Kodieren über das Kodiervorgehen ergaben sich weitere Hinweise. Da meist nur einzelne Transkriptabsätze als Belege genutzt werden konnten, wurden die Hinweise vom Autor bei der Kodierung des weiteren Materials intensiver geprüft. Es ergibt sich mit Relevanz für die Interpretation der Kodierungen: Die Anzahl der Äußerungen, die mit Ideen-Kategorien kodiert wurden, ist mit Vorsicht zu genießen. Manchmal sind es vielfache Äußerungen, die direkt nacheinander folgen und nur von einem Einwurf einer anderen Person unterbrochen wurden (auf den teilweise nicht einmal eingegangen wird), in manchen Fällen sind auch zu ein und derselben (Teil-)Aufgabe an zeitlich deutlich getrennten Stellen Einschätzungen vorhanden (z.B., weil beim Aufschreiben nochmals diskutiert wird), so dass die Anzahl von Einschätzungen (und auch anderen Kategorien) in wenigen Fällen stark durch mehrfaches Diskutieren über dieselbe (Teil-)Aufgabe in die Höhe schnellte. (Diese Ausführungen gelten in ähnlicher Weise für solche Äußerungen mit Kategorien zum Modus, die das Selektionskriterium nicht erfüllen und daher ohne Ideen-Kategorie sind.) In den Auswertungen wird diesem Umstand dadurch begegnet, dass pro (Teil-)Aufgabe ausgewertet wird (und nicht pro Äußerung).

Tabelle 6.7: Übersicht über die generierten Ideen-Kategorien für die verschiedenen Personen bei der Überprüfung der Beurteilerübereinstimmung.

Person 1	Person 2	Autor der Arbeit
IdeeUnklar	IdeeUnklar	IdeeUnklar
Individuell	IndividuelleAntwort	IndividuelleErgebnisse
Messen	Messbar	Messen
Beweis	Beweise	Beweisen
Versuch	Versuch	Versuch
Allgemeine/normale	AllgemeineFrage	AllgemeinVsNaturw...
Frage	NormaleFrage	NormaleFrage
Naturw...Eigenschaften	NaWiEigenschaften	NaWiEigenschaften
Naturw...Methoden	NaWiMethoden	
BeeinflussungVon		
Größen		
AbgrenzungVonPhysik-		
alischenBegriffen		
		HatZuTunMitNawi
		HörtSichSoAn

Tabelle 6.8: Übersicht über die nach Teilen des Kategoriensystems getrennt berechneten κ -Werte der Beurteilerübereinstimmung im Vergleich der beteiligten Personen (κ nach Brennan & Prediger, 1981; Details in Unterabschnitt 5.1.2 ab S. 76).

Personen	Ideen	Formulierung	Modus	Aufgabe
P1 vs. Autor	.79	.75	.72	.70
P2 vs. Autor	.78	.70	.70	.66
P1 vs. P2	.77	.81	.70	.77

6.1.7 Generieren von Ideen-Oberkategorien

Nach Abschluss des Kodierens wurde zur Aufbereitung der Daten eine Zusammenführung der Ideen-Kategorien in Ideen-Obergruppen vorgenommen. Die Zusammenführung basiert in Teilen auf den generierten Daten wie den Häufigkeiten der Ideen-Kategorien bei der analysierten Stichprobe, sie lässt sich allerdings bei Erweiterung der Stichprobe vergleichsweise einfach erneut durchführen bzw. ergänzen. Das Bilden von Ideen-Oberkategorien kann als weiterer Schritt der qualitativen Inhaltsanalyse angesehen werden, bei dem die Ideen-Kategorien induktiv-deduktiv auf einem höheren Abstraktionsniveau zusammengefasst werden.

Die Ideen-Oberkategorien wurden dabei als eine *inhaltliche* Zusammenführung durchgeführt. Das heißt, die Schlagwörter und einzelnen zugehörigen Äußerungen wurden auf ihre inhaltlichen Gemeinsamkeiten hin untersucht und demnach gebündelt. Was als eher ähnlich angesehen wurde, hängt zum Teil auch von theoretischen Überlegungen ab, weshalb das Vorgehen oben als induktiv-deduktiv bezeichnet ist. Ein Fokus auf das Inhaltliche bedeutet auch, dass bei der Bildung von Ideen-Oberkategorien Äußerungen nicht nach Personen oder Teams zusammengelegt wurden. Im Gegenteil wurde dadurch, dass auch wörtlich nicht übereinstimmende, aber dem Inhalt der Vorstellung aus Beobachtersicht nach nah beieinander liegende Äußerungen gruppiert wurden, dafür gesorgt, dass nicht die innerhalb eines Teams ggf. geteilte Sprache der einzige Gruppierungsgrund ist.⁸⁹

Zur Referenzierung der Ideen-Oberkategorien wurden während der inhaltlichen Zusammenführung Schlagworte vergeben, die als Namen der einzelnen Ideen-Oberkategorien dienen. Sie bilden die beim Zusammenführen identifizierte inhaltliche Gemeinsamkeit ab. Die Ideen-Kategorien +Persönlich, +Meinung, +IndividuelleErgebnisse und +Vorlieben zu Kennzeichen von naturwissenschaftlichen Fragen wurden beispielsweise unter dem Schlagwort Persönliches zusammengeführt. Zu jeder Ideen-Oberkategorie wurde eine Beschreibung formuliert, die Aufschluss über die Inhalte der gruppierten Ideen-Kategorien gibt. Für Persönliches lautet die Beschreibung beispielsweise: »Eine nicht-naturwissenschaftliche Frage fragt nach persönlichen Ansichten oder Meinungen.« Zu den erstellten Beschreibungen sind zwei Anmerkungen zu machen:

- *Erstens sind die Beschreibungen nicht als umfassende Definitionen zu lesen.* Lernende, bei denen eine (oder mehrere) Äußerung(en) mit einer Ideen-Kategorie kodiert sind, die der Ideen-Oberkategorie Persönliches zugehörig ist,

⁸⁹Dass Personen innerhalb eines Teams dieselbe Sprache nutzen, aber ggf. eine andere Vorstellung dabei zu kodieren ist, ist bereits auf Ebene der Ideen-Kategorien berücksichtigt.

haben beispielsweise nicht notwendigerweise nicht-naturwissenschaftliche Fragen dadurch definiert, dass diese (immer) nach persönlichen Ansichten oder Meinungen fragen. Treffendere, aber kompliziertere und daher nicht gewählte, Formulierungen wären beispielsweise: »Ein Merkmal einer naturw. Frage ist, dass sie ...« oder »Geht es um Meinungen, so handelt es sich um eine nicht-naturw. Frage«.

- *Zweitens geben die genauen Formulierungen der Ideen-Oberkategorien auch Aufschluss über die typischerweise den zugehörigen Äußerungen zugewiesenen Kategorien der Formulierung.* Die Beschreibungen wurde basierend auf Analysen zum Zusammenauftreten von Ideen-Oberkategorien und Formulierungskategorien formuliert. Das Zurückgreifen auf die Formulierungskategorien der konkreten Äußerungen, die letztlich hinter den Ideen-Oberkategorien stehen, welches für die Formulierung der Beschreibung der Ideen-Oberkategorien getätigt wird, führt unter anderem dazu, dass die Beschreibungen der Ideen-Oberkategorien nicht nur auf den Ideen-Kategorien basieren, sondern stärker an den tatsächlichen Äußerungen der Lernenden sind. Dazu wurde für jede Ideen-Oberkategorie gezählt, wie häufig jede der verschiedenen Formulierungskategorien (allgemein: p+, p−, m+, m−, siehe Tabelle 6.1 auf S. 251; spezifischer für Kennzeichen von naturwissenschaftlichen Fragen in Tabelle 6.2 auf S. 253) mit einer ihr zugehörigen Äußerung einhergeht. Basierend darauf wurde dann auch die Formulierung der Beschreibung der Ideen-Oberkategorie gestaltet.

Beispielsweise wurde für *Persönliches* zu Kennzeichen von naturwissenschaftlichen Fragen eine zugehörige Äußerung mit der Formulierungskategorie n+ versehen, 8 Äußerungen wurden mit n−, 41 Äußerungen mit u+ und 6 Äußerungen mit u− (siehe Abschnitt F.2). Der Schwerpunkt liegt deutlich auf der Gegendiagonalen von Tabelle 6.2 (u+, n−) und dabei zusätzlich auf u+. Da diese Kategorie der Formulierung dafür steht, dass es miteinander einhergeht, wenn eine Frage naturwissenschaftlich ist und wenn ein genannter Ideen-Aspekt nicht zutrifft, wurde die Beschreibung der Ideen-Oberkategorie *Persönliches* so formuliert, dass es um eine nicht-naturwissenschaftliche Frage (und nicht um eine naturwissenschaftliche Frage) geht und dass der Ideen-Aspekt des Persönlich(keits)-Bezugs direkt (und nicht negiert) aufgegriffen wird: »Eine nicht-naturwissenschaftliche Frage fragt nach persönlichen Ansichten oder Meinungen.« Wäre beispielsweise die häufigste Formulierung n− gewesen, so hätte die Beschreibung »Eine naturwissenschaftliche Frage fragt nicht nach persönlichen Ansichten oder Meinungen« gelautet.

Sind Einträge, die auf verschiedenen Diagonalen der Tabelle liegen (Haupt- und Gegendiagonale) ähnlich stark ausgeprägt (oder ist der häufigste Formulierungstyp der unklare), so wurde versucht, die Beschreibung ausgeglichen zu formulieren, etwa: »Ob eine Frage eine naturwissenschaftliche Fragestellung ist, hängt davon ab, ob sie ...« In den Daten der Arbeit ergeben sich fast nur eindeutige Entscheidungen für die gewählten Beschreibungen (siehe auch Ergebnisse zu F-Vor1 ab S. 283).

Die meisten Analysen in der Arbeit beziehen sich auf Ideen-Oberkategorien und nicht auf einzelne Ideen-Kategorien. Diese Entscheidung wurde unter anderem getroffen, um die Nachvollziehbarkeit der Arbeit nicht durch eine sehr große Anzahl verschiedener Ideen-Kategorien zu schmälern, die sonst beim Lesen im Gedächtnis gehalten werden müssten (selbst die Anzahl der Ideen-Oberkategorien ist nicht klein, siehe Ergebnisse zu F-Vor1). Auch die gute Passung von Ideen-Oberkategorien und Formulierungen, die im vorigen Absatz geschildert ist, spricht dafür, dass eine Betrachtung von Ideen-Oberkategorien ein angemessener Ansatz zum Umgang mit Vorstellungen von Lernenden ist. Die Ergebnisse zu F-Vor1 können als weitere Argumente für die Betrachtung von Ideen-Oberkategorien angesehen werden. Gleichwohl werden die einzelnen Ideen-Kategorien in der Arbeit dort differenziert, wo Ergebnisse erst auf dieser Ebene vorgefunden werden oder besonders gut auf dieser Ebene illustriert werden können.

Im weiteren Verlauf der vorliegenden Arbeit wird verkürzt davon gesprochen, dass die Ideen-Oberkategorien den Äußerungen zugewiesen sind. Genau genommen hat eine solche Zuordnung niemals direkt stattgefunden, sondern die Ideen-Oberkategorien sind gewissen Ideen-Kategorien zugewiesen (oder andersherum formuliert: Ideen-Kategorien sind Ideen-Oberkategorien zugeordnet bzw. in Ideen-Oberkategorien zusammengefasst) und Ideen-Kategorien sind Äußerungen zugewiesen.

6.1.8 Einschätzen der Angemessenheit der Ideen-Oberkategorien

In den Ideen-Oberkategorien sind Ideen-Kategorien und damit die Äußerungen der Lernenden so gebündelt, dass sich die Ideen-Oberkategorien als Vorstellungen interpretieren lassen, die den Ideen der Lernenden zugrunde liegen. Damit lässt sich auch einschätzen, ob die so rekonstruierten Vorstellungen fachlich angemessen sind. Fachliche Angemessenheit bezieht sich dabei darauf, ob einzuschätzende Vorstellungen in der Literatur dokumentierten *fachmethodischen* Konzepten (insb. den in der Instruktion adressierten, vgl. Abschnitt 4.1 ab S. 52) entsprechen oder in sinnvoller Weise kompatibel mit diesen sind.

6.1 Qualitative Inhaltsanalyse zur Rekonstruktion von Vorstellungen der Lernenden

Die Einschätzung der Angemessenheit der Ideen-Oberkategorien erfolgte, indem für jede einzelne Ideen-Oberkategorie vorrangig anhand der zugehörigen Beschreibung die Ähnlichkeit zu und Kompatibilität mit dokumentierten fachlich angemessenen (bzw. unangemessene) Konzepten (siehe Unterabschnitt 2.4.3 ab S. 38) eingeschätzt wird. Als Beschreibung einer Ideen-Oberkategorie wird in der vorliegenden Arbeit der Satz bezeichnet, der die Inhalte der zugehörigen Äußerungen basierend auf den Ideen-Kategorien-Beschreibungen (die inhaltliche Zusammenfassungen der einzelnen Äußerungen sind, siehe Punkt 6.1.1, insb. ab S. 248) und den jeweils zugewiesenen Kategorien der Formulierung (siehe Unterabschnitt 6.1.2, ab S. 250) zusammenfasst. Die Beschreibung der Ideen-Oberkategorie **Persönliches** (zu Kennzeichen von naturwissenschaftlichen Fragen) lautet beispielsweise: »Eine nicht-naturwissenschaftliche Frage fragt nach persönlichen Ansichten oder Meinungen« (siehe Tabelle 6.10). Zur Einschätzung der Angemessenheit werden zusätzlich zur Beschreibung der Ideen-Oberkategorie auch die Beschreibungen der jeweils zusammengefassten Ideen-Kategorien auf Ähnlichkeit zu und Kompatibilität mit dokumentierten fachlichen Konzepten verglichen. Das Urteil wird dann aber – bei inhaltlicher statt quantitativ-anteiliger Gewichtung der Ideen-Kategorien und vorrangigem Fokus auf die Beschreibung der Ideen-Oberkategorie – auf Ebene der Ideen-Oberkategorien und mithilfe einer vierstufigen (Rating-)Skala formuliert: (–2) deutlich unangemessen, (–1) eher unangemessen, (1) eher angemessen, (2) deutlich angemessen. Eine Diskussion der jeweiligen Urteile erfolgt im Rahmen der Ergebnisdarstellung; die Reduktion auf Werte der vierstufigen Skala dient der Weiterarbeit mit handhabbaren (stärker als nominal skalierten) Daten.

Mit der Einschätzung der Ideen-Oberkategorien geht auch einher, dass Angemessenheit nur in fachmethodischer Hinsicht beurteilt wird. So können fachinhaltlich angemessene Äußerungen beispielsweise gleichzeitig mit fachmethodisch unangemessenen Ideen-Oberkategorien kodiert sein. Das folgende Beispiel dient der Illustration.

A2: »Ei, wenn du da auf dem Berg bist (lehnt sich zurück) (.) oder sowas und der Luftdruck niedriger wird, (gestikuliert) dann kocht es viel schneller. Dann kriegt man da keinen heißen Kaffee.«

A1 ergänzt, den Diskurs des Teams zusammenfassend: »Also ist das auch eine naturwissensch/«.

(Team A, Abs. 18+20 u. 21, Details im Anhang, siehe Hägele, 2022, S. 4)

Die Aussage, dass der Luftdruck einen Einfluss darauf hat, wie schnell (genauer: bei welcher Temperatur) Wasser kocht, ist grundsätzlich eher angemessen – aus

fachinhaltlicher Perspektive betrachtet. Sie wird aber von A2 in diesem Kontext als Argument dafür verwendet, dass die Frage als naturwissenschaftliche Frage einzuschätzen ist.⁹⁰ Das stellt zwar aus fachmethodischer Sicht eine korrekte Entscheidung dar (siehe im Ergebnisteil zu F-Vor4, Analysen dazu ab S. 312), geht aber mit einer aus fachmethodischer Sicht unangemessenen Vorstellung (kodiert mit der Ideen-Oberkategorie **Effekt**, die dokumentiert ist in Tabelle 6.10 auf S. 284) einher.

⁹⁰Kontext und Intonation legen dies sehr deutlich nahe. Explizit verbalisiert wird das Argument erst durch das Also von A1.

6.2 Zusammenstellung der analysierten Instruktionsextrakte

Die Analyse der Vorstellungen der Lernenden zum experimentbezogenen Denken und Arbeiten während den Prozessen der Bearbeitung des eingesetzten Lernmaterials wird in der Arbeit auf zwei fachmethodische Inhaltsbereiche und dazu zusammengestellte Instruktionsextrakte beschränkt. Damit wird das Ziel verfolgt, zeitökonomisch zu arbeiten sowie eine angemessene Detailliertheit und Rigorosität der Analysen zu ermöglichen und somit auch als exemplarische Grundlage für weitere Analysen zu anderen fachmethodischen Inhaltsbereichen dienen zu können. Aus der Zusammenstellung von Instruktionsextrakten ergeben sich die analysierten Videosequenzen, indem die Ausschnitte aus den Videos gewählt werden, in denen die Bearbeitung der Instruktionsextrakte erfolgt.

Die Wahl der Inhaltsbereiche ist zum einen durch die Möglichkeit der Anknüpfung an die Forschungslandschaft (siehe Unterabschnitt 2.4.3) begründet, zum anderen hängt sie auch mit der Beschaffenheit des Lernmaterials und der damit z. T. eingeschränkten Möglichkeit zur Zusammenstellung von Instruktionsextrakten zusammen. Die zusammengestellten Instruktionsextrakte sollten im wesentlichen das Kriterium erfüllen, dass sie auf ein fachmethodisches Konzept (oder eine klar und eng umrissene Gruppe von Konzepten) fokussieren, im Aufbau zusammenhängend sind und klar sowohl Explorations- als auch Übungsphasen enthalten, in die instruktionale Informationen erkennbar eingebettet sind. Die instruktionale Information sollte dabei nicht nur die Nennung des fachmethodischen Konzeptes (oder der klar und eng umrissene Gruppe von Konzepten), sondern auch eine Erläuterung und ggf. Beispiele umfassen.

Aus dem Material der explizit-fachmethodischen Instruktionsvariante wurden zwei Instruktionsextrakte zusammengestellt, die die genannten Kriterien erfüllen. Die Sequenzen beziehen sich auf Kennzeichen von naturwissenschaftlichen Fragen und die Unterscheidung von Beobachtung und Deutung und sind nachfolgend ausführlich dargestellt. Weitere Instruktionsextrakte, die die genannten Kriterien erfüllen, könnten zusammengestellt werden (z. B. aus den Karten 1-4 bis 1-10 zum Formulieren von präzisen und allgemeinen Fragen; Konzepte dazu finden sich in Tabelle 4.2 auf S. 54). Sie wurden allerdings nicht analysiert.

6.2.1 Darstellung des Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen

Das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen adressiert das Konzept, dass naturwissenschaftliche Fragen dadurch gekennzeichnet sind, dass sie sich mithilfe naturwissenschaftlicher Methoden untersuchen lassen. Es besteht

6 Vorstellungen von Lernenden

aus einem (Instruktions-)Segment mit sieben Karten, und zwei (Instruktions-)Segmenten mit jeweils zeitlichem Abstand, die einzelne Aufgaben von je maximal einer Karte umfassen. Die zum Instruktionsextrakt gehörigen Karten sind vollständig im Anhang dargestellt (Anhang C, ab S. 541). In Abbildung 6.6 findet sich eine Übersicht über die Segmente des Instruktionsextrakts und über die jeweils zugehörigen Karten.

Segment	Seg. 1			Seg. 2	Seg. 3
Karten	1-21	...	1-27	1-31	2-12

Abbildung 6.6: Übersicht über die Segmente der Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen.

In Abbildung 6.7 findet sich für das erste Segment eine Darstellung der zu den Karten jeweils gehörenden (Teil-)Aufgaben. Auszüge aus diesem Segment sind bereits in Abbildung 4.2 (auf S. 56) im Kapitel zur Anlage der Instruktion enthalten. Als Teilaufgaben werden alle einzeln bearbeitbaren Elemente auf den Karten bezeichnet wie etwa Beispielsätze, die eingeschätzt werden müssen oder Fragen, die beantwortet werden sollen, auch dann, wenn ein gemeinsamer Arbeitsauftrag für diese Elemente vorhanden ist. Die inkonsistente Benennung der Teilaufgaben ist teilweise der Passung zur Benennung der (Teil-)Aufgaben im Lernmaterial geschuldet.

Die ersten zwei Karten des ersten Segments sind bereits in weiter oben gegebenen Beispielen größtenteils vorgestellt (Abbildung 6.4 auf S. 258 im Unterabschnitt zum Kodieren der Aufgaben und Abbildung 6.5 auf S. 259 im Unterabschnitt zum Kodieren der Korrektheit der Einschätzungen). Sie sind der Explorationsphase zuzuordnen, weil die Lernenden ohne vorherige Instruktion bei der Bearbeitung zunächst Beispielfragen einschätzen sollen (1-21-i bis 1-21-vii) und dann die zugehörigen Lösungen ohne Begründung erhalten (1-22-i-i bis 1-22-i-vii). Ferner ist der Explorationsphase auch die

Phase	Explorationsphase				Explizierung	Übungsphase			
Karte	1-21		1-22		1-23	1-24	1-25	1-26	1-27
Teil-aufgabe	1-21-i ... 1-21-vii		1-22-i-i ... 1-22-i-vii 1-22-ii			1-24-i 1-24-ii	1-25-A 1-25-B	1-26-C 1-26-D	

Abbildung 6.7: Übersicht über die Karten und Teilaufgaben von Segment 1 aus dem Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen.

folgende Aufgabe zuzurechnen, die sich unter den Lösungen auf der Karte 1-22 findet: »Stellen Sie sich vor, Sie sollten jemandem erklären, was eine naturwissenschaftliche Fragestellung von einer nicht-naturwissenschaftlichen Fragestellung unterscheidet. Was würden Sie sagen?« (1-22-ii). Auf der Karte 1-23 findet sich ein Informationstext dazu, was naturwissenschaftliche Fragen kennzeichnet, der durch zwei Beispielsätze komplementiert ist, für die anhand der Kriterien begründet entschieden ist, ob sie naturwissenschaftliche Fragestellungen sind oder nicht. Nach der Explizierung der Konzepte zu Kennzeichen von naturwissenschaftlichen Fragen finden sich Karten, die im Sinne einer Übungsphase Aufgaben zum Einschätzen und Begründen enthalten. Auf Karte 1-24 sollen die Schülerinnen und Schüler für eine bereits als naturwissenschaftlich klassifizierte (1-24-i) und eine bereits als nicht-naturwissenschaftlich klassifizierte Fragestellung (1-24-ii) begründen, warum die jeweiligen vorgegebenen Einschätzungen angebracht sind. Auf den Karten 1-25 und 1-26 finden sich vier Beispielfragen (1-25-A bis 1-26-D), für die die Lernenden einschätzen sollen, ob sie naturwissenschaftlich sind, sowie Begründungen für die Einschätzungen formulieren sollen. Das Segment wird durch eine abschließende Lösung für die vier Beispielfragen sowie eine Aufgabe abgerundet, bei der die Lernenden spezifische Methoden zur Untersuchung fehlerhaft eingeschätzter Beispielfragen nennen sollen (1-27).

Im zweiten Segment (Karte 1-30), welches durch einen kleinen Exkurs zum gebrauchten Fragewort (Karte 1-28) und eine Zusammenfassung des Abschnitts (Karte 1-29) sowie ein Zwischendeckblatt (1-30) von dem vorigen Segment getrennt ist, sollen die Lernenden für eine weitere Beispielfrage einschätzen, ob diese präzise oder allgemein (1-31-i) und ob diese naturwissenschaftlich (1-31-ii) ist. Auch wenn die Einschätzung, ob die Frage präzise oder allgemein ist, nicht von direkter Relevanz für Analysen zu Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen scheint, werden die Videoaufzeichnungen zur gesamten Karte 1-31 analysiert (die auch eine weitere Teilaufgabe 1-31-iii zum Formulieren von Vermutungen enthält).

Einschub. Die Trennung des ersten und des zweiten Segments kann insofern kritisch hinterfragt werden, als dass Karte 1-29 als eine Zusammenfassung der Konzepte zum Formulieren naturwissenschaftlicher Fragen auch die Kennzeichen von naturwissenschaftlichen Fragen aufgreift und Karte 1-30 ein (in den Videoaufzeichnungen in der Tat) schnell umgeblättertes Zwischendeckblatt ist. Gleichzeitig wird durch Karte 1-28 der Fokus auf etwas anderes als das in Segment 1 adressierte Konzept zu Kennzeichen von naturwissenschaftlichen Fragen gelegt und auch auf Karte 1-31 wird mit der Teilaufgabe 1-31-i zunächst etwas anderes thematisiert als dieses Konzept. Aus den Videoaufzeichnungen zu 1-31-i geht hervor, dass die Schülerinnen

und Schüler teilweise nicht auf präzise und allgemeine Fragen eingehen, sondern trotz der Frage direkt auf die kurz vorher (in Segment 1) behandelte Unterscheidung von naturwissenschaftlichen und nicht-naturwissenschaftlichen Fragestellungen eingehen. Diese Beobachtung stützt die Kritik, dass Segmente 1 und 2 eventuell als ein zusammenhängendes Segment angesehen werden sollten. Trotzdem wird in der Arbeit eine Trennung vorgenommen, um den zwischenzeitlichen Wechsel des inhaltlichen Fokus abzubilden. Gleichzeitig wurde der Kritik insofern begegnet, als dass die Videoaufzeichnungen der Karten 1-28 bis 1-30 für die analysierten Gruppen (siehe Abschnitt 6.3, ab S. 278) nach der qualitativen Inhaltsanalyse von Segment 1 durchgesehen wurden. Es ergab sich, dass hier von den Lernenden keine substantiellen Bezüge zu Kennzeichen von naturwissenschaftlichen Fragen hergestellt werden.

Ein drittes Segment findet sich in der nächsten Einheit und ist damit zeitlich deutlich von den anderen Segmenten getrennt (etwa eine Woche; auch bezüglich der reinen Bearbeitungszeit folgen die Segmente 2 und 3 nicht direkt aufeinander). Es umfasst die Karte 2-12, auf der – neben der Beantwortung einer Frage zu Kontrollvariablen (2-12-i) – für eine gegebene Beispielfrage begründet eingeschätzt werden soll, ob sie naturwissenschaftlich ist (2-12-ii). Diese Beispielfrage entstammt einer Versuchsplanung, die die Lernenden an dieser Stelle durchführen, und die begründete Einschätzung kann als kleiner Exkurs an dieser Stelle eingeschätzt werden.

6.2.2 Darstellung des Instruktionsextrakts zur Unterscheidung von Beobachtung und Deutung

Das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung adressiert das Konzept, dass Beobachtungen intersubjektiv überprüfbare Aussagen sind, wohingegen Deutungen dies nicht sind, sondern über Beobachtungen hinausgehen. Sie besteht aus einem einzigen Segment, welches sechs Karten umfasst (Übersicht in Abbildung 6.8; die vollständigen Karten sind in Anhang D, ab S. 545 auffindbar).



Abbildung 6.8: Übersicht über das erste Segment des Instruktionsextrakts zur Unterscheidung von Beobachtung und Deutung (analog zu Abbildung 6.6).

In Abbildung 6.9 findet sich eine Darstellung der zu den Karten jeweils gehörenden (Teil-)Aufgaben. Als Teilaufgaben werden – wie bereits beim Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen – alle einzeln bearbeitbaren

6.2 Zusammenstellung der analysierten Instruktionsextrakte

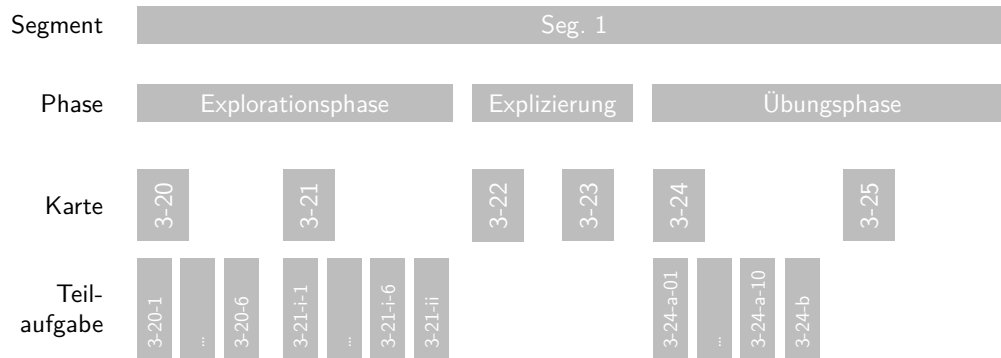


Abbildung 6.9: Übersicht über die Karten und Teilaufgaben des Instruktionsextrakts zur Unterscheidung von Beobachtung und Deutung.

Elemente auf den Karten bezeichnet wie etwa Beispielsätze, die eingeschätzt werden müssen oder Fragen, die beantwortet werden sollen, auch dann, wenn ein gemeinsamer Arbeitsauftrag für diese Elemente vorhanden ist.

Auf der ersten Karte des Segments werden die Lernenden dazu aufgefordert, sechs Sätze (3-20-1 bis 3-20-6) eines fiktiven Gesprächs über einen Versuch zu lesen, bei dem eine Holzkugel und eine Feder fallen gelassen werden. Die Lernenden sollen alle Beobachtungen identifizieren. Auf den nachfolgenden beiden Karten finden sich Informationstexte dazu, was Beobachtungen (3-21) und was Deutungen (3-22) ausmacht und unterscheidet. Beispiele für Beobachtungen und Deutungen sind jeweils angegeben und Begründungen für die Klassifikation sind aufgeführt. Nach der Explizierung sollen die Schülerinnen und Schüler in dem Transkript eines fiktiven Streitgesprächs über einen Falschparker alle Deutungen identifizieren (3-24-a). Die Sätze des Transkriptes sind dabei als Teilaufgaben 3-24-a-01 bis 3-24-a-10 nummeriert. Auf der Karte findet sich zudem der Arbeitsauftrag für die Lernenden, zu begründen, warum es in dem Gespräch sinnvoll wäre, zwischen Beobachtung und Deutung zu unterscheiden (3-24-b). Das Segment wird durch die Karte 3-25 abgerundet, auf der sich ein Informationstext findet, in dem erklärt wird, dass Naturwissenschaftlerinnen und Naturwissenschaftler sich selten über die Beobachtungen uneinig sind, sondern vielmehr über die Deutungen Uneinigkeit unter ihnen besteht. Dieser Informationstext dient zugleich als Überleitung zu den anschließenden Karten, die einen neuen Inhaltsbereich behandeln.

Im Vergleich zum Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen findet sich vor der Explizierung für die Unterscheidung von Beobachtung und Deutung nur ein einziger Arbeitsauftrag zum Einschätzen. Die Lernenden

werden nicht aufgefordert, selbstständig Kriterien von Beobachtungen oder Deutungen zu formulieren, wohingegen sie bei 1-22-ii aufgefordert werden, Kennzeichen von naturwissenschaftlichen Fragen zu formulieren. Die Explizierung zur Unterscheidung von Beobachtung und Deutung findet sich nicht auf einer, sondern auf zwei Karten und kann als umfangreicher eingeschätzt werden als die Explizierung zu Kennzeichen von naturwissenschaftlichen Fragen. Nach der Explizierung findet sich eine einzige Übungsaufgabe mit zehn durch den Gesprächsfluss zusammengehörigen Gelegenheiten zum Unterscheiden von Beobachtungen und Deutungen, was sich von den sechs klar voneinander getrennten Beispielfragen im Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen unterscheidet. Trotz der genannten Unterschiede ist die Grundstruktur vom jeweils ersten Segment der Instruktionsextrakte gut vergleichbar und beinhaltet die als für die Forschungsfragen zentral anzusehenden Phasen der Exploration, der Explizierung und der Übung.

6.2.3 Einschränkungen und einordnende Hinweise zu den erstellten Instruktionsextrakten

Die für die Auswahl der Inhaltsbereiche und zugehörigen Instruktionsextrakte angelegten Kriterien stellen notwendigerweise eine Einschränkung der Generalisierbarkeit der Ergebnisse dar. Sie sind auch in weiterer Hinsicht als Limitation aufzufassen, weil beispielsweise die Forschungsfrage zu Vorstellungen im zeitlichen Umfeld der Explizierung (F-Vor6, formuliert auf S. 50, Analysen dazu ab S. 353) aufgrund der Zusammenstellung der Instruktionsextrakte vorrangig auf ein sehr direktes zeitliches Umfeld der Explizierung bezogen beantwortet wird. Dieser Einschränkung wurde begegnet, indem für einen Inhaltsbereich (Kennzeichen von naturwissenschaftlichen Fragen) Aufgaben mit zeitlichem Abstand hinzugenommen wurden (für den anderen Inhaltsbereich lagen derartige Aufgaben nicht vor). Ferner kann eine möglichst isolierte Betrachtung von einzelnen Konzepten Einschränkungen bezüglich der ökologischen Validität mit sich bringen, weil in Realsituationen des experimentbezogenen Denkens und Arbeitens für gewöhnlich die Vernetzung mehrere Konzepte von Bedeutung zu sein scheint (Schreiber, 2012). Da allerdings die Vorstellungen der Schülerinnen und Schüler in den Prozessen des Kompetenzaufbaus zu eben den einzelnen Konzepten untersucht werden sollen, kann diese Einschränkung als ein Hinweis auf Desiderata bezüglich der Vernetzung von Konzepten beim Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten angesehen werden, welche nicht im Rahmen der Arbeit angegangen werden.

Eine wesentliche Einschränkung, die sich aus der Wahl zweier Inhaltsbereiche ergibt, ist, dass nicht alle drei Teilprozesse des experimentbezogenen Denkens und Arbeitens in den Blick genommen werden. Zwar ist die Dreiteilung, wie in Unterabschnitt 2.2.1 (ab S. 15) diskutiert, in gewisser Hinsicht als artifiziell anzusehen, weshalb nicht starr versucht werden sollte, sie in Analysen zu Vorstellungen der Lernenden zum experimentbezogenen Denken und Arbeiten abzubilden. Allerdings stellt die Dreiteilung eine sinnvolle Heuristik dar, um zu überprüfen, ob das experimentbezogene Denken und Arbeiten in seiner Vielfältigkeit abgebildet ist. Die erstellten Instruktionsextrakte beziehen sich auf die Inhaltsbereiche *Kennzeichen von naturwissenschaftlichen Fragen* und *Unterscheidung von Beobachtung und Deutung*, die sich dem Formulieren von Fragen und Hypothesen sowie dem Auswerten und Interpretieren zuordnen lassen. Es wird somit in den Analysen kein Instruktionsextrakt zum Planen von Untersuchungen betrachtet.

Dass der Teilprozess des Planens von Untersuchungen für die Analysen zu Vorstellungen der Lernenden nicht untersucht wird, hängt damit zusammen, dass sich im Material kein in sich geschlossenes Segment zu einem zugehörigen Inhaltsbereich findet, welches als Basisbaustein eines Instruktionsextrakts dienen könnte. Dabei ist es auch von Bedeutung, dass sich im Material nach Einschätzung des Autors der vorliegenden Arbeit keine dem angelegten Kriterium entsprechende hinreichend isolierte Adressierung einzelner Konzepte oder eng umrissener Gruppen von Konzepten findet. Das heißt, dass sich keine aufeinanderfolgenden Aufgaben zu einem spezifischen Inhaltsbereich zum Planen von Untersuchungen finden, die einen eng umrissenen konzeptuellen Fokus aufweisen. Vielmehr finden sich zu den verschiedenen Konzepten zum Planen von Untersuchungen (Passung der Untersuchung zur Frage/Hypothese, Variation der unabhängigen Variable, Betrachtung der abhängigen Variable, Konstanthaltung der Kontrollvariablen, Abdeckung des verfügbaren Messbereichs; siehe Tabelle 4.2 auf S. 54) über das gesamte Material verteilte Aufgaben, die verschiedene Fähigkeiten betreffen (*Identifizieren*, z. B. von richtig geplanten Versuchen; *Planen*, z. B. von variablenkontrollierten Versuchen; ...) und anhand verschiedener Versuche bzw. Versuchsreihen (z. B. zur Kochzeit eines Eies, zum Fallenlassen von Kegeln mit verschiedenen Durchmessern) strukturiert sind. Diese Ausführungen stellen keine vollständige Materialanalyse dar, sondern dienen dazu, die Schwierigkeit der Auswahl eines zeitlich begrenzten Segments zu einem inhaltlich klar umrissenen Bereich zu begründen.

Für die Suche nach brauchbaren Instruktionsextrakten stellen die im Material vorhandenen Explizierungen einen Ankerpunkt dar. Als Illustration der Schwierigkeit der Zusammenstellung eines Instruktionsextrakts zum Planen von Untersuchungen,

welches ein zusammenhängendes Segment enthält, kann die Explizierung zu den Benennungen der Variablen und ihren Funktionen beim Experimentieren dienen, die in Abbildung 6.10 dargestellt ist. Karte 2-10 umfasst Informationstexte dazu, was eine unabhängige Variable ist, was abhängige Variablen und Kontrollvariablen sind, aber auch dazu, welche Art von Untersuchungen als fair bezeichnet werden sollen und wie die zu einem Versuch passende Frage formuliert wird. Sie findet sich relativ zu Beginn der zweiten Einheit des Lernmaterials und spätere Aufgaben nehmen Bezug auf die Definitionen von dieser Karte. Teilweise wird nur auf einzelne Aspekte Bezug genommen, allerdings dürfte beispielsweise die Aufgabe, zu prüfen, ob Versuche fair sind, unweigerlich damit verbunden sein, das auf der Karte 2-10 dargestellte Netzwerk von Konzepten zumindest intuitiv nutzen zu können. Die Illustration kann auch als Hinweis darauf interpretiert werden, dass die Adressierung einzelner Konzepte für das Planen von Untersuchungen (anders als beispielsweise für das Konzept zu Kennzeichen von naturwissenschaftlichen Fragen im Bereich des Formulierens von Fragen und Hypothesen) instruktional nicht in gleicher Weise umgesetzt werden kann wie zu den anderen Teilprozessen des experimentbezogenen Denkens und Arbeitens, gegebenenfalls weil die fachmethodischen Inhalte anders beschaffen sind. Im Rahmen der Arbeit wird diesem didaktisch-rekonstruktiven und instruktionalen Aspekt nicht weiter nachgegangen. Allerdings sind die Ausführungen zu instruktionalen Aspekten insofern relevant, als dass daraus folgen dürfte, dass die ausgewählten Inhaltsbereiche eventuell das experimentbezogene Denken und Arbeiten in seiner Vielfalt nicht

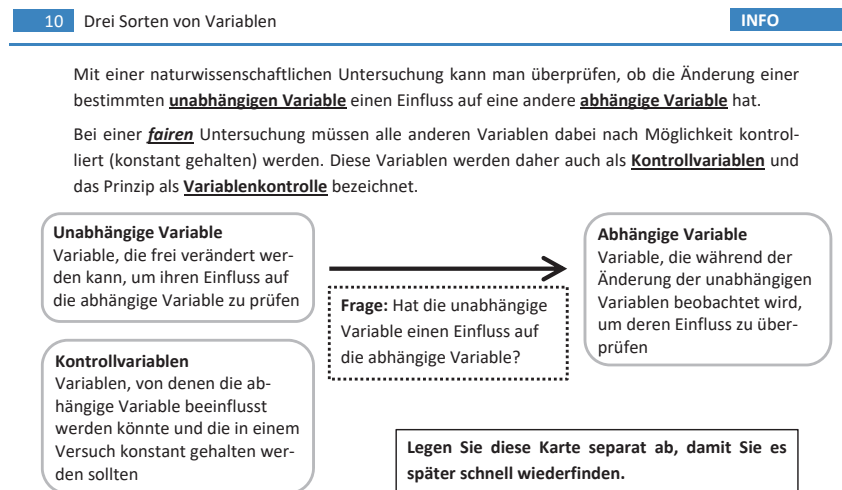


Abbildung 6.10: Karte 2-10 zur Illustration der Vernetzung von Konzepten zum Planen von Untersuchungen.

hinreichend repräsentieren. Nichtsdestotrotz kann die Schwerpunktsetzung der Arbeit als Vorteil angesehen werden, weil in vielen anderen Studien vor allem das Planen von Untersuchungen, insbesondere die Variablenkontrolle, untersucht wird.

Der Einschränkung, dass sich kein in sich geschlossenes Segment zu einem Konzept aus dem Bereich des Planens von Untersuchungen im Material findet, könnte dadurch begegnet werden, dass das Instruktionsextrakt nicht ausgehend von einem zusammenhängenden Segment zusammengestellt wird, sondern nur einzelne Aufgaben betrachtet werden, die zu unterschiedlichen Zeitpunkten bearbeitet werden und nur dadurch verbunden sind, dass sie sich einem ausgewählten Konzept zuordnen lassen. Das methodische Vorgehen der qualitativen Inhaltsanalyse (geschildert in Abschnitt 6.1, ab S. 242) ließe sich bei geringer Modifikation auch auf unzusammenhängende Aufgaben übertragen. Allerdings müssten die Interpretationen der Daten, die durch die qualitative Inhaltsanalyse generiert werden, in anderer Weise erfolgen, weil sie eben nicht einer zusammenhängenden Abfolge von Aufgaben entspringen. Gegebenenfalls müsste auch für nicht-zugehörige Aufgaben eine qualitative Inhaltsanalyse erfolgen, um die Relevanz der dort vorliegenden Bearbeitungsprozesse abschätzen und einordnen zu können. Ferner müsste die Beantwortung der Forschungsfragen, insbesondere zur Stellung der Explizierung in den Instruktionsextrakten, in anderer Weise erfolgen, weil gegebenenfalls keine Aufgabenbearbeitungen direkt vor und direkt nach der Explizierung vorliegen, die mit dem betrachteten Konzept in Verbindung stehen. Angesichts des Umfangs und der Komplexität des Vorgehens für die recht gut vergleichbaren Instruktionsextrakte zu Kennzeichen von naturwissenschaftlichen Fragen und zur Unterscheidung von Beobachtung und Deutung liegt es nahe, Analysen zum Planen von Untersuchungen aufgrund des Nichtvorliegens eines zusammenhängenden Instruktionssegments daher in der Arbeit nicht vorzunehmen und darzustellen.

Zuletzt lässt sich kritisch anmerken, dass die zusammengestellten Instruktionsextrakte sich nur auf einzelne Konzepte beziehen, die wiederum nicht notwendigerweise repräsentativ für den zugehörigen Teilprozess sein müssen. Insbesondere bezüglich des Teilprozesses des Auswertens und Interpretierens ist davon auszugehen, dass die Unterscheidung von Beobachtungen und Deutungen nur einen kleinen Bruchteil der zugehörigen Konzepte repräsentiert (siehe z. B. Tabelle im Theorieteil oder Betrachte Lernmaterial von Vorholzer). Trotzdem betreffen die ausgewählten Instruktionsextrakte zwei Inhaltsbereiche, die für sich genommen relevante Forschungsgegenstände darstellen (siehe Unterabschnitt 2.2.2, ab S. 16).

6.3 Wahl der analysierten Lernenden und zugehörigen Videosequenzen und Transkripte

Ähnlich wie bei den Analysen zu den Aktivitäten sind auch für Vorstellungen nicht alle Lernenden, die sich in der Instruktionsvarianten-Gesamtstichprobe zur explizit-fach-methodischen Instruktionsvariante befinden, für Analysen nutzbar (vgl. Abschnitt 5.2, ab S. 88). Basierend auf den zusammengestellten Instruktionsextrakten wurden die Videosequenzen, auf denen die in der Arbeit berichteten Ergebnisse basieren, dadurch bestimmt, dass mehrere Lernende teamweise ausgewählt wurden. Die Auswahl ist weiter unten detaillierter erläutert. Für die ausgewählten Lernenden ergaben sich die analysierten Videosequenzen dann dadurch, dass die zeitlichen Videoausschnitte gewählt wurden, in denen die Lernenden die dem Instruktionsextrakt zugehörigen Karten bearbeiten. Dafür wurden – soweit sie vorlagen – die Kodierungen der Karten aus Kapitel 5 verwendet und sicherheitshalber im Umfeld dieser Kodierungen eine bis zwei weitere Minuten Vor- sowie Nachlauf hinzugenommen; ansonsten wurden die Videos komplett gesichtet, um die Bearbeitungen der zum Instruktionsextrakt gehörigen Karten zu finden.

Die Videosequenzen wurden für die qualitative Inhaltsanalyse transkribiert, weil der Notwendigkeit der starken Fokussierung auf den genauen Wortlaut der Äußerungen der Lernenden beim Vorliegen von Transkripten insbesondere bei längeren und dicht mit Äußerungen besetzten Videosequenzen besser nachgekommen werden konnte. Segment 1 aus der Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen ist parallel zum Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung, und beide Abschnitte stellen jeweils das Kernstück der Analysen zu den Inhaltsbereichen dar. Segmente 2 und 3 zu Kennzeichen von naturwissenschaftlichen Fragen sind hingegen nur sehr kurz und oft liegen für sie nur wenige Äußerungen von Lernenden vor. Daher wurden Transkriptionen nur für den Videoausschnitt zu Segment 1 aus dem Instruktionsextrakt für Kennzeichen von naturwissenschaftlichen Fragen sowie für das komplette Instruktionsextrakt für die Unterscheidung von Beobachtung und Deutung erstellt. Für diese Ausschnitte erfolgte die qualitative Inhaltsanalyse am Transkript unter ständiger Nutzung der Videoaufzeichnungen (in MAXQDA parallel möglich). Die qualitative Inhaltsanalyse der Segmente 2 und 3 zu Kennzeichen von naturwissenschaftlichen Fragen wurde hingegen direkt am Video vorgenommen (in MAXQDA am Zeitstrahl möglich), nachdem die Analyse von Segment 1 beendet war.

Die Transkriptionen wurden von einem professionellen Transkriptionsbüro nach einheitlichen Transkriptionsregeln erstellt (Dresing & Pehl, 2017, S. 25–32; für einen

Überblick über die Regeln und für alle Transkripte siehe Anhang E, ab S. 549). Das Transkriptionsbüro erhielt die Karten der zugehörigen Sequenz, damit diese im Sinne einer Wortliste verwendet werden konnten. Über die Transkribierer liegen aus Datenschutzgründen keine Informationen vor, außer dass diese Personen nicht in der Physik oder Physikdidaktik ausgebildet wurden oder tätig sind. In einem zweiten Schritt wurden die vom Transkriptionsbüro erstellten Transkripte vom Autor der Arbeit gesichtet. Einige wenige Stellen wurden korrigiert, u. a. weil physikalische Fachbegriffe in den Transkripten nicht korrekt abgebildet waren.

Zwei Transkripte wurden (zeitlich vor den anderen Transkripten) von physikalisch grundlegend geschulten, aber mit den konkreten Inhalten des Materials (Fachmethoden) nicht vertrauten Hilfskräften des Instituts erstellt. Auch in diesen Transkripten ergaben sich einige wenige Stellen, an denen physikalische Fachbegriffe (insbesondere zu Fachmethoden) vom Autor der Arbeit korrigiert werden mussten. Die Hilfskräfte folgten Transkriptionsregeln, die leicht von denen des Transkriptionsbüros abweichen, bei denen allerdings dieselben Elemente der Videoaufzeichnungen abgebildet werden. Eine Vergleichbarkeit der Transkripte ist demnach nicht gefährdet, wie auch der Vergleich von Teilstücken eines Transkriptes zeigte, das zunächst von den Hilfskräften begonnen wurde, dann aber aus organisatorischen Gründen dem Transkriptionsbüro übergeben wurde. Zur Darstellung in der Arbeit wurden Zitate aus den beiden von den Hilfskräften erstellten Transkripten vom Autor der Arbeit an die Transkriptionsregeln des Transkriptionsbüros angeglichen. Überlappungen von Äußerungen mehrerer Personen sind in Zitaten nur dann dargestellt, wenn die Äußerungen beider Personen im Zitat enthalten sind.

Von Interesse für die Analysen zu den Vorstellungen von Lernenden im Bearbeitungsprozess ist bei der Auswahl der Lernenden besonders, dass einerseits ein Spektrum an Personen ausgewählt wird, bei dem naheliegend ist, dass sich Vorstellungen zum experimentbezogenen Denken und Arbeiten aus den Äußerungen im Lernprozess rekonstruieren lassen, und dass andererseits das ausgewählte Spektrum der Personen auch Schlüsse darauf erlaubt, wie (groß) die Varianz der Vorstellungen in den Bearbeitungsprozessen ist oder sein könnte. Daher wurden die Analysen zunächst für Teams begonnen, für die der Mittelwert der Kompetenzzuwächse zum experimentbezogenen Denken und Arbeiten über die drei Personen hoch ist. Dabei wurde versucht, sowohl eher homogene als auch eher heterogene Teams hinsichtlich der Testwerte zum experimentbezogenen Denken und Arbeiten aus dem Prätest und zu fachinhaltlichen Kompetenzen auszuwählen. Später wurden weitere Teams hinzugenommen, die niedrigere mittlere Zuwächse aufweisen, wobei unter anderem darauf

geachtet wurde, auch Teams auszuwählen, für die Analysen zu den Aktivitäten (siehe Kapitel 5) vorlagen oder geplant waren. Teilweise werden durch dieses Verfahren Teams ausgewählt, für die sich Lücken in den Datensätzen finden. Beispielsweise bearbeitet das Team K nur die ersten beiden der zum Instruktionsextrakt für Kennzeichen von naturwissenschaftlichen Fragen gehörigen Karten aus Einheit 1, weil es nicht bis zum Ende der Instruktionseinheit arbeitet; es wurde allerdings bereits für die Unterscheidung von Beobachtung und Deutung analysiert, bevor dies entdeckt wurde. In einem letzten Schritt wurden Teams hinzugenommen, die möglichst wenige Lücken in den eigenen Datensätzen aufweisen oder die in den Gesamtdatensätzen zu Kennzeichen von naturwissenschaftlichen Fragen oder zur Unterscheidung von Beobachtung und Deutung eine sinnvolle Ergänzung darstellen.⁹¹

Eine Übersicht über die analysierten Teams und die für die jeweiligen Teams vorhandenen Videoausschnitte findet sich in Tabelle 6.9 (auf S. 281). Die Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen wurde für 33 Schülerinnen und Schüler (11 Teams) analysiert und umfasste je nach Team bis zu drei Segmente. Die Sequenz zu Unterscheidung von Beobachtung und Deutung wurde für 28 Schülerinnen und Schülern (10 Teams) analysiert und besteht aus einem Segment. Aus Tabelle 6.9 geht zudem hervor, welche Sequenzen (und Segmente) für die Teams jeweils analysiert wurden. Für sieben Teams sind alle Tabellenspalten gefüllt, allerdings liegt aufgrund abwesender Personen nur für zwei Teams ein vollständiger Datensatz vor. In Tabelle 6.9 ist außerdem dargestellt, wie lang die jeweiligen Sequenzen für die analysierten Teams sind. Der Mittelwert über die Teams beträgt für Kennzeichen von naturwissenschaftlichen Fragen etwas mehr als 12 Minuten, liegt aber bei Vernachlässigung des Ausreißers Team I bei 13 Minuten (mit einer deutlich kleineren Standardabweichung von 1 Minute und 16 Sekunden). Der Mittelwert für die Sequenz zur Unterscheidung von Beobachtung und Deutung liegt bei 4 Minuten und 40 Sekunden und ist damit etwa halb so groß wie der Mittelwert von Segment 1 für Kennzeichen von naturwissenschaftlichen Fragen (welches einen angemessenen Vergleichspunkt darstellt, weil beide Segmente strukturell in ähnlicher Weise aufgebaut sind, siehe Abschnitt 6.2).

⁹¹Da die Analysen zu den Vorstellungen und die Analysen zu den Aktivitäten (für letztere siehe Kapitel 5) während der Laufzeit des der Arbeit zugrundeliegenden Projekts parallel prozessiert wurden, ergeben sich zusätzlich Deckungsungleichheiten bezüglich der jeweils ausgewählten Lernenden.

Tabelle 6.9: Zeitumfang der jeweils analysierten Sequenzen für jedes Team (in Minuten und Sekunden).

Team			KnF				BuD
S1	S2	S3	Seg. 1	Seg. 2	Seg. 3	Summe	
A3	A2	A1 ^a	09:33	02:59	01:00	13:32	05:39
B6 ^b	B5	B4	11:22	02:18	00:45	14:25	04:25
C8 ^c	C9	C7	13:57	–	02:01	13:57	03:50
F16	F18	F17 ^b	09:40	02:07	00:35	12:22	04:51
G21	G19	G20 ^b	10:58	02:07	00:26	13:31	02:45
H23	H22	H24	10:58	02:41	00:45	14:24	07:40
I25	I26	I27	04:58	01:28	00:06	06:32	
K33	K31	K32					04:55
L34	L35	L36					02:51
M37	M38	M39	07:27	02:41	02:40	12:48	04:58
N42	N40 ^c	N41	08:56	01:06	00:37	10:39	
O43	O44	O45	09:59	02:18	00:07	12:24	
P48	P47 ^b	P46	08:12	02:44	–	10:56	04:48
Summe			1:46:00	22:29	07:01	2:15:30	46:42
Mittelwert			09:38	02:15	00:47	12:19	04:40
Standardabweichung			02:13	00:34	00:43	02:11	01:20

Anmerkungen. Werte im Tabellenkörper sind in Minuten und Sekunden (mm:ss) bzw. in Stunden, Minuten und Sekunden (h:mm:ss) angegeben. S1/2/3 = Schülerin bzw. Schüler. KnF = Kennzeichen von naturwissenschaftlichen Fragen. BuD = Unterscheidung von Beobachtung und Deutung. Seg. 1/2/3 = Videoausschnitte zu den Segmenten des Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen (siehe Unterabschnitt 6.2.1). Werte sind nur bei analysierten Sequenzen und Segmenten angegeben; – bedeutet, dass keine Analyse möglich war. Die Dauern für die Segmente 2 und 3 von KnF entsprechend den Dauern der Karten 1-31 und 2-12; die reale Zeit der Beschäftigung mit KnF ist in der Regel kürzer. Videoausschnitte, bei denen keine Äußerungen kodiert wurden, sind grau gesetzt.

^a Bei BuD ist eine weitere Schülerin (E14) anwesend.

^b Diese Person ist nicht anwesend bei BuD.

^c Diese Person ist nicht anwesend bei Segment 3 von KnF.

6.4 Auswertungen und Ergebnisse zur inhaltlichen Vielfalt der Vorstellungen

Die Analysen der Verständnisse beziehen sich auf eine Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen und eine Sequenz zur Unterscheidung von Beobachtung und Deutung (siehe Unterabschnitte 6.2.1, ab S. 269, und 6.2.2, ab S. 272). Die Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen wurde für 33 Schülerinnen und Schüler (11 Teams) analysiert und umfasste je nach Team bis zu drei Segmente. Die Sequenz zur Unterscheidung von Beobachtung und Deutung wurde für 28 Schülerinnen und Schülern (10 Teams) analysiert und besteht aus einem Segment. Eine Übersicht über die analysierten Lernenden findet sich in Tabelle 6.9 auf Seite 281. Nachfolgend werden die Analysen und Ergebnisse zu den einzelnen Forschungsfragen dargestellt. Der vorliegende Abschnitt ist auf die erste Forschungsfrage bezogen:

F-Vor1 Welche Vorstellungen entwickeln Lernende zum experimentbezogenen Denken und Arbeiten während der Bearbeitung der explizit-fachmethodischen Instruktionsvariante?

Das methodische Vorgehen der qualitativen Inhaltsanalyse liefert Ideen-Kategorien, die in Ideen-Oberkategorien gebündelt werden. Anhand dieser wird auf die Vorstellungen der Lernenden geschlossen wie nachfolgend erläutert.

6.4.1 Kennzeichen von naturwissenschaftlichen Fragen

Für die Analyse von Vorstellungen der Lernenden zu Kennzeichen von naturwissenschaftlichen Fragen wurden 53 Ideen-Kategorien induktiv generiert und in 16 Ideen-Oberkategorien zusammengefasst. In Tabelle 6.10 (auf S. 284) findet sich für jede Ideen-Oberkategorie eine Beschreibung (generiert anhand von Tabelle F.3 auf S. 555, gemäß des in Unterabschnitt 6.1.7, ab S. 264, geschilderten Vorgehens). Eine Auflösung der in den Oberkategorien jeweils zusammengefassten Ideen-Kategorien findet sich im Anhang in Tabelle F.1 (auf S. 552). Ferner finden sich in Tabelle 6.10 die Anzahlen der Personen und Teams, bei denen eine jeweilige Ideen-Oberkategorie mindestens einmal zugewiesen wurde (Spalten # P und # T). Diese Anzahlen können als grobes Maß dafür angesehen werden, wie dominant die zugehörigen Ideen bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante sind.⁹²

⁹²Genauere Angaben, z. B. dazu, für wie viele Aufgaben sich die Ideen-Oberkategorien finden, finden sich weiter unten, u. a. ab S. 326.

Tabelle 6.10: Beschreibungen der Ideen-Oberkategorien zu Kennzeichen von naturwissenschaftlichen Fragen (und Anzahl der Personen (# P) bzw. Teams (# T) mit jeweiliger OK).

Ideen-Oberkategorie	Beschreibung	# P	# T
Empfindungen	In nicht-naturwissenschaftlichen Fragen geht es um Empfindungen, Gefühle oder Geschmack.	10	5
Persönliches	Eine nicht-naturwissenschaftliche Frage fragt nach persönlichen Ansichten oder Meinungen.	22	10
Subjektiv	Eine nicht-naturwissenschaftliche Frage behandelt Subjektives.	10	5
Objektiv	Eine naturwissenschaftliche Frage behandelt Objektives.	9	5
MethodeSpezifisch	Eine naturwissenschaftliche Frage kann mit einer spezifischen [benannten] Methode (z. B. Experiment, Messung, Zählen, Beobachtung) beantwortet werden.	20	8
MethodeUnspezifisch	Eine naturwissenschaftliche Frage kann mit (naturwissenschaftlichen) Methoden [allgemein formuliert, nicht spezifisch benannt] beantwortet werden.	16	8
NawiEigenschaften	Eine naturwissenschaftliche Frage enthält naturwissenschaftliche Größen oder Eigenschaften.	16	8
Zusammenhang	In einer naturwissenschaftlichen Frage wird nach dem Zusammenhang von Größen gefragt.	6	4
Effekt	Für eine naturwissenschaftliche Frage ist es wichtig, dass ein Einfluss einer Größe (tatsächlich) vorliegt.	8	7
Prüfen	Bei einer naturwissenschaftlichen Frage kann die Antwort geprüft werden.	13	8
Belegen	Bei einer naturwissenschaftlichen Frage können Belege für die Antwort angegeben werden.	7	4
BeweisenNachweisen	Bei einer naturwissenschaftlichen Frage kann die Antwort bewiesen bzw. nachgewiesen werden.	7	4
AntwortAusschlag	Die (Art der) Antwort auf eine Frage bestimmt, ob die Frage naturwissenschaftlich ist (z. B. wenn die Frage beantwortbar ist oder wenn es sich um eine Ja-Nein-Frage handelt).	7	5
PF/AF	Ob eine Frage (nicht-)naturwissenschaftlich ist, hängt davon ab, ob sie eine präzise oder allgemeine Frage ist.	4	2
Fachbezug	Eine (nicht-)naturwissenschaftliche Frage enthält Bezüge zu einer (nicht-)naturwissenschaftlichen Fachdisziplin.	19	9
Klang	Der Klang der Frage (z. B. klingt gut, klingt unsinnig) ist ausschlaggebend dafür, ob die Frage naturwissenschaftlich ist oder nicht.	6	6

Anmerkungen. Die Beschreibungen sind nicht so zu lesen, dass sie vollständige Definitionen geben. Treffendere (aber kompliziertere) Formulierungen wären derart: »Ein Merkmal einer naturw. Frage ist, dass sie ...« oder »Geht es um Empfindungen, so handelt es sich um eine nicht-naturw. Frage«.

Die Tabelle hat drei Blöcke: Block 1 umfasst Kriterien, die von nicht-naturwissenschaftlichen Fragen erfüllt (bzw. von naturwiss. Fragen nicht erfüllt) werden, und Block 2 Kriterien, die von naturwiss. Fragen erfüllt (bzw. von nicht-naturwiss. Fragen nicht erfüllt) werden. Block 3 umfasst Obergruppen, in denen sowohl Ideen-Kategorien enthalten sind, die eher Block 1, als auch solche, die eher Block 2 zugehörig sind.

Ad **MethodeSpezifisch**, **MethodeUnspezifisch**: In beiden Kategorien wird etwas Methodisches als Kriterium für naturwiss. Fragen genutzt; der Unterschied liegt darin, ob eine konkrete Methode angeführt wird.

Ad **PF/AF**: Diese Ideen-OK bezieht sich darauf, dass Schülerinnen und Schüler zwei Kategorien aus der Instruktion nutzen, die sich eigentlich auf das Formulieren präziser Fragen (siehe Konzepte in Tabelle 4.2) beziehen.

6.4.1.1 Vielfalt an Vorstellungen

Die große Anzahl von Ideen-Oberkategorien zu auf Kennzeichen von naturwissenschaftlichen Fragen bezogenen Äußerungen legt gemeinsam mit dem durch die zugehörigen Beschreibungen abgedeckten breiten Spektrum an Inhalten nahe, dass vielfältige Vorstellungen zu diesem vergleichsweise eng umrissenen Konzept zum experimentbezogenen Denken und Arbeiten vorliegen. Wird zusätzlich bedacht, dass für jede der Ideen-Oberkategorien mehrere Ideen-Kategorien vorliegen, die als Aktualisierung, als situativer Ausdruck der zu einer Ideen-Oberkategorie gehörigen Vorstellung angesehen werden können, ergibt sich, dass gleiche Vorstellungen (als Ideen-Oberkategorien) jeweils einer hohen Varianz im aktuellen Handeln unterliegen.

Eine Alternative zur Unterscheidung vielfältiger Vorstellungen bestünde darin, von der Annahme weniger handlungsleitender Vorstellungen auszugehen, die hinter mehreren Ideen-Kategorien oder Ideen-Oberkategorien stehen. Die eigentlichen Vorstellungen der Lernenden könnten dann bei dem aktuell genutzten Vorgehen erst aus der Kombination von Ideen-Kategorien oder Ideen-Oberkategorien abgeleitet werden. Beispielsweise könnte ein Schüler die Vorstellung haben, dass eine naturwissenschaftliche Frage sowohl mit einem einfachen Ja oder einem einfachen Nein beantwortbar als auch thematisch mit Naturwissenschaften verbunden sein muss. Dann würde keine der beiden Ideen-Oberkategorien **AntwortAusschlag** oder **Fachbezug** (und auch keine der ihnen jeweils zugehörigen Ideen-Kategorien) für sich genommen die Vorstellung abbilden. Es wäre daher durchaus möglich, dass einige wenige Cluster von Ideen-Oberkategorien (oder Ideen-Kategorien) vorliegen, die einigen wenigen Vorstellungen entsprechen, so dass keine Vielfalt bezüglich der Vorstellungen vorläge. Dann bliebe zumindest die Aussage angemessen, dass die identifizierten bzw. rekonstruierten Bestandteile der Vorstellungen der untersuchten Schülerinnen und Schüler vielfältig sind.

Um anhand der generierten Ideen-Kategorien und Ideen-Oberkategorien zu prüfen, wie plausibel es ist, dass latente Konstrukte hinter mehreren Ideen-Kategorien oder Ideen-Oberkategorien stehen, wurden die generierten Kodierungen betrachtet. Es ergibt sich, dass vereinzelt plausibel ist, von hinter mehreren Ideen-Kategorien stehenden handlungsleitenden Vorstellungen auszugehen, dass allerdings das Analogon auf Ebene der Ideen-Oberkategorien wenig plausibel ist:

- Beispielsweise scheint es für einige Personen möglich zu sein, dass die Ideen-Kategorien **+Persönlich** und **+Meinung** als eine Vorstellung angesehen werden sollten. Ein Großteil der Äußerungen, die mit **+Persönlich** kodiert wurden, beinhalten Bezüge zu Meinungen von Individuen (6 von 9); etwa: »[Das] hat auch immer was mit persönlicher Meinung auch zu tun« (B5, Abs. 42) oder »[1-24-ii

ist keine naturwissenschaftliche Frage, denn sie betrifft die] Meinung des Einzelnen« (H22, Abs. 129). Andersherum enthalten nur wenige der Äußerungen mit der Ideen-Kategorie +**Meinung** Bezüge dazu, dass in nicht-naturwissenschaftlichen Fragen personenspezifische oder persönliche Aspekte thematisiert werden. Eine Erklärung könnte sein, dass die Vorstellung, die mit +**Persönlich** kodiert ist, umfassender ist als die Vorstellung, die mit +**Meinung** kodiert ist, wie beispielsweise aus folgender Äußerung abgeleitet werden könnte: »Ja und vor allen Dingen, weil es nicht irgendwie was mit den persönlichen Dingen zu tun hat, wie oder Geschmack oder so Vorlieben« (O43, Abs. 64). Geschmack und Vorlieben scheinen in der Äußerung als Beispiele für persönliche Dinge aufgezählt zu werden, und Vorlieben und Meinungen liegen inhaltlich dicht beieinander, weshalb auch Meinung durchaus in der Reihe der Beispiele für persönliche Dinge aufgezählt werden könnte. Zu beachten ist allerdings, dass auf Ebene der Ideen-Oberkategorien diese beiden Ideen-Kategorien ohnehin unter **Persönliches** fallen, so dass auf dieser Ebene das Problem der Trennung nicht auftritt. Dieses Beispiel illustriert, dass es auf Ebene der Ideen-Kategorien vereinzelt dazu kommt, dass für gewisse Personen zwei Ideen-Kategorien erst gemeinsam als eine Vorstellung aufgefasst werden können.

- Auf Ebene der Ideen-Oberkategorien sprechen zwei Ergebnisse gegen das Vorliegen von Clustern, die jeweils als *eine* Vorstellung interpretiert werden könnten. (1) Clusteranalysen der Personen legen nahe, dass es keine typischen Ideen-Oberkategorien-Profile gibt (dokumentiert in Anhang F.2, ab S. 555). Es gibt also keine Typen von Personen, deren geteilte Ideen-Oberkategorien gegebenenfalls insgesamt als eine Vorstellung interpretieren werden müssten. Allerdings ist es möglich, dass bei mehreren Personen jeweils zwei Ideen-Oberkategorien erst gemeinsam als eine Vorstellung interpretiert werden können, weil bei anderen Personen eventuell andere Kombinationen von Ideen-Oberkategorien gemeinsam als Vorstellungen aufgefasst werden sollten. Zugehörige Teilcluster scheinen zwar nicht vorzuliegen, allerdings wurde diese Aussage nicht im Detail untersucht, weil die Teilcluster bei der Größe der analysierten Stichprobe schwer auszumachen wären. (2) Aus der Kreuztabelle in Tabelle 6.11 geht hervor, dass für Kennzeichen von naturwissenschaftlichen Fragen vor allem die Ideen-Oberkategorien **Subjektiv** und **Objektiv** sowie **Beweis** und **MethodeUnspezifisch** mit recht großer Häufigkeit gleichzeitig kodiert (d. h. ein und derselben Äußerung zugewiesen) werden. Ein ähnliches Bild ergibt sich auch, wenn anstelle einer Betrachtung davon, wie häufig mehrere Ideen-Oberkategorien gleichzeitig denselben Äußerungen zugewiesen sind, eine Betrachtung davon vorgenommen wird,

Tabelle 6.11: Häufigkeiten, mit denen verschiedene Ideen-Oberkategorien zu Kennzeichen von naturwissenschaftlichen Fragen gleichzeitig zugewiesen wurden.

Ideen-Oberkategorien	Fachbezug	NawiEigenschaften	MethodeSpezifisch	MethodeUnspezifisch	Prüfen	Belegen	BeweisenNachweisen	Objektiv	Subjektiv	Empfindungen	Persönlich	AntwortAusschlag	Effekt	Zusammenhang	Klang	PF	Sonstiges	Summe
Fachbezug	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
NawiEigenschaften	0	0	0	0	0	0	0	0	0	0	0	0	1	4	0	0	0	5
MethodeSpezifisch	0	0	0	3	0	0	3	3	0	0	3	1	0	0	0	0	0	13
MethodeUnspezifisch	0	0	3	0	2	2	16	0	0	0	1	0	0	0	0	0	0	24
Prüfen	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2
Belegen	0	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	3
BeweisenNachweisen	0	0	3	16	0	0	0	0	0	0	0	1	0	0	0	0	0	20
Objektiv	0	0	3	0	0	1	0	0	5	0	3	4	0	0	0	0	0	16
Subjektiv	0	0	0	0	0	0	0	5	0	2	3	2	0	0	0	0	0	12
Empfindungen	0	0	0	0	0	0	0	0	2	0	2	0	0	1	0	0	0	5
Persönlich	0	0	3	1	0	0	0	3	3	2	0	0	0	0	0	1	0	13
AntwortAusschlag	1	0	1	0	0	0	1	4	2	0	0	0	0	0	0	0	0	9
Effekt	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Zusammenhang	0	4	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	5
Klang	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PF	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
Sonstiges	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Summe	1	5	13	24	2	3	20	16	12	5	13	9	1	5	0	1	0	130

wie häufig Ideen-Oberkategorien innerhalb eines Transkriptabsatzes gemeinsam auftreten. Dadurch kann ausgeschlossen werden, dass andere Oberkategorien grundsätzlich gemeinsam auftreten und als eine Vorstellung zu interpretieren sind. Eine Zusammenfassung von Beweis und MethodeUnspezifisch ist inhaltlich nicht plausibel, nur für Subjektiv und Objektiv könnte eventuell von *einer* latenten Vorstellung ausgegangen werden. Eine Trennung wurde trotzdem aufrechterhalten, um positiv formulierte Vorstellungen von negativ formulierten zu unterscheiden. Auf Kohortenebene ergeben sich somit keine Hinweise auf das Vorliegen von Clustern von Ideen-Oberkategorien, die jeweils als eine Vorstellung interpretiert werden sollten. Auf Ebene der Individuen könnten analog zum ersten Ergebnis wiederum Cluster vor Vorstellungen vorliegen, die sich von den bei anderen Individuen vorliegenden Clustern unterscheiden und daher

nicht auf Kohortenebene erkennbar sind. Genauere Analysen wurden aufgrund des eingeschränkten zeitlichen und inhaltlichen Umfangs der Äußerungen der einzelnen Schülerinnen und Schüler nicht vorgenommen.

6.4.1.2 Abgleich mit dokumentierten Vorstellungen

Inhaltlich kann festgestellt werden, dass die in der Literatur dokumentierten Vorstellungen (siehe Unterabschnitt 2.4.3, ab S.38) auch in den vorliegenden Ideen-Oberkategorien zu Kennzeichen von naturwissenschaftlichen Fragen vorgefunden werden können. Beispielsweise liegen inhaltliche Parallelen zwischen Effekt und der berichteten Tendenz von Schülerinnen und Schülern vor, Effekte herstellen zu wollen. Gleichwohl ist diese Ideen-Oberkategorie nicht so prominent (nur 8 Personen aus insgesamt 7 verschiedenen Teams), wie vor dem Hintergrund der Literaturlage erwartbar wäre (siehe Unterabschnitt 2.4.3). Auch das bei Leach et al. (1996) dokumentierte Kriterium, welche Domäne bzw. Fachdisziplin die Frage in den Blick nimmt, findet sich in der Ideen-Oberkategorie **Fachbezug** wieder. Ferner finden sich mit **MethodeSpezifisch** und **MethodeUnspezifisch** (mindestens) zwei Kategorien, die auf die empirische Prüfbarkeit Bezug nehmen, die auch bei Leach et al. (1996) dokumentiert ist. Allerdings findet sich keine Ideen-Kategorie, die explizit Bezug auf Experimente nimmt (und das Wort wird von den Schülerinnen und Schülern auch nur an drei Stellen verwendet), sondern nur eine Ideen-Kategorie **+Versuch** (Teil der Ideen-Oberkategorie **MethodeUnspezifisch**). Was die Schülerinnen und Schüler als einen Versuch ansehen, ist allerdings unklar (und auch in der fachdidaktischen Terminologie nicht einheitlich, siehe z.B. Scharfenberg, 2005). Sicherlich spielt auch der Kontext des Lernmaterials eine Rolle dafür, welchen Begriff die Schülerinnen und Schüler verwenden. Insgesamt kann nur mit Einschränkung von einer Übereinstimmung mit den Befunden von Grygier (2008) gesprochen werden, die bereits für Schülerinnen und Schüler der vierten Klasse die Vorstellung berichtet, dass naturwissenschaftliche Fragen grundsätzlich mit einem Experiment beantwortet werden können. Weitere Analysen im Verlauf der Arbeit zeigen, dass die geringe Übereinstimmung vorrangig durch die stärkere Differenzierung der Vorstellungen der Lernenden aus der elften Klassenstufe bedingt zu sein scheint, in der Grundaussage der empirischen Prüfbarkeit (durch Experimente/Versuche) allerdings ähnliche inhaltliche Tendenz vorzuliegen scheint.

An dieser Stelle kann zudem bereits angemerkt werden, dass weder die Ideen-Oberkategorien noch die Ideen-Kategorien feine Unterschiede in den Vorstellungen abbilden können, die teilweise für die Einordnung oder Einschätzung der Vorstellung (wie etwa hinsichtlich der Angemessenheit) relevant sind. Beispielsweise bezieht

sich der vorletzte Satz des vorigen Absatzes darauf, dass naturwissenschaftliche Fragen mit Experimenten *beantwortet* werden können. Das Wort wurde in der obigen Formulierung gewählt, weil es der Terminologie bei Grygier (2008) entspricht. In den Aussagen der Schülerinnen und Schüler finden sich bei der Studie der vorliegenden Arbeit allerdings auch Formulierungen wie: Naturwissenschaftliche Fragen lassen sich mit Versuchen *beweisen* (gemeint ist dann vermutlich, dass die Antworten auf naturwissenschaftliche Fragen mit Versuchen bewiesen werden können), sie lassen sich mit Versuchen *prüfen*, sie lassen sich mit Versuchen *untersuchen*. Diese Vielfalt spiegelt sich partiell in den in vielen Fällen parallel vergebenen zweiten Ideen-Oberkategorien. Wenn die Aussage allerdings »Weil du dazu einen Versuch machen kannst« o. ä. lautete, wurde hingegen nur die Ideen-Oberkategorie **MethodeUnspezifisch** vergeben. Manche der vorliegenden Unterschiede, die nicht auf Ebene der Ideen-Oberkategorien aufgelöst sind, sind durch die Ideen-Kategorien abgebildet. Für einen Großteil der Nuancen liegt eine Abbildung durch Zweitkodierungen von Ideen-Oberkategorien oder durch feinere Gliederung innerhalb der Ideen-Kategorien allerdings eher nicht vor. Damit verbunden sind Einschränkungen hinsichtlich der Genauigkeit der Wiedergabe der Vorstellungen verbunden, die insbesondere im Kontext der Einschätzung der Angemessenheit der Vorstellungen zu beachten sind, weshalb sie dort ausführlicher diskutiert werden.

Zu Kennzeichen von naturwissenschaftlichen Fragen finden sich einige Ideen-Oberkategorien, die zwar mit der Literatur in Einklang gebracht werden können, aber bisher nicht dokumentiert sind. Insbesondere zählen dazu Ideen-Oberkategorien, die negativ formuliert sind und beschreiben, wie naturwissenschaftlichen Fragen nicht aussehen dürfen (z. B. **Empfindungen**, **Persönlich**). Insgesamt scheinen die Ideen-Oberkategorien die Verständnisse differenzierter aufzulösen als bisherige Studien oder zumindest bisher undokumentierte Aspekte (insbesondere auch angemessene Vorstellungen) zu ergänzen. Anzumerken ist, dass die Ideen-Oberkategorie **PF/AF** (mit der Äußerungen kodiert wurden, in denen Lernende Fragen auf ihre Naturwissenschaftlichkeit einschätzen, indem sie auf die Unterscheidung von präzisen und allgemeinen Fragen zurückgreifen) vermutlich dem Kontext des Materials entspringt. Aufgabenkarten vor der analysierten Sequenz (1-04 bis 1-09) führen in eine Unterscheidung von präzisen und allgemeinen Fragen ein, wobei darunter verstanden wird, ob die Fragen die zu untersuchenden Merkmale eindeutig benennen (im Sinne von Variablen) oder nicht. Es wird im Rahmen der vorliegenden Arbeit nicht beantwortet, inwiefern derartige Überlegungen außerhalb des Materials der untersuchten explizit-fachmethodischen Instruktionsvariante vorkommen, weil keine zugehörigen Analysen (bspw. für die implizit-fachmethodische Instruktionsvariante) vorgenommen wurden.

6.4.2 Unterscheidung von Beobachtung und Deutung

Für die Analyse von Vorstellungen der Lernenden zur Unterscheidung von Beobachtung und Deutung wurden 28 Ideen-Kategorien induktiv generiert und in 12 Ideen-Oberkategorien zusammengefasst. In Tabelle 6.12 findet sich für jede Ideen-Oberkategorie eine Beschreibung (generiert anhand von Tabelle F.3 auf S. 555 im

Tabelle 6.12: Beschreibungen der Ideen-Oberkategorien zur Unterscheidung von Beobachtung und Deutung

Ideen-Oberkategorie	Beschreibung	# P	# T
Sehen	Beobachtungen beziehen sich darauf, was man sehen kann.	11	6
Beobachten	Beobachtungen beziehen sich darauf, was man beobachten kann.	8	4
FaktenSicher	Beobachtungen beziehen sich auf Fakten bzw. etwas, was sicher ist.	5	5
ObjektivSubjektiv	Beobachtungen beziehen sich auf Objektives	3	1
Beschreiben	Beobachtungen beschreiben (nur).	2	2
(Messen)	Beobachtungen geben wieder, was gemessen wurde. Oder: Beobachtungen sind nur das, was nicht gemessen wurde.	1	1
BegründungErklärung	Deutungen beinhalten Begründungen oder Erklärungen oder z. B. das Wort <i>weil</i> .	16	8
Vermutung	Deutungen sind Vermutungen.	9	5
Interpretation	Deutungen beinhalten, dass interpretiert wird.	6	5
AuswertungAntwort	Deutungen sind Auswertungen bzw. geben eine Antwort (auf eine Frage).	4	3
Meinung	Deutungen beinhalten Meinungsäußerungen.	3	2
Sonstiges	Sonstige Unterscheidungskriterien.	5	2

Anmerkungen. Die Beschreibungen sind nicht so zu lesen, dass sie vollständige Definitionen geben. Treffendere (aber kompliziertere) Formulierungen wären derart: »Wird eine Begründung oder Erklärung gegeben, so handelt es sich um eine Deutung« oder »Ein Merkmal anhand dem zwischen Beobachtung und Deutung unterschieden werden kann ist, dass Beobachtungen eher ...«.

Die Tabelle hat drei Blöcke: Block 1 umfasst Kriterien, die vorrangig auf Beobachtungen bezogen sind (teilweise aber auch genutzt werden, um etwas als Nicht-Deutung zu identifizieren), und Block 2 Kriterien, die vorrangig auf Deutungen bezogen sind (teilweise aber auch genutzt werden, um etwas als Nicht-Beobachtung zu identifizieren). Block 3 umfasst die Obergruppe **Sonstiges**, in der sowohl Ideen-Kategorien enthalten sind, die eher Block 1, als auch solche, die eher Block 2 zugehörig sind, aber nur selten auftreten und schwer zugeordnet werden können.

Ad (Messen): Tritt nur an einer einzigen Stelle auf, scheint aber zu bedeutsam, als dass es unter **Sonstiges** subsubmiert werden könnte. Bei der Einschätzung von Aufgabe 3-20-3 sagt L34: »Äh, da wäre ich mir nicht sicher. Ich b/ Hier steht ja nicht, ob die das gemessen haben.«. Es ist unklar, wie das gemeint ist, daher werden zwei Deutungen in der Tabelle angegeben.

Anhang gemäß der Schilderungen in Unterabschnitt 6.1.7, ab S. 264) sowie die Anzahl der Personen und Teams, bei denen sie mindestens einmal zugewiesen wurde. Eine Auflösung der in den Oberkategorien jeweils zusammengefassten Ideen-Kategorien findet sich im Anhang in Tabelle F.2 (auf S. 554).

6.4.2.1 Vielfalt an Vorstellungen

Analog zum Vorgehen für Kennzeichen von naturwissenschaftlichen Fragen in Unterabschnitt 6.4.1.1 werden Überschneidungen von Ideen-Oberkategorien betrachtet, um festzustellen, inwiefern sich die einzelnen Ideen-Oberkategorien für die Unterscheidung von Beobachtung und Deutung als einzelne Vorstellungen interpretieren lassen (ergänzend ebenfalls analog: Anhang F.2, ab S. 555). Es finden sich nur sehr vereinzelt Überschneidungen von Ideen-Oberkategorien. Wird das Auftreten von Ideen-Oberkategorien innerhalb desselben Transkriptabsatzes betrachtet, so findet sich (abgesehen von vereinzelt Instanzen) nur für die Ideen-Oberkategorien **BegründungErklärung** und **Sehen** substantielles Zusammenauftreten wie in Tabelle 6.13 dokumentiert. Die zugehörigen Instanzen (die sich allesamt bei der Teilaufgabe 3-21-ii finden, siehe

Tabelle 6.13: Häufigkeiten, mit denen verschiedene Ideen-Oberkategorien zu Unterscheidung von Beobachtung und Deutung innerhalb desselben Transkriptabsatzes zugewiesen wurden.

Ideen-Oberkategorien	Meinung	Vermutung	Interpretation	BegründungErklärung	AuswertungAntwort	FaktenSicher	Beschreiben	Messen	Beobachten	Sehen	Sonstiges	ObjektivSubjektiv	Summe
Meinung	0	0	2	0	2	0	0	0	0	0	0	0	4
Vermutung	0	0	0	4	0	2	0	0	2	9	0	0	17
Interpretation	2	0	0	0	2	0	0	0	0	0	0	0	4
BegründungErklärung	0	4	0	0	2	0	2	0	3	14	0	0	25
AuswertungAntwort	2	0	2	2	0	0	0	0	0	0	0	0	6
FaktenSicher	0	2	0	0	0	0	0	0	0	2	0	0	4
Beschreiben	0	0	0	2	0	0	0	0	2	0	0	0	4
Messen	0	0	0	0	0	0	0	0	0	0	0	0	0
Beobachten	0	2	0	3	0	0	2	0	0	6	2	0	15
Sehen	0	9	0	14	0	2	0	0	6	0	4	0	35
Sonstiges	0	0	0	0	0	0	0	0	2	4	0	0	6
ObjektivSubjektiv	0	0	0	0	0	0	0	0	0	0	0	0	0
Summe	4	17	4	25	6	4	4	0	15	35	6	0	120

Anhang D ab S. 545) bestehen aus allgemeinen Darlegungen des Verständnisses von Beobachtungen und Deutungen, bei denen zum einen erläutert wird, dass eine Deutung eine Begründung bzw. Erklärung ist (genauer: eine Ideen-Kategorien vergeben wurde, die zur Ideen-Oberkategorie **BegründungErklärung** gehört), und zum anderen eine Beobachtung als etwas beschrieben wird, was gesehen werden kann (d. h. die Ideen-Oberkategorie **Sehen** wurde vergeben). Auch die zu anderen Ideen-Oberkategorien gehörigen Instanzen bestehen zum Großteil aus zusammenhängenden Gesprächsabschnitten einer Person (die daher als ein Transkriptabsatz vorliegen), in denen sowohl Deutungen als auch Beobachtungen definiert oder umrissen werden. Die Ideen-Oberkategorien werden also aus guten Gründen in den Fällen jeweils nicht als *eine* Vorstellung interpretiert, weil sie sie einerseits Beobachtungen und andererseits Deutungen betreffen.

Die Anzahl der Vorstellungen zur Unterscheidung von Beobachtung und Deutung ist innerhalb der Sequenzen zu Beobachtungen beziehungsweise zu Deutungen jeweils recht gering. Inhaltlich ist die durch Ideen-Oberkategorien jedoch ein breites Spektrum – für Deutungen beispielsweise von Vermutungen über Erklärungen bis hin zu auswertenden Antworten – gegeben. Es ist daher sinnvoll, sowohl bei Kennzeichen von naturwissenschaftlichen Fragen (siehe oben) als auch bei der Unterscheidung von Beobachtung und Deutung davon auszugehen, dass Schülerinnen und Schüler im Verlauf einer kurzen Sequenz vielfältige Vorstellungen äußern.

6.4.2.2 Abgleich mit dokumentierten Vorstellungen

Die zur Unterscheidung von Beobachtung und Deutung dokumentierte Vorstellungen finden sich ebenfalls zu großen Teilen im vorliegenden Datensatz. Es finden sich scheinbar auch für Schülerinnen und Schüler der elften Klassenstufe Antworten, die wenig aussagekräftig sind (in ähnlicher Weise dokumentiert als »tautologische Antworten« für Kennzeichen von naturwissenschaftlichen Fragen bei Leach et al., 1996), was sich beispielsweise in der Häufigkeit der Ideen-Oberkategorie **Beobachten** widerspiegelt. Zugehörige Aussagen scheinen nicht tautologisch in dem Sinne zu sein, dass sie vollständig selbstreferentiell nichtsaussagend wären. Vielmehr wurde zur Definition von Beobachtungen das zugehörige Verb *beobachten* genutzt. Welche Vorstellung von Beobachten bei den Schülerinnen und Schülern allerdings damit verbunden ist, geht aus solchen Aussagen nicht hervor – eventuell weil keine spezifischere Vorstellung vorhanden ist (vgl. von Aufschnaiter & Rogge, 2010a). Ferner sind in Tabelle 6.14 auf Seite 293 den von Petermann (2017) dokumentierten empirisch erhobenen Vorstellungen (siehe Abbildung 2.2, S. 43) die Ideen-Oberkategorien aus

Tabelle 6.12 (S. 290) zugeordnet. Dort nicht aufgeführt ist die Ideen-Oberkategorie **FaktenSicher**, die der ersten bei Petermann (2017) in der Literaturzusammenführung (siehe Tabelle 2.1, S. 42) notierten Vorstellung zu Beobachtungen (»eine Sammlung von gewissenhaft erhobenen Daten führt zu sicherem Wissen«) ähnlich ist.

Zur Unterscheidung von Beobachtung und Deutung findet sich einzig die Ideen-Oberkategorie **Meinung** nicht im Spektrum der dokumentierten Vorstellungen. Diese hohe Übereinstimmung mit der Literatur kann als Bestätigung für das methodische Vorgehen angesehen werden. Eine weitere Abweichung von der Literatur liegt darin, dass in der Arbeit bezüglich der Vorstellungen zu Beobachtungen eine feinere Unterteilung vorgestellt wird, als sie in den Kategorien von Petermann (2017) zu finden ist. Aus den Äußerungen, die bei Petermann (2017) im Anhang zu finden sind, lässt sich die feinere Unterteilung in **Sehen**, **Beobachten**, **Beschreiben** (und ggf. **Messen**) ebenfalls ableiten. Die Zusammenfassung zur Vorstellung »Beobachtung = Wahrgenommenes« scheint aber ihre Berechtigung zu haben, weil die klare Trennung der Bedeutungen verschiedener Schüleräußerungen nicht immer möglich ist, sondern die Ideen-Oberkategorien bezüglich ihrer Inhalte Überlapp haben könnten. In der Arbeit wird anders als bei Petermann (2017) trotzdem eine Unterscheidung vorgenommen, weil – zum Beispiel wegen der Kürze der Schüleräußerungen – für gewöhnlich aus den analysierten Bearbeitungsprozessen keine klaren Hinweise über die tatsächlich von den Lernenden intendierten Inhalte abgeleitet werden konnten (anders als im Interviewkontext bei Petermann, 2017).

Tabelle 6.14: Abgleich der Ideen-Oberkategorien mit dokumentierten Vorstellungen

Dokumentierte Vorstellung	Passende Ideen-Oberkategorie
Beobachtung = Wahrgenommenes	Sehen, Beobachten, Beschreiben, Messen
Deutung = Erklärung	BegründungErklärung
Deutung = Verallgemeinerung	eventuell AuswertungAntwort
Deutung = Verknüpfung mit Bekanntem	
Deutung = Interpretation	Interpretation
Deutung = Vermutung	Vermutung

Anmerkungen. Die Beschreibungen der dokumentierten Vorstellungen sind Kurzformen von Tabelle 2.1 auf S. 42.

6.5 Auswertungen und Ergebnisse zur Allgemeinheit der Vorstellungen

Die zweite Forschungsfrage zu den Vorstellungen der Lernenden lautet:

F-Vor2 Welche Vorstellungen zum experimentbezogenen Denken und Arbeiten sind bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante auf konkrete Situationen, welche auf Verallgemeinerungen bezogen?

Anhand ausgewählter Kategorien zum Modus (siehe Unterabschnitt 6.1.3 ab S. 254) kann untersucht werden wie sehr die Lernenden die Äußerungen, anhand derer auf ihre Vorstellungen geschlossen wird, auf die konkreten Situationen aus der Instruktion beziehen. Mithilfe der Modus-Kategorien *Allgemeine Regel* und *Einschätzung mit Hinweis* (sowie z. T. *Einschätzung ohne Hinweis*) wird dies nachfolgend untersucht. Die Ergebnisse werden jeweils mit Blick auf die mit den Äußerungen verbundenen Vorstellungen gedeutet.

Für die Analysen werden Anzahlen von *Instanzen* verglichen. Eine (Segment-)Instanz ist dabei eine Kombination aus einem kodierten Segment und der zugewiesenen Ideen-Oberkategorie. Demnach wurden, aus Gründen der vereinfachten Zählung, Segmente, denen zwei Ideen-Oberkategorie gleichzeitig zugewiesen sind, als zwei Instanzen gezählt. Analog gilt diese Definition von Instanz auch für Modus-Kategorien. Allerdings gilt die Beschreibung von Instanz nur innerhalb derselben Ebene des Kodiersystems; das heißt, dass ein Segment, welches mit einer Ideen-Oberkategorie und einer Modus-Kategorie kodiert ist, als eine Instanz betrachtet wird. Ein kodiertes Segment, welches mit zwei Ideen-Oberkategorien und einer Modus-Kategorie kodiert ist, wird hingegen als zwei Instanzen gezählt.

6.5.1 Allgemeinheit der Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen

Für Kennzeichen von naturwissenschaftlichen Fragen liegen insgesamt 440 Instanzen mit einer der drei Modus-Kategorien *Allgemeine Regel* oder *Einschätzung mit Hinweis* oder *Einschätzung ohne Hinweis* vor. Bei 368 dieser Instanzen sind ebenfalls Ideen-Oberkategorien zugewiesen; davon entfallen mit 305 Instanzen mehr als 80 % auf Einschätzungen (mit Hinweis) und 63 Instanzen auf Äußerungen, die einen allgemeineren Bezug haben, der über die je spezifische Situation hinaus geht. Die Äußerungen der Lernenden sind also gemäß dieser Operationalisierung sehr deutlich vorrangig auf konkrete Situationen bezogen und nur selten allgemeiner formuliert. In Abbildung 6.11 ist für jede Ideen-Oberkategorie dargestellt, wie groß der relative Anteil

6 Vorstellungen von Lernenden

der allgemeinen Regeln (an allen der jeweiligen Ideen-Oberkategorie zugehörigen Instanzen) ist. Der absolute Anteil ist ebenfalls notiert. Zu jeder Ideen-Oberkategorie findet sich mindestens eine allgemeine Formulierung, wenngleich Klang, Effekt und Prüfen nur jeweils ein einziges Mal bei gleichzeitiger Kodierung als allgemeine Regel vergeben wurden. Die Ideen-Oberkategorien **NawiEigenschaften**, **Objektiv** und **Fachbezug** sind absolut am häufigsten bei allgemeiner Formulierung vertreten.

Zur Beantwortung der Forschungsfrage F-Vor3 wurden die Ideen-Oberkategorien nach ihrer Angemessenheit beurteilt. Genauere Ergebnisse dazu finden sich ab S. 306.

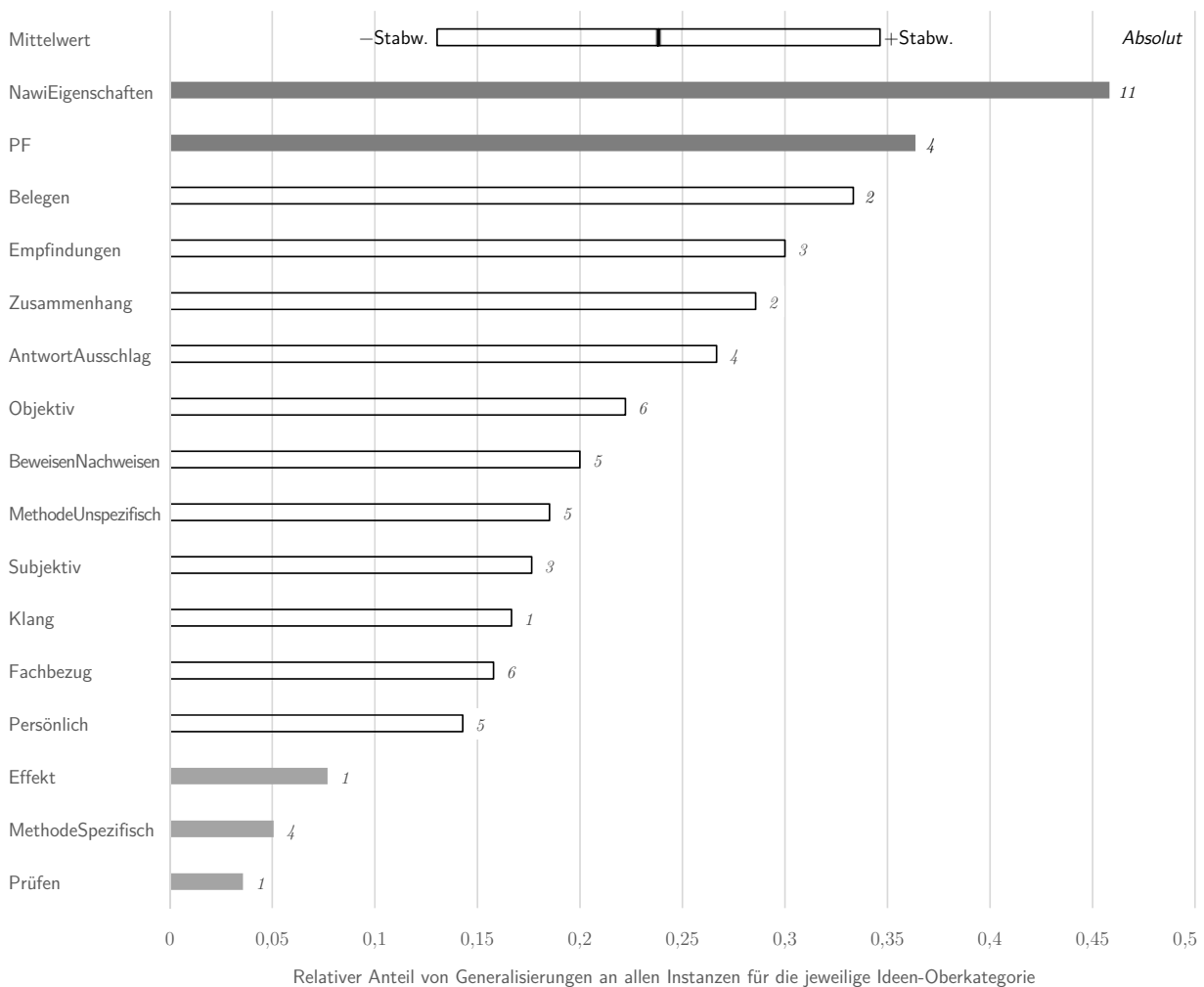


Abbildung 6.11: Relativer Anteil von allgemeinen Formulierungen an allen Instanzen aufgelöst für einzelne Ideen-Oberkategorien zu Kennzeichen von naturwissenschaftlichen Fragen. Abweichungen vom Mittelwert der relativen Anteile, die größer als eine Standardabweichung sind, sind durch gefüllte Balken hervorgehoben. Der absolute Anteil von allgemeinen Formulierungen mit jeweiliger Ideen-Oberkategorie ist am Ende der Balken notiert.

Bezüglich der Frage, wie allgemein die Formulierungen der Äußerungen sind, aus denen die Vorstellungen rekonstruiert werden, lässt sich allerdings bereits an dieser Stelle anmerken: Eher angemessene und eher unangemessene Ideen-Oberkategorien sind fast über das gesamte Spektrum verteilt, wenngleich sie nicht innerhalb der relativ stark generalisiert geäußerten Ideen-Oberkategorien (dunkelgrau markierter, oberer Teil des Spektrums) vorkommen. Der Grad der Generalisierung scheint also eher nicht mit der Angemessenheit der Vorstellungen einherzugehen.

In den bisher berichteten Ergebnissen ist nicht aufgelöst, bei welchen Aufgaben die allgemeinen Formulierungen oder Einschätzungen vorkommen. Es ist anzumerken, dass von den 63 allgemeinen Formulierungen 38 (ca. 60 %) bei der Teilaufgabe 1-22-ii vorkommen, in der nach allgemeinen Kriterien für die Unterscheidung von naturwissenschaftlichen und nicht-naturwissenschaftlichen Fragen gefragt wird: »Stellen Sie sich vor, Sie sollten jemandem erklären, was eine naturwissenschaftliche Fragestellung von einer nicht-naturwissenschaftlichen Fragestellung unterscheidet. Was würden Sie sagen?« Eine gesonderte Betrachtung dieser Teilaufgabe scheint daher sinnvoll.

Bei 1-22-ii finden sich kaum Einschätzungen, so dass eine Betrachtung der relativen Anteile von allgemeinen Formulierungen wenig sinnvoll ist. Stattdessen sind in Abbildung 6.12 die absoluten Häufigkeiten von allgemeinen Formulierungen und Einschätzungen bei der Teilaufgaben 1-22-ii dargestellt. Die Darstellung der Kategorien erfolgt entsprechend der Reihenfolge aus Abbildung 6.11 (S. 296). Eine Betrachtung liefert, dass bei 1-22-ii fast alle Ideen-Oberkategorien vertreten sind. Einige Ideen-Oberkategorien treten nur ein einziges Mal auf. Die 11 allgemeinen Formulierungen zur Ideen-Oberkategorie **NawiEigenschaften** treten vollständig bei 1-22-ii auf. Da die Aufgabe vor der Explizierung eingesetzt wurde, liegt – wie bei der vorigen Forschungsfrage bereits erwähnt – die Vermutung nahe, dass Äußerungen, die mit **NawiEigenschaften** kodiert wurden, unspezifisch formulierte Verbalisierungen des intuitiven Verständnisses zu Kennzeichen von naturwissenschaftlichen Fragen sind. Läge nach der Explizierung eine mit 1-22-ii vergleichbare Aufgabe vor, wäre ein Vergleich der vorliegenden mit allgemeinen Formulierungen verbundenen Ideen-Oberkategorien ein Ansatz zur genaueren Untersuchung der Entwicklung des Verständnisses (auf anderem Wege untersucht bei F-Vor6 ab S. 353).

Werden die allgemeinen Formulierungen, die bei der Teilaufgabe 1-22-ii vorkommen, aus der Analyse ausgeschlossen ergibt sich Abbildung 6.13 (erneut mit der Sortierung aus Abbildung 6.11). Es gilt, dass der Anteil an allgemeinen Formulierung zu jeder einzelnen Ideen-Oberkategorie deutlich geringer ist als bei Einbezug der Aufgabe. Auch die absoluten Häufigkeiten von allgemeinen Formulierungen sind außerhalb von 1-22-ii sehr gering – wenngleich trotzdem bei sehr vielen Ideen-Oberkategorien nach wie vor allgemeine Formulierungen vorliegen. Absolut vergleichsweise häufig

6 Vorstellungen von Lernenden

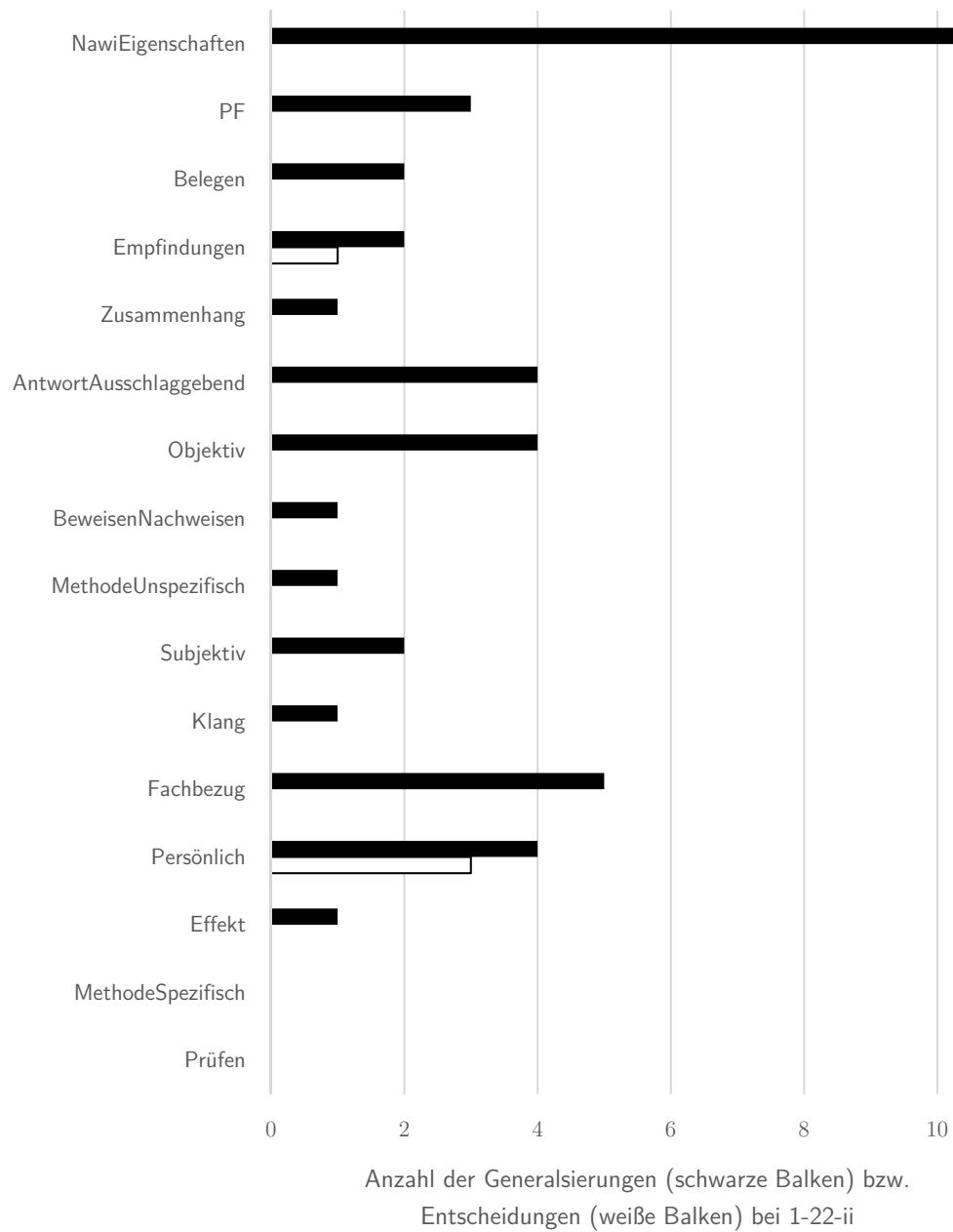


Abbildung 6.12: Anzahl der allgemeinen Formulierungen (schwarze Balken) und Einschätzungen (weiße Balken) bei Instanzen, die 1-22-ii zugehörig sind, aufgelöst für einzelne Ideen-Oberkategorien.

6.5 Auswertungen und Ergebnisse zur Allgemeinheit der Vorstellungen (F-Vor2)

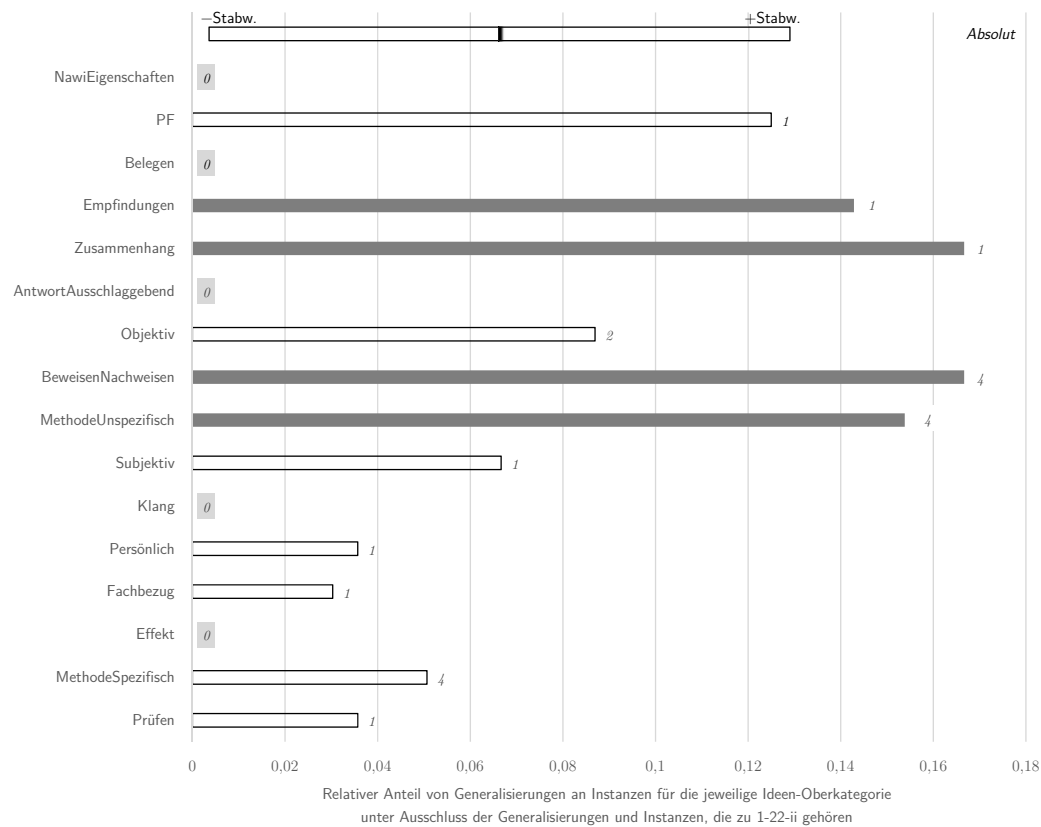


Abbildung 6.13: Relativer Anteil von allgemeinen Formulierungen an Instanzen abgesehen von solchen, die 1-22-ii zugehörig sind, aufgelöst für einzelne Ideen-Oberkategorien. Abweichungen vom Mittelwert der relativen Anteile, die größer als eine Standardabweichung sind, sind durch gefüllte Balken hervorgehoben. Der absolute Anteil von allgemeinen Formulierungen mit jeweiliger Ideen-Oberkategorie ist am Ende der Balken notiert.

vertretene Ideen-Oberkategorien sind auch zumeist solche, die einen hohen relativen Anteil an allgemeinen Formulierungen aufweisen. Davon ausgenommen ist **MethodeSpezifisch**. Dies erklärt sich aus der Bedeutung der Ideen-Oberkategorie: Sie wurde solchen Äußerungen zugewiesen, die sich auf spezifische Methoden beziehen. Es ist daher nicht erwartbar, dass diese Ideen-Oberkategorie bei der Formulierung von allgemeinen Unterscheidungskriterien vorkommt. Die Ideen-Oberkategorie wurde entsprechend auch deutlich häufiger auf Fallebene genutzt. Trotzdem auftretende allgemeine Formulierungen zu **MethodeSpezifisch** sind entweder als unangemessener Gebrauch einzuschätzen (weil eine spezifische Methode als definierendes Unterscheidungskriterium eingesetzt wird) oder im Kontext als Beispiele zu der dann ebenfalls vergebenen Ideen-Oberkategorie **MethodeUnspezifisch** anzusehen. Die absolut und relativ prominenteste Ideen-Oberkategorie ist **MethodeUnspezifisch**, bei der 4 von 5 allgemeine Formulierungen außerhalb von 1-22-ii auftreten. Wenn berücksichtigt wird, dass **MethodeUnspezifisch** vorrangig bei Aufgaben nach 1-22-ii auftaucht, scheint es ferner der Fall zu sein, dass die dort hervorstechende Ideen-Oberkategorie **NawiEigenschaften** (unter anderem) von **MethodeUnspezifisch** abgelöst wird. Dazu mehr bei F-Vor6 zu Vorstellungen im zeitlichen Verlauf ab S. 353.

Eine Auflösung der absoluten und relativen Anteile von allgemeinen Formulierungen pro Person findet sich im Anhang in Tabelle F.7 (S. 561). Anhand dieser Tabelle lässt sich vorläufig ersehen, dass Personen, die außerhalb von 1-22-ii als allgemeine Regeln kodierte Äußerungen tätigen, auch insgesamt einen höheren Anteil an allgemeinen Formulierungen aufweisen. Dies dürfte auch der Fall sein, wenn die Anzahl der Transkriptabsätze pro Person oder die Anzahl der Einschätzungen kontrolliert wird. Die entsprechenden Analysen (mittels Kovarianzanalysen oder partiellen Korrelationsanalysen) wurden aus Zeitgründen nicht mehr durchgeführt. Basierend auf den Schätzungen könnte es also sein, dass es Personen gibt, die grundsätzlich eher zu allgemeinen Formulierungen und Generalisierungen neigen (beispielsweise, weil sie höhere allgemein-kognitive Fähigkeiten haben, dazu mehr bei F-Vor7 ab S. 397). Andererseits muss diese Deutung dadurch eingeschränkt werden, dass die Anzahlen von allgemeinen Formulierungen davon abhängen, wie viele Ideen-Oberkategorien vergeben wurden. Insbesondere bei 1-22-ii kann daher durch gleichzeitige Zuweisen zweier Ideen-Oberkategorien (tritt 11 mal bei 1-22-ii auf, drei gleichzeitig zugewiesene Ideen-Oberkategorien finden sich nur zweimal im gesamten Datensatz) der Eindruck entstehen, es lägen mehr allgemeine Formulierungen vor, als tatsächlich vorgenommen wurden.

6.5.2 Allgemeinheit der Vorstellungen zur Unterscheidung von Beobachtung und Deutung

Für die Unterscheidung von Beobachtung und Deutung finden sich 24 allgemeine Formulierungen bei 72 analysierbaren Instanzen mit zugewiesenen Ideen-Oberkategorien. Damit entspricht der Anteil der Einschätzungen zwei Drittel. Die zugehörigen Äußerungen sind also vorrangig auf Einschätzungen bezogen und nicht verallgemeinernde Aussagen. Die Ideen-Oberkategorien **Messen** und **Sonstiges**, die insgesamt nur einmal vergeben wurden, treten nicht als allgemeine Formulierungen auf. Die Ideen-Oberkategorie **Beschreiben** tritt insgesamt dreimal auf, ebenfalls niemals als allgemeine Regeln. Die relativen Anteile von allgemeinen Formulierungen an Instanzen, bei denen eine Ideen-Oberkategorie vergeben wurde, sind für die einzelnen Ideen-Oberkategorien in Abbildung 6.14 dargestellt.

Wie bei Kennzeichen von naturwissenschaftlichen Fragen findet sich auch in der Sequenz zur Unterscheidung von Beobachtung und Deutung mit der Aufgabe 3-21-ii eine Aufgabe, die direkt zum Verallgemeinern auffordert. Die Aufgabe lautet »Haben Sie eine Idee, woran man erkennt, dass es sich um eine Deutung handelt?« und ist somit klar auf Deutungen und nicht auf Beobachtungen bezogen. In Tabelle 6.15 sind die Anzahlen von Einschätzungen und allgemeine Formulierungen insgesamt und für 3-21-ii gesondert dargestellt. Einschätzungen liegen bei 3-21-ii einzig für **Beobachten** vor (2 Stück von insgesamt 4 Einschätzungen zu dieser Ideen-Oberkategorie) und sind

Tabelle 6.15: Anzahl der allgemeinen Formulierungen pro Ideen-Oberkategorie für die Unterscheidung von Beobachtung und Deutung.

Ideen-Oberkategorie	Einschätzungen	Allgemeine Formul.	
		insgesamt	davon bei 3-21-ii
Messen	1	0	0
Beobachten	3	0	0
Sonstiges	1	0	0
AuswertungAntwort	4	1	0
Vermutung	10	3	2
BegründungErklärung	17	7	7
Meinung	2	1	0
Interpretation	4	3	2
Sehen	4	4	4
FaktenSicher	2	3	2
Beschreiben	0	1	1
ObjektivSubjektiv	0	1	0

6 Vorstellungen von Lernenden

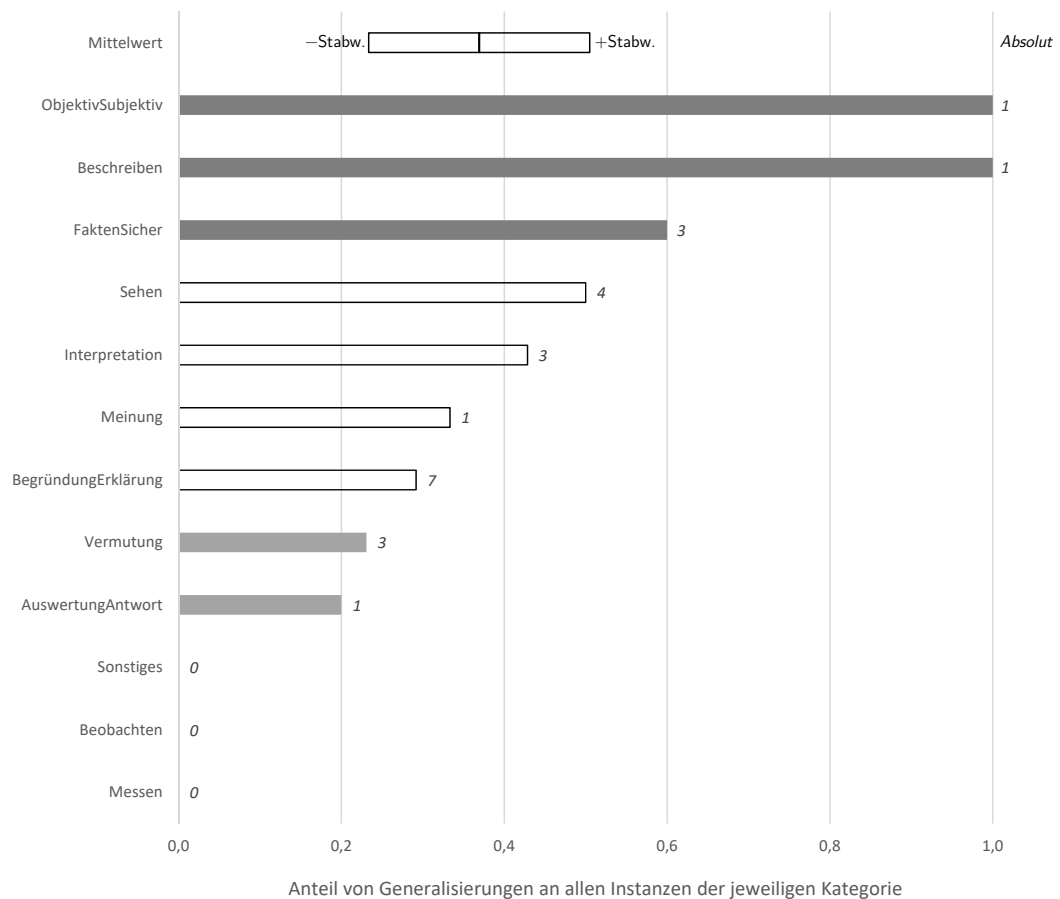


Abbildung 6.14: Relativer Anteil von allgemeinen Formulierungen an allen Instanzen aufgelöst für einzelne Ideen-Oberkategorien zur Unterscheidung von Beobachtung und Deutung. Abweichungen vom Mittelwert der relativen Anteile, die größer als eine Standardabweichung sind, sind durch gefüllte Balken hervorgehoben. Mittelwert und Standardabweichung beziehen sich auf eine korrigierte Stichprobe, aus der die Ideen-Oberkategorien mit 0 %-igem und 100 %-igem Anteil ausgeschlossen sind. Der absolute Anteil von allgemeinen Formulierungen mit jeweiliger Ideen-Oberkategorie ist am Ende der Balken notiert.

6.5 Auswertungen und Ergebnisse zur Allgemeinheit der Vorstellungen (F-Vor2)

nicht in der Tabelle aufgeführt. Die Einschätzungen bei 3-21-ii kommen daher, dass in beiden Fällen die weiter oben auf der Karte aufgeführte Lösung (siehe Abbildung 6.5) sprachlich einbezogen wird: »Und DAS ist alles, was du bloß sehen kannst« (H22, Abs. 50; wobei das betonte »DAS« ziemlich klar für die Deutung als Bezug zur Lösung spricht) und »Und bei [den aufgeführten] Beobachtungen ist halt nur, was man sieht« (P46, Abs. 16; wobei der eingefügte Bezug zur aufgeführten Lösung auch eine Fehlinterpretation sein kann).

Indirekt aus Tabelle 6.15 ersichtlich ist, dass außerhalb von 3-21-ii allgemeine Formulierungen nur jeweils einmal pro Ideen-Oberkategorie vorkommen – und zwar für die Ideen-Oberkategorien **ObjektivSubjektiv** (eine allgemeine Formulierung bei insgesamt einer Instanz), **FaktenSicher** (1 von 3), **Interpretation** (1 von 5), **Meinung** (1 von 3), **Vermutung** (1 von 11) und **AuswertungAntwort** (1 von 5). Ferner kommt eine allgemeine Formulierung zur Ideen-Oberkategorie **Beschreiben** bei der Teilaufgabe 3-21-ii vor, obwohl die Ideen-Oberkategorie auf Beobachtungen bezogen ist, die Aufgabe aber auf Deutungen. Aus der zugehörigen Aussage von C8 wird deutlich, dass er sich Deutungen im Kontrast zu Beobachtungen vorstellt. Daher beinhaltet seine Äußerung zur nur auf Deutungen bezogenen Aufgabe zugleich auch Hinweise über seine Vorstellung zu Beobachtungen. Er sagt: »Ja, eine Deutung ist, wenn man versucht, was zu erklären und Beobachtung ist, wenn man das beschreibt« (C8, Kennzeichen von naturwissenschaftlichen Fragen, Abs. 25; an dem »das« kann die benannte Kontrastierung sprachlich festgemacht werden). Als Randbemerkung lässt sich hinzufügen, dass diese Aussage aus fachmethodischer Sicht als angemessen eingeschätzt werden kann.

Eine Auflösung der absoluten und relativen Anteile von Generalisierungen pro Person findet sich im Anhang in Tabelle F.8 (auf S. 562). Personen formulieren keine oder eine allgemeine Äußerung (und diese taucht grundsätzlich bei 3-21-ii auf). Nur drei Personen machen mehr als eine Äußerung, aus der sich auf eine Vorstellung schließen lässt, in Form einer allgemeinen Formulierung. Der Schüler K33 sagt bei der Teilaufgabe 3-21-ii (Abs. 34):

Ja, also bei Deutung, dafür (?musst du?) schon ein bisschen begründen.

Zusätzlich zu dieser Äußerung findet sich auch bei der Teilaufgabe 3-24-b eine Äußerung, die als allgemeine Formulierung eingeschätzt wurde (Abs. 67):

Warum wäre es bei dieser Diskussion hilfreich, zwischen/ (..) Da man nur sicher

6 Vorstellungen von Lernenden

sagen kann, was man beobachtet. Deutung, das wäre dann wieder so ein Vielleicht und das kann man nicht hundertprozentig sagen.

Die Aussage zu Deutungen bzw. zum Deuten wird so gedeutet, dass sie nicht auf die vorliegende Situation bezogen ist, sondern vielmehr von der Person K33 eine allgemeine Aussage angeführt wird, um Deutungen zu charakterisieren. Die benannten Kennzeichen von Deutungen scheinen dann implizit die Antwort auf die Frage darzustellen, warum es hilfreich wäre, zwischen Beobachtungen und Deutungen zu unterscheiden. Wie stark die Charakterisierung von Deutungen allerdings auf der konkreten Situation der Teilaufgabe beruht, ist nicht auszumachen: Das »wieder« scheint darauf hinzudeuten, dass K33 eine ähnliche Vorstellung schon in einem anderen Kontext aktiviert hat (was aus dem Transkript nicht hervorgeht); der Inhalt der Äußerung – zusammengefasst, dass Deutungen unsicher sind – hat allerdings durchaus einen starken Bezug zu der spezifischen Aufgabenstellung und der als Beispiel dienenden Diskussion über falsches Parken.

Das mehrfache Auftreten von allgemeinen Formulierungen lässt sich bei K33 (siehe voriger Absatz) und M38 durch die Antworten auf die Teilaufgabe 3-24-b nachvollziehen. Bei M38 kommt hinzu, dass sie bei 3-21-ii zwei zeitlich getrennte Äußerungen macht, wodurch auch das zweifache Auftreten von allgemeinen Formulierungen für B4 plausibilisiert werden kann.

6.5.3 Gesamtschau zur Allgemeinheit der Vorstellungen

Insgesamt lässt sich festhalten, dass sowohl für Kennzeichen von naturwissenschaftlichen Fragen als auch für die Unterscheidung von Beobachtung und Deutung die Ideen-Oberkategorien vorrangig mit Äußerungen einhergehen, die auf konkrete Fälle bezogen sind: Einschätzungen. Wenn Ideen-Oberkategorien für Äußerungen vergeben wurden, die zugleich allgemein formuliert sind, geschieht dies fast komplett bei den Aufgaben, die konkret zum Formulieren von Generalisierungen auffordern (1-22-ii und 3-21-ii). Bei diesen Aufgaben sind dann auch die Äußerungen der Lernenden vorrangig allgemein formuliert, und Einschätzungen kommen nur sehr vereinzelt vor.

Das Ergebnis der Häufung von allgemeinen Formulierungen bei den Aufgaben 1-22-ii und 3-21-ii ist vor dem Hintergrund der angewandten Kodiervorschrift zu betrachten. Die Kodiervorschrift schreibt nicht vor, dass auf die konzeptuelle Formulierung geachtet wird; vielmehr wird die im Kontext als plausibel anzusehende rekonstruierte Gesamtaussage auf ihren Allgemeinheitsanspruch hin eingeschätzt. Somit wurde

6.5 Auswertungen und Ergebnisse zur Allgemeinheit der Vorstellungen (F-Vor2)

nicht nur »Naturwissenschaftliche Fragen kann man immer faktisch belegen« (B5, Abs. 29; konzeptuell formuliert) oder »Eine Deutung ist, sobald (..) man Gründe nennt, warum das so sein könnte« (A1, Abs. 48), sondern auch »Es hat naturwissenschaftliche Eigenschaften in sich« (C8, Abs. 54, als direkte Antwort auf die Frage »was eine naturwissenschaftliche Fragestellung von einer nicht naturwissenschaftlichen Fragestellung unterscheidet«) oder »Da sind schon Erklärungen dabei.« (M38, Abs. 20, als Antwort auf die Frage »Haben Sie eine Idee, woran man erkennt, dass es sich um eine Deutung handelt?«) als allgemeine Formulierung kodiert. Es scheint so, dass Sätze der zweiten Art für Kennzeichen von naturwissenschaftlichen Fragen deutlich häufiger auftreten – für die Unterscheidung von Beobachtung und Deutung war es schwierig, überhaupt Beispiele für die zweite Art zu finden (und das gegebene Beispiel passt auch nur bedingt, weil es nahezu konzeptuell wäre, wäre es ein vollständiger Satz). Dass bei den beiden Aufgaben also fast nur allgemeine Formulierungen auftauchen, ist demnach nicht so zu verstehen, dass die Schülerinnen und Schüler alle Vorstellungen in konzeptueller Form äußern, sondern so, dass sie ihre Vorstellungen im Kontext der beiden Aufgaben nicht verbal auf konkrete Situationen beziehen.

6.6 Auswertungen und Ergebnisse zur Angemessenheit der Vorstellungen

Die dritte Forschungsfrage fokussiert auf die Angemessenheit der Vorstellungen der Lernenden:

F-Vor3 Wie angemessen sind die Vorstellungen der Lernenden zum experimentbezogenen Denken und Arbeiten?

Die Angemessenheit der generierten Ideen-Oberkategorien, die als Vorstellungen der Lernenden interpretiert werden, wurde vorrangig anhand ihrer Beschreibungen (siehe Tabelle 6.10 auf S. 284) eingeschätzt (genauer in Unterabschnitt 6.1.8 ab S. 266).

6.6.1 Angemessenheit der Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen

Die Einschätzungen der einzelnen Ideen-Oberkategorien für Kennzeichen von naturwissenschaftlichen Fragen finden sich in Tabelle 6.16. Die Urteile zur Angemessenheit werden für Kennzeichen von naturwissenschaftlichen Fragen exemplarisch ausführlicher dargestellt: Um darzustellen, welche Faktoren die wertende Einschätzung beeinflussten, werden zunächst für einen Großteil der Ideen-Oberkategorien Gründe für die vorgenommenen Einschätzungen diskutiert. Anschließend werden die Ergebnisse in der Zusammenschau diskutiert. Im Einzelnen ist Folgendes anzumerken:

- Für **Empfindungen** wurde *deutlich angemessen* vergeben. Einschränkung ist allerdings anzumerken, dass der Inhalt dieser Vorstellung (ob es in einer Frage um Gefühle, Empfindung oder Geschmack geht) zur Klassifizierung von naturwissenschaftlichen Fragen nicht definitorisch ist. Das heißt, dass es zwar genügt, dass eine Frage nach Empfindungen, Gefühlen oder Geschmack fragt, um sie als nicht-naturwissenschaftlich einzuschätzen, aber es nicht genügt, dass eine Frage nichts mit Empfindungen, Gefühlen oder Geschmack zu tun hat, um sie als naturwissenschaftlich zu bezeichnen. Beispielsweise ist die Frage »Verstehen Studierende diese Dissertation?« nicht auf Empfindungen bezogen, aber sie ist keineswegs naturwissenschaftlich. Die zur Ideen-Oberkategorie **Empfindungen** gehörigen Äußerungen der Lernenden wurden jedoch deshalb als deutlich angemessen eingeschätzt, weil entweder eine weitere Ideen-Oberkategorie zugewiesen wurde, die als Ergänzung dient, oder sich die Äußerung auf konkrete Fälle bezieht, in denen **Empfindungen** als Ausschlusskriterium dienen kann.

6.6 Auswertungen und Ergebnisse zur Angemessenheit der Vorstellungen (F-Vor3)

Tabelle 6.16: Einschätzung der Angemessenheit der Ideen-Oberkategorien zu Kennzeichen von naturwissenschaftlichen Fragen.

Ideen-Oberkategorie	Angemessen
Empfindungen	Deutlich ja
Persönliches	Eher ja
Subjektiv	Eher ja
Objektiv	Eher ja
MethodeSpezifisch	Deutlich ja
MethodeUnspezifisch	Eher ja
NawiEigenschaften	Eher nein
Zusammenhang	Eher ja
Effekt	Deutlich nein
Prüfen	Eher nein
Belegen	Eher nein
BeweisenNachweisen	Eher nein
AntwortAusschlag	Deutlich nein
PF/AF	Deutlich nein
Fachbezug	Eher nein
Klang	Deutlich nein

- Die Ideen-Oberkategorien Subjektiv und Objektiv wurden als *eher angemessen* eingeschätzt. Zum einen wurde Objektiv auch für den Text der Informationskarte 1-23 vergeben. Zum anderen steht auf der Karte »objektiv messbar«, was nicht den typischen Äußerungen der Schülerinnen und Schüler entspricht, die davon sprechen, dass die Frage objektiv ist (z. B. »Es ist keine naturwissenschaftliche Fragestellung, weil die Frage nicht objektiv ist oder so was«; G21, Abs. 102). Insgesamt sind die beiden Ideen-Oberkategorien allerdings häufig Ausdruck von angemessenen Vorstellungen, dass naturwissenschaftliche Fragen unabhängig von spezifischen Personen untersuchbar sein müssen.
- Für MethodeSpezifisch wurde das Urteil *deutlich angemessen* vergeben. Das gewählte Urteil ist im Einklang mit der Explizierung auf der Informationskarte 1-23 (Abbildung C.2 auf S. 542), bei der verschiedene spezifische Methoden aufgezählt werden (mehr dazu bei F-Vor6). Einschränkend ist anzumerken, dass diese Ideen-Oberkategorie u. a. die Ideen-Kategorie +NawiMethode umfasst, welche Äußerungen zugewiesen wurde, denen die Methode, mithilfe derer die Frage beantwortet werden allgemein als naturwissenschaftliche Methode bezeichnen. Damit geben die Äußerungen (und demnach die Ideen-Kategorie und z. T. eben auch die Ideen-Oberkategorie) nur wenig Aufschluss über die konkrete Vorstellung (und damit die tatsächliche Angemessenheit der Vorstellung) der Lernenden.

- Wie bei den Ergebnissen zu F-Vor1 erwähnt, weist Effekt deutliche Ähnlichkeit zu dokumentierten problematischen Vorstellungen auf und wurde daher als *deutlich unangemessen* beurteilt. Die Ideen-Oberkategorie PF/AF bezieht sich, wie ebenfalls zu F-Vor1 angemerkt, auf eine Unterscheidung aus der Instruktion. Ob eine Frage die unabhängige und die abhängige(n) Variable(n) präzise benennt, ist aber kein Unterscheidungskriterium für die Naturwissenschaftlichkeit. Die Ideen-Oberkategorie AntwortAusschlag bezieht sich vorrangig auf Ja-Nein-Antworten, die zwar bei präzise formulierten Fragen einen zentralen Antworttyp darstellen (neben Zahlen, Formeln für die gefundenen Zusammenhänge u. Ä.), aber im Kontext der Unterscheidung von naturwissenschaftlichen und nicht-naturwissenschaftlichen Fragen kein angemessenes Kriterium sind – daher die Einschätzung als *deutlich unangemessen*.
- In Bezug auf die mit der Ideen-Oberkategorie BeweisenNachweisen verbundenen Vorstellung, dass die Antwort auf eine naturwissenschaftliche Frage bewiesen oder nachgewiesen werden kann, ist anzumerken, dass in der Literatur der Gebrauch des Wortes *beweisen* im naturwissenschaftlichen Sprachregister teilweise kritisch gesehen wird (vgl. z. B. Hofheinz, 2008; Perla & Carifio, 2008; anders Werner & Kremer, 2010). Die Funktion des Wortes im Sprachgebrauch der Lernenden scheint in vielen Fällen dem Wort *nachweisen* nahezukommen (welches in einigen anderen Äußerungen verwendet wird). Die vermutlich mit all diesen Fällen verbundene Vorstellung ist im allgemein als ähnlich brauchbar einzuschätzen wie die mit den Ideen-Oberkategorien Prüfen und Belegen verbundenen Vorstellungen. Daher wurde auch BeweisenNachweisen als *eher unangemessen* (und nicht als *deutlich unangemessen*) beurteilt.
- Zur Ideen-Oberkategorie Fachbezug ist anzumerken, dass in Übereinstimmung mit dem Text der Aufgabenkarte 1-21 – »Nicht jede Fragestellung, die etwas mit Physik, Chemie oder Biologie zu tun hat, ist automatisch eine naturwissenschaftliche Fragestellung« – das Urteil *eher unangemessen* vergeben wurde. Gleichzeitig ist grundsätzlich einleuchtend und auch im Text der Aufgabenkarte auffindbar, aber nicht in den Kodierungen abgebildet, dass eine Zuordnung zu einer Fachdisziplin – für die betrachteten Äußerungen finden sich beispielsweise Ideen-Kategorien zu +HatMitPhysikZuTun, +h.m.Chemie.z.t., +h.m.Psychologie.z.t. und +h.m.Pädagogik.z.t. – in den meisten Fällen eine angemessene Heuristik ist, die durch weitere Kriterien ergänzt werden sollte. So kann beispielsweise eine gegebene Beispielfrage nach Überprüfung darauf, ob sie pädagogische Inhalte thematisiert, gegebenenfalls bereits als nicht-naturwissenschaftlich eingeschätzt

werden (wie durch die Ideen-Kategorie +h.m.Pädagogik.z.t. suggeriert). Wenn Lernende also die Ideen-Oberkategorie **Fachbezug** mit angemessenen anderen Ideen-Oberkategorien kombinieren könnte dies also durchaus als angemessenes Muster von Ideen-Oberkategorien angesehen werden. Obwohl jede Ideen-Oberkategorie für sich genommen eingeschätzt wurde, wurde geprüft, ob ein deartiges Muster der Kombination von **Fachbezug** mit anderen Ideen-Oberkategorien vorliegt: In den vorliegenden Daten tritt **Fachbezug** häufig alleine auf (siehe Tabelle 6.11). Die Beurteilung als *eher unangemessen* ist daher nicht nur vor dem Hintergrund des Instruktionsmaterials (siehe Anfang des Absatzes), sondern auch vor dem Hintergrund der tatsächlich vorliegenden Daten über die Lernenden begründet.

Insgesamt lässt sich zusätzlich zur Einschätzung der Ideen-Oberkategorien festhalten, dass die Angemessenheit der *Äußerungen* der Lernenden nicht notwendigerweise mit der für die entsprechende Ideen-Oberkategorie zugewiesenen Angemessenheit übereinstimmen muss. Zum einen können eher angemessene Ideen-Oberkategorien auch eher unangemessene Ideen-Kategorien beinhalten (siehe oben). Zum anderen können aber auch Äußerungen, die grundsätzlich eher unangemessen sind, in spezifischen Aufgabenkontexten als angemessene Äußerungen gelten. Beispielsweise sagt H23 (Abs. 153): »Hm, kann man Vollmond messen? (.) Ich meine Astronomie ist doch auch eine Naturwissenschaft.« Der hintere Teil der Aussage ist der Ideen-Oberkategorie **Fachbezug** zugehörig, die grundsätzlich als eher unangemessen eingeschätzt wurde. Hier nutzt H23 allerdings die Fachzugehörigkeit als einen Hinweis darauf, ob in der Frage benannte Größen als messbar angesehen werden können, was als eher angemessen eingeschätzt wird.

In der Zusammenschau ist deutlich, dass Lernende in ihren Vorstellungen nicht nur inhaltlich große Varianz zeigen, sondern auch hinsichtlich der Angemessenheit ein breites Spektrum an Vorstellungen vorhanden ist. Daher kann davon ausgegangen werden, dass beispielsweise auch im unterrichtlichen Kontext keine Positivauswahl an Vorstellungen vorliegt und mit einem großen Teil unangemessener Vorstellungen von Lernenden umgegangen werden muss.

Als weiterer Faktor bei der Einschätzung von Gesamtaussagen zur Angemessenheit ist zu berücksichtigen, dass die Urteile als zusammenfassende Urteile von Ideen-Oberkategorien notwendigerweise oberflächlich und daher zusätzlich mit Vorsicht zu betrachten sind. Dennoch können sie erste Näherungen zur Angemessenheit der Vorstellungen von Lernenden ermöglichen. Für F-Vor7 (Auswertungen ab S. 397) werden die Angemessenheiten der Ideen-Oberkategorien für jede der Personen bzw.

für jedes der Teams einzeln aufgelöst. In den folgenden Zeilen wird eine Betrachtung der gesamten Kohorte berichtet:

- Den Urteilen werden gemäß Unterabschnitt 6.1.8 Zahlen zugewiesen:
 -2 = deutlich unangemessen, -1 = eher unangemessen, 1 = eher angemessen, 2 = deutlich angemessen.
- Alle Ideen-Oberkategorien treten für mehrere Personen und für mehrere Teams auf (siehe Tabelle 6.10). Wird über alle Ideen-Oberkategorien gemittelt, ergibt sich für die Angemessenheit ein Mittelwert von -0.25 mit einer Standardabweichung von 1.44 . In der Tendenz sind die Vorstellungen also eher unangemessen. Aufgrund der großen Streuung ist es allerdings genauso plausibel, davon auszugehen, dass die Vorstellungen im Mittel weder angemessen noch unangemessen sind.
- Die Ideen-Oberkategorien treten für unterschiedlich viele Personen bzw. Teams auf (siehe Tabelle 6.10). Wird dies berücksichtigt und ein nach Personen- bzw. Teamzahl gewichteter Mittelwert gebildet,⁹³ ergeben sich für die Angemessenheit ein Mittelwert von 0.02 mit einer Standardabweichung von 0.51 bei Gewichtung nach Personen und ein Mittelwert von -0.08 mit einer Standardabweichung von 0.81 bei Gewichtung nach Teams. Diese Werte stützen die Deutung, dass die Angemessenheit der Vorstellungen breit gestreut um Null – den Wert der unklaren Angemessenheit herum – ist.⁹⁴

Die Standardabweichungen zu den gewichteten Maßen sind schwer zu interpretieren, weil sie neben der (stark) variierenden Angemessenheit auch abbilden, dass die Ideen-Oberkategorien (mäßig stark variierend) verschieden vielen Personen bzw. Teams zugewiesen sind und nur der Vollständigkeit halber angegeben. Insgesamt verdeutlichen die Werte die bereits oben gegebene Deutung, dass die Angemessenheit der Vorstellungen ein großes Spektrum abdeckt.

⁹³Der gewichtete Mittelwert wird aus der Angemessenheit der jeweiligen Ideen-Oberkategorie, der Anzahl der Personen mit der jeweiligen Ideen-Oberkategorie und der Anzahl der insgesamt analysierten Personen bzw. Teams folgendermaßen berechnet:

$$\frac{\text{Angemessenheit der jeweiligen Ideen-Oberkat.} \times \text{Anzahl der P/T zur jeweiligen Ideen-Oberkat.}}{\text{Anzahl der analysierten Personen bzw. Teams}}$$

⁹⁴Analog zu den Aufzählungspunkten werden im späteren Verlauf der Arbeit auf Eben der Personen bzw. Teams ebenfalls ein ungewichteter und ein gewichteter Mittelwert gebildet. Der gewichtete Mittelwert wichtet anhand der Anzahl der (Teil-)Aufgaben, für die der jeweiligen Person bzw. für die dem jeweiligen Teams die entsprechenden Ideen-Oberkategorien zugewiesen sind. Eine Wichtung anhand der (Teil-)Aufgaben wäre auch auf Ebene der Kohorten denkbar, wird allerdings als am wenigsten aufschlussreiche Wichtung nicht vorgenommen.

6.6.2 Angemessenheit der Vorstellungen zur Unterscheidung von Beobachtung und Deutung

Die Vorstellungen der Lernenden zur Unterscheidung von Beobachtung und Deutung sind grundsätzlich eher angemessen, was der formulierten Hypothese entspricht. Die Parallelität der Ideen-Oberkategorien mit den bei Petermann (2017) dokumentierten Kategorien ist hilfreich, weil die dort vorgenommenen Einschätzungen in der Mehrheit übernommen werden können. Daher ergibt sich, dass die meisten Ideen-Oberkategorien als eher angemessen oder deutlich angemessen eingeschätzt wurden, wobei die Trennung zwischen *eher angemessen* und *deutlich angemessen* sich für die Unterscheidung von Beobachtung und Deutung als sehr unscharf erwiesen hat und daher in der Arbeit nicht dargestellt ist. Die Ideen-Oberkategorie **FaktenSicher** ist nicht angemessen, sondern stellt eine dokumentierte problematische Vorstellung dar (siehe bei F-Vor1 ab S. 283). Daher wird das Urteil *deutlich unangemessen* vergeben. Als *eher unangemessen* sind die Ideen-Oberkategorien **Beobachten** (weil eventuell tautologisch, siehe bei F-Vor1) und **Meinung** (weil Deutungen nicht immer Meinungen und auch Meinungen nicht immer Deutungen sein müssen) anzusehen. Bei Petermann (2017) wird insbesondere die Anknüpfbarkeit der Vorstellungen herausgestellt, die auch hier für fast alle Vorstellungen gegeben zu sein scheint.

Analog zum Vorgehen für Kennzeichen von naturwissenschaftlichen Fragen werden auch für die Unterscheidung von Beobachtung und Deutung ein ungewichteter und ein gewichteter Mittelwert zur Angemessenheit gebildet. Es ergibt sich unter der Setzung des Wertes für angemessene Ideen-Oberkategorien auf 1.5 (da nicht zwischen *eher* und *deutlich angemessen* unterschieden wird, siehe oben) und unter Ausschluss der Ideen-Oberkategorie **Sonstiges**:

- Als Mittelwert über die Angemessenheit aller Ideen-Oberkategorien ergibt sich 0.72 mit einer mit Standardabweichung von 1.28.
- Wird anhand der Personen gewichtet, ergibt sich ein Mittelwert von 0.18 mit einer Standardabweichung von 0.35; wird anhand von Teams gewichtet, ergibt sich ein Mittelwert von 0.27 mit einer Standardabweichung von 0.61.

Die Standardabweichungen zu den gewichteten Maßen sind schwer zu interpretieren, weil sie neben der (gering) variierenden Angemessenheit auch abbilden, dass die Ideen-Oberkategorien (stark variierend) verschieden vielen Personen bzw. Teams zugewiesen sind und nur der Vollständigkeit halber angegeben. Alle Werte zusammengefasst verdeutlichen die positive Tendenz der Gesamteinschätzung und die Breite der verschiedenen Vorstellungen zur Unterscheidung von Beobachtung und Deutung.

6.7 Auswertungen und Ergebnisse zu Vorstellungen und zur Korrektheit von Entscheidungen

Die vierte Forschungsfrage setzt die Vorstellungen der Lernenden und ihre Entscheidungen im Prozess des experimentbezogenen Denkens und Arbeitens miteinander in Beziehung:

F-Vor4 Mit welchen Vorstellungen gehen korrekte und fehlerhafte Entscheidungen zum experimentbezogenen Denken und Arbeiten einher?

Äußerungen, die sowohl eine Einschätzung als auch einen Hinweis auf eine Vorstellung enthalten, sind sowohl mit einer Kodierung zur Korrektheit der Einschätzung (*korrekt*, *inkorrekt*, *unklar*, siehe Unterabschnitt 6.1.5 auf S. 259) als auch einer Ideen-Oberkategorie versehen. Für die zugewiesenen Ideen-Oberkategorien wurde zudem die Angemessenheit eingeschätzt (siehe Unterabschnitt 6.1.8 auf S. 266). Die entsprechend vorliegenden Daten zur Korrektheit der Einschätzung und zur Angemessenheit der Ideen-Oberkategorie können miteinander in Beziehung gesetzt werden.

Es kann beispielsweise die Äußerung »Klar, du kannst es irgendwie beweisen oder nicht« (Abs. 180) von C8 betrachtet werden. Diese Äußerung stellt eine Einschätzung der Beispielfrage »Gibt es einen Zusammenhang zwischen der Farbe einer Flüssigkeit und ihrer Dichte?« (1-25-A, siehe Abbildung C.3 auf S. 543) daraufhin, ob sie eine naturwissenschaftliche Fragestellung ist, dar. Die Beispielfrage wird von der Person korrekterweise als naturwissenschaftliche Frage eingeschätzt (»Klar«⁹⁵). Gleichzeitig geht aus der Äußerung eine Vorstellung zu Kennzeichen von naturwissenschaftlichen Fragen hervor, die mit der Ideen-Oberkategorie **Beweisen** kodiert wurde. Diese Ideen-Oberkategorie ist eher unangemessen. In diesem Fall geht eine korrekte Einschätzung also mit einer eher unangemessenen Ideen-Oberkategorie einher.

Nachfolgend wird einerseits danach gefragt, ob das Aktivieren angemessener Vorstellungen eher im direkten zeitlichen und inhaltlichen Kontext von korrekten Entscheidungen stattfindet, andererseits danach, ob Lernende mit in der Gesamttendenz eher angemessenen Vorstellungen auch insgesamt eher korrekte Entscheidungen treffen. Beide Fragen stellen verschiedene Herangehensweisen an einen potentiellen Zusammenhang von Fähigkeiten und Verständnissen (oder von dem Aufbau von Fähigkeiten und dem Aufbau von Verständnissen) dar.

Bevor die spezifischen Analysen zu Kennzeichen von naturwissenschaftlichen Fragen (ausführlich) und zur Unterscheidung von Beobachtung und Deutung (weniger ausführ-

⁹⁵Erinnerung: Nicht immer ist eindeutig entscheidbar, ob ein Ja/Klar oder ein Nein sich auf die Einschätzung der Beispielfrage bezieht oder eine Antwort auf die Beispielfrage darstellt.

lich) dargestellt werden, wird ein Ergebnis berichtet, das sich als Nebenprodukt der Analysen finden ließ, aber relevant für die Interpretation der anderen Ergebnisse ist:

Korrektes Entscheiden und Verbalisieren von Vorstellungen scheinen zwei eher voneinander unabhängige Vorgänge zu sein. Eine Betrachtung der Transkripte zeigt, dass die Teile der Äußerungen, die die Hinweise auf die zugehörigen Ideen-Oberkategorien enthalten, meist auf die Teile der Äußerungen folgen, in denen die Entscheidungen verbalisiert wird. Beispielsweise wird eher »Ne, ist keine [naturwissenschaftliche Frage], weil es persönlich ist« gesagt als »Hier geht es um persönliche Meinungen, also ist es keine [naturwissenschaftliche Frage]«. Ferner gibt es viele Fälle, in denen im Material erst in einem zweiten Schritt nach einer Begründung gefragt wird (bei 1-25 und 1-26, siehe Abbildung C.3 im Anhang, sind z. B. dem Lesefluss nach zuerst Kreuze zur Entscheidung zu setzen und dann nachgelagert Begründungen für die Entscheidung zu verschriftlichen); in diesen Fällen ist die Entscheidung also schon getroffen, bevor die Lernenden überhaupt eine Verbalisierung treffen, anhand derer auf ihre Vorstellungen geschlossen werden kann. Zwar ist jede Verbalisierung als einzelne Entscheidung anzusehen (und auch so kodiert), weil die Möglichkeit zum Umentscheiden unter Berücksichtigung der Begründung besteht (und die Lernenden real auch Gebrauch davon machen), aber es bleibt ungeklärt, inwiefern die Prozesse des Entscheidens und des Verbalisierens von Vorstellungen einander beeinflussen, und ist vermutlich plausibel, dass in vielen Fällen die Prozesse unverbunden bleiben. Zuletzt lässt sich herausstellen, dass bei Aufgaben, in denen nicht explizit zum Begründen aufgefordert wird, auch weniger Ideen-Oberkategorien kodiert sind als bei den anderen Aufgaben. Für solche Aufgaben wird allerdings nicht grundsätzlich schlechter entschieden (siehe weiter unten in Unterunterabschnitt 6.7.1.2).

Insgesamt sprechen die genannten Ergebnisse dafür, dass die Prozesse des Entscheidens und des Verbalisierens von Gründen/Begründungen oder anderer Äußerungen, anhand derer auf die Vorstellungen der Lernenden geschlossen werden kann, relativ unabhängige Prozesse sind. Die in den theoretischen Grundlagen (Kapitel 2) vorgenommene Unterscheidung von Fähigkeiten (hier: die Fähigkeit, korrekt zwischen naturwissenschaftlichen und nicht-naturwissenschaftlichen Fragen bzw. Beobachtungen und Deutungen zu unterscheiden) und Verständnissen (hier: das Verständnis, was naturwissenschaftliche Fragen von nicht-naturwissenschaftlichen Fragen unterscheidet bzw. was Beobachtungen und Deutungen unterscheidet) als zwei zusammengehörige, aber wechselseitig unabhängig betrachtbare Facetten einer bestimmten Kompetenz erweist sich daher an dieser Stelle als hilfreich. Einschränkend ist allerdings anzumerken, dass gerade die Struktur der Aufgaben der Instruktion, in der häufig zuerst

nach Entscheidungen gefragt wird, das Ergebnis als Artefakt erzeugen könnte.

6.7.1 Korrektheit einzelner Entscheidungen und Angemessenheit zugehöriger Vorstellungen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen

Die Frage, ob das Aktivieren angemessener Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen eher im direkten zeitlichen und inhaltlichen Kontext von korrekten Entscheidungen zu diesem Themenkomplex stattfindet, wird einerseits in der Gesamtschau über alle Aufgaben des Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen hinweg, andererseits für jede einzelne Aufgabe beantwortet.

6.7.1.1 Gesamtschau für Kennzeichen von naturwissenschaftl. Fragen

Korrekte Entscheidungen treten mit größerer Häufigkeit bei Verbalisierung von Vorstellungen auf als bei fehlender Verbalisierung. Wird die Gesamtheit aller Einschätzungen – unabhängig davon, ob mit oder ohne Hinweis auf eine Idee – betrachtet, ergibt sich: Der Anteil der inkorrekten Einschätzungen für Einschätzungen ohne Verknüpfung zu Ideen-Oberkategorien (etwa 36 %) ist deutlich größer als der Anteil inkorrektur Einschätzungen für Einschätzungen mit Verknüpfung zu Ideen-Oberkategorien (etwa 5 %). In Tabelle 6.17 sind die entsprechenden absoluten Anteile dargestellt. Es ergibt sich aufgrund eines χ^2 -Tests hochsignifikant ($\chi^2(1) = 32.766$, $p < .001$; unklare Einschätzungen wurden nicht berücksichtigt) und gemessen mit dem Quotenverhältnis (engl. *odds ratio*), dass Einschätzungen, bei

Tabelle 6.17: Korrektheit der Einschätzungen und Angabe, ob Ideen-Oberkategorie bei der Einschätzung kodiert wurde, für Kennzeichen von naturwissenschaftlichen Fragen.

Ideen-OK bei Einschätzung	Korrektheit der Einschätzung			Summe
	Richtig	Falsch	Unklar	
vorhanden	275	14	10	299
nicht vorhanden	123	36	3	159
Summe	398	50	13	461

Anmerkung. In 29 Fällen wurden mehrere Ideen-Oberkategorien zusammen mit einer einzelnen Kodierung für Einschätzungen vergeben. Diese Fälle wurden aus technischen Gründen als zwei Einschätzungen ausgezählt.

denen eine Ideen-Oberkategorie kodiert wurde, 5.75-mal so häufig mit korrekten Einschätzungen einhergehen wie mit inkorrekten.

Wird davon ausgegangen, dass Entscheiden und Verbalisieren von Vorstellungen zwei unabhängige Prozesse sind (siehe oben), dann ist das Ergebnis der größeren Korrektheit bei Vorliegen von Verbalisierungen in der Gesamtschau über alle Aufgaben eventuell in folgender Weise plausibilisierbar: Es könnte sein, dass für einige der falschen Entscheidungen schlicht die Fähigkeit fehlt, überhaupt irgendwas zu verbalisieren, weil die Lernenden möglicherweise stark auf lebensweltliche Intuitionen zugreifen. Der Validität einer solchen Interpretation wurde allerdings nicht weiter nachgegangen.

Korrekte Entscheidungen gehen sowohl mit angemessenen als auch unangemessenen Vorstellungen einher. Inkorrekte Entscheidungen gehen eher mit unangemessenen Vorstellungen einher. In Tabelle 6.18 ist für Kennzeichen von naturwissenschaftlichen Fragen dargestellt, wie sich bei Äußerungen, denen sowohl eine Ideen-Oberkategorie als auch eine Korrektheit der Entscheidung zugewiesen ist, die Korrektheit der Entscheidungen zur Angemessenheit der Ideen-Oberkategorien verhält. Aus einem χ^2 -Test ergibt sich zwar, dass angemessenere Einschätzungen hoch-signifikant mit angemesseneren Ideen-Oberkategorien zusammenhängen ($\chi^2(1) = 15.716$, $p < .001$; unklare Einschätzungen wurden ausgeschlossen, auf eine Unterscheidung von deutlich vs. eher angemessenen bzw. unangemessenen Ideen-Oberkategorien wurde verzichtet), und gemäß des Quotenverhältnisses gehen angemessene Ideen-Oberkategorien zwanzigmal so häufig mit korrekten Einschätzungen einher wie mit unangemessenen. Die Bedeutung dieser Aussage kann jedoch vor

Tabelle 6.18: Korrektheit der Einschätzung sowie Angemessenheit der je Instanz zugehörigen Ideen-Oberkategorie für Kennzeichen von naturwissenschaftlichen Fragen.

Angemessenheit der Ideen-OK	Korrektheit der Einschätzung			Summe
	Richtig	Falsch	Unklar	
(eher) angemessen	167	1	3	171
(eher) unangemessen	108	13	7	128
Summe	275	14	10	299

Anmerkung. In 29 Fällen wurden mehrere Ideen-Oberkategorien zusammen mit einer einzelnen Kodierung für Einschätzungen vergeben. Diese Fälle wurden aus technischen Gründen als zwei Einschätzungen ausgezählt.

dem Hintergrund des relativ hohen Anteils korrekter Einschätzungen (inhaltlich wie statistisch) kritisch gesehen werden. Bedeutsam aus fachdidaktischer Sicht scheinen jedoch auf jeden Fall zwei direkt aus Tabelle 6.18 ersichtliche Tatsachen zu sein. Erstens gibt es insgesamt nur sehr wenige inkorrekte Einschätzungen.⁹⁶ Zweitens finden sich für die wenigen vorliegenden inkorrekten Einschätzungen vorrangig unangemessene Ideen-Oberkategorien. Andersherum gehen die unangemessenen Ideen-Oberkategorien jedoch nicht grundsätzlich mit inkorrekten Einschätzungen einher. Nur für die als deutlich unangemessen eingeschätzte Ideen-Oberkategorie **Effekt** zeigt sich in einer Auflösung der einzelnen Ideen-Oberkategorien (Tabelle F.9 im Anhang auf S. 572), dass sie typischerweise mit inkorrekten Einschätzungen einherzugehen scheint.

Ob eine Verbalisierung Hinweise auf aus fachmethodischer Sicht eher angemessene oder eher unangemessene Vorstellungen enthält, scheint im Bereich der Kennzeichen von naturwissenschaftlichen Fragen nicht oder nur in geringem Maße von Bedeutung zu sein. Ferner lassen sich die Ergebnisse auch so deuten, dass aus korrekten Einschätzungen keineswegs folgt, dass diese mit angemessenen Vorstellungen einhergehen. Wenn eine Aufgabe also im intendierten Sinne gelöst wird, kann noch lange nicht auf intendierte Vorstellungen geschlossen werden. Für inkorrekte Einschätzungen ist zwar aus den Daten ableitbar, dass sie häufig mit unangemessenen Vorstellungen einhergehen (allerdings gehen unangemessene Vorstellungen nicht vorrangig mit inkorrekten Einschätzungen einher); dennoch ist in ähnlicher Weise wie bei korrekten Einschätzungen auch bei inkorrekten Einschätzungen Vorsicht geboten, von den Einschätzungen auf eine nicht vorhandene Angemessenheit der Vorstellungen der Lernenden zu schließen.

6.7.1.2 Betrachtung der einzelnen Aufgaben des Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen

Aus einer Betrachtung der korrekten und inkorrekten Einschätzungen in Bezug auf die einzelnen Aufgaben geht hervor, dass nur bei wenigen Aufgaben überhaupt inkorrekte Einschätzungen vorgenommen wurden: 1-21-ii (2 inkorrekte Einschätzungen von insgesamt 27 Einschätzungen), 1-21-iv (13 von 21), 1-21-vi (3 von 18), 1-21-vii (1 von 18), 1-25-A (2 von 35), 1-25-B (2 von 34), 1-26-C (15 von 48) und 1-31-2 (4 von 28). Unklare Einschätzungen finden sich bei den meisten der aufgezählten Aufgaben und drei anderen Aufgaben. Dieses Ergebnis legt nahe, dass einzelne

⁹⁶Das gilt auch, wenn solche Einschätzungen hinzugenommen werden, die nicht mit Vorstellungen verknüpft sind (siehe Summenzeile von Tabelle 6.17).

Aufgaben besondere Schwierigkeiten für Lernenden hinsichtlich der Einschätzungen darstellen. Insbesondere stellt sich daran anschließend die Frage, ob die inkorrekten Einschätzungen im Kontext der beiden Aufgaben – 1-21-iv und 1-26-C – mit eher unangemessenen Ideen-Oberkategorien einhergehen.

Die Teilaufgabe 1-21-iv

»Bei welchen der folgenden Fragestellungen handelt es sich Ihrer Einschätzung nach um naturwissenschaftliche Fragestellungen? [...]
Beeinflusst die Größe eines Hundes, wie oft er im Durchschnitt pro Stunde bellt?«

wird von den Lernenden vor der Explizierung bearbeitet. Es wird nicht zum Begründen aufgefordert, weshalb es nicht verwunderlich ist, dass bei dieser Aufgabe nur wenige Ideen-Oberkategorien kodiert wurden. Nur vier Einschätzungen gehen mit Ideen-Oberkategorien einher – alle mit der Ideen-Oberkategorie **Fachbezug**, davon drei korrekte und eine inkorrekte. Obwohl die Ideen-Oberkategorie als eher unangemessen eingeschätzt wurde, finden sich dort, wo sie im Kontext der Aufgabe 1-21-iv auftaucht, vorrangig korrekte Entscheidungen, wohingegen zur Aufgabe insgesamt vorrangig inkorrekte Entscheidungen vorliegen. Im Kontext der Aufgabe 1-21-iv wird daher das in der Gesamtschau berichtete Ergebnis reproduziert, dass die bloße Verbalisierung von fachmethodischen Überlegungen unabhängig von der Angemessenheit mit korrekten Entscheidungen korreliert.

Für die Aufgabe 1-26-C, die

»Bei welchen der folgenden Fragestellungen handelt es sich Ihrer Einschätzung nach um naturwissenschaftliche Fragestellungen?
Versuchen Sie, Ihre Einschätzung kurz zu begründen. [...]
Muss man ein Ei bei Vollmond länger kochen, damit es hart wird?«

lautet und nach der Explizierung bearbeitet wird, finden sich deutlich mehr Ideen-Oberkategorien: 23 von 32 korrekten, 7 von 15 inkorrekten und 0 von 1 unklaren Einschätzung gehen mit Ideen-Oberkategorien einher. In Tabelle 6.19 (S.318) ist dargestellt, mit welchen Ideen-Oberkategorien die korrekten und inkorrekten Einschätzungen einhergehen. Im Kontext dieser Aufgabe bestätigt sich die oben formulierte Deutung, dass inkorrekte Einschätzungen zwar vorrangig mit (eher) unangemessenen Ideen-Oberkategorien einhergehen, unangemessene Ideen-Oberkategorien aber nicht grundsätzlich auch im Zusammenhang mit inkorrekten Einschätzungen auftauchen.

Tabelle 6.19: Korrektheit der Einschätzungen und zugehörige Ideen-Oberkategorien im Kontext der Aufgabe 1-26-C.

Ideen-Oberkategorie	Einschätzung		
	Korrekt	Inkorrekt	Gesamt
Fachbezug	2	0	2
NawiEigenschaften	0	0	0
MethodeSpezifisch	10	0	10
MethodeUnspezifisch	2	0	2
Prüfen	3	1	4
Belegen	0	0	0
BeweisenNachweisen	2	0	2
Objektiv	0	0	0
Subjektiv	1	0	1
Empfindungen	0	0	0
Persönlich	0	0	0
AntwortAusschlag	1	0	1
Effekt	1	5	6
Zusammenhang	0	0	0
Klang	0	1	1
PF/AF	1	0	1
Summe	23	7	30

Allerdings finden sich für die als deutlich unangemessenen Ideen-Oberkategorien Klang und Effekt (fast) nur inkorrekte Einschätzungen.⁹⁷

Eine Betrachtung der Aufgaben, bei denen inkorrekte Entscheidungen getroffen wurden, zeigt, dass ein Großteil der dort gegebenen Beispielfragen Zusammenhänge thematisiert, die als unplausibel angesehen werden können. So ist etwa nicht grund-

⁹⁷Eine einzige Äußerung wird als korrekte Einschätzung kodiert und ist zugleich mit der Ideen-Oberkategorie **Effekt** versehen. Sie lautet: »Weil man sich ja erstmal denkt, das hat gar nichts damit zu tun – aber das hat ja nichts mit der Frage zu tun.« Die Kodierung dieser Äußerung ist deutlich deutungsbehafteter als die in der Arbeit vorgenommene Kodierung anderer Äußerungen. Unabhängig davon, ob eine Kodierung mit **Effekt** geteilt wird oder nicht, lässt sich aber feststellen, dass die Formulierung »Weil man sich ja erstmal denkt« als eine selbstreflexive Aussage angesehen werden kann, die eine Distanz zur instantanen eigenen Vorstellung aufbauen soll. In der Kategorie *Formulierung* wurde daher auch nn vergeben. Die korrekte Einschätzung geht also vermutlich deshalb mit der Aussage einher, weil die Aussage im Prinzip lautet, dass die Vorstellung, der Effekt spiele eine Rolle für die Naturwissenschaftlichkeit der Frage, eine unsinnige Vorstellung ist. Die korrekte Einschätzung geht also mit der Ablehnung der unangemessenen Vorstellung einher.

sätzlich anzunehmen, dass die Größe eines Hundes einen Einfluss auf die Häufigkeit seines Bellens hat (1-21-iv, siehe oben). Auch der Einfluss der Mondphase auf die benötigte Zeit zum Kochen eines Eies scheint der alltäglichen Erfahrung gemäß nicht gegeben zu sein (1-26-C, siehe oben). Aufgrund der fehlenden Plausibilität der empirischen Zusammenhänge ist es einsichtig, dass im Kontext der Aufgaben mit gehäuften inkorrekten Entscheidungen auch vielfach die Ideen-Oberkategorie Effekt vorkommt, weil die Lernenden im Kontext dieser Aufgaben den (hier nicht vorhandenen) Effekt fälschlicherweise zu einem Kennzeichen naturwissenschaftlicher Fragen deklarieren. Solch ein Deklarieren geschieht, wie weiter oben erwähnt, für gewöhnlich nach der Entscheidung, was der Aufgabenstruktur entspricht (zunächst soll ein Kreuz gesetzt werden, dann soll begründet werden).

Daran, dass bei Aufgaben mit vielen inkorrekten Einschätzungen häufig Zusammenhänge thematisiert werden, die als unplausibel angesehen werden dürften, schließt gleichermaßen die Frage an, inwiefern die Lernenden über die Zusammenhänge selbst – also die fachinhaltlichen Aspekte der Zusammenhänge – sprechen. Darauf bezieht sich der nachfolgende Unterunterabschnitt.

6.7.1.3 Entscheidungen und fachinhaltliche Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen

Im Fokus der Arbeit stehen das experimentbezogene Denken und Arbeiten und die zugehörigen Vorstellungen. Analysen zu den Einschätzungen verschiedener Beispielfragen aus dem Lernmaterial führen allerdings unweigerlich auch zur Aktivierung von fachinhaltlichem Wissen bzw. von fachinhaltlichen Vorstellungen. Da eine ausführliche Analyse der verschiedenen fachinhaltlichen Vorstellungen nicht vorgenommen wurde, bleiben die Bemerkungen an dieser Stelle darauf beschränkt, ob und wie viel fachinhaltliche Vorstellungen im Kontext der Einschätzung von Beispielfragen auftreten. Dafür wird die Kategorie FachinhaltlicheÜberlegung aus dem Teil des Kodiersystems zum Modus genutzt.

Insbesondere stellt sich die Frage, ob und wie inkorrekte Entscheidungen mit dem Rückgriff auf fachinhaltliche Vorstellungen verbunden sind. Die Bedeutsamkeit dieser Frage kann auch dadurch verdeutlicht werden, dass bei der Aktivierung von *fachmethodischen* Vorstellungen die Angemessenheit nur eine untergeordnete Rolle für die Korrektheit der Entscheidungen zu haben scheint.

Im Anhang in Tabelle F.10 (auf S. 573) ist dargestellt, wie viele richtige, falsche und unklare Einschätzungen bei den Teilaufgaben auftreten. Außerdem ist aufgeführt, wie häufig die Kategorie FachinhaltlicheÜberlegung bei der jeweiligen Teilaufgabe vergeben

wurde. Es ist ersichtlich, dass bei Aufgaben mit vielen inkorrekten Einschätzungen auch gehäuft fachinhaltliche Beiträge aufzufinden sind. Eine mögliche Deutung ist, dass diese Aufgaben ein besonderes fachinhaltliches Klärungspotential aufweisen, welches die fachmethodische Einschätzung erschwert. So könnten etwa Begriffe oder Variablen unbekannt sein und zunächst fachinhaltlich geklärt werden. Von der Konzentration auf eine fachmethodische Einschätzung könnte aber auch das für Lernende unerwartete Zusammenbringen zweier Variablen ablenken, weil sie darauf zunächst fachinhaltlich reagieren wollen. Die zweite Deutung würde davon unterstützt, dass alle Beispiele als Fragen formuliert sind. Daher provozieren sie auf einer basalen Ebene zunächst (inhaltliche) Antworten, bevor Lernende den Gesamtkontext der Aufgabenkarte (wieder) in den Fokus nehmen. Das wird auch in der Betrachtung der Videos anhand instantaner inhaltlicher Antworten auf die Frage deutlich (bspw. sagt A2 »Nein. Aber man kann es ja messen, ne, (lachend) und beobachten, irgendwie«, Abs. 130, wobei das Nein (u. a. aufgrund der Intonation) vermutlich als inhaltliche Antwort gemeint ist und danach sofort auf eine fachmethodische Ebene gewechselt wird).

Die Ausführungen legen auch nahe, dass die Aufgaben vermutlich in den wenigsten Fällen missverstanden wurden. Vielmehr scheint es plausibel, dass die Lernenden die Aufgabe grundsätzlich so verstehen, dass eine Einschätzung, ob die Frage eine naturwissenschaftliche Frage ist, gefordert ist. Sie scheinen aber im Prozess diesen Fokus bewusst oder unbewusst zu verlassen, um fachinhaltliche Aspekte anzusprechen. Eine spannende Anschlussfrage wäre, ob Lernende mit eher inhaltlichen Interessen auch stärker bzw. Lernende mit eher methodischen Interessen weniger stark dazu neigen, von dem durch die explizit-fachmethodische Instruktion vorgegebenen fachmethodischen Fokus abzuweichen. Mithilfe der implizit-fachmethodischen Instruktion könnte dann auch gefragt werden, ob Personen mit unterschiedlichen Interessen auch eher oder eher nicht zu fachmethodischen Beiträgen im Kontext dieser Instruktion neigen.

6.7.2 Vorstellungen und Entscheidungen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen auf Ebene der Personen und auf Ebene der Teams

Es stellt sich auch die Frage, ob die insgesamt für eine Person rekonstruierten Vorstellungen und die im Mittel erreichte Korrektheit der Entscheidungen in einem Zusammenhang stehen. Dies ist insbesondere daher relevant, weil die jeweils einzelne Entscheidung scheinbar nachträglich begründet wird (siehe oben), aber beispielsweise der insgesamt vorliegende Rahmen an Vorstellungen hinter ad hoc getroffenen Entscheidungen stehen könnte. Um möglichst offen zu halten, ob mentale Konstrukte innerhalb einzelner Lernender vorliegen, wird auch auf Ebene der Teams nach Zusam-

menhängen zwischen insgesamt rekonstruierten Vorstellungen und im Mittel erreichter Korrektheit gefragt – denn ggf. treffen einzelne Lernende eher korrekte Entscheidungen, wenn in ihrem Umfeld viele angemessene Vorstellungen verbalisiert werden.⁹⁸

Eine Analyse unter Berücksichtigung der einzelnen Personen ist in Abbildung 6.15 dargestellt. Für jede einzelne Person ergibt sich dabei der Wert für die Angemessenheit der Ideen-Oberkategorien als Summe der Angemessenheiten ($-2 =$ deutlich unangemessen, $-1 =$ eher unangemessen usw.) aller Instanzen der jeweiligen Person, bei denen Ideen-Oberkategorien vergeben wurden, im Verhältnis zur Anzahl der Instanzen. Beispielsweise wurde für die fünf Äußerungen der Person M39 2-mal die Ideen-Oberkategorie **Effekt** mit der Angemessenheit -2 vergeben, 2-mal Ideen-Oberkategorien mit der Angemessenheit -1 (einmal **Fachbezug**, einmal **Beweisen-Nachweisen**), 0-mal Ideen-Oberkategorien mit Angemessenheit 1 und 1-mal die Ideen-Oberkategorie **MethodeSpezifisch** mit Angemessenheit 2. Dann ergibt sich der Wert für die Angemessenheit bei M39 als Verhältnis von $2 \cdot (-2) + 2 \cdot (-1) + 0 \cdot 1 + 1 \cdot 2 = -4$ und der Anzahl der Äußerungen (fünf) zu $-4/5 = -0.8$. Der Wert für die Korrektheit der Einschätzungen ergibt sich für jedes Individuum als Summe der Korrektheiten

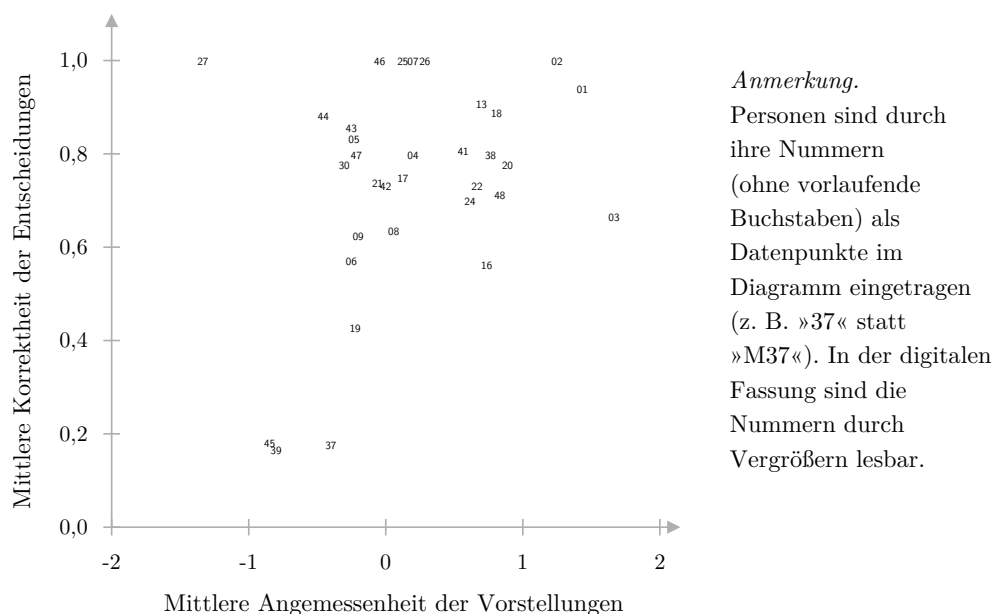


Abbildung 6.15: Streudiagramm zur Korrektheit der Einschätzungen und Angemessenheit der Ideen-Oberkategorien für jede einzelne Person ($N = 33$).

⁹⁸Genau genommen könnte auch danach gefragt werden, ob einzelne Personen, die sich in Teams mit hoher Angemessenheit der Vorstellungen befinden, im Mittel korrekter entscheiden, sowie danach, ob einzelne Personen, die eine hohe Angemessenheit der Vorstellungen aufweisen sich eher in Teams befinden, die im Mittel korrekter entscheiden. In ähnlicher Weise wurde bspw. in Kapitel 5 vorgegangen. Derart feingliedrige Analysen sind aber dem Gegenstand des vorliegenden Kapitels und der verwendeten Methode nicht optimal angemessen und sind daher nicht berichtet.

von allen getroffenen Einschätzungen (-1 = unkorrekt, 0 = unklar, 1 = korrekt) geteilt durch die Anzahl aller Einschätzungen. Die genauen Werte für die Personen, die Abbildung 6.15 zugrunde liegen, und zugehörige Zwischenergebnisse der Berechnungen finden sich im Anhang in den Tabellen F.11 (S. 574) und F.12 (S. 575).

Grafisch wird deutlich, dass es mit I27 (ganz oben links) einen Ausreißer gibt. Diese Person hat nur drei verschiedene zugewiesene Ideen-Oberkategorien und damit am wenigsten von allen Lernenden (Mittelwert der Ideen-Oberkategorien pro Person: 15.09, Standardabweichung: 7.74). Nur eine der vier vorgenommenen Einschätzungen geht mit einer Einschätzung mit Hinweis einher (kodiert mit der eher unangemessenen Ideen-Oberkategorie Prüfen). Ein Aussortieren scheint daher eine Berechtigung zu haben.

Eine Korrelationsanalyse (unter Ausschluss von I27) ergibt einen signifikanten Zusammenhang der personenweise berechneten Angemessenheit aller Instanzen mit Ideen-Oberkategorien und der ebenfalls personenweise berechneten Korrektheit aller Einschätzungen ($r = .462$, $p = .008$). Es liegt demnach ein positiver Zusammenhang von Angemessenheit der Ideen-Oberkategorien einer Person und Korrektheit der Einschätzungen derselben Person vor. Gleichzeitig sind allerdings auch die in Abbildung 6.15 ersichtliche Häufung von Personen um den Null-Wert für die Angemessenheit sowie die insgesamt eher hohe Korrektheit als Einschränkungen dieser Deutung zu erwähnen.

In Abbildung 6.16 ist für jedes Team die Korrektheit der Einschätzungen sowie die Angemessenheit der Ideen-Oberkategorien dargestellt. Eine Korrelationsanalyse nach Pearson zeigt einen vorsignifikanten Zusammenhang von Korrektheit der Einschätzungen und Angemessenheit der Ideen-Oberkategorien ($r = .443$, $p = .199$, unter Ausschluss von Team M, in der Abb. unten mittig). Aufgrund der geringen Fallzahl ist hier nicht von einer Signifikanz auszugehen, so dass die Werte gemeinsam mit der grafischen Betrachtung für eine Interpretation im Sinne eines Zusammenhangs auszureichen scheinen. Detaillierte Betrachtungen und Diskussionen zur Position einzelner Teams im Gesamtbild sind in Anhang F.4.1 (ab S. 563) aufgeführt.

Die Analysen zum Zusammenhang der Korrektheit von Entscheidungen mit der Angemessenheit von Vorstellungen lassen sich auf Ebene der Personen und Teams so zusammenfassen, dass eine positive Korrelation vorliegt. Im Gegensatz zu dem Ergebnis, dass einzelne Entscheidungen nicht grundsätzlich korrekter sind, wenn sie mit Verbalisierungen angemessenerer Vorstellungen einhergehen, scheinen also Personen und Teams, die eher angemessene Entscheidungen treffen auch eher angemessene Vorstellungen zu haben. Die Art des Zusammenhangs ist allerdings anhand

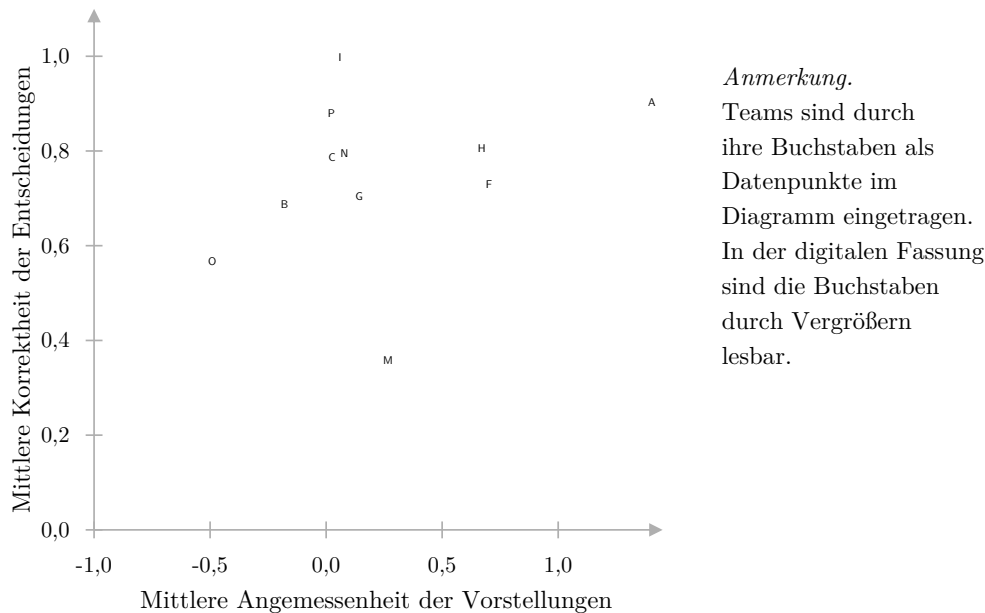


Abbildung 6.16: Streudiagramm zur Korrektheit der Einschätzungen und Angemessenheit der Ideen-Oberkategorien für jedes Team ($N = 11$).

der vorgenommenen Analysen nicht auszumachen, so dass ein Kausalzusammenhang von angemessenen Vorstellungen einer Person auf korrekte Entscheidungen derselben Person nicht als zu favorisierende Deutungsmöglichkeit dargestellt werden kann.

Für die Personen findet sich kein Zusammenhang von der Anzahl geäußerter *fachinhaltlicher* Vorstellungen und dem Maß der Korrektheit der Einschätzungen ($r = .075$, $p = .683$, basierend auf Tabelle F.13 im Anhang auf S. 576). Dieses Ergebnis ist insofern plausibel, als dass – anders als für die (Teil-)Aufgaben, bei denen von spezifischen Problemen auszugehen ist, die gleichzeitig die Entscheidungen und die Aktivierung fachinhaltlicher Vorstellungen betreffen, weshalb das dazu berichtete Ergebnis auch einen Zusammenhang von fachinhaltlichen Vorstellungen und korrekten Entscheidungen nahelegt – es eher nicht erwartbar wäre, dass Personen mit einer Neigung zu fachinhaltlichen Beiträgen auch eine Neigung zu inkorrekten Entscheidungen haben. Für die Teams findet sich kein Zusammenhang von der Anzahl geäußerter *fachinhaltlicher* Vorstellungen und dem Maß der Korrektheit der Einschätzungen ($r = 0$, $p = 1$, basierend auf Tabelle F.14 im Anhang auf S. 577). Dies ist in Übereinstimmung mit dem Befund zu den einzelnen Personen.

6.7.3 Entscheidungen und Vorstellungen zur Unterscheidung von Beobachtung und Deutung

In Tabelle 6.20 ist dargestellt, wie viele Einschätzungen ohne einen Hinweis auf eine zugehörige Ideen-Oberkategorie oder mit einem Hinweis auf eine zugehörige Ideen-Oberkategorie vorgenommen wurden und welche Ideen-Oberkategorie im zweiten Fall jeweils vergeben wurde. Ein sehr großer Teil der Einschätzungen ist korrekt. Gleichzeitig sind die meisten Einschätzungen – sowohl korrekte als auch inkorrekte – ohne Hinweis auf eine Ideen-Oberkategorie. Im Gegensatz zur Tabelle 6.17 (auf S. 314) zu Kennzeichen von naturwissenschaftlichen Fragen scheint ein Zusammenhang zwischen der Korrektheit der Einschätzungen und der Vergabe von Ideen-Oberkategorien nur bedingt vorzuliegen. Es lässt sich zwar feststellen, dass keine der 44 Einschätzungen mit Ideen-Oberkategorie inkorrekt ist, aber auch der Anteil inkorrektur Entscheidungen ohne Ideen-Oberkategorie an allen Einschätzungen ohne Ideen-Oberkategorie ist sehr gering (9 von 95).

Die Angemessenheit der Ideen-Oberkategorien geht in keiner Weise mit der Korrektheit der Einschätzungen einher. Da dies aus einem direkten Vergleich der absoluten Zahlen ersichtlich ist, sind an dieser Stelle keine statistischen Analysen zum Zusammenhang von Angemessenheit der Ideen-Oberkategorien und Korrektheit der

Tabelle 6.20: Korrektheit der Einschätzungen, die mit den einzelnen Ideen-Oberkategorien einhergehen, für die Unterscheidung von Beobachtung und Deutung.

Ideen-OK		Einschätzungen				
Name	Angemessenheit	Über Erwartung	Richtig	Falsch	Unklar	SUMME
Meinung	-1	0	1	0	0	1
Vermutung	1	1	9	0	0	10
Interpretation	1	0	2	0	0	2
BegründungErklärung	1	5	13	0	0	18
AuswertungAntwort	1	0	1	0	1	2
FaktenSicher	-1	1	1	0	0	2
Beschreiben	1	0	0	0	0	0
Messen	1	0	1	0	1	2
Beobachten	-1	0	4	0	0	4
Sehen	1	0	2	0	0	2
Sonstiges	1	0	0	0	1	1
ObjektivSubjektiv	1	0	0	0	0	0
Ohne Ideen-OK		8	95	10	8	121
SUMME mit Ideen-OK		7	33	0	3	44
SUMME mit und ohne Ideen-OK		15	128	10	11	165

Einschätzungen aufgeführt. Das Maß zur Korrektheit wird in Abschnitt 6.10 (ab S. 397) erneut verwendet.

Statt weitere statistische Analysen durchzuführen, wurde die Zeit darauf verwendet, die inkorrekten und unklaren Einschätzungen genauer zu betrachten. Es finden sich keine Hinweise zur Beantwortung der Forschungsfrage (also darauf, mit welchen Vorstellungen korrekte und fehlerhafte Entscheidungen zum experimentbezogenen Denken und Arbeiten einhergehen). Im Anhang sind allerdings Ausführungen hinterlegt, die verschiedene Ergebnisse zum Diskurskontext bei inkorrekten Entscheidungen benennen (siehe Unterabschnitt F.4.2 ab S. 567). Außerdem ist dort diskutiert, dass einige Lernende die Aufgaben vermutlich missverstehen und daher Einschätzungen vornehmen, die feingliedriger sind als in der Instruktion intendiert. Die in diesem Anhang aufgeführten Transkriptausschnitte illustrieren gut, was schon im Kontext der Darlegung des methodischen Vorgehens angemerkt wurde: Einschätzungen sind keine gehäuft auftretenden Konstrukte, sondern können innerhalb weniger Sekunden geändert werden. Die endgültige Einschätzung einer Person geht nicht zwingend aus der letzten formulierten Einschätzung hervor (z. B. weil sie durch implizite Zustimmung zur Äußerung einer anderen Person die Einschätzung noch ändert) und zudem ist nicht klar, ob die letzte Einschätzung überhaupt die im Denken der Person prominente Einschätzung ist. Die methodische Entscheidung, jede Einschätzung einzeln zu kodieren, hat also den Vorteil, diese Fluktuation abbilden zu können; allerdings wird dieses Potential in der vorliegenden Arbeit nur zum Teil ausgeschöpft, weil zeitliche Verläufe oder Reihenfolgen von Entscheidungen innerhalb einzelner (Teil-)Aufgaben nicht in den Blick genommen werden.

6.8 Auswertungen und Ergebnisse zur Stabilität und Variabilität der Vorstellungen

F-Vor5 Wie stabil bzw. variabel sind die Vorstellungen der Lernenden über Situationen hinweg?

Die betrachteten Instruktionsextrakte sind so angelegt, dass alle Situationen und Aufgaben die als angemessen einzuschätzenden, mit den adressierten Konzepten im Einklang stehenden, Vorstellungen *und gleichzeitig* in großen Teilen auch eigene, ideosynkratische oder sachunangemessene Vorstellungen genutzt werden können (z. B. weil über weite Teile nicht suggeriert wird, welche Vorstellungen zentral sind). Als Einstieg in die Untersuchung von der Stabilität und Variabilität der Vorstellungen wird geprüft, inwieweit diese Anlage der Instruktion sich auch in der Varianz der Ideen-Oberkategorien über die Aufgaben hinweg zeigt (Unterabschnitt 6.8.1). Anschließend werden unterschiedliche Analysen dazu dokumentiert, wie stabil oder situativ Vorstellungen über mehrere Aufgaben hinweg vorkommen, wobei zunächst auf über alle Aufgaben hinweg gehäuft auftretende Vorstellungen eingegangen wird (Unterabschnitte 6.8.2 und 6.8.3) und danach untersucht wird, welche Vorstellungen direkt aufeinanderfolgend mehrfach auftreten und welche stärker variabel sind (Unterabschnitt 6.8.4).

6.8.1 Vorstellungen der Kohorte über die Aufgaben hinweg

Insgesamt ergibt sich, dass die Situationen aus dem Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen unterschiedliche Spektren von Ideen-Oberkategorien abdecken, die einander aber deutlich überlappen. Für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung decken die Situationen sehr ähnliche Spektren ab. Es ist daher für beide Instruktionsextrakte davon auszugehen, dass prinzipiell vielfältige Vorstellungen pro Situation angebracht werden können – wenngleich die Lernenden in verschiedenen Situationen, insbesondere im Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen, verschiedene Vorstellungen favorisieren (ggf., weil unterschiedliche Situationen jeweils unterschiedliche Vorstellungen stärker anregen).

Der Zusammenhang von Vorstellungen und Situationen wird auf Kohortenebene betrachtet, indem analysiert wird, *welche Ideen-Oberkategorien bei welchen Aufgaben für wie viele Personen* kodiert wurden. Hierfür werden die Teilaufgaben einzeln dargestellt, die jeweils spezifische Situation in den Blick nehmen (Abbildung 6.17 für Kennzeichen von naturwissenschaftlichen Fragen, Abbildung 6.18 für die Unterscheidung von Beobachtung und Deutung). Wurden Personen Ideen-Oberkategorien für Äußerungen zugewiesen, die sich nur einer kompletten Karte zuordnen lassen, sind die entsprechenden Anzahlen in schwarz hinterlegten Feldern eingetragen.

6.8 Ausw. u. Ergebnisse zur Stabilität u. Variabilität der Vorstell. (F-Vor5)

Ideen-Oberkategorien		Ideen-Unterkategorien																														
		1-21	1-21-i	1-21-ii	1-21-iii	1-21-iii	1-21-iv	1-21-iv	1-21-v	1-22	1-22-i	1-22-ii	1-22-iii	1-22-iv	1-22-v	1-22-vi	1-22-vii	1-22-ii	1-23	1-24	1-24-i	1-24-ii	1-25	1-25-A	1-25-B	1-26	1-26-C	1-26-D	1-27	1-31	1-31-1	1-31-2
deutlich	Empfindungen	1	2	2													5	1		1	5											
angemessen	MethodeSpezifisch																	1	3	12	10	2	8	11		8	5			6		4
	MethodeUnspezifisch			1													1	4		7	2		2	1		2	2			2		6
eher	Subjektiv					1											3				9		3			1						1
angemessen	Objektiv									1							4		1	5	1		4	4	1		1			2		
	Zusammenhang	1	1	1				2					1				3															1
	Persönlich					1			2	1							9	1	1	1	16	1	1	3								
	BeweisenNachweisen																1	3		2	2		3	2		1	1	1		3		4
eher	Belegen								2								2				1									1		1
unangemessen	Prüfen																			2			1	4		4	4		2	3		1
	NawiEigenschaften			1								1					9			7	1									1		1
	Fachbezug	2	3	4	2	4	1		1								8			5	1		1			1	4					
	PF/AF	1						2	1								3				1	1				2						1
deutlich	Klang				1				1		1		1				1						1			3						
unangemessen	Effekt				1							1					1						1			4						
	AntwortAusschlag		1														4			2	1		3	1		1	1			1		

Abbildung 6.17: Anzahl der Personen, denen eine Ideen-Oberkategorie bei einer Teilaufgabe zugewiesen wurde (Kennzeichen von naturwissenschaftlichen Fragen, insgesamt 33 Personen).

Für die nach rechts abgetragenen (Teil-)Aufgaben siehe Unterabschnitt 6.2.1, ab S. 269. Schwarz hinterlegte Felder bedeuten, dass die Kodierungen keiner spezifischen (Teil-)Aufgabe zugeordnet werden konnte, sondern nur der aktuell bearbeiteten Karte.

[illegible]

Abbildung 6.18: Anzahl der Personen, denen eine Ideen-Oberkategorie bei einer Teilaufgabe zugewiesen wurde (Unterscheidung von Beobachtung und Deutung, insges. 28 Personen)

Für die nach rechts abgetragenen (Teil-)Aufgaben siehe Unterabschnitt 6.2.2, ab S. 272.

Zunächst fällt auf, dass es Teilaufgaben mit breitem und schmalem Spektrum an Ideen-Oberkategorien gibt, was ein Hinweis darauf sein könnte, dass in den verschiedenen Situationen unterschiedlich viele verschiedene Vorstellungen aktiviert werden.

Bei **Kennzeichen von naturwissenschaftlichen Fragen** haben für gewöhnlich die späteren Teilaufgaben ein breiteres Spektrum (dazu mehr bei F-Vor6, Analysen zum zeitlichen Verlauf). Allerdings scheint die noch vor der instruktionalen Information platzierte Teilaufgabe 1-22-ii das Gesamtspektrum am ehesten abzubilden. Die Aufgabe lautet: »Stellen Sie sich vor, Sie sollten jemandem erklären, was eine naturwissenschaftliche Fragestellung von einer nicht-naturwissenschaftlichen Fragestellung unterscheidet. Was würden Sie sagen?« Es ist also inhaltlich nicht verwunderlich, dass bei dieser offenen Frage eine große Bandbreite an Vorstellungen auftritt. Gleichzeitig ist es bemerkenswert, dass schon vor der instruktionalen Information (1-23) ein so breites Spektrum an Vorstellungen vorliegt (dazu mehr bei F-Vor6).

Zeitweise dominieren die Ideen-Oberkategorien **Fachbezug** (Teilaufgaben zur Aufgabenkarte 1-21) und **MethodeSpezifisch** (Teilaufgaben von 1-23 bis 1-26) stark, was darauf hinweisen könnte, dass zumindest diese beiden Vorstellungen unabhängig von den Situationen genutzt werden (können). Da es ein Kohortenergebnis ist, lässt es sich allerdings nicht so deuten, dass tatsächlich ein Individuum eine der Vorstellungen unabhängig von den Situationen nutzt. Erst in den weiteren Analysen zu dieser Forschungsfrage werden die Individuen genauer in den Blick genommen, um hierzu Hinweise zu erlangen. Inwiefern das Dominieren verschiedener Ideen-Oberkategorien in verschiedenen Zeitintervallen (d. h. in zu gewissen Teilaufgaben gehörigen Abschnitten) als Entwicklung auf Kohorten- oder Individualebene gedeutet werden kann, wird später im Kontext von F-Vor6 diskutiert.

Bei der **Unterscheidung von Beobachtung und Deutung** ist zunächst augenfällig, dass nach der instruktionalen Information (d. h. bei Teilaufgaben der Aufgabenkarte 3-24 und 3-25) insgesamt nur sehr wenige Ideen-Oberkategorien kodiert wurden. Wenn Aufgabe 3-24-b auñenvor gelassen wird (in der danach gefragt wird, warum es im Kontext der Aufgabenkarte hilfreich sein dürfte, zwischen Beobachtungen und Deutungen zu unterscheiden), finden sich zudem nach der instruktionalen Information nur drei verschiedene Ideen-Oberkategorien. Dieses Ergebnis könnte der Anlage der Aufgabenkarte 3-24 geschuldet sein, die ein fiktives Gespräch über alltägliche Dinge präsentiert und dazu zunächst die Aufgabe stellt, Beobachtungen und Deutungen in diesem Gespräch zu identifizieren (3-24-a). Zum einen ist es für eine Identifikation nicht notwendig, zu beschreiben oder zu erläutern, was eine Beobachtung oder Deutung ausmacht. Zum anderen sind die als Teilaufgaben benannten Entitäten

eigentlich Sätze des fiktiven Gesprächs, die gegebenenfalls von den Schülerinnen und Schülern anders behandelt werden als die isoliert betrachtbaren Beispielfragen im Kontext von Kennzeichen von naturwissenschaftlichen Fragen (Aufgabenkarte 1-21), was sich darin widerspiegeln könnte, dass nicht zu jedem Satz eine Entscheidung getroffen wird. Beide möglichen Erklärungen scheinen vor dem Hintergrund einer erfolgten Videoeinsicht plausibel zu sein.

Eine Zusammenfassung der Teilaufgaben 3-24-a-01 bis 3-24-a-10 (und der Aufgabe 3-24-a selbst) zu einer Aufgabe 3-24-a ist vor dem Hintergrund des vorigen Absatzes naheliegend. An dem Ergebnis ändert dies allerdings nur wenig, weil für solche ein zusammengefasste Aufgabe 3-24-a nach wie vor nur drei verschiedene Ideen-Oberkategorien vorliegen. Bei einer Zusammenfassung für Gesamtaufgabe 3-24-a würde für zwei Personen die Ideen-Oberkategorie **Vermutung**, für zwei Personen **Interpretation** und (dadurch dass die Kodierungen bei 3-24-a-03 und -04 zu derselben Person gehören) für drei Personen **BegründenErklären** notiert werden.

Ferner fällt – auch im Vergleich zu Kennzeichen von naturwissenschaftlichen Fragen – für die Unterscheidung von Beobachtung und Deutung auf, dass ein großer Anteil der Ideen-Oberkategorien bei mehreren Teilaufgaben kodiert wurde. Die Ideen-Oberkategorien **Vermutung** und **BegründungErklärung** finden sich bei zehn oder mehr Teilaufgaben. Die Ideen-Oberkategorien **Interpretation** (fünfmal), **Sehen** (viermal) und **Beobachten** (dreimal) wurden häufiger als zweimal vergeben. Von den bei zwei Teilaufgaben vergebenen Ideen-Oberkategorien wurden mit **FaktenSicher** und **AuswertungAntwort** zudem die Hälfte für mehr als zwei Personen kodiert, was vor dem Hintergrund der wenigen überhaupt kodierten Ideen-Oberkategorien bedeutsam zu sein scheint. Daher könnte davon gesprochen werden, dass die Vorstellungen zur Unterscheidung von Beobachtung und Deutung nicht allzu stark mit den Situationen zusammenhängen – insbesondere wenn dies im Vergleich zu Kennzeichen von naturwissenschaftlichen Fragen berichtet wird. Weil sich die betrachteten Situationen teilweise nicht eignen, um Aussagen über Vorstellungen zu machen, könnte aber insofern ein Zusammenhang von Situationen und Vorstellungen vorliegen, als dass manche Situationen die Äußerung von Vorstellungen nicht zulassen oder anregen.

Auf einer erklärenden Ebene kann berücksichtigt werden, dass Beobachtungen und Deutungen *zwei* inhaltliche Kategorien sind, naturwissenschaftliche und nicht-naturwissenschaftliche Fragen aber nur *eine* Kategorie und ihre Negation. Das könnte für Abgrenzungsüberlegungen, die Lernende vornehmen, und ihre darauf basierenden Vorstellungen durchaus einen Unterschied machen. Wenn beispielsweise zu Beobachtungen aufgrund von Alltagserfahrungen bereits Vorstellungen vorliegen, lässt sich dies von Lernenden als Startpunkt nutzen, um auch Deutungen in Abgrenzung

von Beobachtungen zu erschließen. Für Kennzeichen von naturwissenschaftlichen Fragen dürfte es für Lernende schwieriger sein, einen solchen Startpunkt auszumachen. Diese Überlegungen zeigen, dass es – den berichteten Ergebnissen entsprechend – im Kontext der Unterscheidung von Beobachtung und Deutung plausibler ist, dass Vorstellungen über verschiedene Situationen hinweg genutzt werden, wohingegen es im Kontext von Kennzeichen von naturwissenschaftlichen Fragen zumindest dann, wenn zu Beginn der Bearbeitung nur intuitive Alltagsvorstellungen vorliegen, wahrscheinlicher scheint, dass konkrete Situationen (wie etwa Beispielfragen) als Ausgangspunkt zur Generierung von Vorstellungen genutzt werden.

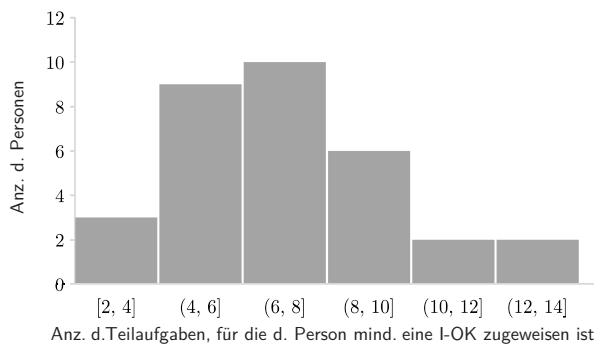
Zusammenfassend lässt sich anmerken, dass die im weiteren Verlauf berichtete Betrachtung der Stabilität und Variabilität von Vorstellungen über Aufgaben hinweg der Einschränkung der unterliegt, dass in den tatsächlich vorliegenden Daten nicht für alle Aufgaben auch alle Ideen-Oberkategorien kodiert sind. Dieser Einschränkung kann insofern begegnet werden, dass die Ergebnisse nicht auf alle Vorstellungen in gleicher Weise bezogen werden. Bei einer deutlich größeren Stichprobe ist zwar nicht davon auszugehen, dass in allen Situationen alle Ideen-Oberkategorien vorliegen, aber vom dem Hintergrund der Anlage der Instruktion ist es zumindest naheliegend, dass für alle Aufgaben breitere Spektren vorliegen dürften. Analysen dazu, welche Aufgaben welche Vorstellungen besonders stark anregen (oder genauer: im Kontext welcher Aufgaben Lernende mit gewissen Voraussetzungen stärker zur Verbalisierung welcher spezifischen Vorstellungen neigen), sollten also auf größeren Stichproben beruhen.

6.8.2 Über mehrere Aufgaben hinweg gehäuft auftretende Vorstellungen der Lernenden

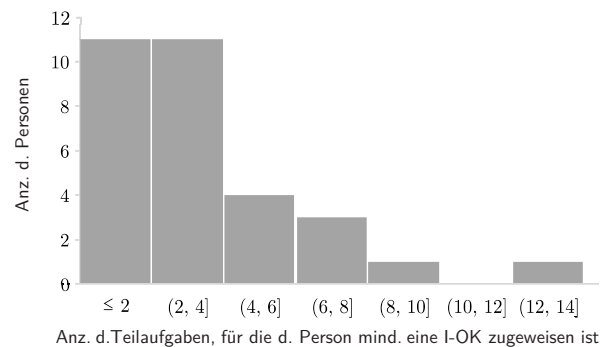
Die folgende Betrachtung von über mehrere Aufgaben hinweg auftretenden Vorstellungen bezieht ein, dass sich das *Zusammenspiel von Vorstellungen und Situationen für verschiedene Personen unterschiedlich* gestalten kann. Dazu wird die Anzahl der verschiedenen *gehäuft auftretenden* Ideen-Oberkategorien pro Person betrachtet. Was genau darunter verstanden wird, wird weiter unten ab Seite 332 erläutert. Zunächst wird diskutiert, inwiefern eine Untersuchung von Häufungen von Ideen-Oberkategorien interpretativ in Verbindung mit Stabilität und Variabilität von Vorstellungen über Aufgaben hinweg gebracht werden kann.

Die Analysen zu den gehäuft auftretenden Ideen-Oberkategorien (und auch später der Wechsel von Ideen-Oberkategorien, siehe Unterabschnitt 6.8.4) dienen dazu, auf die Stabilität und Variabilität der Vorstellungen der Lernenden über die Aufgaben hinweg zu schließen. Dieser Schluss ist allerdings mit Einschränkungen verbunden,

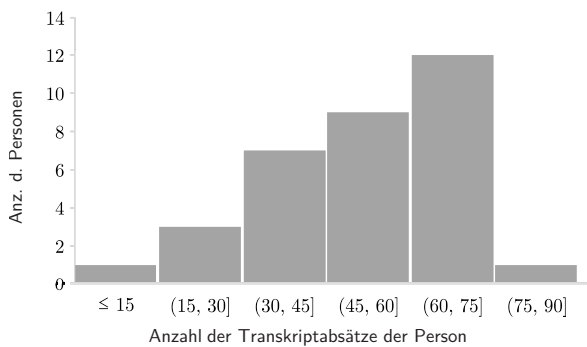
unter anderem aus dem Grund, dass verschiedene Ideen-Oberkategorien bei Betrachtung der gesamten Kohorte für verschiedene (Teil-)Aufgaben unterschiedlich häufig auftreten und dass unterschiedliche Lernende unterschiedlich viel verbal beitragen. Der Umfang des Beitrags kann für eine Person beispielsweise durch die Anzahl der Teilaufgaben eingeschätzt werden, bei denen einer Person auch eine Ideen-Oberkategorie zugewiesen ist. Die zugehörigen Verteilungen sind in Abbildung 6.19 dargestellt. Für das detaillierter untersuchte Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen ist ebenfalls die Verteilung der Transkriptabsätze pro Person abgebildet, weil auch diese nachfolgend als Maß für den Umfang des Beitrags genutzt wird. Insgesamt ergibt sich, dass sehr wohl Streuung vorliegt, für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen allerdings plausibel davon ausgegangen werden kann, dass alle Lernenden in einem ausreichenden Umfang beitragen. Für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung sind die Einschränkungen deutlich größer, weil viele Personen nur bei wenigen Teilaufgaben einen Beitrag machen. Das Vorgehen zu gehäuft auftretenden Ideen-Oberkategorien wird daher stellenweise entsprechend modifiziert und diskutiert.



(a) Kennzeichen von naturwissenschaftlichen Fragen



(b) Unterscheidung von Beobachtung und Deutung



(c) Transkriptabsätze

Abbildung 6.19: Verteilung der Anzahl der Teilaufgaben, denen pro Person mindestens eine Ideen-Oberkategorie zugewiesen ist (a+b) sowie Verteilung der Anzahl der pro Person existierenden Transkriptabsätze für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen (c).

Als *über mehrere Aufgaben hinweg gehäuft auftretende Ideen-Oberkategorie* (nachfolgend kurz: *gehäuft auftretende Ideen-Oberkategorie*) wird eine Ideen-Oberkategorie bezeichnet, die einer Person (bzw. einem Team) bei g oder mehr Teilaufgaben zugewiesen wurden. Das heißt, dass die Ideen-Oberkategorie bei mindestens g Teilaufgaben für eine spezifische Person kodiert sein muss, um als für diese Person gehäuft auftretende Ideen-Oberkategorie angesehen zu werden. Die (Teil-)Aufgaben müssen jedoch nicht direkt aufeinander folgen (Analysen dazu finden sich weiter unten in Unterabschnitt 6.8.4 ab S. 343). Die Wahl von g ist an das untersuchte Instruktionsextrakt angepasst, um als Ursachen für die Deklaration als gehäuft auftretende Ideen-Oberkategorie möglichst gut sowohl Anregungen durch die Aufgaben als auch einfaches Wiederholen derselben Begründung bei direkt aufeinanderfolgenden Teilaufgaben ausschließen zu können. Gleichzeitig sollte die geforderte Vorkommenshäufigkeit nicht zu hoch angesetzt werden, um keine ständige Nutzung einer Vorstellung zu verlangen, obwohl die Aufgaben diese Nutzung gegebenenfalls nicht zulassen. Als sinnvolle Setzungen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen hat sich $g = 4$ ergeben, für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung $g = 3$.

In der Definition von gehäuft auftretenden Ideen-Oberkategorien wurden Teilaufgaben als Zählgrundlage gewählt. Zum einen liegt dies darin begründet, dass Aufgaben oder Aufgabenkarten in den meisten Fällen die spezifische Situation (etwa den konkreten Beispielsatz, der betrachtet wird) nicht hinreichend von der nächsten spezifischen Situation trennen. Zum anderen wurden nicht einzelne Äußerungen der Personen genommen, da diese ebenfalls nicht die Situation abbilden, sondern eine der Reaktionen auf die Situation. Lernende können außerdem unterschiedliche viele Äußerungen machen, so dass ein weiteres Maß zur Herstellung der Vergleichbarkeit von Personen eingeführt werden müsste. Zudem würde ein Zählen einzelner Äußerungen dazu führen, dass wiederholtes Kodieren einer Ideen-Oberkategorie zu einer spezifischen Situation (etwa weil eine Person bei einer einzelnen Aufgabe ihre Position mehrfach kundtut, um sich Gehör zu verschaffen) schon dazu führt, dass eine Ideen-Oberkategorie als stabil angesehen wird, obwohl sie nur in wenigen Situationen auftaucht. An der Begründung wird allerdings auch deutlich, dass die Wahl von Teilaufgaben als Zählgrundlage zu einer mit einer Unterschätzung der Wiederholung von Ideen-Oberkategorien einhergeht.

Die Anzahl der verschiedenen gehäuft auftretenden Ideen-Oberkategorien pro Person sind für Kennzeichen von naturwissenschaftlichen Fragen in Tabelle 6.21 auf S. 334, für die Unterscheidung von Beobachtung und Deutung in Tabelle 6.23 auf S. 337 notiert. Als Vergleichswerte sind zudem die Anzahlen der verschiedenen Ideen-Oberkategorien aufgeführt, die einer jeweiligen Person innerhalb der gesamten Se-

quenz zugewiesen wurden. Es ist beispielsweise denkbar, dass eine Person, die über die Aufgaben hinweg gehäuft eine spezifische Ideen-Oberkategorie zugewiesen bekommt, nur selten Äußerungen macht, denen andere Ideen-Oberkategorien zugewiesen sind, etwa deshalb, weil die Person sich stark auf eine Vorstellung konzentriert und diese z. B. bei allen Teilaufgaben anbringt, bei denen es ihr möglich erscheint. Es wäre dann spannend, für wie viele Lernende welche Vorstellungen in dieser Weise stabil rekonstruiert werden können, weil dies Aufschluss über die in der Forschungsfrage thematisierte Stabilität der Vorstellungen über die Situationen hinweg geben könnte.

Für **Kennzeichen von naturwissenschaftlichen Fragen** haben Individuen mit überdurchschnittlich vielen verschiedenen Ideen-Oberkategorien mehr gehäuft auftretende Ideen-Oberkategorien als solche mit unterdurchschnittlich vielen verschiedenen Ideen-Oberkategorien, wie Mann-Whitney-U-Tests zeigen. Das Ergebnis ist bei Betrachtung der *absoluten* Anzahl gehäuft auftretender Ideen-Oberkategorien signifikant mit mittlerer Effektstärke (Mann-Whitney-U-Test für 19 Personen mit $Mdn_{OK>MW} = 1$ gegen 14 Personen mit $Mdn_{OK<MW} = 0$; $U = 82$, $z = -2.012$, $p = .049$, $r = .35$, wobei der Index $OK<MW$ bedeutet, dass die Gruppe aller Personen betrachtet wird, die weniger verschiedene Ideen-Oberkategorien aufweisen als der Mittelwert über alle Personen beträgt und $OK>MW$ analog ist). Bei Betrachtung des *relativen* Anteils (gehäuft auftretender Ideen-Oberkategorien einer Person an verschiedenen Ideen-Oberkategorien derselben Person) ist die Tendenz nur noch mit kleiner Effektstärke vorhanden und zudem nicht signifikant (Mann-Whitney-U-Test für 19 Personen mit $Mdn_{OK>MW} = 0.125$ gegen 14 Personen mit $Mdn_{OK<MW} = 0$; $U = 100.5$, $z = -1.159$, $p = .213$, $r = .202$). Das Ergebnis spiegelt sich auch innerhalb der einzelnen Teams wider: Für Individuen, denen weniger Ideen-Oberkategorien als ihren Gruppenmitgliedern zugewiesen wurden, treten auch (absolut) weniger gehäuft auftretende Ideen-Oberkategorien auf (siehe die folgenden Personen im Kontext ihrer Teams: A3, C8, G20, I25, M37, P48; aber anders: G21 und ggf. O44).⁹⁹

Eine mögliche Erklärung für die Ergebnisse des vorigen Absatzes könnte sein, dass Personen, die weniger verschiedene Vorstellungen äußern, dies nicht aus dem Grund tun, dass sie relativ betrachtet auf mehr ihrer Vorstellungen gehäuft, also über die Aufgaben hinweg wiederholt, zurückgreifen. Es ist stattdessen naheliegender, dass gehäuft auftretende Ideen-Oberkategorien vor allem bei Personen auftreten, die einen höheren Redeanteil haben, der auch mit mehr verschiedenen aus den Äußerungen rekonstruierbaren Vorstellungen einhergeht. Auch Korrelationsanalysen

⁹⁹Bei einer größeren Stichprobe wären eventuell Mehrebenenanalysen denkbar. Die vorliegenden Daten erfüllen aber die Voraussetzungen nicht (siehe bspw. Field, 2013).

Tabelle 6.21: Anzahl der verschiedenen Ideen-Oberkategorien und gehäuft auftretenden Ideen-Oberkategorien, die einer Person zugewiesen wurden (Kennzeichen von naturwissenschaftlichen Fragen).

Person	Verschiedene OK	gehäufte OK
A1	6	3
A2	6	1
A3	2	1
B4	4	0
B5	7	0
B6	6	0
C7	5	0
C8	6	1
C9	4	0
F16	8	0
F17	6	0
F18	6	2
G19	3	2
G20	5	0
G21	4	1
H22	9	1
H23	8	1
H24	8	1
I25	6	2
I26	5	0
I27	3	0
M37	8	1
M38	3	0
M39	2	0
N40	9	2
N41	5	2
N42	2	0
O43	8	0
O44	6	2
O45	6	1
P46	7	1
P47	7	0
P48	4	0
<i>Mittelwert</i>	<i>5.58</i>	<i>0.76</i>
<i>Median</i>	<i>6</i>	<i>1</i>

Anmerkungen. Eine gehäuft auftretende Ideen-Oberkategorie liegt für eine Person dann vor, wenn die entsprechende Ideen-Oberkategorie der Person bei mindestens vier verschiedenen (Teil-)Aufgaben zugewiesen wurde.

von der Anzahl der Transkriptabsätze pro Person mit den gehäuft auftretenden bzw. verschiedenen Ideen-Oberkategorien legen dies nahe: Der Korrelationskoeffizient für verschiedene Ideen-Oberkategorien ($r = .525$, $p = .002$) ist größer als der für gehäuft auftretende Ideen-Oberkategorien ($r = .418$, $p = .016$) oder für den relativen Anteil gehäuft auftretender Ideen-Oberkategorien ($r = .330$, $p = .061$). Weil insbesondere die Zusammenhänge für verschiedene sowie für gehäuft auftretende Ideen-Oberkategorien signifikant positiv sind, scheint die Erklärung plausibel, dass der Redeanteil einen maßgeblichen Einfluss hat. Die Deutung, dass eine größere Anzahl der verschiedenen Ideen-Oberkategorien auch mit einer größeren Anzahl stabiler Ideen-Oberkategorien einhergeht, kann also vor dem Hintergrund der Daten genauso wenig aufrechterhalten (oder verworfen) werden wie die Deutung, dass Personen, die über die Aufgaben hinweg gehäuft je spezifische Ideen-Oberkategorien zugewiesen bekommen, seltener Äußerungen machen, denen andere Ideen-Oberkategorien zugewiesen sind.

Ein weiteres, zentrales Ergebnis ist, dass es gemäß Tabelle 6.21 (auf S. 334) für etwa die Hälfte der Schülerinnen und Schüler einzelne gehäuft auftretende Ideen-Oberkategorien gibt. Weil die Lernenden beim Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen im Mittel etwa bei 7 Teilaufgaben Ideen-Oberkategorien zugewiesen bekommen haben und jede Ideen-Oberkategorie maximal dreimal auftreten darf, damit sie nicht als gehäuft auftretend angesehen wird, können für Lernende dieser groben Rechnung gemäß im Mittel etwa 3 verschiedene Vorstellungen aus den Äußerungen zu 7 verschiedenen Teilaufgaben rekonstruiert werden. Insgesamt kann also davon die Rede sein, dass sich viel Varianz entlang der Situationen zeigt.

Es kann auch gefragt werden, ob Personen, die miteinander in einem Team arbeiten, auf die Vorstellungen der anderen Teammitglieder rekurren. Denkbar wäre dann gegebenenfalls, dass einzelne Personen keine gehäuft auftretenden Vorstellungen vorweisen, weil bereits andere Teammitglieder bei manchen der Teilaufgaben eine der entsprechenden Vorstellung Ausdruck verleihende Äußerung gemacht haben. Ein Vergleich auf Teamebene (siehe Tabelle 6.22 auf S. 336) zeigt, dass bei nur geringer Streuung im Mittel 8.81 verschiedene Ideen-Oberkategorien pro Team vorliegen. Ein Vergleich von Teams mit überdurchschnittlich vielen bzw. wenigen Ideen-Oberkategorien wäre daher schwer zu interpretieren. Für jedes Team liegen gehäuft auftretende Ideen-Oberkategorien vor: für 2 Teams eine, für 4 Teams zwei und für fünf Teams drei gehäuft auftretende Ideen-Oberkategorien. Werden die Personen nicht aufgelöst, sondern wird das Team als Gesamtes betrachtet, wird also scheinbar tatsächlich eher von den gehäuft auftretenden Ideen-Oberkategorien aus agiert. Das zeigt sich auch dann, wenn die Anzahl der Teilaufgaben, bei denen eine gehäuft auftretende Ideen-Oberkategorie vergeben wurde, ins Verhältnis zu allen Teilaufgaben, bei denen irgend-

eine Ideen-Oberkategorie vergeben wurde, gesetzt wird. Dieser Anteil ist für viele Teams (teilweise deutlich) größer als 50 %, so dass es plausibel erscheint, dass diese Teams im Diskurs immer wieder auf die gehäuft auftretenden Ideen-Oberkategorien zurückgreifen.

Tabelle 6.22: Anzahl der verschiedenen Ideen-Oberkategorien und gehäuft auftretenden Ideen-Oberkategorien, die einem Team zugewiesen wurden, sowie Anzahl der Teilaufgaben, bei denen einem Team irgendwelche bzw. gehäuft auftretende Ideen-Oberkategorien zugewiesen wurden (Kennzeichen von naturwissenschaftlichen Fragen).

Team	Ideen-Oberkategorien			Teilaufgaben mit OKs		
	Verschied.	gehäuft aufr.	n = 3	Gesamt	gehäuft	Verhältnis
A	9	3	1	51	41	0.8
B	8	1	1	25	6	0.24
C	8	2	1	34	19	0.56
E	11	3	1	61	39	0.64
F	10	2	3	39	16	0.41
G	6	2	0	35	26	0.74
I	8	2	1	31	12	0.39
M	9	1	1	28	10	0.36
N	9	3	0	56	41	0.73
O	10	3	1	54	34	0.63
P	9	3	1	43	25	0.58
<i>Mittelwert</i>	<i>8.82</i>	<i>2.27</i>	<i>1</i>	<i>41.55</i>	<i>24.45</i>	
<i>Median</i>	<i>9</i>	<i>2</i>	<i>1</i>	<i>39</i>	<i>25</i>	

Anmerkungen. Verschied. = Anzahl verschiedener Ideen-Oberkategorien, die einem Team über die gesamte Bearbeitung des Instruktionsextrakts hinweg zugewiesen wurden. Gehäuft aufr. = Eine gehäuft auftretende Ideen-Oberkategorie liegt für ein Team dann vor, wenn die entsprechende Ideen-Oberkategorie bei mindestens vier verschiedenen (Teil-)Aufgaben irgendeiner Person aus dem Team zugewiesen wurde. Die Spalte zu n = 3 bezieht sich auf das Auftreten bei drei (Teil-)Aufgaben. Gesamt/Gehäuft = Anzahl der Teilaufgaben, bei denen dem jeweiligen Team irgendwelche Ideen-Oberkategorien bzw. irgendwelche gehäuften I-OKs zugewiesen wurden.

Für die **Unterscheidung von Beobachtung und Deutung** ergibt sich ein anderes Bild (Tabelle 6.23 auf S. 337). Nur für zwei Personen finden sich gehäuft auftretende Ideen-Oberkategorien. Ein Blick auf die Anzahl der verschiedenen Ideen-Oberkategorien pro Person legt nahe, dass insgesamt wenige *verschiedene* Vorstellungen pro Person vorliegen. Wird außerdem beachtet, dass – konsistent mit den beiden vorigen Ergebnissen – Individuen für gewöhnlich nur bei sehr wenigen Teilaufgaben (typischerweise etwa einer bis vier, siehe Abbildung 6.19 auf S. 331) Ideen-Oberkategorien zugewiesen bekommen haben, lässt sich dies so zusammenfassen, dass Individuen im Rahmen der analysierten Sequenz zur Unterscheidung von Beobachtung und Deutung wenige Äußerungen machen, aus denen sich Vorstellungen rekonstruieren lassen. Zum einen finden sich nur kurze Diskurse zum untersuchten Instruktionsextrakt, was sich bereits in Unterabschnitt 6.7.3 (ab S. 324) darin widerspiegelte, dass nur wenige

Einschätzungen formuliert werden. Zum anderen lassen sich nur aus wenigen der wenigen Äußerungen, die getätigt werden, Hinweise auf Vorstellungen ableiten, weil viele der Äußerungen Einschätzungen ohne Hinweis auf eine Idee zur Unterscheidung von Beobachtung und Deutung sind (siehe Modus-Kategorien in Unterabschnitt 6.1.3, ab S. 254). Insgesamt könnte dies zusätzlich zur bereits oben erwähnten vermutlich geringen Notwendigkeit der Äußerung von Vorstellungen im Kontext einiger (Teil-)Aufgaben auch dadurch erklärt werden, dass Teammitglieder einander nicht mehr ergänzen oder wiederholen, wenn bereits eine Vorstellung genutzt wurde. Dafür spräche beispielsweise, dass es nur bei Teilaufgabe 3-21-ii (»Haben Sie eine Idee, woran man erkennt, dass es sich um eine Deutung handelt?«, oben auf der Karte findet sich eine Zuweisung von Beispielsätzen zu den Kategorien Beobachtungen und Deutungen) vorkommt, dass innerhalb desselben Teams zwei verschiedene Personen dieselbe Ideen-Oberkategorie zugewiesen bekommen. Es könnte auch daran liegen, dass für viele Entscheidungen keine Begründungen gefordert wurden (anders als bei Kennzeichen von naturwissenschaftlichen Fragen) und die (Teil-)Aufgaben insgesamt zusammenhängender waren (eine Vorstellung für mehrere Entscheidungen statt für jede Entscheidung eine Vorstellung wie bei Kennzeichen von naturwissenschaftlichen Fragen).

Tabelle 6.23: Anzahl der Ideen-Oberkategorien und gehäuft auftretenden Ideen-Oberkategorien, die einer Person zugewiesen wurden (Unterscheidung von Beobachtung und Deutung). E14 arbeitet bei Einheit 3 in Team A mit.

Person	#OK	Gehäuft(4)	Gehäuft(3)	Person	#OK	Gehäuft(4)	Gehäuft(3)
A3	0	0	0	I25	0	0	0
A1	3	0	0	I26	0	0	0
E14	5	0	0	I27	0	0	0
B4	0	0	0	K31	3	0	0
B5	4	0	1	K32	3	0	1
B6	0	0	0	K33	5	0	0
C7	3	0	0	L34	3	0	0
C8	0	0	0	L35	5	0	0
C9	0	0	0	L36	2	0	0
F16	4	0	0	M37	2	0	0
F17	0	0	0	M38	6	0	0
F18	3	0	0	M39	0	0	0
G19	0	0	0	P46	4	1	0
G20	0	0	0	P47	0	0	0
G21	0	0	0	P48	4	0	0
H22	7	2	0	Anmerkungen. #OK = Anzahl der Ideen-Oberkategorien; #Gehäuft(x) = Anzahl der x-mal im zeitlichen Verlauf (bei verschiedenen Teilaufgaben) auftretenden Ideen-Oberkategorien.			
H23	3	0	0				
H24	8	0	0				

Aufgrund der geringen Anzahl wiederholten Auftretens von Ideen-Oberkategorien kann über die zeitliche Stabilität der Vorstellungen von Individuen zur Unterscheidung von Beobachtung und Deutung keine genauere Aussage getroffen werden, als dass die Personen in den wenigen Situationen, in denen sie Vorstellungen äußern, nur wenige verschiedene Vorstellungen nutzen. Erst in Unterabschnitt 6.8.4 (ab S. 343), in dem der Wechsel zwischen Ideen-Oberkategorien auf Individualebene betrachtet wird, kann diese Aussage etwas spezifiziert werden.

6.8.3 Inhalte von über mehrere Aufgaben hinweg gehäuft auftretenden Vorstellungen

Neben der in Unterabschnitt 6.8.2 vorgenommenen Untersuchung zur Stabilität der Vorstellungen, die die Lernenden einzeln und insgesamt betrifft, stellt sich auch die Frage, ob einige der rekonstruierten Vorstellungen eher stabil und andere eher variabel auftreten. Zur inhaltsspezifischen Betrachtung der Stabilität und Variabilität wird nachfolgend aufgelöst, *welche Ideen-Oberkategorien für wie viele Individuen über die Aufgaben hinweg gehäuft auftreten*.¹⁰⁰

Für die einzelnen Ideen-Oberkategorien zum Bereich **Kennzeichen von naturwissenschaftlichen Fragen** ist in Tabelle 6.24 (auf S. 339) dargestellt, für wie viele Personen bzw. Teams die Ideen-Oberkategorie über die Aufgaben hinweg gehäuft auftritt (d. h., viermal oder häufiger zugewiesen wurde). Zunächst ist anzumerken, dass die Ideen-Oberkategorien **Fachbezug** und **MethodeSpezifisch** für 4 bzw. 9 Personen gehäuft kodiert wurden. Vor dem Hintergrund der Dominanz dieser Kategorien in der ersten Analyse (Abbildung 6.17, S. 327) ist dieser Befund nicht verwunderlich. Ferner tritt die Ideen-Oberkategorie **Objektiv** für relativ viele Personen gehäuft auf (für F18 gleichzeitig auch **Subjektiv**).

Vor dem Hintergrund der Diskussion zur Angemessenheit ist es eventuell als für das Lernen suboptimal anzusehen, dass die als fachmethodisch weniger angemessen eingestuften Ideen-Oberkategorien **PF/AF** und **AntwortAusschlag** für zwei Personen gehäuft auftreten. Gleichzeitig ergeben Detailbetrachtungen der Daten, dass die ebenfalls als weniger angemessen eingestuften Ideen-Oberkategorien **Effekt** (deutlich unangemessen) und **Fachbezug** (eher unangemessen) für viele der Personen nur bei einer einzigen Teilaufgabe vorkommen (für 6 von insgesamt 8 Teams mit **Effekt** und für 5 von 9 Teams mit **Fachbezug**). Hieraus kann zwar nicht geschlossen werden, dass

¹⁰⁰Die zugehörigen Deutungen lassen sich nicht anhand von Abbildung 6.17 (auf S. 327) oder Abbildung 6.18 (auf S. 327) ableiten, weil in diesen nur auf Kohortenebene dargestellt ist, bei wie vielen Aufgaben die entsprechenden Ideen-Oberkategorien zugewiesen wurden. Vielmehr wurden die Analysen aus Unterabschnitt 6.8.2 (ab S. 330) verwenden, aber nicht nach Personen, sondern nach Ideen-Oberkategorien zusammengefasst.

Tabelle 6.24: Anzahl der Personen bzw. Teams, für die eine Ideen-Oberkategorie stabil vorliegt (Kennzeichen von naturwissenschaftlichen Fragen).

Ideen-Oberkategorie	Gehäuft (Gesamt)	
	# S	# T
Empfindungen	1 (10)	1 (5)
Persönliches	0 (22)	2 (10)
Subjektiv	1 (10)	1 (5)
Objektiv	3 (9)	3 (5)
MethodeSpezifisch	9 (20)	5 (8)
MethodeUnspezifisch	0 (16)	2 (8)
NawiEigenschaften	0 (16)	1 (8)
Zusammenhang	1 (6)	1 (4)
Effekt	0 (8)	0 (7)
Prüfen	1 (13)	1 (8)
Belegen	0 (7)	0 (4)
BeweisenNachweisen	1 (7)	2 (4)
AntwortAusschlag	2 (7)	2 (5)
PF/AF	2 (4)	1 (2)
Fachbezug	4 (19)	3 (9)
Klang	0 (6)	0 (6)

Anmerkungen. #S = Eine gehäuft auftretende Ideen-Oberkategorie liegt für eine Person dann vor, wenn die entsprechende Ideen-Oberkategorie der Person bei mindestens vier verschiedenen (Teil-)Aufgaben zugewiesen wurde. #T = Eine gehäuft auftretende Ideen-Oberkategorie liegt für ein Team dann vor, wenn diese bei mindestens vier verschiedenen (Teil-)Aufgaben irgendeiner Person aus dem Team zugewiesen wurde.

die zugehörigen Vorstellungen nicht oder nur instabil vorliegen, weil die Personen entsprechende Äußerungen in das Lernmaterial ergänzenden Situationen eventuell gemacht hätten. Dennoch könnte dies ein Hinweis darauf sein, dass die Schülerinnen und Schüler keine Äußerungen machen, aus denen sich diese Vorstellungen rekonstruieren ließen, weil ihnen im Kontext der sonstigen bearbeiteten Aufgaben bewusst ist, dass sie aus fachmethodischer Sicht eher unangemessen bzw. nicht hinreichend wären.

Die nur selten zugeschriebenen Ideen-Oberkategorien sind insofern über den vorigen Absatz hinaus interessant, als dass sie bei allen Personen bzw. Teams jeweils mit denselben (Teil-)Aufgaben verbunden sein könnten. In diesem Fall wäre die im letzten Satz des vorigen Absatzes gegebene Erklärung des seltenen Auftretens der beiden tendenziell unangemessenen Ideen-Oberkategorien **Effekt** und **Fachbezug** zusätzlich plausibilisiert, wenn sich Gemeinsamkeiten dieser (Teil-)Aufgaben finden ließen,

anhand derer ein dort vermehrtes Auftreten begründet werden könnte. Dies ist mit Einschränkungen der Fall, wie in den nachfolgenden Absätzen diskutiert:

Für die Teams, bei denen die Ideen-Oberkategorie **Fachbezug** nur bei einer einzigen Aufgabenkarte vorkommt, geschieht dies bei der Aufgabenkarte 1-21 (meist bei früheren Teilaufgaben). Dieses Ergebnis scheint sich nicht aus der Natur der Beispielfragen erklären zu lassen, sondern vielmehr mit dem Antext der Aufgabenkarte zu tun zu haben, welcher den folgenden Satz enthält: »Nicht jede Fragestellung, die etwas mit Physik, Chemie oder Biologie zu tun hat, ist automatisch eine naturwissenschaftliche Fragestellung.« Im Informationstext wird verneint, dass **Fachbezug** ein passendes Kriterium zur Entscheidung ist. Dass die Schülerinnen und Schüler dennoch auf dieses Kriterium zurückgreifen, könnte daran liegen, dass sie keine anderen Vorstellungen aktivieren (können) – eventuell weil der Antext eine gedankliche Fokussierung nahelegt, die für die Entwicklung weiterer Vorstellungen nicht förderlich ist.

Eine weitere Erklärung für das vermehrt einmalige Vorkommen der Ideen-Oberkategorie **Fachbezug** bei der Aufgabenkarte 1-21 könnte darin bestehen, dass einfach die Aussage des Antextes (siehe voriger Absatz) sinngemäß wiederholt wird. Dies ist allerdings insgesamt betrachtet nicht der Fall. Zum einen würden Aussagen, die dem Antext entsprechen, bei der Kategorie zur Formulierung der Ideen-Oberkategorien mit nn kodiert werden, weil die Aussage des Antextes ist, dass eben gerade kein Bezug von Naturwissenschaftlichkeit und Vorhandensein des Fachbezugs vorliegt. Aus Tabelle 6.25 geht aber hervor, dass Zuordnungen von Naturwissenschaftlichkeit und Vorhandensein des Fachbezugs grundsätzlich vorgenommen werden. Zum anderen wirft die Tabelle 6.25 die Frage auf, warum sowohl die Hauptdiagonale (Fachbezug hängt positiv mit Naturwissenschaftlichkeit zusammen) als auch die Gegendiagonale (Fachbezug ist ein Ausschlusskriterium für Naturwissenschaftlichkeit der Frage) nicht unwesentlich besetzt sind. Eine Betrachtung der Ideen-Kategorien, die zur Ideen-Oberkategorie **Fachbezug** gehören, zeigt, dass die Schülerinnen und Schüler bei manchen Teilaufgaben Bezüge zu Naturwissenschaften benennen (z. B. die Ideen-

Tabelle 6.25: Kodierte Formulierungen für **Fachbezug** bei der Aufgabenkarte 1-21.

Fragestellung ist naturwissenschaftlich	Fachbezug	
	liegt vor	liegt nicht vor
Ja	n+: 6	n–: 2
Nein	u+: 4	u–: 4
Unklar	nn: 4	

Kategorie +Physik) und einen positiven Bezug zur Naturwissenschaftlichkeit der Frage herstellen, bei anderen Teilaufgaben Bezüge zu nicht-naturwissenschaftlichen Fachdisziplinen (z. B. +Pädagogik) herstellen und einen negativen Bezug zur Naturwissenschaftlichkeit der Frage herstellen. Die Zuordnung der Ideen-Kategorien zur Haupt- bzw. Gegendiagonalen sind sehr eindeutig. Dieses Ergebnis stützt die Deutung, dass die Schülerinnen und Schüler keine Vorstellungen im Sinne des Antextes der Aufgabenkarte 1-21 äußern. Eine inhaltliche Betrachtung der vier mit nn kodierten Instanzen zu **Fachbezug** im Kontext der Aufgabenkarte zeigt, dass sogar hier nur in einem Fall eine Äußerung vorliegt, die dem Antext nahekommt: »Aber das ist halt so, dass du auch im Alltag Sachen hast, so wie, ähm/ (.) wie so was (zeigt auf Wasserbehälter auf dem Tisch) (.) und das ist dann auch schon Naturwissenschaftliches, was aber auch im Alltag vorkommen kann« (G21, Abs. 11).

Die Ideen-Oberkategorie **Effekt** findet sich, wenn sie in den Bearbeitungsprozessen eines jeweiligen Teams nur ein einziges Mal auftritt, bei 1-22-i-iv, 1-22-ii, 1-25-A (jeweils für ein Team), 1-26-C (für zwei Teams). Die Ideen-Oberkategorie **NawiEigenschaften** bei einmaligem Auftreten jeweils bei 1-21-ii (für ein Team), 1-22-ii (für zwei Teams) und 1-24-i (für ein Team). Beide Ideen-Oberkategorie scheinen also nicht auf eine bestimmte Aufgabe bezogen zu sein, weil die einmalige Aktivierung für verschiedene Teams bei verschiedenen (Teil-)Aufgaben stattfindet.

Obwohl die Ideen-Kategorien **Effekt** und **NawiEigenschaften** in den Daten nicht deutlich vermehrt bei spezifischen Aufgaben kodiert sind, lassen sich die existenten minimalen Häufungen inhaltlich gut nachvollziehen und wären bei größerer Stichprobe demnach vermutlich stärker ausgeprägt: Die Ideen-Oberkategorie **Effekt** tritt bei 1-26-C für zwei Teams auf, bei anderen Teilaufgaben nur für ein Team. Bei der Teilaufgabe 1-26-C geht es darum die Beispielfrage »Muss man ein Ei bei Vollmond länger kochen, damit es hart wird?« einzuschätzen. Eine typische Einschätzung lautet: »Weil es einfach nicht so ist« (O43). Vorerfahrungen der Schülerinnen und Schüler scheinen dazu zu führen, dass häufig fachinhaltliche Überlegungen stattfinden (mehr dazu im Kontext von F-Vor4), die dann auch die Begründung prägen. Das scheint sich in der Ideen-Oberkategorie **Effekt** widerzuspiegeln. Die Ideen-Oberkategorie **NawiEigenschaften** tritt bei 1-22-ii für zwei Teams auf, bei anderen Teilaufgaben nur für ein Team. Teilaufgabe 1-22-ii ist die bereits oben erläuterte Aufgabenstellung, in der ganz offen nach Begründungen gefragt wird. Ein mehrfaches Vorkommen von **NawiEigenschaften** ist also nicht verwunderlich, da diese Ideen-Oberkategorie eine sehr allgemeine Beschreibung eines vagen, intuitiven Verständnisses darzustellen scheint.

Für die **Unterscheidung von Beobachtung und Deutung** sind die über die Aufgaben hinweg gehäuft auftretenden Ideen-Oberkategorien (d. h., solche, die für mindestens eine der Personen bei mindestens drei Teilaufgaben kodiert sind) in Tabelle 6.26 (auf S. 343) dargestellt. Nur die beiden Ideen-Oberkategorien **BegründungErklärung** und **Vermutung** kommen vor. Die Teilaufgaben 3-20-3 bis -5 (siehe Abbildung D.1 im Anhang auf S. 545) scheinen sich besonders anzubieten, um Äußerungen zu machen, aus denen sich die zu den Ideen-Oberkategorien **BegründungErklärung** und **Vermutung** gehörigen Vorstellungen rekonstruieren lassen. Die von den Lernenden als Beobachtungen oder Deutungen zu klassifizierenden Beispielsätze in 3-20-3 und -5 beinhalten jeweils ein Weil bzw. ein Wegen. In 3-20-4 wird als Beispielsatz eine Beobachtung formuliert. Eventuell stellt diese Reihung der Sätze einen Kontext bereit, in dem sich die Unterscheidung von Beobachtung und Deutung besonders gut erörtern lässt. Spannend ist außerdem, dass für H22 und B5 keine der Ideen-Oberkategorien bei der allgemein formulierten Frage 3-21-ii vorkommt, obwohl sie bei vorigen Aufgaben kodiert wurden.

Die beiden Ideen-Oberkategorien **BegründungErklärung** und **Vermutung** stellen gemeinsam mit **Sehen** und **Beobachten** auch insgesamt die häufigsten Ideen-Oberkategorien dar. Allerdings kommen **Sehen** und **Beobachten** nicht über die Aufgaben gehäuft für einzelne Personen vor. Vielmehr treten für eine Vielzahl an Personen, aber bei nicht vielen Teilaufgaben auf (insgesamt treten sie bei den Teilaufgaben 3-20-3 bis -6 und 3-21 sowie 3-21-ii auf). Um zu überprüfen, dass dieses Ergebnis nicht aufgrund eines Ungleichgewichts in den Aufgaben entstanden ist, wurde untersucht, wie viele Aufgaben nach Beobachtungen und wie viele nach Deutungen fragen. Außerdem wurde nachgezählt, wie viele Beispielsätze als Beobachtungen und wie viele als Deutungen einzuschätzen sind. Die Arbeitsaufträge sind ausgeglichen (3-20: Beobachtungen; 3-21: Deutungen; 3-24: Beobachtungen und Deutungen). Bis zur Aufgabenkarte 3-24, auf der deutlich mehr Deutungen als Beispielsätze zu finden sind, sind die Beispielsätze ebenfalls ausgeglichen. Die Häufung von Vorstellungen zu Deutungen haben daher zumindest keinen direkten Zusammenhang zu den untersuchten Oberflächenmerkmalen des Materials. Auch weitere Materialaspekte (z. B. Inhalte der Informationstexte) scheinen keinen Zusammenhang mit dem Ungleichgewicht zu haben. Vielmehr wird in Analysen außerhalb der vorliegenden Arbeit auf Seite der Lernenden zu untersuchen sein, ob grundsätzlich mehr Vorstellungen zu Beobachtungen vorhanden sind oder welche Faktoren bei Einschätzungsaufgaben dazu führen, dass vorrangig Vorstellungen zu Deutungen aus den Aussagen der Schülerinnen und Schüler hervorgehen.

Tabelle 6.26: Gehäuft auftretende und dreifach auftretende Ideen-Oberkategorien für die Unterscheidung von Beobachtung und Deutung.

Person Ideen-Oberkategorie	Teilaufgabe									
	3-20-3	3-20-4	3-20-5	3-20-6	3-21-i-4	3-21-i-5	3-21-i-6	3-21-ii	3-24-a	3-24-b
<i>Schülerin H22</i>										
Vermutung					x	x	x		x	
BegründungErklärung	x				x	x	x			
<i>Schüler P46</i>										
Vermutung										
BegründungErklärung	x	x	x	x				x		
<i>Schülerin B5</i>										
Vermutung	x		x							x
BegründungErklärung										
<i>Schüler K32</i>										
Vermutung										
BegründungErklärung								x	2x	

Insgesamt zeigt die inhaltliche Betrachtung der gehäuft auftretenden Ideen-Oberkategorien, dass vorrangig die auch insgesamt häufig vorkommenden Ideen-Oberkategorien über die Aufgaben hinweg gehäuft auftreten. Einige Ideen-Oberkategorien treten nur für eine oder zwei Personen gehäuft vor. Es treten bei Kennzeichen von naturwissenschaftlichen Fragen sowohl eher angemessene als auch eher unangemessene Ideen-Oberkategorien gehäuft auf. Bei der Unterscheidung von Beobachtung und Deutung scheinen vorrangig angemessene Ideen-Oberkategorien gehäuft vorzukommen, was aber aufgrund der Datenlage (fast nur angemessene Ideen-Oberkategorien) nicht optimal aussagekräftig ist.

6.8.4 Variabilität von Vorstellungen im Verlauf der Aufgabebearbeitung

Zuletzt wird in den Blick genommen, welche Hinweise auf Stabilität und Variabilität der Vorstellungen sich *im Verlauf der Aufgabebearbeitung* für die einzelnen Personen finden lassen. Dazu werden direkt aufeinander folgende Teilaufgaben miteinander abgeglichen. Eine hohe Stabilität der Vorstellungen in Bezug auf die Situationen würde sich darin widerspiegeln, dass nur an wenigen (idealerweise inhaltlich nachvollziehbaren) Stellen die Lernenden einen Wechsel von Verbalisierungen, die sich einer Vorstellung zuordnen lassen, zu Verbalisierungen vornehmen, die sich einer anderen Vorstellung zuordnen lassen. Eine hohe Variabilität und Situationsbezogenheit würde

hingegen darin Ausdruck finden, dass von (Teil-)Aufgabe zu (Teil-)Aufgabe vorrangig unterschiedliche Vorstellungen rekonstruiert werden.

Eine Einschränkung der Analysen ist an dieser Stelle bereits vorwegzunehmen: Nicht in jeder Situation, d. h. nicht zu jeder Aufgabe, liegen Äußerungen der Lernenden vor, die auf ihre spezifischen Vorstellungen schließen lassen. Dieser Einschränkung wurde auf zwei Weisen begegnet, die nachfolgend beschrieben sind. Es werden zwei unterschiedliche Maße für die Variabilität definiert.¹⁰¹

Die beiden Maße für die Variabilität lassen sich am Beispiel der Kodierung für N41 erläutern. In Abbildung 6.20 sind in den Zeilen die unterschiedlichen Ideen-Oberkategorien, in den Spalten die (Teil-)Aufgaben vermerkt. Die Diagramme für alle untersuchten Lernenden finden sich in Anhang F.5 (ab S. 577); für diskutierte Personen sind die Diagramme im weiteren Verlauf des Textes eingebunden. Grau gefüllte Kästen stehen dafür, dass für die Person bei der jeweiligen (Teil-)Aufgabe die jeweilige Ideen-Oberkategorie kodiert wurde. Schwarz gefüllte Kästen symbolisieren analog, wenn es die Kodierung zur jeweils kompletten Karte betrifft, weil keine genauere Zuordnung der Äußerung zu einer Teilaufgabe vorgenommen werden konnte (siehe Unterabschnitt 6.1.4 auf S. 257). Horizontale graue Linien, die mit Punkten enden, stehen für sogenannte *Ketten*, die im zweiten Aufzählungspunkt erläutert sind.

- Zum einen wurde für jede Person gezählt, wie häufig bei einer Teilaufgabe keine der Ideen-Oberkategorien kodiert wurde, die bei der letzten vorhergehende Teilaufgabe mit Ideen-Oberkategorie zugewiesen wurde. Für N41 wurde beispielsweise bei 1-21-iv die Ideen-Oberkategorie **Fachbezug** vergeben (siehe Abbildung 6.20). Bei 1-21-v wurde keine Ideen-Oberkategorie kodiert. Erst bei 1-21-vi wurde wieder eine Ideen-Oberkategorie vergeben – nämlich PF/AF. Da diese beiden Ideen-Oberkategorien verschieden sind, liegt ein sogenannter *harter Wechsel* vor. Für einen harten Wechsel dürfen die Ideen-Oberkategorien zweier Teilaufgaben keine Schnittmenge haben. Von Teilaufgabe 1-21-vii zu Teilaufgabe 1-22-ii liegt für N41 hingegen kein harter Wechsel vor, weil zwischen diesen Teilaufgaben keine weiteren Ideen-Oberkategorien vergeben wurden und bei 1-22-ii die Ideen-Oberkategorien **NawiEigenschaften** und **PF/AF** vergeben

¹⁰¹Es gäbe auch noch eine dritte – plausible, aber eher auf die Instruktion als auf die Lernenden gerichtete – Möglichkeit: Für jedes Paar von aufeinanderfolgenden Aufgaben werden alle Personen betrachtet, die bei beiden Aufgaben Ideen-Oberkategorien zugewiesen bekommen haben. Die Anzahl der Personen, die zu den beiden Aufgaben unterschiedliche Ideen-Oberkategorien zugewiesen bekommen haben, wird ins Verhältnis zur Anzahl der Personen gesetzt, die irgendwelche Ideen-Oberkategorien für beide Aufgaben zugewiesen bekommen haben. Dieses Verhältnis könnte Aufschluss darüber geben, wie unterschiedlich die aufeinanderfolgenden Aufgaben aus Sicht einer Welche-Vorstellungen-finden-sich-Brille sind. Allerdings müsste dafür vermutlich die Stichprobengröße deutlich größer sein als in der vorliegenden Arbeit.

6.8 Ausw. u. Ergebnisse zur Stabilität u. Variabilität der Vorstell. (F-Vor5)

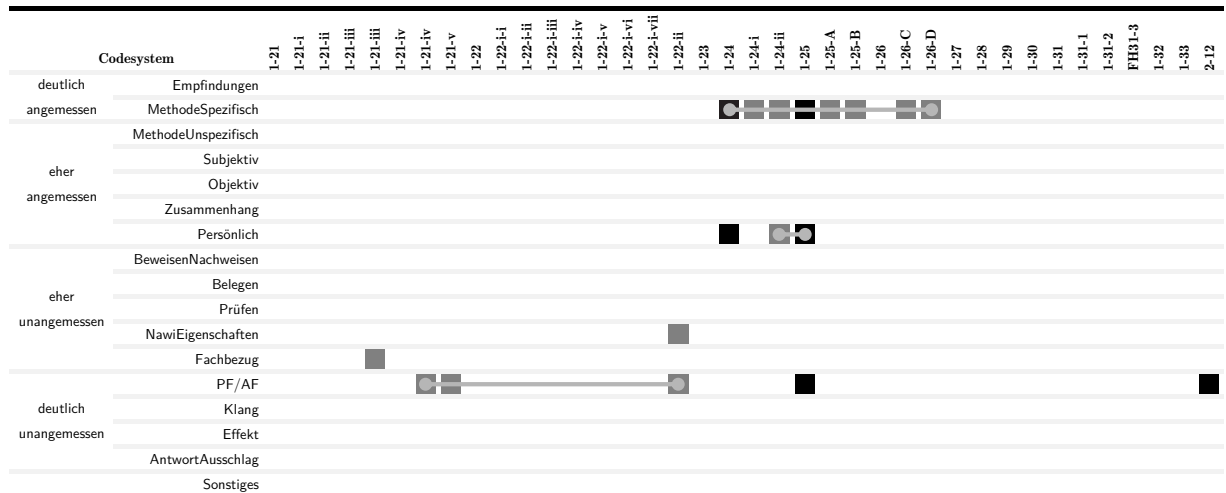


Abbildung 6.20: Ideen-Oberkategorien, die dem Schüler N41 bei den (Teil-)Aufgaben zugewiesen wurden (Ketten = hellgraue Linien).

wurden und somit mindestens eine der Ideen-Oberkategorien mit der Ideen-Oberkategorie bei Teilaufgabe 1-21-vii übereinstimmt. Dieser Wechsel wird als *weicher Wechsel* bezeichnet. Weiche Wechsel liegen bereits dann vor, wenn die Ideen-Oberkategorien zweier Teilaufgaben nicht vollständig übereinstimmen. Das heißt, dass zu Bestimmung weicher Wechsel gesucht wurde, wo bei einer Teilaufgabe mindestens eine andere Ideen-Oberkategorie kodiert wurde als bei der letzten vorhergehende Teilaufgabe *mit* Ideen-Oberkategorie (da Teilaufgaben *ohne* Ideen-Oberkategorie für die jeweils betrachtete Person in der Definition der Wechsel unbeachtet geblieben sind). Harte Wechsel sind somit immer auch weiche Wechsel, andersherum gilt dies nicht. Für nail8m liegen weiche Wechsel sowohl von Aufgabe 1-21-iv zu 1-21-vi als auch von Aufgabe 1-21-vii zu 1-22-ii vor.

- Zum anderen wurde der Einschränkung, dass Lernende nicht bei jeder Teilaufgabe durch ihre Äußerungen Rückschlüsse auf ihre Vorstellungen erlauben, so begegnet, dass *Ketten* von Vorstellungen gebildet wurden. Für eine fest gewählte Person besteht eine Kette von Teilaufgaben aus allen aufeinander folgenden Teilaufgaben, denen dieselbe Ideen-Oberkategorie zugewiesen ist, wobei Teilaufgaben, bei denen *zusätzlich* weitere Ideen-Oberkategorien vorliegen, als Bestandteile der Kette gewertet werden und Ketten auch die zwischenliegenden Teilaufgaben ohne Kodierung überdecken. Eine Kette geht also von einem harten Wechsel bis zum nächsten harten Wechsel. Beispielsweise liegt für N41 eine PF/AF-Kette von 1-21-vi bis 1-22-ii vor. Wäre allerdings zum Beispiel bei

1-22-i-iv die Ideen-Oberkategorie **Fachbezug** und nur diese vorhanden, würde die PF/AF-Kette nur von 1-21-vi bis 1-21-vii gehen.

Ketten werden unter zwei verschiedenen Gesichtspunkten analysiert. Als *Aufgaben-Länge* einer Kette wird die Anzahl der Teilaufgaben, die von der Kette überdeckt werden, bezeichnet. Dazu gehören also auch Teilaufgaben, bei denen keine Ideen-Oberkategorie kodiert wurde.¹⁰² Die *Instanzen-Länge* einer Kette zu einer Ideen-Oberkategorie X ist hingegen dadurch gegeben, wie häufig die Ideen-Oberkategorie X innerhalb der Kette vergeben wurde. Beispielsweise ist die Aufgaben-Länge der Kette, die für nai18 von 1-21-vi bis 1-22-ii geht, 10 (die Spalte für die Aufgabenkarte 1-22 wird nicht mitgezählt, die Kette überdeckt dann 10 Spalten); die Instanzenlänge derselben Kette ist 3 (weil drei graue Kästen von der Kette überdeckt werden).

Für **Kennzeichen von naturwissenschaftlichen Fragen** sind im Anhang in Tabelle F.15 (auf S. 578) die Anzahlen der weichen und harten Wechsel pro Person und die Anzahl der Aufgaben mit Ideen-Oberkategorien aufgeführt. Die daraus abgeleiteten *relativen Maße* für die Variabilität von Ideen-Oberkategorien berechnen sich als Verhältnisse von weichen bzw. harten Wechseln zur Anzahl der möglichen Wechsel (d. i. Anzahl der Aufgaben mit Ideen-Oberkategorien abzüglich der ersten Aufgabe, bei der kein Wechsel stattfinden kann). Der Mittelwert über das relative Maß zu weichen Wechseln ist 0.72 (mit einer Standardabweichung von 0.21). Dies lässt sich so deuten, dass die Personen im Mittel in etwa 70 % der Fälle die genutzte Ideen-Oberkategorie von einer Aufgabe zur nächsten Aufgabe (mit Vorstellungsäußerung) wechseln. Für harte Wechsel ergibt sich immerhin noch ein Anteil von mehr als 50 %. Diese stark zusammenfassenden Ergebnisse sprechen für eine mittlere bis hohe Variabilität der Vorstellungen der Lernenden im Kontext der untersuchten Aufgaben.

Eine Übersicht über die Längen der vorliegenden Ketten findet sich in Abbildung 6.21. Im Anhang ist aufgeführt, wie viele Ketten für jede einzelne Person vorliegen (Tabelle F.16 auf S. 579) und zu welchen Ideen-Oberkategorien diese gehören (Tabelle F.17 auf S. 580). Der Mittelwert der Anzahl von Ketten pro Person liegt knapp unter 2. Ein Zusammennehmen der beiden Kettenlängen-Maße spricht

¹⁰²Hierbei wurde eine technische Festlegung vorgenommen: Genau genommen gehören zu den zu gezählten Elementen nur echte (Teil-)Aufgaben, also nicht die Aufgabenkarten selbst. Fällt der Anfang oder das Ende einer Kette auf eine Aufgabenkarte, auf der auch bei einer (Teil-)Aufgabe die Ideen-Oberkategorie vergeben wurde, wird die Aufgabenkarte nicht mitgezählt; auch Aufgabenkarten innerhalb der Kette werden nicht mitgezählt; nur im Sonderfall, dass Anfang oder Ende auf eine Aufgabenkarte fällt, bei der keine (Teil-)Aufgabe mit der entsprechenden Ideen-Oberkategorie kodiert ist (bspw. 1-23), wird die Aufgabenkarte als eine Aufgabe gezählt, die das Maß für die Aufgaben-Länge der Kette um eins erhöht.

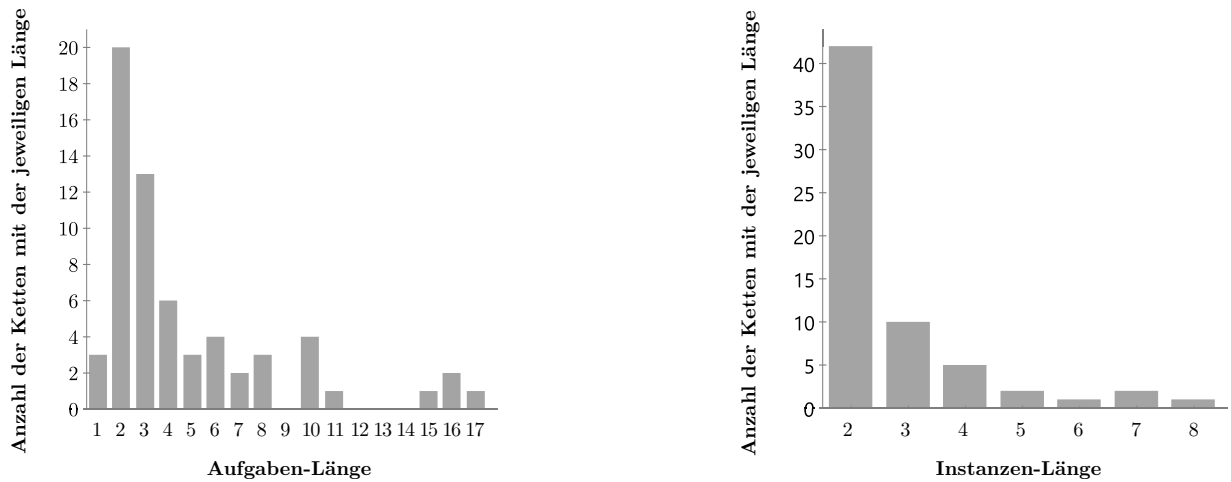


Abbildung 6.21: Häufigkeiten der beiden Kettenlängen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen. (Je länger die Ketten, desto höher wird die Stabilität der Vorstellungen der Lernenden eingeschätzt.)

dafür, dass die Ketten meist sehr kurz sind (nur wenige Instanzen, sehr häufig nur 2, enthalten und sich zudem über wenige Aufgaben, meist nur 2 oder 3, ausstrecken). Ein Vorliegen von Ketten ist generell eher als Hinweis auf Stabilität von Vorstellungen über die Aufgaben hinweg anzusehen; in den Daten finden sich aber wie bisher (z. T. nur exemplarisch) dargestellt eher Hinweise, dass die Stabilität der Vorstellungen über die untersuchten Situationen hinweg mäßig bis gering ist.

Dass sich für die Personen nur wenige Ketten finden und dass diese zumeist kurz sind, könnte ein Artefakt davon sein, dass die Instruktion so angelegt ist, dass gleiche Verständnisse nur über wenige Aufgaben hinweg genutzt werden können und vielmehr von Aufgabe zu Aufgabe unterschiedlich agiert werden muss. Personen, für die sich lange Ketten finden, wären also solche, die eine Vorstellung sehr stark vertreten und zusätzlich dazu keine Äußerungen machen, die auf irgendwelche weiteren Vorstellungen schließen lassen. Dies ist aber vor dem Hintergrund der Ergebnisse aus den Abschnitten 6.8.1 und 6.8.2 nicht plausibel, weil die Spektren an Ideen-Oberkategorien für viele Aufgaben groß sind und weil Lernenden zumeist mehrere Ideen-Oberkategorien zugewiesen wurden.

Die kurzen Ketten, derer für gewöhnlich nur wenige pro Person vorliegen, könnten so zu deuten zu sein, dass Personen aktivierte Vorstellungen noch eine Weile aktiv halten. Bei empfundener Nicht-Passung bei der nächsten Aufgabe werden die Vorstellungen dann zeitweise aufgegeben (und ggf. später reaktiviert). Für diese Deutung spricht auch, dass bei den Personen, die mehr Ketten als typisch vorweisen, auch häufig mehrere direkt aufeinander folgende Ketten auftreten. Es findet sich also in diesen

Fällen oft nicht nur eine kurze Kette, sondern mehrere, weil die entsprechenden Personen dazu neigen, ihre Vorstellungen länger als in der je spezifischen Situation aktiv zu halten und in nachfolgenden Situationen versuchen, diese erneut anbringen.

Fallbezogene Analysen der Ketten mit großer Aufgaben-Länge zeigen, dass viele davon Bezug zu den Aufgabenkarten 1-21 und 1-22 haben. Ein Beispiel findet sich in Abbildung 6.22. Die Aufgabenkarte 1-21 enthält mehrere einzuschätzende Beispielsätze. Die Aufgabenkarte 1-22 enthält Lösungen (1-22-i-i bis 1-22-i-vii), die als einzelne Teilaufgaben behandelt wurden. Diesem Umstand können einige der Aufgaben mit großer Aufgaben-Länge geschuldet sein. Dass die Teilaufgaben auf einer Karte verzeichnet sind, könnte dazu geführt haben, dass die Personen in der Realität die Teilaufgaben nicht einzeln bearbeitet haben, sondern alle Teilaufgaben mit denselben Ideen zu bearbeiten versuchten. Gleichwohl gibt es auch andere Ketten mit größerer Aufgaben-Länge, welche dann häufig auch große Instanzen-Länge haben. Ein Großteil dieser Ketten betrifft die Ideen-Oberkategorie *MethodeSpezifisch* und dabei die Aufgabenkarten 1-24 bis 1-31 (oder sogar bis 2-12). Das heißt, dass die Ideen-Oberkategorie, die am stärksten mit der im Lernmaterial intendierten Vorstellung verbunden ist, nach der instruktionalen Information in einer gewissen Stabilität vorgefunden werden kann (mehr dazu in Abschnitt 6.9, ab S. 353).

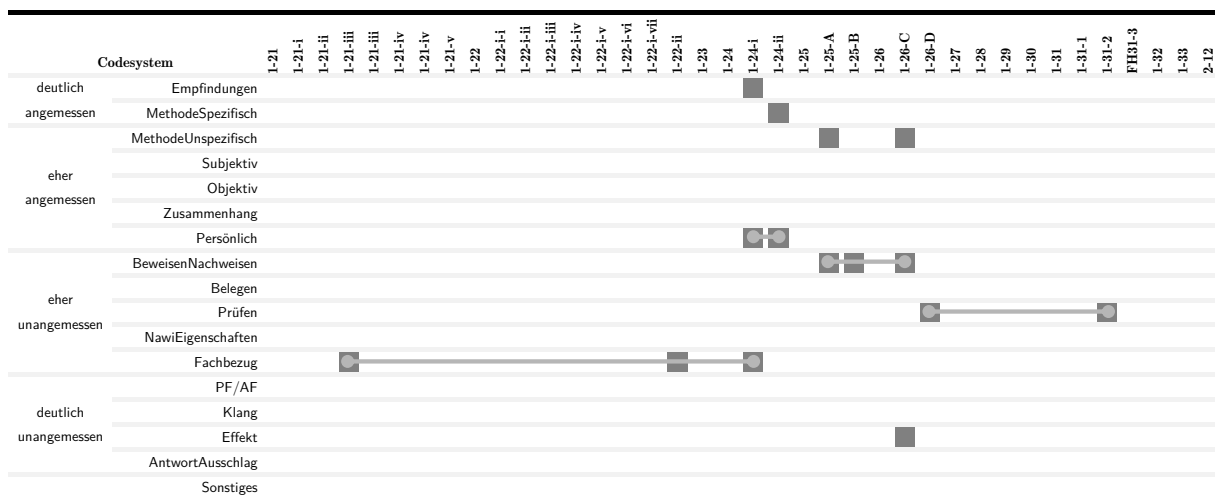


Abbildung 6.22: Ideen-Oberkategorien, die dem Schüler O43 bei den einzelnen (Teil-)Aufgaben zugewiesen wurden.

Neben dem Ergebnis der kurzen Ketten ist auch zu berichten, dass für viele Personen eine große Anzahl *freier Ideen-Oberkategorien* (also solche, die nicht in Ketten gebunden sind) vorhanden ist (siehe Tabelle F.16 im Anhang auf S.579). Diese Ideen-Oberkategorien, die nicht in Ketten eingebunden sind, wurden entweder bei einzelnen Aufgaben zusätzlich zu den geketteten Ideen-Oberkategorien kodiert

oder sie füllen Zwischenräume zwischen Ketten. Dieser Umstand spricht zusätzlich zu den Analysen der Ketten dafür, dass die Vorstellungen aller einzelnen Personen über die untersuchten Aufgaben hinweg nur mäßig bis wenig stabil sind. Für vier Personen liegen sogar einzig und alleine freie Ideen-Oberkategorien vor.

Für die **Unterscheidung von Beobachtung und Deutung** sind im Anhang in Tabelle F.18 (auf S. 581) die Anzahlen der harten Wechsel pro Person sowie die Anzahl der Aufgaben mit Ideen-Oberkategorien aufgeführt. Als Vorbemerkung ist zu machen, dass für neun Personen überhaupt keine Ideen-Oberkategorien vergeben wurden. Zusammenfassend lässt sich das Ergebnis festhalten, dass die Anzahl der harten Wechsel pro Person stark mit der Anzahl der Aufgaben einhergeht, für die den Personen jeweils Ideen-Oberkategorien zugewiesen wurden: Treten überhaupt Ideen-Oberkategorien auf, ist fast immer die Anzahl der harten Wechsel um genau eins geringer als die Anzahl der Aufgaben mit Ideen-Oberkategorien. Das liegt daran, dass für die meisten Personen *von Aufgabe zu Aufgabe unterschiedliche Ideen-Oberkategorien* zugewiesen wurden. Personen, die aus diesem Muster herausfallen – und die demnach mehr Stabilität in ihren Vorstellungen aufweisen als sich im Rest der Kohorte findet –, werden im Folgenden einzeln betrachtet.

Codesystem		AI20	AI20-1	AI20-2	AI20-3	AI20-4	AI20-5	AI20-6	AI21	AI21-i-1	AI21-i-2	AI21-i-3	AI21-i-4	AI21-i-5	AI21-i-6	AI21-ii	AI22	AI23	AI24	AI24-a	AI24-a-01	AI24-a-02	AI24-a-03	AI24-a-04	AI24-a-05	AI24-a-06	AI24-a-07	AI24-a-08	AI24-a-09	AI24-a-10	AI24-b	AI25
angemessen	Messen																															
	ObjektivSubjektiv																															
	Sehen																															
	Beschreiben																															
	AuswertungAntwort																															
	BegründungErklärung																															
	Interpretation																															
eher unangemessen	Vermutung																															
	Sonstiges																															
	Meinung																															
deutlich unang.	Beobachten																															
	FaktenSicher																															

Abbildung 6.23: Ideen-Oberkategorien, die der Schülerin H22 bei den einzelnen (Teil-)Aufgaben zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden.

Die Schülerin H22 (siehe Abbildung 6.23) hat für die drei direkt aufeinanderfolgenden Aufgaben 3-21-4, -5 und -6 die Ideen-Oberkategorien **Vermutung** und **BegründungErklärung** zugewiesen bekommen. Für B5 (Diagramm nur im Anhang, siehe Hägele, 2022, S. 104, Abb. F.35) wurde bei den Aufgaben 3-20-3 und 3-20-5 die Ideen-Oberkategorie **Vermutung** vergeben. Die jeweils kurzen Kettenlängen scheinen nicht dafür zu sprechen, dass hier stabile Vorstellungen vorliegen. Allerdings sollten die Längen der Ketten im Kontext der Unterscheidung von Beobachtung und Deutung

6 Vorstellungen von Lernenden

nicht zu starkes Gewicht in einem derartigen Argument erhalten, weil nahezu alle Ketten kurz sind. Außerdem geht aus den Diagrammen für die Personen hervor, dass für beide Schülerinnen die Ideen-Oberkategorie **Vermutung** bei einer späteren Aufgabe erneut zugewiesen wurde. Zusammengekommen spricht dies dafür, für H22 und B5 tatsächlich von einer eher stabilen Vorstellung auszugehen, die mit der kodierten Ideen-Oberkategorie **Vermutung** assoziiert werden kann.

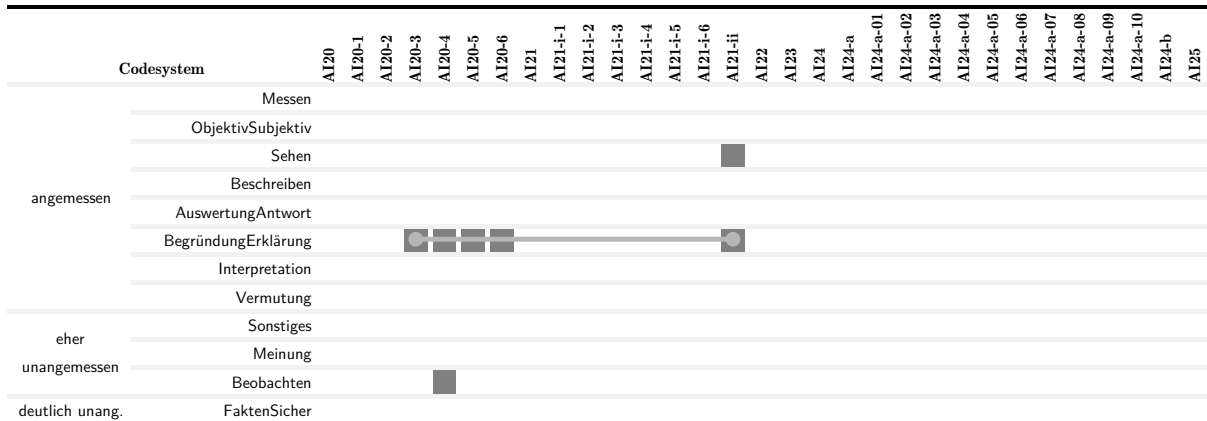


Abbildung 6.24: Ideen-Oberkategorien, die dem Schüler P46 bei den einzelnen (Teil-)Aufgaben zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden.

Bei P46 (siehe Abbildung 6.24) treten nur weiche Wechsel auf (zwei Stück: von 3-20-3 zu 3-20-4 und von 3-20-6 zu 3-21-ii). Weiche Wechsel treten im Übrigen auch nur bei nae13m auf und nicht bei anderen Personen. Die weichen Wechsel kommen dadurch zustande, dass P46 bei allen Aufgaben, bei denen ihm Ideen-Oberkategorien zugewiesen wurden, mindestens die Ideen-Oberkategorie **BegründungErklärung** erhalten hat. Zu dieser Ideen-Oberkategorie liegt daher auch eine Kette der Aufgaben-Länge 11 und der Instanzen-Länge 5 vor. Für P46 könnte von einer über die analysierten Aufgaben hinweg gehäuft auftretenden Vorstellung gesprochen werden, die zur Ideen-Oberkategorie **BegründungErklärung** gehört. Gleichwohl bringt P46 die Vorstellung nicht mehr bei späteren Aufgaben an. (Der Schnitt entspricht genau der Informationskarte – mehr dazu bei F-Vor6 ab Seite 50.)

Für L36 (Diagramm nur im Anhang, siehe Hägele, 2022, S. 109, Abb. F.49) treten nur zwei Aufgaben mit Ideen-Oberkategorien auf (3-20-3 und 3-21-ii). In beiden Fällen wurde **Sehen** zugewiesen. Da zwischen dem Auftreten der Ideen-Oberkategorie die Aufgaben zur Einschätzung des Textes bearbeitet wurden (wobei jeder Satz als eigene Teilaufgabe gezählt wurde), ist die Aufgaben-Länge der Kette deutlich überschätzt. Im Zusammenspiel damit, dass die Kette nur eine Instanzen-Länge

6.8 Ausw. u. Ergebnisse zur Stabilität u. Variabilität der Vorstell. (F-Vor5)

von 2 hat, wird daher eher davon ausgegangen, dass L36 nur Vorstellungen zum Beobachten verbalisiert hat oder verbalisieren konnte.

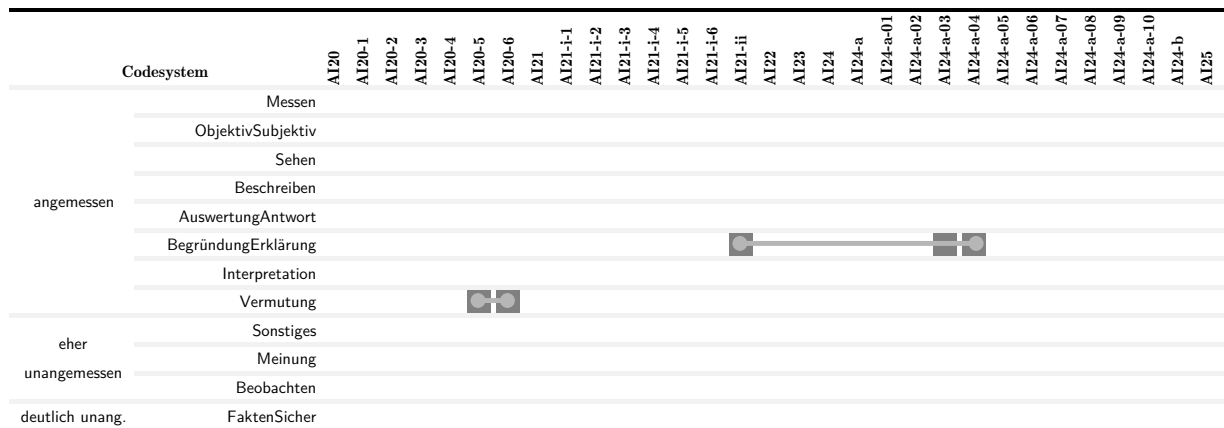


Abbildung 6.25: Ideen-Oberkategorien, die dem Schüler K32 bei den einzelnen (Teil-)Aufgaben zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden.

Der Schüler K32 (siehe Abbildung 6.25) stellt einen besonderen Fall dar, weil zwei Ketten vorliegen: Eine Kette zu **Vermutung** von 3-20-5 bis 3-20-6 und eine Kette zu **BegründungErklärung** von 3-21-ii bis 3-24-a-04. Damit liegt das einzige Mal eine im Sinne der vorgenommenen Definitionen direkt auf eine andere Kette folgende Kette innerhalb der Analysen zur Unterscheidung von Beobachtung und Deutung vor (weil zwischen den Ketten keine freie Ideen-Oberkategorie kodiert ist). Zwar ist die **Vermutung**-Kette recht kurz, dennoch könnte (mit Vorsicht und mit Einschränkungen) von einem Tausch einer gehäuft auftretenden (quasi stabilen) Vorstellung mit einer nächsten gesprochen werden.

6.8.5 Zusammenschau zu Stabilität und Variabilität der Vorstellungen über Aufgaben hinweg

Die Vorstellungen der Lernenden zum experimentbezogenen Denken und Arbeiten variieren von Situation zu Situation. Für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen weichen die Vorstellungsspektren, die sich für die einzelnen Teilaufgaben finden lassen, substantiell (aber nicht vollständig) voneinander ab (siehe Unterabschnitt 6.8.1 ab S. 326). Außerdem geht größer Stabilität in den Vorstellungen einer Person nicht mit einer geringeren Anzahl an Vorstellungen einher (was sich als Fokussieren auf spezifische Vorstellungen deuten ließe), sondern ggf. sogar eher mit einer größeren Anzahl an verschiedenen Vorstellungen (siehe

Unterabschnitt 6.8.2 ab S. 330). Ferner finden sich vorrangig kurze Ketten von für je einzelne Personen über mehrere Aufgaben hinweg ununterbrochen rekonstruierten Vorstellungen (siehe Unterabschnitt 6.8.4 ab S. 343). Von Situation zu Situation finden sich außerdem etwas mehr Wechsel als Beibehaltungen. Für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung wurden für einige Personen überhaupt keine Ideen-Oberkategorien zugewiesen; für die übrigen Personen finden sich – auch vor dem Hintergrund der nicht allzu großen Gesamtzahl an Ideen-Oberkategorien von 10 Stück¹⁰³ – eher wenige (meist zwischen 3 und 5) verschiedene Ideen-Oberkategorien (siehe Unterabschnitt 6.8.2 ab S. 330). Besonders eindrücklich ist, dass trotz der geringen Anzahl an verschiedenen Ideen-Oberkategorien pro Person für nahezu jede der Personen (20 von 33) von Aufgabe zu Aufgabe die zugewiesene Ideen-Oberkategorie wechselt (siehe Unterabschnitt 6.8.4 ab S. 343). Die wenigen Ausnahmen sind in Unterabschnitt 6.8.4 dargestellt. Die Ergebnisse zu den beiden Instruktionsextrakten weichen also einerseits substantiell voneinander ab, deuten andererseits allerdings beide darauf hin, dass die aus den Äußerungen der Lernenden rekonstruierbaren Vorstellungen von Situation zu Situation unterschiedlich sind und eine große Variabilität in den Vorstellungen vorhanden ist, was besonders vom dem Hintergrund der jeweils kurzen Instruktionsextrakte bemerkenswert ist.

Das Ergebnis, dass vorrangig Fluktuation vorliegt und nur wenige stabile Elemente für einzelne Personen identifiziert werden können, könnte so gedeutet werden, dass vorrangig statistisches Rauschen abgebildet wird. Dann wäre bei größerer Stichprobe und anderen Untersuchungsmethoden (z. B. längere Zeitspanne, häufigere Vorstellungsabfrage) damit zu rechnen, dass mehr Stabilität aufgefunden wird. Dass das Vorliegen von Stabilität über Situationen hinweg bezüglich der Vorstellungen für die untersuchten Personen nicht der Normalfall ist, kann allerdings vermutlich mit großer Wahrscheinlichkeit als eigener Befund angesehen werden. Dafür spricht unter anderem, dass aus den Daten hervorgeht, dass in fast allen Situationen sehr breite Spektren von Vorstellungen möglich sind (siehe Unterabschnitt 6.8.1 ab S. 326). Lernende könnten also prinzipiell von Situation zu Situation dieselben Vorstellungen nutzen, tun es aber nicht. Eine andere Erhebungsmethode, mit der für jede Person zu jeder Situation ein umfangreicheres Bild über die Vorstellungen generiert würde (bspw. ein Interview), könnte zwar weitere Vorstellungen offenlegen, aber nichts an dem Befund ändern, dass Lernende instantan verschiedene Vorstellungen in verschiedenen Situationen generieren.

¹⁰³Die Ideen-Oberkategorien (Messen) und Sonstiges wurden nicht mitgezählt.

6.9 Auswertungen und Ergebnisse zu Vorstellungen vor und nach instruktionalen Informationen

Die zu den Vorstellungen zum experimentbezogenen Denken und Arbeiten formulierte Forschungsfrage bezüglich der instruktionalen Informationen lautet:

F-Vor6 Welche Unterschiede u. Gemeinsamkeiten gibt es zwischen den Vorstellungen der Lernenden in Phasen vor und nach instruktionalen Informationen zu Konzepten des experimentbez. Denkens und Arbeitens?

Die für die Arbeit betrachteten Instruktionsextrakte beinhalten den Auswahlkriterien entsprechend jeweils eine ausführliche instruktionale Information, die mindestens eine Karte des Instruktionsmaterials ausfüllt und ein eigenständiges Element der Instruktion darstellt. Da auf den jeweiligen Karten das jeweils adressierte Konzept genannt, erläutert und mit Beispielen veranschaulicht wird, werden sie in der vorliegenden Arbeit teilweise auch als (instruktionale) Explizierungen bezeichnet. Für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen findet sich die instruktionale Information im Anhang in Abbildung C.2 (S. 542) und ist außerdem später auf Seite 361, wo sie zur Diskussion der Ergebnisse benötigt wird, als Abbildung 6.26 aufgeführt. Für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung umfasst die instruktionale Information die beiden Karten, die im Anhang in Abbildung D.2 (S. 546) zu finden und außerdem später auf Seite 377 für die Diskussion der Ergebnisse als Abbildung 6.31 dargestellt sind.

So, wie die Instruktionsextrakte gewählt sind, umfasst die analysierte Phase *vor* der instruktionalen Information für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen die Karten 1-21 – 1-22 (im Anhang in Abbildung C.1 auf S. 541) und für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung die Karten 3-20 – 3-21 (im Anhang in Abbildung D.1 auf S. 545). In beiden Fällen umfassen diese Karten Explorationsphasen mit Aufgaben, bei denen Einschätzungen von (systematisch angelegten) Beispielfragen oder Beispielsätzen gefordert werden. Für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen lässt sich die Phase *nach* der instruktionalen Information, die Übungs-, Anwendungs- und Wiederholungsaufgaben umfasst, in drei Zeitfenster unterteilen: Beim Zeitfenster *direkt nach* der instruktionalen Information handelt es sich um die ohne zeitlichen oder inhaltlichen Abstand folgenden Karten 1-24 – 1-27 der explizit-fachmethodischen Instruktionsvariante (also alle Karten des ersten Segments, die auf die Explizierung folgen, siehe Abbildung 6.7 auf S.270). Bei dem Zeitfenster *später nach* der instruktionalen Information handelt es sich um die in einem geringem Abstand innerhalb der ersten Einheit der explizit-fachmethodischen Instruktions-

variante auf das erste Zeitfenster folgende Karte 1-31 (d.i. Segment 2). Bei dem Zeitfenster *deutlich später nach* der instruktionalen Information handelt es sich um die erst in der nächsten Einheit der explizit-fachmethodischen Instruktionsvariante platzierte Karte 2-12, auf der eine kurze Wiederholung zu Kennzeichen von naturwissenschaftlichen Fragen angelegt ist (d.i. Segment 3). Für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung besteht die Phase der Übungs-, Anwendungs- und Wiederholungsaufgaben nach der instruktionalen Information nur aus dem Zeitfenster, das *direkt nach* dieser liegt (Karten 3-24–3-25 aus Segment 1).

In einer ersten Analyse werden die Phasen jeweils zusammengefasst betrachtet (Unterabschnitt 6.9.1). Es wird dabei aufgelöst, in welchen und in wie vielen Phasen welchen Personen welche Ideen-Oberkategorien zugewiesen sind. Nicht aufgelöst wird allerdings die Anzahl der Teilaufgaben innerhalb der Phasen. Beispielsweise wird nicht unterschieden, ob eine Ideen-Oberkategorie einer Person bei zwei verschiedenen Teilaufgaben aus der Phase *direkt nach* der instruktionalen Information zugewiesen ist oder ob sie ihr nur bei einer Teilaufgabe aus derselben Phase zugewiesen ist. Da beide analysierten Instruktionsextrakte im jeweils ersten Segment vergleichbar angelegt sind, werden schwerpunktmäßig die Phasen direkt vor und direkt nach der Explizierung analysiert.

In weiteren Analysen werden die einzelnen Teilaufgaben aufgelöst, indem die zur Beantwortung der Forschungsfrage F-Vor5 eingeführten *gehäuft auftretenden* Ideen-Oberkategorien (Unterabschnitt 6.8.2 ab S. 330) sowie die eingeführten *Ketten* von über mehrere Aufgaben ununterbrochen vorkommenden Ideen-Oberkategorien (Unterabschnitt 6.8.4 ab S. 343) zur Analyse genutzt werden. Die Ergebnisse zu diesen Analysen finden sich in den Unterabschnitten 6.9.2 und 6.9.3 (d. h. ab S. 380).

6.9.1 Vergleich der Phasen vor und nach den instruktionalen Informationen

6.9.1.1 Für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen

In Tabelle 6.27 ist für Kennzeichen von naturwissenschaftlichen Fragen aufgelöst, welchen Personen welche Ideen-Oberkategorien in der Phase vor der instruktionalen Information bzw. in der Phase nach der instruktionalen Information zugewiesen wurden. Die drei Zeitfenster nach der instruktionalen Information wurden dabei (zunächst) als ein Gesamtes, d. h. als eine Phase, betrachtet, um die Komplexität zunächst zu reduzieren. Eine zeitlich detailliertere Auflösung für alle Personen und Ideen-Oberkategorien findet sich im Anhang in Tabelle F.19 auf S. 583; auf ausgewählte Personen und Ideen-Oberkategorien wird weiter unten eingegangen.

Tabelle 6.27: Auflösung der Ideen-Oberkategorien, die den einzelnen Personen jeweils vor bzw. nach der instruktionalen Information zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden (sortiert nach Angemessenheit der Ideen-Oberkategorien sowie nach Team-Zugehörigkeit der Personen).

Person	Antwort/Ausschlag				deutlich unangemessen				eher unangemessen				eher angemessen				Methodespezifisch		Empfindungen	Anzahl Vor	Anzahl Nach													
	Effekt	Klang	PT	Fachbezug	NaWiEigenschaften	Prüfen	Belegen	Beweisen/Nachweisen	Persönlich	Zusammenhang	Objektiv	Subjektiv	MethodenUnspezifisch	deutlich ang.																				
A1				V	V	V			V	N	V			N	V	5	1																	
A2			V	N	V				V	V				N	V	5	3																	
A3						N								N		0	2																	
B4						N			N	N			N			1	4																	
B5					V	N	N	V	N	V			N			3	5																	
B6		N			V	N			N	N			N			1	5																	
C7									N	N			V			2	3																	
C8			V		V				N	N				N		2	3																	
C9					V				N				N			2	2																	
F16	V							N		V	N	V	N	N	V	6	5																	
F17						N	V	N		V	N		N		N	2	4																	
F18						N		N		V	N	V	N			3	5																	
G19			N		V	N								N		1	3																	
G20					N				N		N	N	N		N	0	5																	
G21		N			V	N					N				N	1	4																	
H22					V	N				N			N			3	6																	
H23					V	N				N	N	V	N		V	4	6																	
H24	V				V					N		V	N		V	5	4																	
I25	V	N				N				V			V			4	5																	
I26						N	N			V	N	V	N			3	3																	
I27	V				V	N										2	1																	
M37		N			V	N							N			1	5																	
M38	N					N	N		N		N	N	N		N	0	8																	
M39					V									N		1	1																	
N40	N	N	V	N	V	N			V	N				N		5	6																	
N41			V	N	V		V			N				N		3	3																	
N42				V										N		1	1																	
O43		N			V	N			N	N			N	N	N	1	8																	
O44					V	N	N			N	N		N	N		2	6																	
O45	N	N			V	N	N		N	N			N			1	6																	
P46	V	N			V				V	N			N			3	6																	
P47	V						N		V	V			N	N		3	5																	
P48	V									V	N	N	N	N		2	2																	
Anzahl der Personen	5	4	3	6	3	4	4	2	18	7	10	10	0	13	3	4	1	10	11	16	6	0	4	9	3	9	2	15	0	21	5	6	2,36	4,12
	V: 14 N: 13				V: 22 N: 27				V: 17 N: 28				V: 5 N: 24				1,57	1,86																

Anmerkungen. Durch halbfetten Satz hervorgehoben sind Ideen-Oberkategorien, die auch für den Text der Explizierung kodiert wurden. Die Anzahl der Personen entspricht allen Zeilen, die im entsprechenden Spalten-Bereich (jeweilige Ideen-Oberkategorie bzw. jeweilige Gruppe von Ideen-Oberkategorien zu einer spezifischen Angemessenheit) mindestens einen Eintrag haben.
V= Vor der Explizierung, N = Nach der Explizierung.

Eine zentrale Frage, die sich bezüglich der vor und nach der instruktionalen Information rekonstruierbaren Vorstellungen stellt, ist, wie viel Kontinuität zwischen den beiden Phasen existiert. Es lassen sich die folgenden Ergebnisse berichten:

Vorstellungen, die vor der instruktionalen Information vorliegen, liegen grundsätzlich auch danach vor. Aus Tabelle 6.27 (auf S. 355) geht hervor, dass die Anzahl der Personen, denen eine Ideen-Oberkategorie zugewiesen wurde, zwar nach der instruktionalen Information (teilweise deutlich) geringer sein kann, jedoch nur im Fall von **Zusammenhang** vor der instruktionalen Information größer Null und danach Null ist. Dies lässt sich so deuten, dass Vorstellungen, die vor der instruktionalen Information vorliegen, grundsätzlich auch nach der instruktionalen Information vorgefunden werden.

Dass sich für die Ideen-Oberkategorie **Zusammenhang** abweichend von den anderen Ideen-Oberkategorien nach der instruktionalen Information für keine der Personen rekonstruiert wird, obwohl sie sich vor der instruktionalen Information vorfindet, kann eventuell damit plausibilisiert werden, dass nach der instruktionalen Information andere Ideen-Oberkategorien zugewiesen sind, die aus Sicht der Lernenden ähnliche inhaltliche Akzente setzen. Aus Beobachtersicht liegen für die Ideen-Oberkategorie **Zusammenhang** nach der instruktionalen Information Situationen vor, in denen die zugehörige Vorstellung angebracht werden könnte; allerdings zeigt eine nachträgliche Durchsicht der gesamten Transkripte und Videos aller Personen, bei denen die Ideen-Oberkategorie (vor der instruktionalen Information) vergeben wurde, dass diese Personen nach der instruktionalen Information auch innerhalb der Äußerungen, denen anderen Ideen-Oberkategorien zugewiesen wurden, grundsätzlich nicht auf Aspekte der Zusammenhänge anzuspielen scheinen, auch nicht subtil. Einzig für das Team H lässt sich – auf recht weitem Umweg – auf Team-Ebene in den Äußerungen, die mit **Zusammenhang** kodiert wurden, ein Aspekt der *Beantwortung* von Fragen erkennen, der auch später aufgegriffen wird – allerdings ohne dass dabei auf Zusammenhänge von Variablen oder Größen eingegangen wird. Die entsprechenden Transkriptstücke sehen so aus (Kursivierung von JJH):

H23: Eine naturwissenschaftliche Frage, äh, *sucht* nach (H22 setzt den Stift an und beginnt zu schreiben) einer, (.) ähm, *Erklärung* für

H24: bestimmte Phänomene?

H23: Ähm, naturwissenschaftlich //basierte Phänomene//

H24: //Ja.// H23: *unbeeinflusst* von (.) weiteren Faktoren wie (H22 hört auf zu schreiben)

Team H (Abs. 80–84 im Anhang, siehe Hägele, 2022, S. 36
– vor der instrukt. Info.)

H22: Wie wär es denn mit dem *Sucht*? Ich fand das gut. *Sucht eine Antwort* auf die | auf eine

H24: Auf eine naturwissenschaftliche (..) (lachend) Frage.

H22: Auf ein naturwissenschaftliches Phänomen.

H24: Ja, genau.

Team H (Abs. 100–103, Anhang auf S. 37 – vor der instrukt. Info.)

H22: (Schreibt) (..) Weil (..) man es messen kann. (Lacht)

H23: (Lachend) Ja. (..) Weil es mit zwei einfachen Messungen zu *beantworten* ist, um mal was anderes zu schreiben.

Team H (Abs. 181–182, Anhang auf S. 39 – nach der instrukt. Info.)

In allen drei Transkriptstücken wird darauf eingegangen, dass naturwissenschaftliche Fragen nach Antworten suchen bzw. sich beantworten lassen; nur im ersten Transkriptstück wird die Ideen-Oberkategorie **Zusammenhang** zugewiesen (weil aufgrund des Wortes »unbeeinflusst« darauf geschlossen wird, dass die Schülerin H24 naturwissenschaftliche Fragen mit Zusammenhängen von Größen in Verbindung bringt, die unbeeinflusst von anderen Faktoren untersucht werden; vgl. Beschreibung der Ideen-Oberkategorie in Tabelle 6.12 auf S. 290). Insgesamt scheint es so, dass Vorstellungen, die mit **Zusammenhang** kodiert wurden, nach der instruktionalen Information nicht verbalisiert werden – eventuell weil genügend (besser) passende Alternativen vorliegen. Eine Gemeinsamkeit in den Ideen-Oberkategorien nach der instruktionalen Information (d. h. den gewählten Alternativen) lässt sich für die untersuchten Personen aber aus Tabelle 6.27 nicht ableiten.

Für einen Großteil der Personen lassen sich die Vorstellungen, die sich vor der instruktionalen Information rekonstruieren lassen, nach der instruktionalen Information nicht erneut rekonstruieren. Wird in Tabelle 6.27 (S. 355) markiert, wann einer je spezifischen Person eine Ideen-Oberkategorie vor der instruktionalen Information *und auch* danach zugewiesen wurde (Beispiel: A2 und Klang; Gegenbeispiel: A1 und Effekt), ergibt sich, dass dies 29-mal auftritt. Vor der instruktionalen Information liegen 78 Instanzen vor (mit V gefüllte Zellen in der Tabelle). Personen wird also in ca. 37% all dieser Fälle eine Ideen-Oberkategorie sowohl vor als auch nach der instruktionalen Information zugewiesen.

Das Ergebnis ist vor dem Hintergrund des ab Seite 356 berichteten Ergebnisses zur gesamten Kohorte als eher unerwartet einzuschätzen. Eine mögliche Erklärung für das Ergebnis, dass Personen die vor der instruktionalen Information für sie rekonstruierten Vorstellungen nicht nach der instruktionalen Information zeigen, könnte lauten, dass andere Personen desselben Teams bereits entsprechende Äußerungen gemacht haben,

so dass die Kontinuität sich zwar nicht auf Ebene der Personen, aber auf Ebene der Teams zeigt. Dies ist allerdings nicht der Fall:

Für einen Großteil der Teams lassen sich die Vorstellungen, die sich für eine zugehörige Person vor der instruktionalen Information rekonstruieren lassen, nach der instruktionalen Information nicht erneut für eine der zugehörigen Personen rekonstruieren. Wird analog zur Betrachtung der Personen gezählt, in wie vielen Fällen eine Ideen-Oberkategorie einem *Team* vor bzw. vor und nach der instruktionalen Information zugewiesen ist (d. h. irgendeiner Person im Team), ergibt sich der Anteil auf Team-Ebene zu ca. 39 % (19 von 48).

In den Personen- bzw. Team-Verläufen scheint sich daher das auf Kohorten-Ebene oben berichtete Ergebnis nicht einzustellen: In vielen Fällen liegen Vorstellungen für eine Person bzw. ein Team zwar vor der instruktionalen Information vor, aber nicht danach. Die daran ersichtliche Diskrepanz zwischen Kohorten-Ebene und individuellen (bzw. team-spezifischen) Verläufen wird auch in nachfolgend dargestellten Analysen immer wieder deutlich.

Eine weitere Frage zum Vergleich der Vorstellungen vor vs. nach der instruktionalen Information bezieht sich darauf, *welche* Vorstellungen sich jeweils rekonstruieren lassen. Beispielsweise lässt sich dabei einerseits danach fragen, ob und wie sich die Angemessenheit der Vorstellungen im Vergleich der Phasen darstellt, und andererseits, ob und inwiefern die Vorstellungen nach der instruktionalen Information stärker dem Inhalt der Explizierung entsprechen als vor der instruktionalen Information.¹⁰⁴

Die Vorstellungen der Personen sind nach der instruktionalen Information angemessener als vor der instruktionalen Information. Zur Untersuchung der Angemessenheit werden die Einträge in Tabelle 6.27 bepunktet. Für ein eingetragenes V oder N im Bereich der deutlich unangemessenen Ideen-Oberkategorien erhält eine Person beispielsweise –2 Punkte (eher unangemessen: –1, eher angemessen: 1, deutlich angemessen: 2). Die Punktzahlen der Personen sind im Anhang dargestellt (Tabelle F.20 auf S. 584). Die Angemessenheit der Ideen-Oberkategorien der Personen ist vor der instruktionalen Information im Mittel negativ und signifikant von Null verschieden ($MW = -0.788$, $t(32) = -2.769$, $p = .009$, $r = .193$); nach der instruktionalen Information ist sie im Mittel positiv und ebenfalls signifikant von Null

¹⁰⁴Da die Angemessenheit vor dem Hintergrund weiterer Überlegungen und Literaturrecherchen eingeschätzt wurde, gehen in der vorliegenden Arbeit Angemessenheit einerseits und Übereinstimmung mit der Explizierung andererseits nicht eindeutig miteinander einher. Dies muss auch grundsätzlich nicht der Fall sein, weil eine instruktionale Information nicht immer alle als angemessen einzuschätzenden Vorstellungen adressieren muss, sondern sich nur auf ausgewählte Teile beziehen kann (und sollte, siehe Renkl, 2002).

verschieden ($MW = 0.818$, $t(32) = 2.224$, $p = .033$, $r = .134$). Die Differenzen der Angemessenheiten betragen im Mittel 1.61, was einem signifikant positiven Zuwachs mit mittlerer Effektstärke entspricht (gepaarter t-Test: $t(32) = 3.891$, $p < .001$, $r = .321$). Die Vorstellungen der Personen sind also nach der instruktionalen Information angemessener als vor der instruktionalen Information.

Angemessene Vorstellungen lassen sich nach der instruktionalen Information für mehr Personen rekonstruieren als vor der instruktionalen Information, unangemessene Vorstellungen nehmen allerdings zum Teil ebenfalls zu. Auf Ebene der gesamten Kohorte zeigt sich bezüglich der Angemessenheit ein schwieriger zu interpretierendes Bild. Zwar finden sich für eher angemessene und deutlich angemessene Ideen-Oberkategorien (abgesehen von **Zusammenhang**, bereits oben auf Seite 356 diskutiert) keine Ideen-Oberkategorien, die nach der instruktionalen Information weniger Personen zugewiesen wurden als vorher. Gleichzeitig finden sich aber auch (teilweise deutliche) Zunahmen in den Personenzahlen für eher bzw. deutlich unangemessene Ideen-Oberkategorien (siehe Tabelle 6.27 auf S. 355). Dies spiegelt sich darin wieder, dass nach der instruktionalen Information 27 Personen eine eher unangemessene Ideen-Oberkategorie erhalten haben, obwohl es vor der instruktionalen Information nur 22 Personen waren.

Zusammengenommen sprechen die Ergebnisse zur Angemessenheit dafür, dass die Angemessenheit der Vorstellungen nach der instruktionalen Information zwar tendenziell höher ist als davor, aber die Varianzen bezüglich der Personen und Ideen-Oberkategorien die Gesamtaussage einschränken. Positiv formuliert sind die Verschiedenheiten im Umgang der Personen mit der instruktionalen Information und die Verschiedenheiten der Auswirkung auf die Ideen-Oberkategorien als näher zu beleuchtenden Aspekte anzusehen. Die Diskrepanz der Ergebnisse könnte allerdings auch darin begründet liegen, dass die Gruppierung in eher/deutlich angemessene/unangemessene Ideen-Oberkategorien nicht optimal gelungen ist.

Zur Bestimmung des Grads der Übereinstimmung von den Vorstellungen der Lernenden mit den Informationstexten wurden diese mit dem finalen (induktiv generierten) Kategoriensystem der Ideen-Kategorien kodiert. Entsprechend ist eine Zuordnung von Ideen-Oberkategorien zu Sätzen der instruktionalen Information möglich; um auf diese zu verweisen, wird nachfolgend vereinfacht der Begriff der *explizierten* Ideen-Oberkategorien verwendet. Die Ideen-Oberkategorien werden also hinsichtlich der Tatsache bewertet, ob sie auch für die instruktionalen Information auftreten, d. h. die ihnen zugehörigen Vorstellungen Teil des explizierten Konzepts

sind. In Tabelle 6.27 (auf S. 355) halbfett gesetzte Ideen-Oberkategorien wurden mit einem Punkt bewertet, die Ideen-Oberkategorie **Fachbezug** wurde aufgrund des Antextes von Aufgabenkarte 1-21 mit minus einem Punkt bewertet (vgl. dazu bereits Unterabschnitt 6.6.1, ab S. 306). Damit lässt sich prüfen, ob Ideen-Oberkategorien, die mit den durch die Instruktion explizierten Konzepten übereinstimmen, nach der instruktionalen Information stärker auftreten als davor.

Die Vorstellungen der Lernenden stimmen nach der instruktionalen Information deutlich stärker mit dem explizierten Konzept überein als vor der instruktionalen Information. Vor der instruktionalen Information liegt nur eine sehr geringe Übereinstimmung mit den explizierten Ideen-Oberkategorien vor ($MW = 0.121$, nicht-signifikante Abweichung von Null mit sehr kleinem Effekt bei Einstichproben-t-Test: $t(32) = 0.611$, $p = .545$, $r = .012$), nach der instruktionalen Information eine höhere ($MW = 1.818$, Abweichung von Null mit Einstichproben-t-Test: $t(32) = 12.344$, $p < .001$, $r = .826$); eine genauere Dokumentation für jede einzelne Person findet sich im Anhang in Tabelle F.20 auf S. 584. Die Übereinstimmung mit explizierten Ideen-Oberkategorien ist auch nach der instruktionalen Information noch weit von der Maximalpunktzahl von 5 entfernt (Einstichproben-t-Test gegen 5: $t(32) = -21.602$, $p < .001$, $r = .936$), gleichzeitig liegt aber mit einer mittleren Differenz von 1.7 ein signifikant positiver Zuwachs von vor der instruktionalen Information zu danach vor (gepaarter t-Test: $t(32) = 7.577$, $p < .001$), der einem großen Effekt entspricht ($r = .642$). Diese Ergebnisse deuten darauf hin, dass die instruktionale Explizierung einen deutlichen Effekt auf das Auftreten von durch die Instruktion explizierten Ideen-Oberkategorien hat.

Auch wenn der weiter oben berichtete Wert für die mittlere Differenz der Angemessenheiten mit 1.61 (Stabw. = 2.335) in einer ähnlichen Größenordnung zu liegen scheint wie der im vorigen Absatz berichtete Wert dafür, wie groß die Übereinstimmung zu den instruktional explizierten Ideen-Oberkategorien ist ($MW = 1.7$, Stabw. = 1.267), ist der Effekt für die explizierten Ideen-Oberkategorien als deutlich stärker einzuschätzen. Dafür ist zu berücksichtigen, dass es hinsichtlich der Angemessenheiten keine Ideen-Oberkategorien gibt, die mit null Punkten bewertet wurden; deshalb entspricht eine mittlere Differenz von 1.61 nicht einmal einem Wechsel von einer eher unangemessenen zu einer eher angemessenen Ideen-Oberkategorie. Für explizierte Ideen-Oberkategorien ist hingegen eine mittlere Differenz von 1.7 dem Hinzukommen von ein bis zwei Ideen-Oberkategorien der Explizierung zu vergleichen.

Ähnlich wie die Bewertung der Ideen-Oberkategorien hinsichtlich ihrer Angemessenheit kritisch diskutiert werden kann, ist auch zu diskutieren, was es mit Ideen-

Oberkategorien auf sich hat, die der Explizierung zugewiesen wurden. Das Vorgehen, den Text der instruktionalen Information mit demselben Kodiersystem zu analysieren wie die Äußerungen der Lernenden, sichert die Vergleichbarkeit. Es lässt sich jedoch anmerken, dass die Explizierung (siehe Abbildung 6.26) im Kern nur ein einziges Konzept betont (welches sich auch in einer einzigen Vorstellung spiegeln könnte): *Naturwissenschaftliche Fragen können mit naturwissenschaftlichen Methoden untersucht werden*. Gegebenenfalls gehört zu dem Konzept noch der folgende Nachsatz hinzu: *und beziehen sich auf objektivierbare Ereignisse*. Das Konzept wird im Informationstext mit feineren Beispielen unterstützt. Dieses Konzept ist breiter, d. h. umfassender, als die Vorstellungen der Lernenden, die nur das Messen, nur den Geschmack usw. ansprechen. Ein Grund für die im Vergleich zum explizierten Konzept schmalen Vorstellungen der Lernenden könnte sein, dass diese die Beispiele nicht hinreichend mit dem zentralen Konzept vernetzen, sondern als alleinstehend begreifen. Ein weiterer Grund könnte das methodische Vorgehen sein, bei dem die Äußerungen der Lernenden in einer feingliedrigen Weise kodiert wurden. Betrachtungen der Transkriptausschnitte legen jedoch nahe, dass das methodische Vorgehen für die Schüleräußerungen durchaus angemessen ist, weil diese Äußerungen tatsächlich wenig vernetzt sind (und oft sogar nur alleinstehende Schlagworte enthalten).

Es kann nicht immer eindeutig entschieden werden, ob eine Fragestellung „naturwissenschaftlich“ ist oder nicht. **Ein zentrales Merkmal einer naturwissenschaftlichen Fragestellung ist jedoch, dass sie mit naturwissenschaftlichen Methoden untersucht werden kann.**

Naturwissenschaftliche Methoden sind Messungen/Beobachtungen/Zählungen von objektivierbaren Ereignissen. Nicht objektivierbare Ereignisse sind z. B. Geschmack, Meinung, Einstellung usw.

Beispiel-Frage 1: *Hat der Luftdruck einen Einfluss darauf, bei welcher Temperatur Wasser kocht?*

Der Luftdruck kann mit einem Barometer *gemessen* werden. Die Temperatur von Wasser am Siedepunkt kann mit einem Thermometer *gemessen* werden. Das Ergebnis beider Messungen ist unabhängig davon, wer die Messungen durchführt und deshalb *objektiv*.

✓ **Es handelt sich um eine naturwissenschaftliche Fragestellung.**

Beispiel-Frage 2: *Schmeckt Mineralwasser mit Kohlensäure besser als Leitungswasser?*

Ob etwas gut schmeckt oder nicht, ist *nicht objektiv messbar*, sondern hängt von dem Geschmack der befragten Personen ab.

✗ **Es handelt sich *nicht* um eine naturwissenschaftliche Fragestellung.**

Abbildung 6.26: Text der Explizierung zu Kennzeichen von naturwissenschaftlichen Fragen.

Für das Kodieren der instruktionalen Information hingegen, die ein schriftlicher Erklärungstext ist, scheint das methodische Vorgehen problematischer zu sein (und für Interview-Analysen dürfte es vermutlich ebenfalls problembehaftet sein, so dass zumindest eine Ergänzung durch stärker zusammenhangsbetonenden Rekonstruktionen erforderlich sein dürfte). Diese Ausführungen nahe, dass die Anwendung des Kodiersystems auf schriftliche Texte gewisse Nachteile hat, die die Aussagekraft bezüglich der Übereinstimmung mit explizierten Ideen-Oberkategorien einschränken können. Einerseits müssen für ein deutlich mit der Explizierung übereinstimmendes Verständnis einer Person nicht alle Ideen-Oberkategorien, die der instruktionalen Information zugewiesen wurden, auch der Person zugewiesen werden. Andererseits bildet auch keine der vergebenen Ideen-Oberkategorien für sich genommen das zentrale Konzept der Explizierung hinreichend gut ab (auch nicht *MethodeUnspezifisch* oder *MethodeSpezifisch*; siehe Ausführungen zur Angemessenheit bei F-Vor3), als dass ein Auftreten nur dieser Ideen-Oberkategorie für eine Person bereits als mit der Explizierung übereinstimmendes und in diesem Sinne angemessenes Verständnis gedeutet werden könnte.

Wird für die Untersuchung der Angemessenheit der Vorstellungen der Lernenden sowie der Übereinstimmung der Vorstellungen der Lernenden mit den der instruktionalen Information zugewiesenen Ideen-Oberkategorien für die **Phase nach der instruktionalen Information differenziert**, ob es sich um das Zeitfenster *direkt nach* der instruktionalen Information (Aufgabenkarten 1-24 – 1-26) oder das Zeitfenster mit zeitlichem Abstand *später nach* der instruktionalen Information (später in derselben Einheit des Lernmaterials: 1-31) oder das Zeitfenster *deutlich später* (d. h. in der nachfolgenden Einheit des Lernmaterials: 2-12) handelt, ergibt sich für die Effekte der Explizierung auf die Angemessenheit der Ideen-Oberkategorien sowie das Auftreten von explizierten Ideen-Oberkategorien ein zeitlich differenziertes Bild, das in Abbildung 6.27 für die jeweiligen Mittelwerte über die Personen dargestellt ist (und auf den im Anhang vorzufindenden Tabellen F.19, S. 583, und F.21, S. 585, basiert – analog zu Tabelle 6.27 auf S. 355).

Die instruktionale Information hat auf das Zeitfenster, das direkt auf sie folgt, einen Einfluss bezüglich der Vorstellungen – und allem Anschein nach auch auf die anderen Zeitfenster. Die in Abbildung 6.27 eingetragenen Mittelwerte sind bis auf die Mittelwerte für die Angemessenheit im späteren und im deutlich späteren Zeitfenster und den Mittelwert bei explizierten Ideen-Oberkategorien für die Phase vor der instruktionalen Information signifikant von

6.9 Ausw. u. Ergebnisse zu Vorst. vor u. nach instruktionalen Informationen (F-Vor6)

Null verschieden (dokumentiert in Tabelle F.22 auf S. 586). Die gepaarten t-Tests zu den verschiedenen Zeitfenstern sind in Tabelle 6.28 dokumentiert: Die beiden deutlichen Zunahmen von dem Zeitfenster vor der instruktionalen Information zum Zeitfenster direkt nach der instruktionalen Information (für Angemessenheit und für Übereinstimmung, siehe Abbildung 6.27 a und b) sind signifikant und von mittlerer

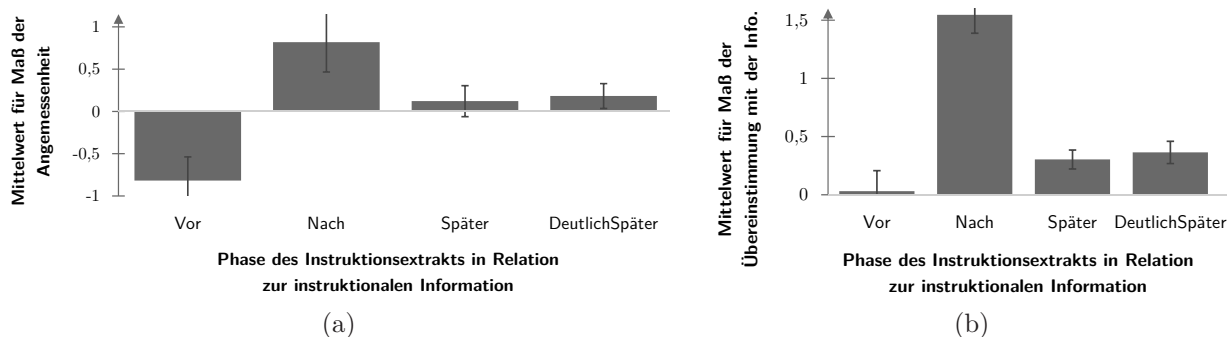


Abbildung 6.27: Mittlere absolute Punktzahlen für die Angemessenheit der Ideen-Oberkategorien (a) und die Übereinstimmungen der Ideen-Oberkategorien mit der Explizierung (b) vor der, direkt nach der, später nach der und deutlich später nach der Explizierung (für Kennzeichen von naturwissenschaftlichen Fragen). Fehlerbalken entsprechen einem Standardfehler des Mittelwerts nach oben und nach unten.

Tabelle 6.28: T-Tests zum Vergleich der absoluten Punktzahlen im zeitlichen Umfeld nach der instruktionalen Information mit der Punktzahl vor der instruktionalen Information (oben für Angemessenheit, unten für Übereinstimmung; für Kennzeichen von naturwissenschaftlichen Fragen).

Vergleich	Messzeitpunkt 1			Messzeitpunkt 2			df	t (df)	p	r	Unterschied	
	N ₁	MW	SD	N ₂	MW	SD					Absolut	Relativ
Angemessenheit der Ideen-Oberkategorien der Lernenden												
Nach – Vor	33	1,00	2,21	33	-0,79	1,63	32	4,411	0,000	-0,615	-1,79	-178,79%
Später – Vor	33	0,18	1,63	33	-0,79	1,63	32	3,379	0,002	-0,513	-0,97	-123,08%
Dtl. sp. – Vor	33	0,09	1,07	33	-0,79	1,63	32	2,922	0,006	-0,459	-0,88	-111,54%
Später – Nach	33	0,09	1,63	33	1,00	2,21	32	-2,222	0,033	0,366	0,91	90,91%
Dtl. sp. – Nach	33	0,18	0,85	33	1,00	2,21	32	-2,046	0,049	0,340	0,82	81,82%
Dtl. sp. – Später	33	0,18	0,85	33	0,09	1,07	32	0,452	0,654	-0,080	-0,09	-50,00%
Übereinstimmung der Ideen-Oberkategorien der Lernenden mit den der Explizierung zugewiesenen Ideen-Oberkategorien												
Nach – Vor	33	1,64	0,90	33	0,12	1,14	32	6,812	0,000	-0,769	-1,52	-92,59%
Später – Vor	33	0,36	1,14	33	0,12	1,14	32	1,092	0,283	-0,190	-0,24	-66,67%
Dtl. sp. – Vor	33	0,30	0,47	33	0,12	1,14	32	0,882	0,385	-0,154	-0,18	-60,00%
Später – Nach	33	0,30	1,14	33	1,64	0,90	32	-8,000	0,000	0,816	1,33	81,48%
Dtl. sp. – Nach	33	0,36	0,55	33	1,64	0,90	32	-6,341	0,000	0,746	1,27	77,78%
Dtl. sp. – Später	33	0,36	0,55	33	0,30	0,47	32	0,529	0,601	-0,093	-0,06	-16,67%

Anmerkungen. Gepaarte t-Test; $N_{1/2}$ =Stichprobengröße für den im zeitlichen Verlauf zuerst/später verorteten Messzeitpunkt (z. B. N_1 : Vor und N_2 : Nach). Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

bzw. großer Effektstärke. Die Zeitfenster später nach und deutlich später nach der instruktionalen unterscheiden sich jeweils signifikant vom Zeitfenster direkt nach der instruktionalen Information und weisen sowohl für die Angemessenheit als auch die Übereinstimmung mit der Explizierung jeweils schlechtere Werte auf als das Zeitfenster direkt nach der instruktionalen Information. Die Zeitfenster später und deutlich später unterscheiden sich nicht signifikant voneinander, in der Tendenz finden sich für das spätere Zeitfenster aber sowohl hinsichtlich der Angemessenheit als auch hinsichtlich der Übereinstimmung mit den aus dem Explizierungstext der instruktionalen Information rekonstruierbaren Vorstellungen schlechtere Werte als für das deutlich spätere Zeitfenster.

Die Ergebnisse bei Differenzierung der Zeitfenster nach der instruktionalen Information lassen sich so deuten, dass die instruktionalen Information vorrangig auf das Zeitfenster direkt nach der instruktionalen Information einen Einfluss hat. Die Angemessenheit der Vorstellungen und die Übereinstimmung der Vorstellungen mit der Explizierung sind in später gelegenen Zeitfenstern schlechter. Allerdings ist zu beachten, dass die Zeitfenster später und deutlich später nur aus einzelnen (Teil-)Aufgaben bestehen – im Gegensatz zu den Zeitfenstern direkt vor und direkt nach der instruktionalen Information, die jeweils mehrere (Teil-)Aufgaben umfassen. Die niedrigere Angemessenheit und die geringere Übereinstimmung mit den der instruktionalen Information zugeordneten Vorstellungen kann also auch dadurch bedingt sein, dass eine geringe Zahl an (Teil-)Aufgaben einen nicht ausreichenden Kontext für eine ausgewogene Angemessenheit oder für eine gute Übereinstimmung mit den der instruktionalen Information zugeordneten Vorstellungen bereitstellt oder dass die spezifischen (Teil-)Aufgaben die Angemessenheit oder die Übereinstimmung mit der Explizierung einschränken.

Insbesondere mit Blick auf die Übereinstimmung mit der Explizierung scheint die **geringe Anzahl der (Teil-)Aufgaben des späteren und des deutlich späteren Zeitfensters** relevant zu sein, weil in beiden Zeitfenstern die Zahl der Ideen-Oberkategorien-Instanzen gering ist und dies im Maß für die Übereinstimmung deutlich ins Gewicht fallen kann. Personen mit einer Ideen-Oberkategorie, die der Explizierung entspricht, würden beispielsweise keine weiteren Punkte erhalten, wenn sie keine weiteren Ideen-Oberkategorien zugewiesen bekommen. Wird pro Team nur eine Ideen-Oberkategorie in jeder der beiden Zeitfenster zugewiesen, ist nur ein Drittel der in den anderen Zeitfenstern üblichen Übereinstimmung zu erwarten. Die Werte im späteren Zeitfenster ($MW = 0.303$) und im deutlich späteren Zeitfenster ($MW = .364$) sind geringer als die beispielsweise bei einer Normalverteilung um

0.667 herum (entspricht: Ideen-Oberkategorien für zwei Personen jedes Teams kodiert) zu erwartenden Werte (für das spätere Zeitfenster: $t(32) = -4.476$, $p < .001$, $r = .385$; für das deutlich spätere Zeitfenster: $t(32) = -3.172$, $p = .003$, $r = .239$; Schätzungen jeweils jeweils durch Einstichproben-t-Tests), liegen in beiden Fällen jedoch im erwartbaren Bereich bei einer Normalverteilung um 0.333 herum (entspricht dem Drittel, also der Annahme, dass in beiden Zeitfenstern jeweils nur eine Ideen-Oberkategorie und nur für eine der drei Person jedes Teams kodiert ist), wie entsprechende t-Tests zeigen (für das spätere Zeitfenster: $t(32) = -0.373$, $p = .712$, $r = .004$; für das deutlich spätere Zeitfenster: $t(32) = 0.317$, $p = .753$, $r = .003$; jeweils Einstichproben-t-Tests).

Die bisher dargestellten Ergebnisse und Überlegungen zu den Zeitfenstern mit zeitlichem Abstand zur instruktionalen Information legen nahe, zusätzlich zu den bisher genutzten Bepunktungen auch **relative Bepunktungen** vorzunehmen, um dem absolut geringeren Auftreten von Ideen-Oberkategorien in manchen Zeitfenstern Rechnung zu tragen. Um eine relativierte Punktzahl zu erhalten, werden die jeweils für die Personen vergebenen Punkte (dokumentiert in Tabelle F.21 auf S. 585) durch die Anzahl aller im jeweiligen Zeitfenster gefüllten Zellen aus Tabelle F.19 (auf S. 583, d. i. das alle vier Phasen auflösende Analogon zu Tabelle 6.27 auf S. 355) geteilt. Beispielsweise gibt es in dem Zeitfenster vor der Explizierung insgesamt 78 Instanzen, für den Eintrag zu Effekt erhält A1 also $-2 \cdot 1/78 = -0.0256$ Punkte für die Angemessenheit; zusammengenommen mit den anderen Einträgen erreicht die Schülerin A1 in dem Zeitfenster vor der instruktionalen Information insgesamt 0.0128 Punkte im Bereich der Angemessenheit. Die vergebenen Punkt pro Person haben keine für sich stehende Aussage, erlauben allerdings die Vergleichbarkeit über die Segmente hinweg. In Abbildung 6.28 ist dargestellt, wie die relativen Punktemittelwerte für die jeweiligen Zeitfenster sich zueinander verhalten (die genauen Werte, die die einzelnen Personen erreichen, sind in Tabelle F.23 dokumentiert). Im Anhang ist in Tabelle F.24 (auf S. 588) aufgeführt, wie sicher die jeweiligen Werte von Null abweichen. In Tabelle 6.29 (auf S. 366) sind die t-Tests zum Vergleich der verschiedenen Zeitfenster dokumentiert.

Bei gemäß der Instanzen der jeweiligen Zeitfenster relativierter Bepunktung ergibt sich: Die Angemessenheit ist in allen Zeitfenstern nach der instruktionalen Information signifikant und deutlich größer als vor der instruktionalen Information (allerdings aufgrund der Streuungen nur direkt nach der instruktionalen Information auch signifikant größer als Null). Die Übereinstimmung mit den Ideen-Oberkategorien der Explizierung ist vor der instruktionalen Information nicht signifikant

6 Vorstellungen von Lernenden

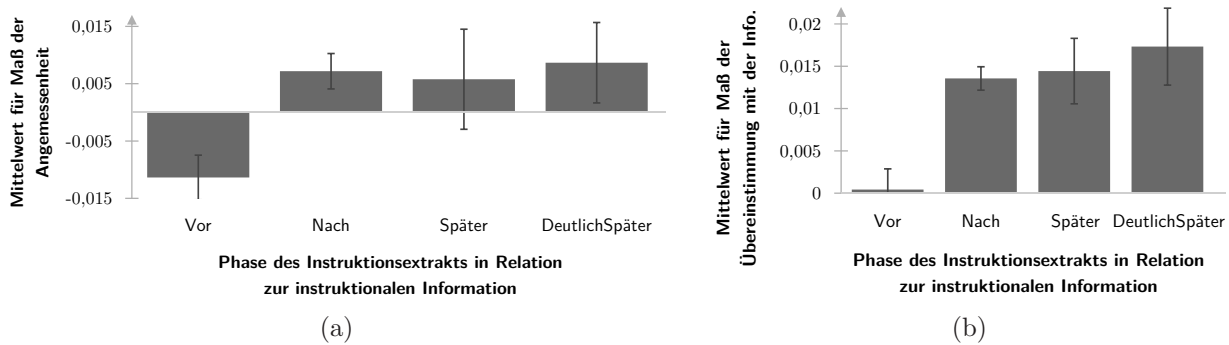


Abbildung 6.28: Mittlere relative Punktzahlen für die Angemessenheit der Ideen-Oberkategorien (a) und die Übereinstimmungen der Ideen-Oberkategorien mit der instruktionalen Information (b) vor der, direkt nach der, später nach der und deutlich später nach der instruktionalen Information (für Kennzeichen von naturwissenschaftlichen Fragen). Fehlerbalken entsprechen einem Standardfehler des Mittelwerts nach oben und nach unten.

Tabelle 6.29: T-Tests zum Vergleich der relativen Punktzahlen im zeitlichen Umfeld nach der instruktionalen Information mit der Punktzahl vor der instruktionalen Information (oben für Angemessenheit, unten für Übereinstimmung; für Kennzeichen von naturwissenschaftlichen Fragen).

Vergleich	Messzeitpunkt 1			Messzeitpunkt 2			df	t (df)	p	r	Unterschied	
	N ₁	MW×10 ⁻²	SD	N ₂	MW×10 ⁻²	SD					Absolut	Relativ
Angemessenheit der Ideen-Oberkategorien der Lernenden												
Nach – Vor	33	0,85	0,02	33	-1,01	0,02	32	4,511	0,000	-0,623	-1,86	-183,84%
Später – Vor	33	0,87	0,02	33	-1,01	0,02	32	2,649	0,012	-0,424	-1,87	-185,80%
Dtl. sp. – Vor	33	0,41	0,05	33	-1,01	0,02	32	1,693	0,100	-0,287	-1,42	-140,87%
Später – Nach	33	0,41	0,02	33	0,85	0,02	32	-0,495	0,624	0,087	0,43	51,25%
Dtl. sp. – Nach	33	0,87	0,04	33	0,85	0,02	32	0,026	0,979	-0,005	-0,02	-2,28%
Dtl. sp. – Später	33	0,41	0,05	0	0,00	0,00	32	0,488	0,629	-0,086	-0,41	-100,00%
Übereinstimmung der Ideen-Oberkategorien der Lernenden mit den der Explizierung zugewiesenen Ideen-Oberkategorien												
Nach – Vor	33	1,39	0,01	33	0,16	0,01	32	4,766	0,000	-0,644	-1,23	-88,76%
Später – Vor	33	1,73	0,01	33	0,16	0,01	32	2,995	0,005	-0,468	-1,58	-91,00%
Dtl. sp. – Vor	33	1,38	0,02	33	0,16	0,01	32	2,871	0,007	-0,453	-1,22	-88,70%
Später – Nach	33	1,38	0,01	33	1,39	0,01	32	-0,019	0,985	0,003	0,01	0,52%
Dtl. sp. – Nach	33	1,73	0,03	33	1,39	0,01	32	0,687	0,497	-0,121	-0,34	-19,92%
Dtl. sp. – Später	33	1,73	0,03	33	1,38	0,02	32	0,657	0,516	-0,115	-0,35	-20,34%

Anmerkungen. Gepaarte t-Test; $N_{1/2}$ =Stichprobengröße für den im zeitlichen Verlauf zuerst/später verorteten Messzeitpunkt (z. B. N_1 : Vor und N_2 : Nach). Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

von Null verschieden und in allen Zeitfenstern nach der instruktionalen Information signifikant größer als vor der instruktionalen Information. Die Zeitfenster nach der instruktionalen Information unterscheiden sich hinsichtlich der Angemessenheit und der Übereinstimmung mit Ideen-Oberkategorien der Explizierung nicht signifikant voneinander; auffällig ist jeweils die deutlich größere Streuung im späteren und im deutlich späteren Zeitfenster.

Die Ergebnisse des vorigen Absatzes lassen sich so deuten, dass die oben berichteten – auf der absoluten Bepunktung basierenden – Ergebnisse hinsichtlich der Abnahme der Angemessenheit und der Übereinstimmung mit der Explizierung wesentlich von der Team-Arbeit mitbestimmt sind. Für die Kohorte findet sich im Mittel ein Zuwachs hinsichtlich beider Beurteilungskriterien, bei den späteren Zeitfenstern ist der Effekt der reduzierten Beteiligung deutlich ersichtlich. Wenn davon ausgegangen würde, dass Personen ohne zugewiesene Ideen-Oberkategorien eine schweigende Zustimmung vornehmen, ließen sich die Ergebnisse bei relativer Bepunktung als klares Zeichen für die positive Wirkung der instruktionalen Information auf die Angemessenheit und auf das Auftreten der explizierten Ideen-Oberkategorien ansehen. Da allerdings nicht grundsätzlich von einer schweigenden Zustimmung ausgegangen werden sollte (wie die ausführlichen Transkripte zu den Zeitfenstern direkt vor und direkt nach der instruktionalen Information nahelegen), ist die Deutung nur mit starker Einschränkung auf die Vorstellungen aller Lernenden zu übertragen. Mindestens die Einschränkung der wenigen (Teil-)Aufgaben in den beiden letzten Zeitfenstern wäre für eine Generalisierung auf die Lernenden der Stichprobe zu beheben.

6.9.1.2 Für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung

In Tabelle 6.30 (auf S. 369) ist für die Unterscheidung von Beobachtung und Deutung aufgelöst, welchen Personen welche Ideen-Oberkategorien im Zeitfenster direkt vor der Explizierung bzw. im Zeitfenster direkt nach der instruktionalen Information zugewiesen wurden. Die Analysen wurden soweit möglich analog zu den Analysen für Kennzeichen von naturwissenschaftlichen Fragen durchgeführt.

Vorstellungen, die vor der instruktionalen Information rekonstruiert werden, finden sich nur zum Teil auch danach. Anhand von Tabelle 6.30 lässt sich ermitteln, dass sich für die Hälfte der insgesamt 10 vor der Explizierung kodierten Ideen-Oberkategorien nach der Explizierung keine Vorkommnisse finden; für 2 finden sich genauso viele Personen mit der Ideen-Oberkategorie nach der Explizierung wie davor; für 2 finden sich weniger Personen nach der Explizierung als vorher; nur

für AuswertungAntwort findet sich eine Zunahme der Personenanzahl mit der Ideen-Oberkategorie von 1 auf 2. Allerdings sind diese Ergebnisse vor dem Hintergrund der Tatsache zu interpretieren, dass nach der Explizierung deutlich weniger Ideen-Oberkategorien pro Person vergeben wurden als vorher (siehe bereits bei F-Vor5 und zwar auf S. 328). Daher ist eine Abnahme in der Personenanzahl pro Ideen-Oberkategorien nicht verwunderlich, sondern erwartbar. Gleichzeitig ist das Nichtauftreten nach der Explizierung von 5 vor der Explizierung kodierten Ideen-Oberkategorien ein einschränkendes Resultat.

Die Tatsache, dass nach der Explizierung deutlich weniger Ideen-Oberkategorien-Instanzen vorliegen, spiegelt sich in den nachfolgenden Resultaten wider. Insbesondere die Betrachtung des Auftretens von Ideen-Oberkategorien vor und nach der Explizierung auf Ebene der einzelnen Personen ist stark davon betroffen:

Für die meisten der Personen lassen sich die Vorstellungen, die sich vor der instruktionalen Information rekonstruieren lassen, nach der Instruk-tionalen Information nicht erneut rekonstruieren. Dies gilt auch, wenn Teams statt Personen betrachtet werden. Wird markiert, wann eine Person eine Ideen-Oberkategorie vor der Explizierung und auch danach erhalten hat (z. B. bei B5 zur Ideen-Oberkategorie *Vermutung*, aber nicht bei E14 zu *Beobachtung*), ergibt sich, dass dies nur 3 mal auftritt. Vor der Explizierung liegen 43 Instanzen (mit V gefüllte Zellen der Tabelle) vor. Personen werden also nur in ca. 7 % vor der Explizierung zugewiesene Ideen-Oberkategorien auch nach der Explizierung zugewiesen. Wird analog gezählt, in wie vielen Fällen eine Ideen-Oberkategorie einem Team vor bzw. vor und nach der Explizierung zugewiesen ist (d. h. irgendeiner Person im Team), finden sich keine zusätzlichen Fälle (der Anteil steigt auf ca. 10 % von 29). Was sich bereits auf der Kohorten-Ebene im vorigen Absatz abgezeichnet hat, scheint also auf Ebene der Personen umso deutlicher zu sein: Weil nach der Explizierung wenige Ideen-Oberkategorien pro Person kodiert wurden, gibt es eine Schieflage in den berichteten Ergebnissen.

Eine Möglichkeit, den geringen Vorkommnissen von Ideen-Oberkategorien nach der Explizierung zu begegnen und gleichzeitig einigermaßen aussagekräftige Ergebnisse zu erhalten, besteht darin, nur die Teilstichprobe all der Personen zu betrachten, für die nach der Explizierung mindestens eine Ideen-Oberkategorien kodiert wurde.¹⁰⁵

¹⁰⁵Eine andere Option besteht in der Betrachtung aller Personen, denen überhaupt irgendwann eine Ideen-Oberkategorien zugewiesen wurde. Um den Rahmen der Arbeit nicht zu sprengen wurde auf die Darstellung dieser Teilstichprobe verzichtet. Alle Daten für eine derartige Analyse liegen

Tabelle 6.30: Auflösung der Ideen-Oberkategorien, die den einzelnen Personen jeweils vor bzw. nach der Explizierung zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden (sortiert nach Angemessenheit der Ideen-Oberkategorien sowie nach Team-Zugehörigkeit der Personen).

Person	FaktenSicher		Beobachten		Meinung		Sonstiges		Vermutung		Interpretation		BegründungErklärung		AuswertungAntwort		Beschreiben		Sehen		ObjektivSubjektiv		Messen		Anzahl d. Pers: Pre		Anzahl d. Pers: Post	
	deutlich				eher unangemessen								N		eher angemessen										2		1	
A0 = E14			V						V																2		1	
A1											V		V												0		0	
A2																									0		0	
A3																									0		0	
B4																									0		0	
B5									V	N		N							V						2		2	
B6																									0		0	
C7	N												N												0		2	
C8																									0		0	
C9																									0		0	
F16													V		V										3		0	
F17																									0		0	
F18						N					V														1		1	
G19																									0		0	
G20																									0		0	
G21	V		V						V	N			V						V		N				1		0	
H22			V										V												4		2	
H23			V										V		V										2		0	
H24			V					V					V						V		N				4		3	
K31									V																2		0	
K32									V				V	N											2		1	
K33	N		V						V				V						V						3		1	
L34																									2		0	
L35	V								V						N				V				V		3		1	
L36																			V						1		0	
M37													V												1		0	
M38						N						N	V		N				V						2		3	
M39													V												0		0	
P46			V																V						3		0	
P47																									0		0	
P48			V										V						V						3		0	
Anzahl der	2	2	6	0	0	2	1	0	7	3	3	3	11	3	1	2	1	0	10	0	0	2	1	0	1,39		0,55	
Personen	s. o.		V: 6 N: 2										V: 18 N: 8												1,31		0,91	

Anmerkungen: Durch halbfetten Satz hervorgehoben sind Ideen-Oberkategorien, die auch für den Text der Explizierung kodiert wurden. Die Anzahl der Personen entspricht allen Zeilen, die im entsprechenden Spalten-Bereich (jeweilige Ideen-Oberkategorie bzw. jeweilige Gruppe von Ideen-Oberkategorien zu einer spezifischen Angemessenheit) mindestens einen Eintrag haben. V = Vor der Explizierung, N = Nach der Explizierung

Ergebnisse für diese Teilstichprobe haben natürlich nicht dieselbe Aussagekraft, sondern sind eben auf all die Personen beschränkt, für die nach der Explizierung Ideen-Oberkategorien kodiert wurden. Welche Faktoren dazu geführt haben, dass sich aus den Äußerungen dieser Personen nach der Explizierung Hinweise auf ihre Vorstellungen ableiten lassen, aus den Äußerungen anderer Personen jedoch nicht, bleibt offen – könnte aber von Relevanz für Schlussfolgerungen aus den Ergebnissen sein.

Werden nur Personen betrachtet, denen nach der Explizierung mindestens eine Ideen-Oberkategorie zugewiesen wurde, so ergibt sich Abbildung 6.30 (auf S. 375). Aus der Tabelle geht hervor, dass 3 der 7 vor der Explizierung auftretenden Ideen-Oberkategorien nach der Explizierung nicht mehr vorliegen; für 2 liegt eine Zunahme der Personenanzahlen vor, für 2 eine Abnahme. Ferner liegen dieselben 3 Fälle vor, in denen einer Person eine Ideen-Oberkategorie sowohl vor als auch nach der Explizierung zugewiesen wurde; es ergibt sich diesmal ein Anteil von 13 % an 23 Ideen-Oberkategorien-Instanzen vor der Explizierung. Eine Betrachtung auf Team-Ebene wurde nicht vorgenommen, weil durch die Aussortierung keine sinnvollen Teamstrukturen mehr vorliegen. Die Ergebnisse auf der Personen-Ebene legen keine substantiell andere Deutung nahe als die oben formulierte: Vorstellungen, die vor der instruktionalen Information rekonstruiert werden, finden sich nur zum Teil auch danach.

Zur Untersuchung der Angemessenheit in den Phasen werden die Einträge in Tabelle 6.30 analog zu Unterunterabschnitt 6.9.1.1 bepunktet (deutlich unangemessen: –2, eher unangemessen: –1, eher angemessen: 2). Für die Übereinstimmung mit der Explizierung wurden Einträge für in Tabelle 6.30 (auf S. 369) halbfett gesetzte Ideen-Oberkategorien mit zwei Punkten bewertet, für die Ideen-Oberkategorien *Interpretation* und *AuswertungAntwort* wurde jeweils ein Punkt vergeben.¹⁰⁶ In Abbildung 6.29 sind die Mittelwerte der Angemessenheiten über alle für die Unterscheidung von Beobachtung und Deutung untersuchten Personen eingetragen. Die Punktzahlen der einzelnen Personen sind im Anhang in Tabelle F.25 (auf S. 589) dargestellt; analog zur Betrachtung bei Kennzeichen von naturwissenschaftlichen Fragen wird zusätzlich

aber in Tabelle 6.30 (auf S. 369) vor; es müssen nur alle leeren Personen-Zeilen aussortiert werden.

Anzumerken ist an dieser Stelle, dass für Kennzeichen von naturwissenschaftlichen Fragen bei einer Aufteilung in zwei Phasen das Bilden einer Teilstichprobe in beiden Fällen zur Gesamtstichprobe führt, weil keine Personen aussortiert werden müssen. Eine Betrachtung der Teilstichproben bei einer Aufteilung in vier Phasen wird aufgrund des geringen Mehrwerts nicht berichtet. Für die Unterscheidung von Beobachtung und Deutung ist der Mehrwert einer Analyse der Teilstichprobe hingegen deutlich gegeben.

¹⁰⁶ Dem Text der Explizierung wurden weitere Ideen-Oberkategorien zugewiesen; in Anhang F.7.3 (ab S. 590) ist erläutert, warum nur die im Text aufgeführten Ideen-Oberkategorien bepunktet wurden.

Tabelle 6.31: Auflösung der Ideen-Oberkategorien, die den einzelnen Personen jeweils vor bzw. nach der Explizierung zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden (sortiert nach Angemessenheit der Ideen-Oberkategorien sowie nach Team-Zugehörigkeit der Personen), nur für Personen, denen nach der Explizierung mindestens eine Ideen-Oberkategorie zugewiesen wurde.

Person	FaktenSicher		eher unangemessen					eher angemessen					BegründungErklärung					AuswertungAntwort					ObjektivSubjektiv		Messen	Anzahl d. Pers: Pre	Anzahl d. Pers: Post
	deutlich		Beobachten	Meinung	Sonstiges	Vermutung	Interpretation																				
A0 = E14			V			V			N															2	1		
B5						V	N																	2	2		
C7	N								N															0	2		
F18				N				V																1	1		
H22			V			V	N		V															4	2		
H24			V		V	N		N	V															4	3		
K32						V			V	N														2	1		
K33	N		V			V			V															3	1		
L35	V					V												N						3	1		
M38				N				N	V									N						2	3		
Anzahl der Personen	1	2	4	0	0	2	1	0	6	3	1	3	5	3	0	2	0	0	0	2	0	0	0	2,30	1,70		
s. o.																							1,19	0,78			
Anmerkungen. Diese Tabelle geht aus der Basistabelle für die Unterscheidung von Beobachtungen und Deutungen hervor, wenn nur die Personen betrachtet werden, die nach der Explizierung mindestens eine gefüllte Zelle aufweisen. Durch halbierten Satz hervorgehoben sind Ideen-Oberkategorien, die auch für den Text der Explizierung kodiert wurden. Die Anzahl der Personen entspricht allen Zeilen, die im entsprechenden Spalten-Bereich (jeweilige Ideen-Oberkategorie bzw. jeweilige Gruppe von Ideen-Oberkategorien zu einer spezifischen Angemessenheit) der Basistabelle mindestens einen Eintrag für eine der betrachteten Personen haben. V = Vor der Explizierung, N = Nach der Explizierung																											
V: 4 N: 2													V: 9 N: 8														

eine relative Bepunktung (geteilt durch Anzahl aller gefüllten Zellen vor bzw. nach der Explizierung) vorgenommen.

Die Vorstellungen der Lernenden sind konstant von hoher Angemessenheit.

Die Angemessenheit der Ideen-Oberkategorien der Personen ist vor der Explizierung im Mittel positiv und signifikant von Null verschieden ($MW = 1.871$, Einstichproben-t-Test: $t(30) = 5.177$, $p < .001$, $r = .472$); nach der Explizierung ist sie ebenfalls im Mittel positiv und signifikant von Null verschieden ($MW = 0.645$, Einstichproben-t-Test: $t(30) = 2.185$, $p = .037$, $r = .137$). Die Differenzen der Angemessenheiten betragen im Mittel -1.226 , was einer signifikanten Abnahme der Angemessenheit entspricht (gepaarter t-Test: $t(30) = -3.236$, $p = .003$, $r = .259$). Demnach wären die Vorstellungen der Personen entgegen der Hypothese nach der Explizierung nicht angemessener als vor der Explizierung, sondern sogar weniger angemessen. Bei relativer Bepunktung zeigt sich allerdings, dass keine signifikanten Unterschiede zwischen den beiden Phasen vorliegen (gepaarter t-Test: $t(30) = -0.343$, $p = .734$, $r = .004$) und die Angemessenheit in beiden Phasen signifikant von Null verschieden ist (t-Tests ergeben gleiche Werte wie bei absoluter Bepunktung). Dieser Befund unterstreicht das Ergebnis aus Abschnitt 6.6 (ab S. 306), dass die Vorstellungen zur Unterscheidung von Beobachtung und Deutung grundsätzlich angemessen sind.

Auf Kohorten-Ebene zeigt sich eine Abnahme nahezu aller Ideen-Oberkategorien von vor der Explizierung zu nach der Explizierung und damit auch bezüglich der Angemessenheit ein vergleichbares Bild (siehe Tabelle 6.30). Einzig für die Ideen-Oberkategorien *Meinung*, die als eher unangemessen eingeschätzt wurde, und *ObjektivSubjektiv* und *AuswertungAntwort*, die als eher angemessen eingeschätzt wurden, finden sich Zunahmen. Die grundsätzlich geringere Beteiligung spiegelt sich auch darin wieder, dass nach der Explizierung sowohl in der Gruppe der eher unangemessenen als auch der Gruppe der eher angemessenen Ideen-Oberkategorien deutliche Abnahmen in der Personenzahlen zu verzeichnen sind. Nur die Anzahl der Personen, für die die deutlich unangemessene Ideen-Oberkategorie *FaktenSicher* kodiert wurde, bleibt konstant – interessanterweise aber wird sie nach der Explizierung für andere Personen (aus anderen Teams) zugewiesen als davor. Von einer bedeutsamen Abnahme der Angemessenheit kann also nicht die Rede sein – von einer Zunahme allerdings auch nicht.

Die Vorstellungen der Lernenden stimmen nach der instruktionalen Information ähnlich gut mit dem explizierten Konzept überein wie davor.

Die grafisch in Abbildung 6.29 zusammengefassten Werte (Details im Anhang in Tabelle F.25) zeigen, dass vor der Explizierung eine niedrige bis mittlere Übereinstimmung mit den explizierten Ideen-Oberkategorien vorliegt ($MW = 0.903$, signifikante

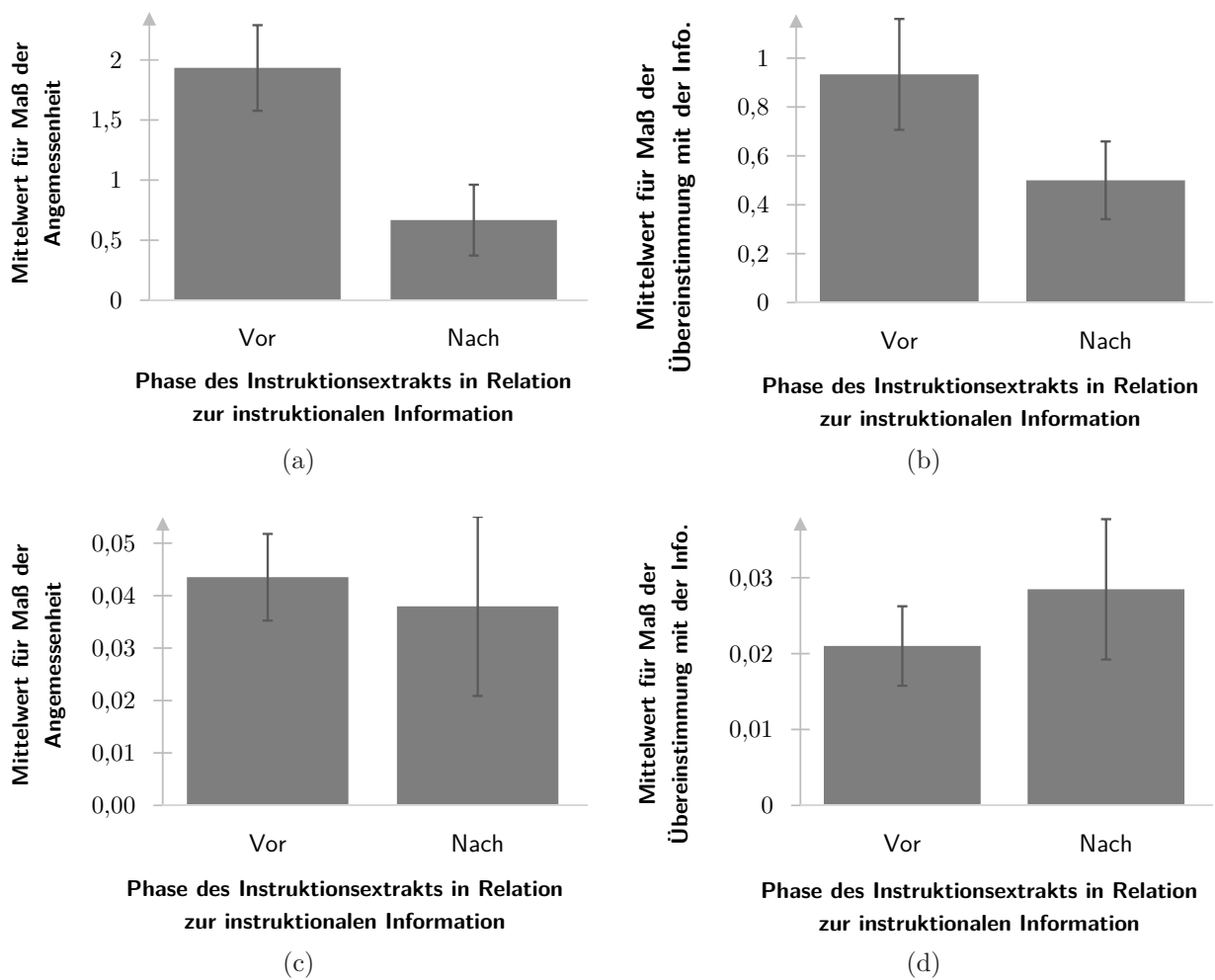


Abbildung 6.29: Mittlere absolute (a, b) bzw. relative (c, d) Punktzahlen für die Angemessenheit der Ideen-Oberkategorien (a, c) und die Übereinstimmungen der Ideen-Oberkategorien mit der Explizierung (b, d) vor und nach der Explizierung (für die Unterscheidung von Beobachtung und Deutung).

Abweichung von Null mit mittlerem Effekt bei Einstichproben-t-Test: $t(30) = 3.947$, $p < .001$, $r = .342$), nach der Explizierung eine geringere ($MW = 0.484$, Einstichproben-t-Test: $t(30) = 3.028$, $p = .005$, $r = .234$). Der Unterschied zwischen den beiden Phasen ist nicht signifikant und entspricht einem kleinen Effekt (gepaarter t-Test: $t(30) = -1.604$, $p < .119$, $r = .079$). Die Übereinstimmung mit explizierten Ideen-Oberkategorien liegt in beiden Fällen weit unter der der Maximalpunktzahl von 8. Eine Betrachtung der relativen Bepunktung liefert ebenfalls signifikant von Null verschiedene Mittelwerte in beiden Phasen (t-Tests ergeben gleiche Werte wie bei absoluter Bepunktung). Der Unterschied zwischen der Phase vor der Explizierung und der Phase nach der Explizierung hinsichtlich der Mittelwerte bei relativer Bepunktung für die Übereinstimmung mit den explizierten Ideen-Oberkategorien ist zwar nicht signifikant und entspricht nur einem sehr kleinen Effekt ($t(30) = 0.725$, $p = .474$, $r = .017$); es liegt aber in der Tendenz ein Zuwachs der Übereinstimmung mit den explizierten Ideen-Oberkategorien vor. An der Unterschiedlichkeit der Ergebnisse bei absoluter und relativer Bepunktung wird deutlich, welchen Einfluss die Abnahme der Anzahl der Ideen-Oberkategorien auf die untersuchten Werte hat. Von der bei absoluter Bepunktung suggerierten Abnahme der Übereinstimmung mit den explizierten Ideen-Oberkategorien sollte keineswegs gesprochen werden; vielmehr legt die relative Bepunktung sogar in der Tendenz eine Zunahme der Übereinstimmung nahe, die sich in den nachfolgenden Analysen einer Teilstichprobe sowohl absolut als auch relativ noch deutlicher manifestiert.

Anknüpfend an die obigen Überlegungen dazu, dass nach der Explizierung nur wenige Ideen-Oberkategorien-Instanzen vorliegen, werden nachfolgend auch für die Angemessenheit und für die Übereinstimmung der den Lernenden zugewiesenen Ideen-Oberkategorien mit den der instruktionalen Information zugewiesenen Ideen-Oberkategorien Analysen für die Teilstichprobe aller Personen, für die nach der Explizierung mindestens eine Ideen-Oberkategorie kodiert wurde, berichtet.¹⁰⁷ Das Vorgehen ist plausibel, da für Personen, die nach der Explizierung keine Ideen-Oberkategorie zugewiesen bekommen haben, keine Aussage über die Angemessenheit ihrer Vorstellungen oder die Übereinstimmung ihrer Vorstellungen mit den Konzepten der Explizierung getroffen werden kann und sollte. Alternative statistische Analysen, bei denen die Anzahl der zugewiesenen Ideen-Oberkategorien nach der Explizierung als Moderator genutzt wird, sind hier nicht berichtet, weil entsprechende Ergebnisse sich vor dem Hintergrund des vorigen Satzes nur schwer interpretieren lassen.

¹⁰⁷Für die Analysen der Teilstichprobe sind die relativen Maße vor der Explizierung neu berechnet, die relativen Maße nach der Explizierung stimmen mit den bisher genutzten überein. Die entsprechenden Werte sind im Anhang in Tabelle F.25 auf S. 589 dokumentiert.

Für die Teilstichprobe der 10 Personen, für die nach der Explizierung eine Ideen-Oberkategorie kodiert wurde, sind die Ergebnisse in Abbildung 6.30 dargestellt. Es ergibt sich basierend auf den Werten aus Tabelle F.25, dass die Angemessenheit in der Tendenz absolut bepunktet mit kleiner Effektstärke abnimmt ($t(9) = -1.149$, $p = .280$, $r = .128$), aber relativ bepunktet nahezu gleich bleibt ($t(9) = -0.099$, $p = .923$, $r = .001$). Eine Interpretation der Ergebnisse als Konstantbleiben der Angemessenheit scheint plausibel, weil auf Ebene der 10-Personen-Kohorte die Zu- bzw. Abnahmen in den Gruppen deutlich/eher unangemessener und eher angemessener Ideen-Oberkategorien sehr gering sind und nur für zwei Person (F18, K33) nach der Explizierung mehr als 50 % der Ideen-Oberkategorien eher oder deutlich unangemessen sind. Ferner ist in Abbildung 6.30 dargestellt, dass die Übereinstimmung mit den expliziten Ideen-Oberkategorien sowohl absolut bepunktet ($t(9) = 1.000$, $p = .343$, $r = .100$) als auch relativ bepunktet zunimmt ($t(9) = 1.962$, $p = .081$, $r = .300$). Die Effekte der Zunahme sind klein bis mittel, nur bei relativer Bepunktung ist das Ergebnis signifikant. Dennoch sollten die Ergebnisse so interpretiert werden, dass in der Teilstichprobe derjenigen Personen, die nach der Explizierung überhaupt Äußerungen machen, aus denen auf ihre Vorstellungen zur Unterscheidung von Beobachtung und Deutung geschlossen werden kann, die Vorstellungen nach der Explizierung nicht weniger, sondern eher mehr mit der Explizierung übereinstimmen als zuvor.

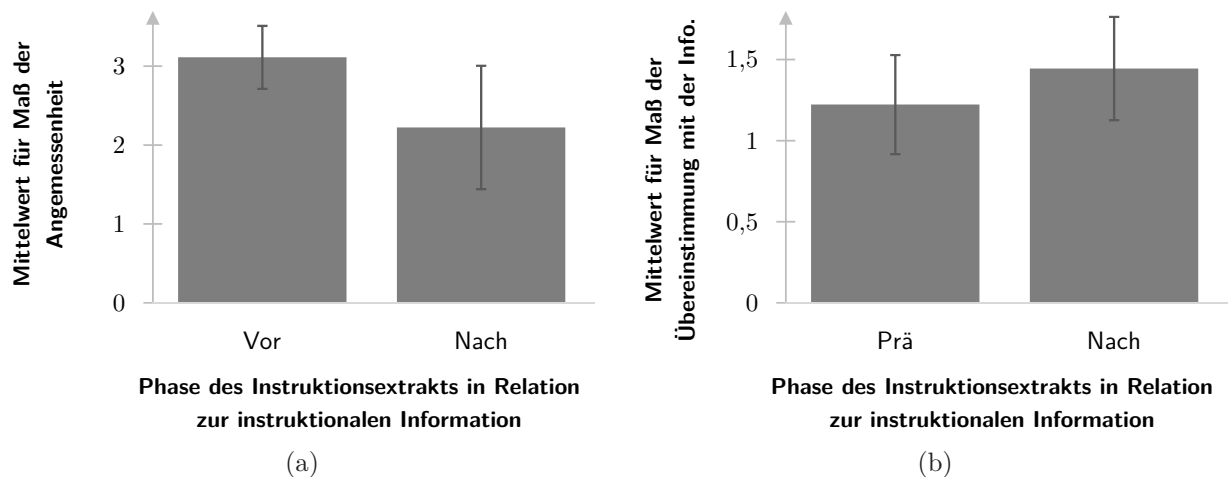


Abbildung 6.30: Mittlere absolute (a, b) bzw. relative (c, d) Punktzahlen für die Angemessenheit der Ideen-Oberkategorien (a, c) und die Übereinstimmungen der Ideen-Oberkategorien mit der Explizierung (b, d) vor und nach der Explizierung bei korrigierter Stichprobe (für die Unterscheidung von Beobachtung und Deutung). *Teil 1 von 2 (Teil 2 auf S. 376).*

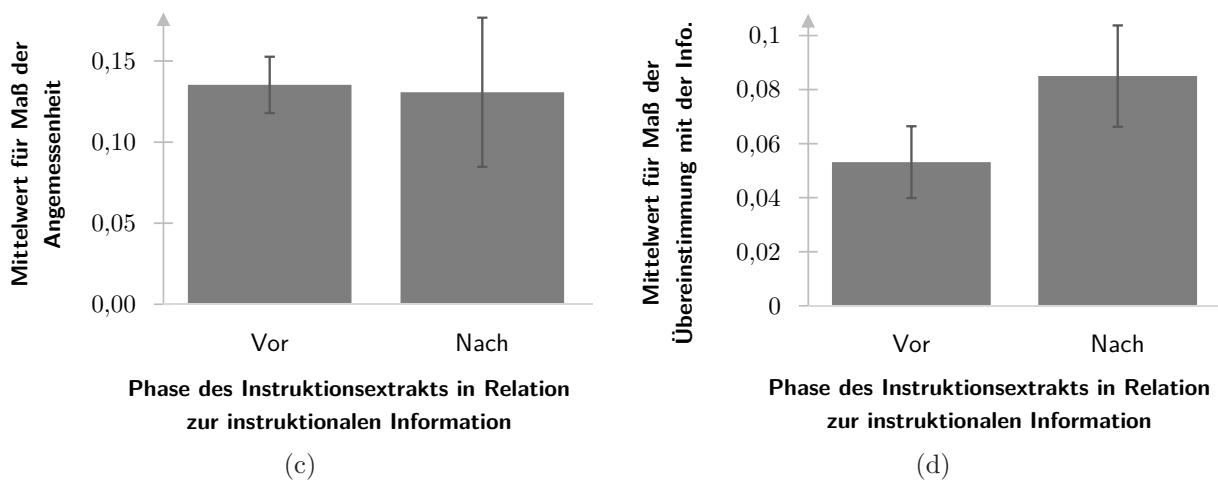


Abbildung 6.30: Mittlere absolute (a, b) bzw. relative (c, d) Punktzahlen für die Angemessenheit der Ideen-Oberkategorien (a, c) und die Übereinstimmungen der Ideen-Oberkategorien mit der Explizierung (b, d) vor und nach der Explizierung bei korrigierter Stichprobe (für die Unterscheidung von Beobachtung und Deutung). *Teil 2 von 2.*

Einschränkungen der Interpretation für die Übereinstimmung mit der Explizierung liegen – wie schon bei Kennzeichen von naturwissenschaftlichen Fragen – nicht nur in den bereits aufgeführten statistischen Unsicherheiten aufgrund der Daten zu den Personen, sondern auch in der Frage der Operationalisierung von Übereinstimmung mit der Explizierung. Die Explizierung zur Unterscheidung von Beobachtung und Deutung (siehe Abbildung 6.31) besteht aus zwei Karten des Lernmaterials und betont im Kern nur zwei Konzepte: *Beobachtungen sind intersubjektiv überprüfbar, Deutungen nicht*; und: *Deutungen gehen über das, was beobachtbar ist hinaus, Beobachtungen nicht*. Die Konzepte werden jeweils im Informationstext mit feineren Beispielen unterstützt. Gleichzeitig wurden für die Explizierung jedoch 7 Ideen-Oberkategorien vergeben: BegründungErklärung (an 5 Stellen), ObjektivSubjektiv (3), Beschreiben (3), AuswertungAntwort (2), Meinung (1), Interpretation (1), Beobachten (1).¹⁰⁸ Ähnlich wie bei Kennzeichen von naturwissenschaftlichen Fragen scheint die Anwendung des Kodiersystems auf den schriftlichen Text der Explizierung mit einer Zerteilung der zentralen Konzepte der Explizierung einherzugehen, die Aussagen dazu, ob das zentrale Ziel der Explizierung erreicht ist, schwierig macht. Trotzdem lassen sich die berichteten Ergebnisse so interpretieren, dass die Aspekte der Explizierung in den

¹⁰⁸Die genauen Kodierungen sind im Anhang in Unterabschnitt F.7.3 auf S. 590 dokumentiert und erläutert.

Vorstellungen der Lernenden in der Phase nach der Explizierung leicht prominenter sind als vorher.

22 Beobachtungen

INFO

Beobachtungen: Bei der Beobachtung geht es ausschließlich darum, zu beschreiben, was passiert und nicht darum, *warum* oder *wieso* es passiert. Beobachtungen müssen daher intersubjektiv überprüfbar sein. Das bedeutet, dass verschiedene Personen, die das gleiche Phänomen beobachten, zu gleichen Aussagen gelangen sollten.

Beispiel: Die Aussage „Ja, und die Holzkugel ist zuerst auf dem Boden aufgekommen“ ist eine Beobachtung, denn...

- ... sie beschreibt ausschließlich was passiert ist.
- ... alle Personen, die das Auftreffen der Holzkugel beobachtet haben, werden mit großer Wahrscheinlichkeit in dieser Aussage übereinstimmen.

23 Deutungen

INFO

Deutungen: Eine Deutung geht immer über die Beschreibung von dem, was passiert ist, hinaus, z. B. durch eine zusätzliche Erklärung oder eine Verallgemeinerung.

- **Erklärung:** Macht eine Aussage darüber, warum etwas passiert sein könnte.
- **Verallgemeinerung:** Macht aus den Beobachtungen eine Regel, die auch für andere (meist ähnliche) Fälle gilt.

Deutungen sind nicht intersubjektiv überprüfbar, weil verschiedene Personen zu unterschiedlichen Erklärungen und Verallgemeinerungen kommen können.

Beispiel: Der zweite Teil der Aussage „Stimmt, die Feder war langsamer, weil sie viel leichter ist als die Holzkugel.“ ist eine Deutung, denn...

- ...sie beschreibt nicht nur, was passiert (Feder ist langsamer als Holzkugel), sondern auch, warum es passiert (Masse).
- ...verschiedene Personen, die diesen Versuch beobachten, können unterschiedlicher Meinung darüber sein, warum die Holzkugel schneller fällt als die Feder.

Hinweis: Deutungen werden manchmal auch als Interpretationen oder Schlussfolgerungen bezeichnet.

Abbildung 6.31: Text der Explizierung zur Unterscheidung von Beobachtung und Deutung.

6.9.1.3 Zusammenschau

Im Vergleich der Angemessenheiten der Vorstellungen sowie der Übereinstimmung der Vorstellungen mit den explizierten Konzepten für Kennzeichen von naturwissenschaftlichen Fragen einerseits und die Unterscheidung von Beobachtung und Deutung andererseits lässt sich zunächst festhalten, dass in den Phasen vor und nach der Explizierung unterschiedlich viele Vorstellungen rekonstruiert wurden. Während sich für Kennzeichen von naturwissenschaftlichen Fragen vor der Explizierung vergleichsweise wenige Ideen-Oberkategorien finden, direkt danach deutlich mehr und in späteren Phasen – vermutlich aufgrund der wenigen passenden Aufgaben – nur vereinzelt, finden sich für die Unterscheidung von Beobachtung und Deutung vor Explizierung viele Ideen-Oberkategorien, nach der Explizierung hingegen deutlich weniger. Insbesondere bei der Unterscheidung von Beobachtung und Deutung dürfte das Ergebnis durch die Aufgaben der Instruktion bedingt sein. Im Vergleich lässt sich ferner hinzufügen, dass die Unterscheidung von Beobachtung und Deutung für die untersuchte Personengruppe, die aus Schülerinnen und Schülern der elften Klassenstufe besteht, als das der beiden Themengebiete angesehen werden kann, für welches sie bereits vor Beginn der Bearbeitung des Lernmaterials angemessenere Vorstellungen haben dürften. Auch diese Tatsache könnte dazu beitragen, dass bei Kennzeichen von naturwissenschaftlichen Fragen eher eine Zunahme der Vielfalt von Vorstellungen durch die Explizierung identifiziert werden kann, bei der Unterscheidung von Beobachtung und Deutung eher eine Abnahme der Vielfalt. Eventuell werden die Lernenden im ersten Fall dazu angeregt, neue Vorstellungen zu konstruieren, wohingegen sie im zweiten Fall bereits vorliegende oder zuvor situativ konstruierte Vorstellungen aufgeben oder zumindest nicht nutzen.

Für beide Explizierungen stellt sich vor dem Hintergrund der Verschränkung von Lehren und Lernen im Kontext von Instruktionen die Frage, wodurch intendierte Lernverläufe gekennzeichnet sein könnten. Ein Aspekt dieser Frage besteht darin, welche Ideen-Oberkategorien im Kontext des betrachteten Materials zurückgehen sollten und welche nach der Explizierung vermehrt auftreten sollten. Die Unterteilung der Ergebnisse in die Angemessenheit der Vorstellungen einerseits und die Übereinstimmung der Vorstellungen mit den explizierten Konzepten andererseits liefert einen Anknüpfungspunkt zur Beantwortung dieser Frage. Eine Abnahme von Vorstellungen, die nicht mit der Explizierung übereinstimmen, scheint nur insofern erwartbar und intendiert, als dass diese im Verhältnis zu den mit der Explizierung übereinstimmenden Vorstellungen betrachtet seltener konstruiert werden dürften. Einzig bei einer Explizierung, die spezifische Vorstellungen als fehlerhaft oder unbrauchbar

deklariert, wäre von einer Abnahme auszugehen. Bei den Bepunktungen für die Übereinstimmung mit den Konzepten der Explizierung wurden entsprechend positive Punkte für affirmativ erwähnte Ideen vergeben – und im Fall der Ideen-Oberkategorie **Fachbezug** bei Kennzeichen von naturwissenschaftlichen Fragen auch ein negativer Punkt vergeben. In den Ergebnissen zeigt sich, dass grundsätzlich von einer relativen Zunahme von zu explizierten Konzepten passenden Vorstellungen gesprochen werden kann. Eine Abnahme der Ideen-Oberkategorie **Fachbezug** bei Kennzeichen von naturwissenschaftlichen Fragen ist ebenfalls wie intendiert und erwartbar zu verzeichnen.

Ein weiterer Aspekt, unter dem die Zu- und Abnahme spezifischer Ideen-Oberkategorien im Kontext der Explizierungen betrachtet werden könnte, ist durch die Unterscheidung von generalisierten und nicht-generalisierten Äußerungen, denen die Ideen-Oberkategorien jeweils zugewiesen sind, gegeben. Allerdings ist aufgrund der Ergebnisse zu F-Vor2, die insbesondere darin bestehen, dass nur sehr wenige Generalisierungen außerhalb von 1-22-ii bzw. 3-21-ii vorliegen, eine Betrachtung der Zu- oder Abnahme von Generalisierungen im Umfeld der Explizierung mit den vorliegenden Daten nicht sinnvoll untersuchbar, sondern bleibt als Desiderat für nachfolgende Arbeiten bestehen. Das Desiderat ist insofern spannend, als dass Explizierungen einerseits grundsätzlich Generalisierungen vermitteln sollen, andererseits unterschiedlich stark auf konkrete Beispiele bezogen sein können. Wodurch (angemessene) Generalisierungen von Lernenden zum experimentbezogenen Denken und Arbeiten gefördert werden, ist aber bisher nicht hinreichend geklärt.

6.9.2 Gehäuft auftretende Vorstellungen zur Beschreibung der individuellen Verläufe von Vorstellungen

Zusätzlich zu einer Betrachtung von Phasen vor und nach der Explizierung, bei der (Teil-)Aufgaben zu Phasen zusammenfasst werden, kann zur Analyse der Vorstellungen im zeitlichen Verlauf auch eine Auflösung der einzelnen Teilaufgaben erfolgen, etwa indem die über mehrere Teilaufgaben hinweg gehäuft auftretenden Vorstellungen (siehe Unterabschnitt 6.8.2 ab S. 330) betrachtet werden. Die für einzelne Personen gehäuft auftretenden Ideen-Oberkategorien sind bereits in Tabelle 6.21 (auf S. 334) und Tabelle 6.23 (auf S. 334) aufgeführt. Die Betrachtung der Verläufe nimmt daher vor allem die zeitliche Auflösung in den Blick, wobei das Zeitliche zum Zwecke der Vergleichbarkeit von Teams durch die Abfolge von (Teil-)Aufgaben operationalisiert ist.

6.9.2.1 Zwei Phasen: Vor und nach der instruktionalen Information

Im einem ersten Schritt der Betrachtung von über mehrere (Teil-)Aufgaben hinweg gehäuft auftretenden Ideen-Oberkategorien wird von der Frage gestartet, ob Instanzen der gehäuft auftretenden Ideen-Oberkategorien (1) nur vor, (2) nur nach oder (3) sowohl vor als auch nach der Explizierung auftreten. Darauf basierend ergeben sich für Personen verschiedene zeitliche Profile von gehäuft auftretenden Ideen-Oberkategorien im zeitlichen Verlauf, die nachfolgend exemplarisch erläutert sind und für die angegeben ist, welchen weiteren Personen das jeweilige Profil zugewiesen ist. Eine schematische Übersicht über die deduktiv hergeleiteten möglichen Profile findet sich in Abbildung 6.32; alle weiteren möglichen Profile weichen nicht substantiell von den dargestellten ab, sondern nur in Anzahlen verschiedener Ideen-Oberkate-

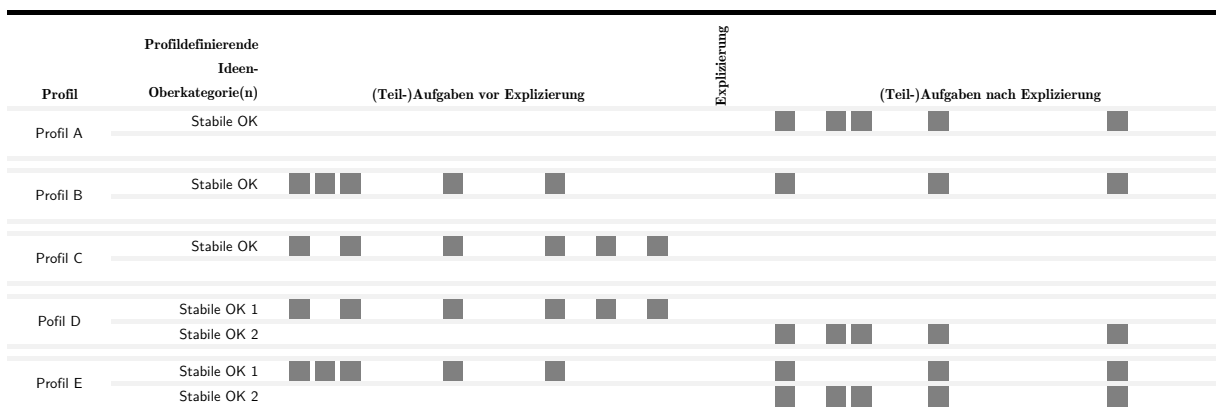


Abbildung 6.32: Schematische Übersicht über die Profile gehäuft auftretender Vorstellungen im zeitlichen Verlauf.

gorien (z. B. wird das Vorkommen einer gehäuft auftretenden Ideen-Oberkategorie *vor* der instruktionalen Information, welches mit dem Vorkommen *zweier* gehäuft auftretender Ideen-Oberkategorien *nach* der Information als Variation von Profil D aufgefasst). Am Ende der Erläuterungen zu den einzelnen Profilen findet sich eine Übersichtstabelle 6.32 (auf S.385), in der dargestellt ist, wie viele Personen den zeitlichen Profilen jeweils zugehörig sind.

Profil A: Alle gehäuft auftretende Ideen-Oberkategorien finden sich nur nach der instruktionalen Information. Hinsichtlich ihrer Verläufe von Vorstellungen über die (Teil-)Aufgaben hinweg lässt sich eine erste Gruppe von Personen identifizieren, für die alle gehäuft auftretenden Ideen-Oberkategorien nach der Explizierung auftreten. Andere Ideen-Oberkategorien können dabei sowohl vor als auch nach der Explizierung mehrfach vorliegen, solange sie nicht viermal oder häufiger vorliegen (denn dann wären sie gehäuft auftretende Ideen-Oberkategorien). In Abbildung 6.33 ist das Verlaufsdigramm für H24 abgebildet (die Verlaufsdigramme für alle Personen sind in Anhang F.5, ab S. 577, vorzufinden). Die nachfolgend als *profil-definierende Ideen-Oberkategorie* bezeichnete Ideen-Oberkategorie ist **MethodeUnspezifisch**, die nur nach der instruktionalen Information auftritt. Die Ideen-Oberkategorien **Fachbezug** und **Empfindungen** liegen beispielsweise bei zwei (Teil-)Aufgaben vor und sind daher nicht stabil im Sinne der Definition gehäuft auftretender Ideen-Oberkategorien. Für andere Personen liegen andere profil-definierende Ideen-Oberkategorien vor und es sind mehr oder weniger Instanzen für andere Ideen-

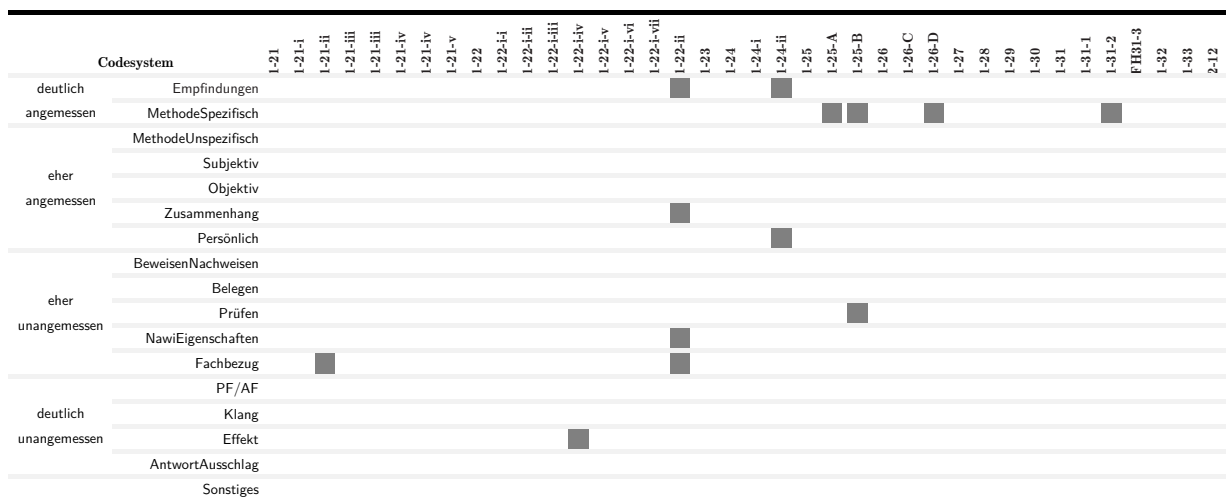


Abbildung 6.33: Ideen-Oberkategorien, die der Schülerin H24 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden. *Profil A*.

Oberkategorien vorhanden. Insgesamt sind dem Profil A, bei dem alle gehäuft auftretenden Ideen-Oberkategorien nach der Explizierung auftreten, für Kennzeichen von naturwissenschaftlichen Fragen 6 Personen zugehörig: A2, A3, H22, H23, H24 (alle mit der profil-definierenden Ideen-Oberkategorie *MethodeSpezifisch*) und C8 (mit *BeweisenNachweisen*). Für die Unterscheidung von Beobachtung und Deutung weisen null Personen das Profil A auf. In ähnlicher Weise wie in Unterabschnitt 6.8.2 (beginnend auf S. 330) werden für die Unterscheidung von Beobachtung und Deutung in einer erweiterten Auffassung von gehäuft auftretenden Ideen-Oberkategorien zusätzlich jeweils auch dreifach auftretende Ideen-Oberkategorien hinzugenommen. Das Profil A weisen dann weiterhin null Personen auf.

Profil B: Alle gehäuft auftretenden Ideen-Oberkategorien finden sich sowohl vor als auch nach der instruktionalen Information. Für eine zweite Gruppe von Personen tritt (mindestens) eine Ideen-Oberkategorie gehäuft und sowohl vor als auch nach der Explizierung auf. In Abbildung 6.34 ist das Verlaufsdiagramm für F18 abgebildet. Zwei Ideen-Oberkategorien sind profil-definierend: *Subjektiv* und *Objektiv*. Insgesamt sind dem *Profil B*, bei dem alle gehäuft auftretenden Ideen-Oberkategorien sowohl vor als auch nach der Explizierung auftreten, für Kennzeichen von naturwissenschaftlichen Fragen 5 Personen zugehörig: F18 (*Subjektiv* und *Objektiv*), G21 (mit der profil-definierenden Ideen-Oberkategorie *Fachbezug*), I25 (*AntwortAusschlag* und *Objektiv*), O45 (*Fachbezug*) und P46 (*AntwortAusschlag*). Für die Unterscheidung von Beobachtung und Deutung weisen null Personen das Profil

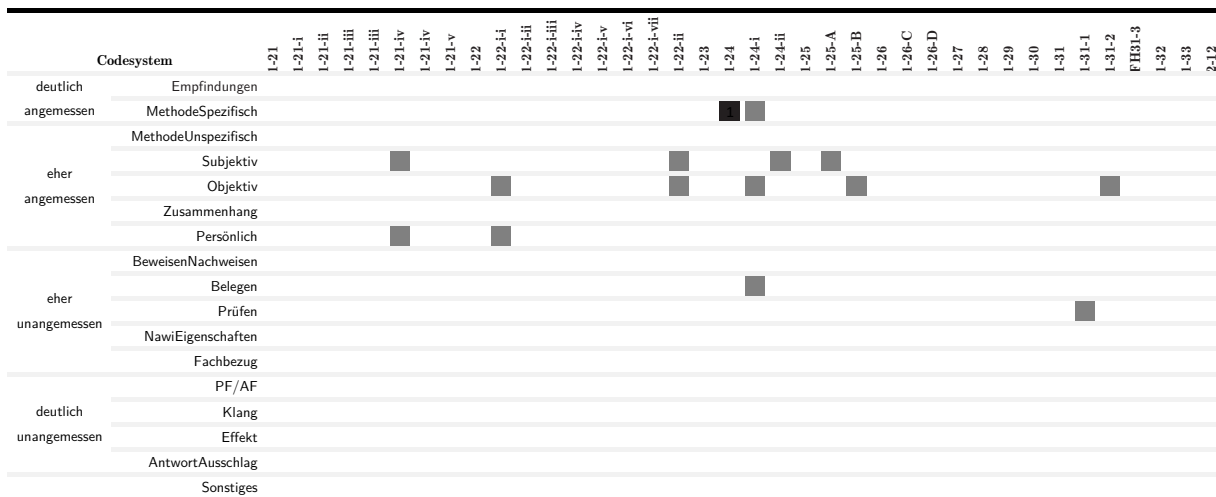


Abbildung 6.34: Ideen-Oberkategorien, die der Schülerin F18 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden. *Profil B*.

auf; werden dreifach auftretende Ideen-Oberkategorien hinzugenommen, so findet sich das Profil B für B5 und K32 (beide mit Vermutung).

Profil C: Alle gehäuft auftretenden Ideen-Oberkategorien finden sich nur vor der instruktionalen Information. Eine dritte Gruppe von Personen hat für alle vorliegenden gehäuft auftretenden Ideen-Oberkategorien nur Instanzen vor der Explizierung. Dieses *Profil C* findet sich für Kennzeichen von naturwissenschaftlichen Fragen nicht. Für die Unterscheidung von Beobachtung und Deutung findet Profil C sich nur für eine Person: P46 (mit der profil-definierenden Ideen-Oberkategorie BegründungErklärung). Das zugehörige Verlaufsdiagramm ist in Abbildung 6.35 dargestellt. Auch bei zusätzlicher Betrachtung dreifach auftretender Ideen-Oberkategorien für die Unterscheidung von Beobachtung und Deutung finden sich nicht mehr Personen in diesem Profil.

Profil D: Manche gehäuft auftretenden Ideen-Oberkategorien finden sich nur vor der, manch andere nur nach der instruktionalen Information. Bei einer vierten Gruppe von Personen ist (mindestens) eine Ideen-Oberkategorie nur vor der Explizierung und (mindestens) eine Ideen-Oberkategorie nur nach der Explizierung stabil vorhanden. Es darf keine Ideen-Oberkategorie vorliegen, die gehäuft und gleichzeitig vor und nach der Explizierung auftritt. Ein Beispiel für das *Profil D* ist A1 bei Kennzeichen von naturwissenschaftlichen Fragen mit den Ideen-Oberkategorien Zusammenhang (vor der Explizierung) und MethodeSpezifisch (nach der Explizierung) – die einzige Person mit diesem Profil.

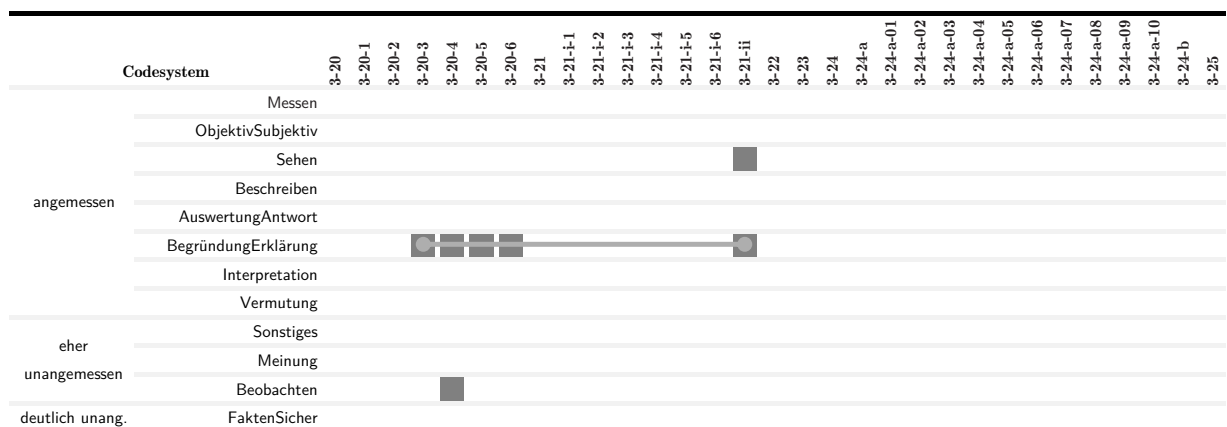


Abbildung 6.35: Ideen-Oberkategorien, die dem Schüler P46 bei den einzelnen (Teil-)Aufgaben zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden. *Profil C*.

Profil E: Manche gehäuft auftretenden Ideen-Oberkategorien finden sich sowohl vor als auch nach der instruktionalen Information, manch andere nur danach. Für Personen mit dem *Profil E* tritt (mindestens) eine Ideen-Oberkategorien gehäuft auf und kommt vor und nach der Explizierung vor. Außerdem tritt (mindestens) eine Ideen-Oberkategorie gehäuft auf, die nur nach der Explizierung auftritt. Diesem Profil sind bei Kennzeichen von naturwissenschaftlichen Fragen 4 Personen zugehörig: G19 (mit der Ideen-Oberkategorie *Fachbezug* vor und nach der Explizierung und der Ideen-Oberkategorie *Objektiv* nur danach), N40 und N41 (beide mit *PF/AF* vor und nach der Explizierung und *MethodeSpezifisch* nur danach) und O44 (mit *Fachbezug* vor und nach der Explizierung und *Prüfen* nur danach). Für die Unterscheidung von Beobachtung und Deutung weist nur die Person H22 das Profil E auf (mit *Vermutung* vor und nach der Explizierung und *BegründungErklärung* nur danach). Werden für die Unterscheidung von Beobachtung und Deutung dreifach auftretende Ideen-Oberkat. hinzugenommen, so finden sich keine weiteren Personen mit dem Profil.

Es ist möglich, dass das letzte Auftreten einer gehäuft auftretenden Ideen-Oberkategorien, die vor und nach der Explizierung vergeben wurde, sehr direkt nach der Explizierung vorliegt (bei 1-24-i bzw. 3-24-a). Eine Zuordnung zum Profil B oder E wurde dann zwar vorgenommen, gleichzeitig könnte der tatsächliche zeitliche Verlauf der Ideen-Oberkategorien aber inhaltlich ggf. so interpretiert werden, dass die Person die zugehörige Vorstellung nach der Explizierung ein letztes Mal versucht zu aktivieren und sie danach aufgibt, was eher dem Profil C (oder dem Profil D) entspricht, bei dem die gehäuft auftretende Ideen-Oberkategorie nur vor der Explizierung vorhanden ist. Analog ist auch das Auftreten einer Ideen-Oberkategorie in der letzten Aufgabe direkt vor der Explizierung (1-21-ii) eine Möglichkeit dafür, dass eine Ideen-Oberkategorie nur knapp über die Explizierung hinweg vorliegt. Die Deutung, dass dies durch die Explizierung induziert sei, ist allerdings in diesem Fall nicht plausibel, weil diese erst zeitlich später gelesen wird (vielmehr dürfte die Ursache in der spezifischen Anlage von Teilaufgabe 1-21-ii liegen; dazu später mehr). Etwas ausführlicher wird auf derartige Spezifika von individuellen Verläufen im nächsten Unterabschnitt 6.9.3 (zu Ketten von Vorstellungen, die über die Explizierung hinweg andauern) eingegangen.

In Tabelle 6.32 ist zusammengefasst, wie viele und welche Personen welchem Profil zugeordnet sind. Außerdem ist dargestellt, welche Personen bei Hinzunahme dreifach auftretender Ideen-Oberkategorien den jeweiligen angepassten Profilen zugeordnet sind. Teilweise wechselt das einer Person zugewiesene Profil, wenn dreifach auftretende Ideen-Oberkategorien betrachtet werden (bspw. A2 von Profil A zu Profil E bei Kennzeichen von naturwissenschaftlichen Fragen, siehe Abbildung 6.36).

Tabelle 6.32: Zuordnung der Personen zu den Profilen gehäuft auftretender Vorstellungen im zeitlichen Verlauf.

Profil	KnF (4)	KnF (3)	BuD (4)	BuD (3)
Profil A	A2, A3, C8, H22, H23, H24	A3, C8, C9, H22, H24, M37, M38, P48	–	–
Profil B	F18, G21, I25, O45, P46	F18, G21, I25, O45, P47	–	B5, K32
Profil C	–	–	P46	P46
Profil D	A1	A1	–	–
Profil E	G19, N40, N41, O44	A2, G19, H23, I26, N40, N41, O43, O44, P46	H22	H22

Anmerkungen. KnF = Kennzeichen von naturwissenschaftlichen Fragen; BuD = Unterscheidung von Beobachtung und Deutung. (x) = Mindeste notwendige Anzahl des mehrfachen Auftretens der profildefinierenden Ideen-Oberkategorien. – = Keine Person.

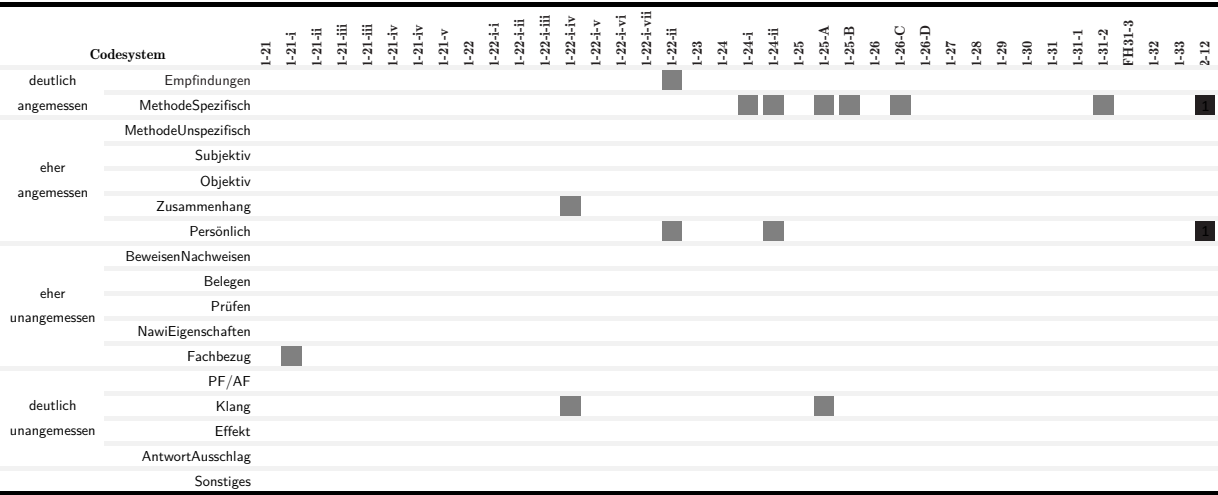


Abbildung 6.36: Ideen-Oberkategorien, die der Schülerin A2 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden. Werden mindestens vierfach vorkommende Ideen-Oberkategorien als gehäufte Ideen-Oberkategorien gewertet, ist die Person dem Profil A zugehörig (nur MethodeSpezifisch nach der instruktionalen Information); werden dreifach vorkommende Ideen-Oberkategorien hinzugenommen, ist die Person dem Profil E zugehörig (Persönlich vor und nach, MethodeSpezifisch nach).

Die Profile lassen sich als Verläufe von Vorstellung in jeweils spezifischer Weise deuten. *Profil A* kann so gedeutet werden, dass vor der Explizierung keine Vorstellungen rekonstruiert werden, weil keine Vorstellungen vorliegen oder keine Notwendigkeit zur Äußerung gesehen wird. Welche der beiden Deutungen plausibler ist, bleibt offen. Zwar ist im Bereich der naturwissenschaftlichen Fragen aufgrund geringen Vorwissens von Schülerinnen und Schülern eher davon auszugehen, dass Vorstellungen nicht oder nur unzureichend abrufbar vorliegen. Gleichwohl finden sich bei 1-21-ii durchaus vielfältige Vorstellungen, obwohl sich diese Aufgabe vor der Explizierung befindet.

Bei den *Profilen B und E* scheinen die Schülerinnen und Schüler an gewissen Vorstellungen festzuhalten, während weitere Vorstellungen hinzukommen. Über beide Themenbereiche hinweg ist das *Profil B* am prominentesten, was dafür spricht, dass Lernende an gewissen Vorstellungen auch über die Explizierung hinweg stabil festhalten. Gleichwohl ist an dieser Stelle zu betonen, dass für 17 (bzw. 10 bei Hinzunahme dreifach auftretender Ideen-Oberkategorien) Personen bei Kennzeichen von naturwissenschaftlichen Fragen beziehungsweise 28 (bzw. 26) Personen bei der Unterscheidung von Beobachtung und Deutung überhaupt keine Profile zugewiesen sind. In der Gesamtstichprobe scheint also eher eine große Fluktuation der Vorstellungen vorzuliegen (siehe auch schon F-Vor5).

Auch im *Profil E* scheinen die Lernenden an gewissen Vorstellungen auch über die Explizierung hinweg stabil festzuhalten. Es ist ebenfalls relativ prominent und liegt bei Hinzunahme der dreifach auftretenden Ideen-Oberkategorien mit deutlichem Abstand für die meisten Personen vor. Anhand der bisher berichteten Daten lässt sich allerdings nicht entscheiden, ob die Vorstellungen situativ gewechselt werden oder ob neue Vorstellungen in ein kohärentes System des Verständnisses integriert werden.

Das *Profil C* findet sich nur für die Unterscheidung von Beobachtung und Deutung und könnte ein Artefakt des untersuchten Instruktionsextrakts sein. Das Profil könnte jedoch auch insofern mit der Explizierung zusammenhängen, als dass die Lernenden nach der Explizierung nur noch eine sehr geringe Notwendigkeit sehen, ihre Vorstellungen zu äußern, weil diese – eventuell zumindest aus ihrer Sicht – stark mit den explizierten Konzepten übereinstimmen.

Im *Profil D* scheint die Explizierung einen klaren Wechsel von Vorstellungen zu induzieren. Gleichzeitig könnten Fälle, bei denen das letzte Auftreten einer Ideen-Oberkategorie bei der Aufgabenkarte direkt nach der Explizierung stattfindet, wie oben geschildert, in einem ähnlichen Sinne gedeutet werden, obwohl sie in einem anderen Profil einsortiert sind.

Die jeweiligen profil-definierenden Ideen-Oberkategorien (das sind die in einer spezifischen Phase des Instruktionsextrakts gehäuft auftretenden Ideen-Oberkategorien) sind in Tabelle 6.33 gelistet. Es wird deutlich, dass für Kennzeichen von naturwissenschaftlichen Fragen die Angemessenheit der Ideen-Oberkategorien in allen Profilen durchmischt ist und nach der Explizierung hinzukommende Ideen-Oberkategorien (Profile A und E) eine mittlere bis hohe Übereinstimmung mit den Ideen-Oberkategorien der Explizierung aufweisen. Die am häufigsten sowohl vor als auch nach der Explizierung vertretene Ideen-Oberkategorie (Profile B und E) ist **Fachbezug**. Das ist vor dem Hintergrund, dass diese Ideen-Oberkategorie zwar eher unangemessen ist, aber gleichzeitig eine plausible Stütze für Klassifizierung von Fragen darstellt, nachvollziehbar. Allerdings treten auch deutlich unangemessene oder eher unangemessene Ideen-Oberkategorien als vor und nach der Explizierung gehäuft auftretende Ideen-Oberkategorien auf, denen durch die Explizierung deutlicher widersprochen wird. Hier scheinen die Lernenden besonders stark an den spezifischen Vorstellungen (insbesondere **AntwortAusschlag**, **PF/AF**) festzuhalten. Ideen-Oberkategorien mit hoher Angemessenheit treten selten als vor und nach der Explizierung gehäuft auftretende Ideen-Oberkategorien auf, nur für **Objektiv** und **Subjektiv** ist dies der Fall. Die Kategorien der Objektivität und Subjektivität scheinen einzelnen Ler-

Tabelle 6.33: Häufigkeit der Ideen-Oberkategorien als profil-definierende Ideen-Oberkategorien aufgelöst nach Profilen gehäuft auftretender Ideen-Oberkategorien im zeitlichen Verlauf.

Profil	KnF (4)	KnF (3)	BuD (4)	BuD (3)
Profil A	MethodeSpezifisch (5), BeweisenNachweisen (1)	MethodeUnspezifisch (2), MethodeSpezifisch (6), BeweisenNachweisen (2)	–	–
Profil B	Fachbezug (2), AntwortAusschlag (2), Objektiv (2), Subjektiv (1)	Fachbezug (2), AntwortAusschlag (1), Objektiv (2), Subjektiv (1), BeweisenNachweisen (1)	–	Vermutung (2)
Profil C	–	–	BegründungErklärung (1)	–
Profil D	Zusammenhang & MethodeSpezifisch	–	–	–
Profil E	PF/AF & MethodeSpezifisch (2), Fachbezug & Prüfen (1), Fachbezug & Objektiv (1)	PF/AF & MethodeSpezifisch (2), PF/AF & Persönlich (1), Fachbezug & Objektiv (1), Fachbezug & Prüfen (1), Fachbezug & BeweisenNachweisen (1), Fachbezug & MethodeSpezifisch (1), Objektiv & Prüfen (1), NawiEigenschaften & MethodeSpezifisch (1), AntwortAusschlag & BeweisenNachweisen (1)	Vermutung & BegründungErklärung (1)	–

Anmerkungen. KnF = Kennzeichen von naturwissenschaftlichen Fragen; BuD = Unterscheidung von Beobachtung und Deutung. (x) = Mindeste notwendige Anzahl des mehrfachen Auftretens der profildefinierenden Ideen-Oberkategorien. – = Keine Person.

nenden bereits vor der Beschäftigung mit der Explizierung zugänglich zu sein und sie scheinen die Explizierung so zu verstehen, dass diese Vorstellungen nicht aufgegeben werden müssen. Bei der Unterscheidung von Beobachtung und Deutung ist aufgrund der Datenstruktur (besonders aufgrund der geringen Anzahl von Instanzen nach der Explizierung) nicht davon auszugehen, dass spezifische Vorstellungen aufgrund der Explizierung aufgegeben werden.

Die in den Profilen A und E nach der Explizierung hinzukommenden Ideen-Oberkategorien sind eher angemessen. Die Übereinstimmung mit der Explizierung ist nicht grundsätzlich gegeben, allerdings scheinen die für die Lernenden kodierten Ideen-Oberkategorien vor dem Hintergrund des Wortlautes der der Explizierung plausibel. Beispielsweise ist die Ideen-Oberkategorie Prüfen vor dem Hintergrund der Formulierung »dass sie [eine naturwissenschaftliche Fragestellung] mit naturwissenschaftlichen Methoden untersucht werden kann« (1-23, siehe Abbildung 6.26) nachvollziehbar. Nach der Explizierung hinzukommende gehäuft auftretende Ideen-Oberkategorien scheinen also tendenziell in Übereinstimmung mit der Explizierung und daher auch von höherer Angemessenheit zu sein.

6.9.2.2 Differenziertere Betrachtung für die Phasen des

Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen

In einem zweiten Schritt der Betrachtung von über mehrere (Teil-)Aufgaben hinweg gehäuft auftretenden Ideen-Oberkategorien wird die Möglichkeit der Differenzierung von zeitlich der instruktionalen Information nachlaufenden Aufgaben in ein Zeitfenster *direkt nach* der Explizierung, ein Zeitfenster *später nach* der Explizierung und ein Zeitfenster *deutlich später nach* der Explizierung genutzt. Bisher wurden alle Aufgaben als äquivalente Bestandteile ein und derselben Phase nach der instruktionalen Information behandelt, das heißt, als schlossen spätere und deutlich spätere Aufgaben ohne zeitlichen Abstand an die Aufgaben direkt nach der Explizierung an. Wird aufgelöst, welche Ideen-Oberkategorien auch bzw. nicht bzw. nur in der späteren und der deutlich späteren Phase nach der Explizierung auftreten, lassen sich einige der bisher dargestellten Profile ausdifferenzieren.

Für Profil A gibt es zum einen das Subprofil *A-nach*, bei dem die profil-definierende Ideen-Oberkategorie das letzte Mal bei oder vor der Aufgabenkarte 1-27 auftritt, zum anderen das Subprofil *A-später*, bei dem die profil-definierende Ideen-Oberkategorie zum letzten Mal bei 1-31 vorkommt, und zuletzt noch das Subprofil *A-deutlichSpäter*, in dem die profil-definierende Ideen-Oberkategorien auch noch bei 2-12 kodiert ist.

Dem Subprofil A-nach lässt sich eine Person zuordnen (C8 mit **BeweisenNachweisen**), dem Subprofil A-später lassen sich 3 Personen zuordnen (H22, H23, H24, alle mit **MethodeSpezifisch**). Das Subprofil A-deutlichSpäter findet sich für 2 Personen (A2 und A3, beide mit **MethodeSpezifisch**).

Analog können die Subprofile B-nach, B-später und B-deutlichSpäter unterschieden werden, bei denen die profil-definierende Ideen-Oberkategorie vor und nach der Explizierung vorliegt und die sich nur darin unterscheiden, bei welcher Aufgabenkarte sie das letzte Mal vergeben wurde. Das Subprofil B-nach lässt sich für 4 Personen finden (F18 mit **Subjektiv**; G21, O45, beide mit **Fachbezug**; I25 mit **Objektiv** und **Subjektiv**). Dem Subprofil B-später lassen sich 2 Personen zuweisen (F18 mit **Objektiv**; P46 mit **AntwortAusschlag**). Das Subprofil B-deutlichSpäter findet sich nicht in den Daten.

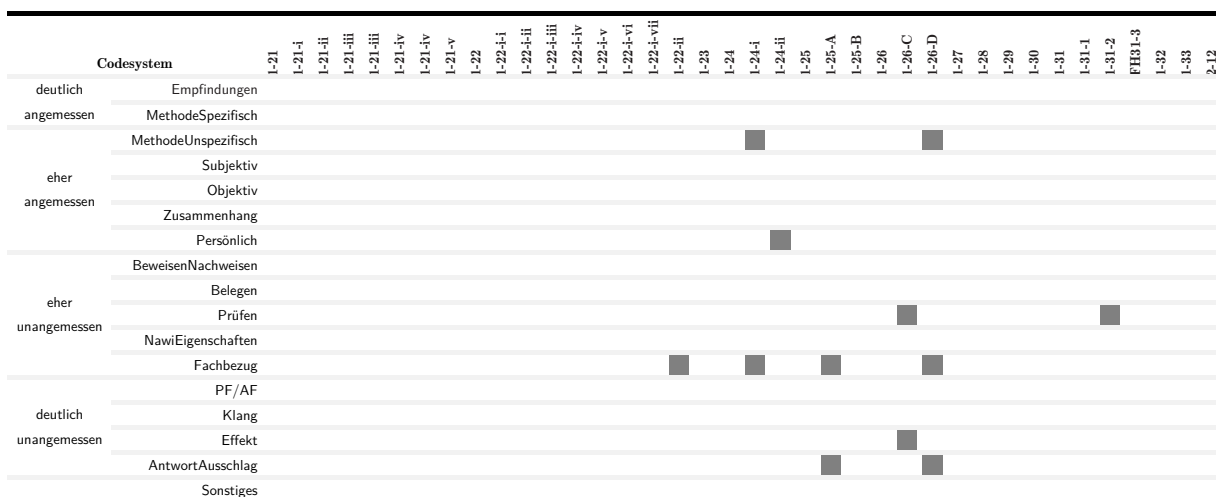
Die Schülerin A1, die als einzige dem Profil C zugeordnet ist, fällt in das Subprofil C-deutlichSpäter, weil die Ideen-Oberkategorie **MethodeSpezifisch** von der Explizierung bis zur Aufgabenkarte 2-12 vergeben wurde. An dieser Stelle kann angemerkt werden, dass eine Zugehörigkeit zu einem Subprofil *-deutlichSpäter nur einschließt, dass die profil-definierende Ideen-Oberkategorie bei 2-12 vergeben wurde. Über das Auftreten der profil-definierenden Ideen-Oberkategorie bei 1-31 sagt die Zugehörigkeit in einem Subprofil *-deutlichSpäter nichts aus. Für A1 tritt die Ideen-Oberkategorie **MethodeSpezifisch** allerdings sowohl bei 1-31 als auch bei 2-12 auf.

Profil D tritt bei Kennzeichen von naturwissenschaftlichen Fragen nicht auf. Für das Profil E finden sich bei dem Subprofil E-nach 3 Personen (G19 mit **Fachbezug**; N40 mit **PF/AF**; N41 mit **MethodeSpezifisch**), im Subprofil E-später findet sich eine Person (O44 mit **Prüfen**), das Subprofil E-deutlichSpäter findet sich nicht in den Daten.

Es lässt sich anhand der dargestellten Ergebnisse zu den Subprofilen feststellen, dass profil-definierende Ideen-Oberkategorien, die erst nach der Explizierung vorkommen (Profil A), relativ häufig in Subprofilen auftreten, die *-später oder *-deutlichSpäter lauten. Dem entgegen findet sich, dass profil-definierende Ideen-Oberkategorien, die bereits vor der Explizierung vorlagen (Profile B und E) vorrangig in den Subprofilen *-nach, nur zu geringen Teilen in den Subprofilen *-später und überhaupt nicht in den Subprofilen *-deutlichSpäter auftreten. Eine Ursache dafür kann in dem methodischen Vorgehen liegen, bei dem eine Ideen-Oberkategorie mindestens viermal vorliegen muss, um als gehäuft auftretende Ideen-Oberkategorie angesehen zu werden. Tritt eine Ideen-Oberkategorie also erst nach der Explizierung auf, ist die Wahrscheinlichkeit größer, dass sie auch bei späteren oder deutlich späteren (Teil-)Aufgaben kodiert

werden muss, um überhaupt gehäuft auftreten zu können (weil die Anzahl verfügbarer Aufgaben kleiner ist als in dem Fall, dass bereits vor der Explizierung Instanzen vorliegen). Diese Tatsache könnte zugleich damit einhergehen, dass für viele Personen mit den Profilen B und E bei späteren oder deutlich späteren (Teil-)Aufgaben überhaupt keine Kodierungen von Ideen-Oberkategorien vorliegen. Eine andere Ursache könnte sein, dass in dem untersuchten kurzen Zeitfenster die Personen nicht genügend Gelegenheiten hatten, alternative Vorstellungen zu nutzen und zu festigen. Aus dem Verlaufsdigramm für O45 (Abbildung 6.37) geht hervor, dass die Schülerin beispielsweise eventuell bei weiteren Aufgaben die Ideen-Oberkategorien Prüfen festigen hätte können.

Insgesamt lässt sich festhalten, dass die Subprofile Hinweise darauf geben können, wie überdauernd die profil-definierenden Ideen-Oberkategorien sind und welche Ideen-Oberkategorie am Ende der Sequenz tatsächlich genutzt wird. Wird ferner betrachtet, welche Ideen-Oberkategorien bei 1-31 und 2-12 überhaupt auftreten und ob diese zu den gehäuft auftretenden Ideen-Oberkategorien gehören, ergeben sich teilweise spannende Befunde. Ein Beispiel dazu ist N41, der nach der Explizierung vorrangig **MethodeSpezifisch** nutzt, aber bei 2-12 wieder zurückkehrt zu PF/AF. Bei H23 (siehe Abbildung 6.38) zeigt sich eigentlich das Profil A mit **MethodeSpezifisch** nach der Explizierung; bei Hinzunahme dreifach auftretender Ideen-Oberkategorien ergibt sich, dass **Fachbezug** und **NawiEigenschaften** gehäuft vor und nach der Explizierung auftreten. Die Ideen-Oberkategorie **Fachbezug** liegt allerdings das letzte Mal bei



390

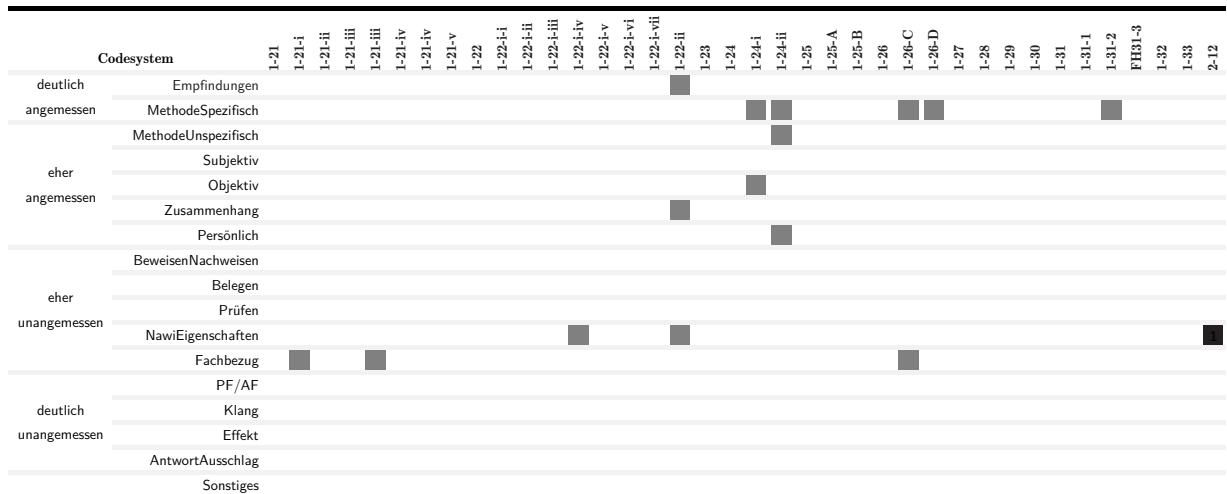


Abbildung 6.38: Ideen-Oberkategorien, die der Schülerin H23 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden.

1-26-C vor, MethodeSpezifisch das letzte Mal bei 1-31-2 und NawiEigenschaften liegt – nachdem es bei 1-22-ii zuletzt vorkommt – noch bei 2-12 vor. Dies könnte so gedeutet werden, dass NawiEigenschaften die gesamte Explizierung überdauert hat. Insbesondere wenn berücksichtigt wird, dass Fachbezug und NawiEigenschaften inhaltlich nicht sehr weit voneinander entfernt sind, liegt die Deutung nahe, dass *eine* Vorstellung von H23 sich in den beiden Ideen-Oberkategorien widerspiegelt und nahezu unbeeinflusst von der Explizierung überdauert. Diese Deutung steht der Zuordnung zum Profil A entgegen und kann nur in Ansätzen aus der Zuordnung zum Subprofil A-nach entnommen werden. Erst Analysen der einzelnen (Teil-)Aufgaben, bei denen nicht nur nach den groben Zeiträumen vor und nach der Explizierung gruppiert wird, sondern auf die tatsächlichen Aufgaben geschaut wird, bei denen die Ideen-Oberkategorien (beispielsweise zum ersten und zum letzten Mal) vergeben werden, liefern derartige Einsichten. Die im nächsten Abschnitt berichteten Analysen finden gezielt auf dieser Ebene statt.

6.9.3 Ketten von Vorstellungen zur Beschreibung der individuellen Verläufe von Vorstellungen

Eine Betrachtung von Verläufen über mehrere (Teil-)Aufgaben hinweg, die zugleich die einzelnen Teilaufgaben als Analyseeinheiten nutzt, kann nicht nur durch Betrachtung der gehäuft auftretenden Ideen-Oberkategorien erfolgen, sondern auch mithilfe der bereits für F-Vor5 in Unterabschnitt 6.8.4 untersuchten Ketten von über mehrere Aufgaben ununterbrochen vorkommenden Ideen-Oberkategorien. Während gehäuft auftretende Ideen-Oberkategorien darauf bezogen sind, wie prominent die zugehörigen Vorstellungen über den gesamten zeitlichen Verlauf hinweg sind, stellen Ketten einen Bezug dazu her, wie sehr Vorstellungen bei der nächsten Gelegenheit (also der nächsten Aufgabe, bei der derselben Person eine Ideen-Oberkategorie zugewiesen ist) wieder aktiviert werden. In Bezug auf die gegebenen instruktionalen Informationen ist daher unter anderem interessant, ob und inwiefern vor der Information aktivierte Vorstellungen auch ohne Unterbrechung nach der Information wieder aktiviert werden.

Für **Kennzeichen von naturwissenschaftlichen Fragen** beginnen 26 von insgesamt 63 Ketten vor der Explizierung. Die Hälfte dieser 26 Ketten überdauert die Explizierung; das heißt, mindestens eine Zuweisung der Ideen-Oberkategorie erfolgt vor und mindestens eine Zuweisung erfolgt nach der Explizierung. In Tabelle 6.34 ist für jede Person aufgeführt, wie viele Ketten vor der Explizierung, nach der Explizierung und über die Explizierung hinweg vorliegen. Für die Ketten, die über die Explizierung hinweggehen (und somit als für Vorstellungen stehend gedeutet werden, die vor der und ohne Unterbrechung nach der instruktionalen Information wieder aktiviert werden), sind zudem die zugewiesenen Ideen-Oberkategorien vermerkt. Die 13 Ketten, die über die Explizierung hinweggehen, sind auf 11 Personen verteilt (für zwei Personen liegen jeweils zwei Ketten vor). Für diese 11 Personen liegen zumeist zusätzlich auch Ketten vor, die auf den Zeitraum nach der Explizierung beschränkt sind. Für alle drei Personen, bei denen nur die jeweils eine Kette über die Explizierung hinweg vorliegt, gehört diese Kette zur Ideen-Oberkategorie **Fachbezug**. Insgesamt betrachtet gehören fünf der Ketten, die über die Explizierung hinweg auftreten, zur Ideen-Oberkategorie **Fachbezug** (zwei verschiedene Teams), jeweils zwei zu **AntwortAusschlag** und **Objektiv** (beide Male bei zwei verschiedenen Teams), jeweils eine Kette gehört zu **NawiEigenschaften**, **Empfindungen**, **BeweisenNachweisen** und **MethodeSpezifisch**. Die häufigste Ideen-Oberkategorie bei Ketten nach der Explizierung ist **MethodeSpezifisch** (12 von 38; gefolgt von **Persönlich** mit 5 Ketten).

Die Ergebnisse lassen sich so deuten, dass die Lernenden für ein breites Spektrum von Vorstellungen diese in direkt folgenden Situationen einsetzen – auch dann, wenn dazwischen eine instruktionale Information gegeben wird. Bei Vorstellungen, die in

Tabelle 6.34: Anzahl der Ketten zum Instruktionsextrakt für Kennzeichen von naturwissenschaftlichen Fragen pro Person, aufgelöst danach, ob alle Instanzen der Kette vor der instruktionalen Information vorliegen (*nur davor*) oder danach (*nur danach*) oder mindestens eine davor und mindestens eine danach (*über*). Für Ketten, die über die instruktionale Information hinweggehen, ist zudem die zugehörige Ideen-Oberkategorie angegeben. Es sind alle 31 analysierten Personen aufgeführt.

Person	Anzahl der Ketten			Ideen-Oberkategorien der Ketten über Info
	nur davor	nur danach	über	
A1	2	1		
A2		1		
A3		1		
B4				
B5	2	1		
B6		1		
C7	1	2		
C8		1		
C9		2		
F16	1	1	1	MethodeSpezifisch
F17		1		
F18	1	2	1	Objektiv
G19		1	1	Fachbezug
G20		2		
G21			1	Fachbezug
H22	1	2		
H23	2	1		
H24		1	1	Empfindungen
I25		2	2	AntwAusschlag, Objektiv
I26		1		
I27				
M37	1			
M38		2		
M39				
N40	1	1		
N41	1	2		
N42		1		
O43		3	1	Fachbezug
O44		1	2	Fachbezug, NawiEig
O45			1	Fachbezug
P46		2	1	AntwAusschlag
P47		1	1	BeweisenNachweisen
P48		1		
Spaltensumme Σ	13	38	13	
Anteil ($\Sigma/64$)	20,31%	59,38%	20,31%	

der Information nicht thematisiert werden, ist einigermaßen nachvollziehbar, dass Personen diese nach der Information erneut anbringen. Ferner ist insbesondere für die Ideen-Oberkategorie **Fachbezug** – die aus fachmethodischer Sicht eine gewisse Berechtigung hat, wenngleich sie alleinstehend eher als problematisch anzusehen ist – ein Auftreten nach der Explizierung als erwartbar einzuschätzen. Allerdings treten in wenigen Fällen auch Ideen-Oberkategorien auf, die als eher oder deutlich unangemessen einzuschätzen sind, für die eine solche Plausibilisierung des Auftretens über die Explizierung hinweg nicht möglich scheint. Beispielhaft sei hier die Ideen-Oberkategorie **AntwortAusschlag** genannt. Die beiden Personen, bei denen **AntwortAusschlag** über die Explizierung hinweg vorliegt, sind nachfolgend detaillierter dargestellt, weil sich an ihnen auch weitere Befunde illustrieren lassen.

Aus Abbildung 6.39 geht hervor, dass für die Schülerin I25 jeweils zwei Ketten zu den beiden Ideen-Oberkategorien **AntwortAusschlag** und **Objektiv** vorliegen (jeweils von 1-22-ii bis 1-24-i und von 1-25-A bis 1-26-B). Die Schülerin I25 redet in Bezug auf drei der zugehörigen Beispielfragen (Teilaufgaben 1-24-i, 1-25-A, 1-26-B) wörtlich davon, dass diese »objektiv beantwortbar« sind. Alle drei Äußerungen wurden mit **AntwortAusschlag** versehen (weil »beantwortbar« einen Bezug dazu herstellen könnte, dass Antworten vorliegen müssen), zugleich aber auch mit **Objektiv**. Vor dem Hintergrund, dass nur die Ideen-Oberkategorie **Subjektiv** zwischen den Ketten kodiert wurde (bei 1-24-ii), diese aber einen starken Bezug zur Ideen-Oberkategorie **Objektiv** hat (siehe Diskussion zu F-Vor1 in Abschnitt 6.4 ab S.283), kann der Verlauf für I25 (anders als zunächst von den Ketten her naheliegend) so gedeutet werden, dass eine eher angemessene Vorstellung bereits vor der Explizierung vorliegt

Codesystem		1-21	1-21-i	1-21-ii	1-21-iii	1-21-iv	1-21-v	1-22	1-22-i	1-22-ii	1-22-iii	1-22-iv	1-22-v	1-22-vi	1-22-vii	1-22-ii	1-23	1-24	1-24-i	1-24-ii	1-25	1-25-A	1-25-B	1-26	1-26-C	1-26-D	1-27	1-28	1-29	1-30	1-31	1-31-1	1-31-2	FH31-3	1-32	1-33	2-12	
deutlich	Empfindungen																																					
angemessen	MethodeSpezifisch																																					
	MethodeUnspezifisch																																					
eher	Subjektiv																																					
angemessen	Objektiv																																					
	Zusammenhang																																					
	Persönlich																																					
	BeweisenNachweisen																																					
eher	Belegen																																					
unangemessen	Prüfen																																					
	NawiEigenschaften																																					
	Fachbezug																																					
	PF/AF																																					
deutlich	Klang																																					
unangemessen	Effekt																																					
	AntwortAusschlag																																					
	Sonstiges																																					

Abbildung 6.39: Ideen-Oberkategorien, die der Schülerin I25 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden.

(situativ aktiviert wird) und auch danach fortgeführt wird. Gegebenenfalls ist diese Vorstellung nur durch ein Konglomerat an Ideen-Oberkategorien abzubilden (Objektiv, Subjektiv, evtl. sogar **AntwortAusschlag**; siehe zu Konglomeraten Abschnitt 6.4). Zu Abbildung 6.39 lässt sich ferner anmerken, dass das (Nicht-Mehr-)Auftreten weiterer Ideen-Oberkategorien nicht notwendigerweise als Folge der Explizierung zu sehen ist, insbesondere dann nicht, wenn vor (oder nach) der Explizierung nur wenige (Teil-)Aufgaben vorliegen, bei denen Ideen-Oberkategorien zugewiesen wurden.

Die Äußerungen des Schülers P46, die mit **AntwortAusschlag** versehen sind, sind nicht nur über einen größeren Zeitraum verteilt als bei I25 (siehe Abbildung 6.40), sondern auch inhaltlich alleinstehend und klar auf die Natur der Antwort bezogen. Zum Beispiel sagt er:

P46: Zum Beispiel ähm (.) schmeckt/ es/ Mineralwasser und Leitungswasser, [eine Anspielung auf die Beispielfrage: »Schmeckt Mineralwasser mit Kohlensäure besser als Mineralwasser ohne Kohlensäure?«] das kann anders spe/ äh jedem anders schmecken.

P48: //Ja.//

P46: //Von daher// ich denke, es muss eine eindeutige Lösung geben und äh/ (.) (P48 nickt) Ja. (Abs. 28–30, Details im Anhang, siehe Hägele, 2022, S. 61–62)

Weitere Äußerungen sind ähnlich oder vergleichbar und die bei denselben Teilaufgaben vorkommenden Ideen-Oberkategorien (insbesondere **BeweisenNachweisen**) sind auf andere Äußerungen desselben Schülers zur selben Teilaufgabe bezogen. Demnach aktiviert P46 zwar mehrere Vorstellungen bei den Teilaufgaben, aber eben auch in den

Codesystem		1-21	1-21-i	1-21-ii	1-21-iii	1-21-iv	1-21-v	1-22	1-22-i-i	1-22-i-ii	1-22-i-iii	1-22-i-iv	1-22-i-v	1-22-i-vi	1-22-i-vii	1-22-ii	1-23	1-24	1-24-i	1-24-ii	1-25	1-25-A	1-25-B	1-26	1-26-C	1-26-D	1-27	1-28	1-29	1-30	1-31	1-31-1	1-31-2	PH31-3	1-32	1-33	2-12		
deutlich	Empfindungen																																						
angemessen	MethodeSpezifisch																																						
	MethodeUnspezifisch																																						
eher	Subjektiv																																						
angemessen	Objektiv																																						
	Zusammenhang																																						
	Persönlich																																						
	BeweisenNachweisen																																						
	Belegen																																						
eher	Prüfen																																						
unangemessen	NawiEigenschaften																																						
	Fachbezug																																						
	PF/AF																																						
deutlich	Klang																																						
unangemessen	Effekt																																						
	AntwortAusschlag																																						
	Sonstiges																																						

Abbildung 6.40: Ideen-Oberkategorien, die dem Schüler P46 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden.

aufeinanderfolgenden Situationen jedes Mal die Vorstellung, die zur Ideen-Oberkategorie **AntwortAusschlag** gehört. Auch nach der Explizierung ändert sich an seinen Formulierungen nicht viel (bspw. »Es gibt eine eindeutige Antwort«; Abs. 70), so dass davon auszugehen ist, dass die Vorstellung unberührt von der Explizierung bestehen bleibt.

Für die **Unterscheidung von Beobachtung und Deutung** findet sich eine Übersicht über alle Ketten in Tabelle 6.35. Es liegen insgesamt nur 8 Ketten vor, davon keine einzige, die nur auf den Zeitraum nach der Explizierung beschränkt ist, und eine einzige, die über die Explizierung hinweggeht – diese findet sich zur Ideen-Oberkategorie **BegründungErklärung** bei dem Schüler K32. Wie in Abschnitt 6.8 bereits ausgeführt, scheinen diese Ergebnisse vor allem mit der Datenlage insgesamt zusammenzuhängen, so dass keine spezifischen Deutungen hinsichtlich der Verläufe von Vorstellungen in den Prozessen des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten möglich sein dürften. Vielmehr liegt vor dem Hintergrund der Ausführungen zu dieser und anderen Forschungsfragen nahe, dass für die Unterscheidung von Beobachtung und Deutung die Vorstellungen der Lernenden angemessen und größtenteils in Übereinstimmung mit der Explizierung sind, weshalb sie auch nach der Explizierung wieder auftauchen müssten. Vermutlich tun sie dies deshalb nicht, weil die Situationen danach das Verbalisieren von Vorstellungen nicht gut genug ermöglichen oder die Lernenden keine Notwendigkeit sehen, die Vorstellungen zu verbalisieren.

Tabelle 6.35: Anzahl der Ketten zum Instruktionsextrakt für die Unterscheidung von Beobachtung und Deutung pro Person, aufgelöst danach, ob alle Instanzen der Kette vor der instruktionalen Information vorliegen (*nur davor*) oder danach (*nur danach*) oder mindestens eine davor und mindestens eine danach (*über*). Für Ketten, die über die instruktionale Information hinweggehen, ist zudem die zugehörige Ideen-Oberkategorie angegeben. Es sind 25 Personen ohne Ketten nicht aufgeführt.

Person	Anzahl der Ketten			Ideen-Oberkategorien der Ketten über Info
	nur davor	nur danach	über	
B5	1			
H22	2			
K32	1		1	BegründungErklärung
L36	1			
P46	1			
P48	1			
Spaltensumme Σ	7	0	1	
Anteil ($\Sigma/8$)	10,94%	0,00%	1,56%	

6.10 Auswertungen und Ergebnisse zu Vorstellungen für Lernende mit unterschiedlichen Kompetenzzuwächsen und Eingangsvoraussetzungen

Die in den vorigen Abschnitten berichteten Analysen und Ergebnisse beziehen sich auf die Vorstellungen der untersuchten Lernenden bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante, wobei die Lernenden einzeln betrachtet sowie nach ihren Teams gruppiert betrachtet wurden. Eine zentrale Frage zum Kompetenzaufbau ist darauf gerichtet, inwiefern die berichteten Ergebnisse sich für Lernenden mit unterschiedlichen Kompetenzzuwächsen und von Lernenden mit unterschiedlichen Eingangsvoraussetzungen darstellen.

F-Vor7 Welche Unterschiede und Gemeinsamkeiten finden sich bezüglich der Forschungsfragen F-Vor1 – F-Vor6 für Lernende mit unterschiedlichen Kompetenzzuwächsen zum experimentbezogenen Denken und Arbeiten von der Prä- zur Post-Erhebung sowie für Lernende mit unterschiedlichen Eingangsvoraussetzungen?

Um dieser Forschungsfrage nachzugehen, werden die analysierten *Lernenden* (33 Personen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen, 31 Personen für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung) für jede von Vorholzer (2016) erhobene Testvariable (siehe bereits in Abschnitt 4.5, beginnend auf S. 63, und überblicksartig in Abbildung 4.1 auf S. 51) *gruppiert*. Die Testvariablen beziehen sich auf

- die fachinhaltlichen Kompetenzen (FCI),
- die allgemeinen kognitiven Fähigkeiten (KFT),
- das naturwissenschaftsbezogene Selbstwirksamkeitskonzept (SWK) sowie
- die physikbezogenen Interessen (INT)

der Lernenden jeweils zum Zeitpunkt vor dem Einsatz der Instruktion.

Ferner wurden

- die fachmethodischen Kompetenzen zum Zeitpunkt vor dem Einsatz der Instruktion (FM-Prä) sowie
- die fachmethodischen Kompetenzen zum Zeitpunkt nach dem Einsatz der Instruktion (FM-Post)

erhoben, aus denen sich als Differenz

- der Kompetenzzuwachs der Lernenden (Zuwachs = FM-Post minus FM-Prä) ergibt. In Unterabschnitt 6.10.1 wird das Vorgehen zur Gruppierung der Lernenden für die einzelnen Testvariablen vorgestellt und eine Übersicht darüber gegeben, welche Gruppen von Lernenden sich ergeben.

Um die Forschungsfrage zu beantworten, werden die Eingangsvoraussetzungen sowie die Kompetenzzuwächse der Lernenden mit den *Prozessmerkmalen* in Beziehung gesetzt, die Gegenstände der Forschungsfragen F-Vor1 bis F-Vor6 sind. Eine erläuternde Übersicht über diese Prozessmerkmale und die zugehörigen Prozessvariablen wird in Unterabschnitt 6.10.2 gegeben. Aus Zeitgründen erfolgte nur der Vergleich anhand der Prozessmerkmale, die sich durch diese Variablen abbilden ließen, aber keine detaillierte Aufarbeitung aller Detailergebnisse zu den Forschungsfragen durch vollständige Reanalyse der Daten für die Substichproben der Lernenden mit hohen sowie niedrigen Kompetenzzuwächsen; denn dafür wären nicht nur weitere Analysen, sondern aufgrund der Stichprobengrößen zum Teil auch weitere Kodierungen notwendig gewesen.

Die nachfolgend aufgeführten generierten Ergebnisse stellen eine *grobe Näherung* an potentielle Zusammenhänge dar. Es wird zunächst eine knappe Übersicht über die Zusammenhänge bei allen Kombinationen von Test- und Prozessvariablen gegeben (Unterabschnitt 6.10.3). Allerdings sind nicht für alle Kombinationen von Test- und Prozessvariablen Zusammenhänge zu erwarten. Ferner wird aufgrund der Stichprobengröße und des notwendigen Zeiteinsatzes auf Analysen verzichtet, in denen mehr als zwei Test- bzw. Prozessvariablen aufeinander bezogen werden (bspw. böte es sich bei größeren Substichproben an, mittels ANOVA statistisch auf die Frage einzugehen, ob Personen mit großen Kompetenzzuwächsen zu Fachmethoden auch dann dadurch charakterisiert werden können, dass für sie signifikant mehr verschiedene Vorstellungen rekonstruiert werden, wenn die Eingangsvoraussetzungen der allgemeinen kognitiven Fähigkeiten oder der fachinhaltlichen Kompetenzen kontrolliert werden). Da nicht alle weiteren Kombinationen einzeln diskutiert werden sollen und können, wird in den weiteren Unterabschnitten 6.10.4, 6.10.5 und 6.10.6 exemplarisch auf ausgewählte Kombinationen von Test- und Prozessvariablen eingegangen.¹⁰⁹

¹⁰⁹ Anhand der vorliegenden Daten könnte mit etwas Zeiteinsatz anhand von Fallvergleichen darauf eingegangen werden, ob und wie mehrere Test- und Prozessvariablen miteinander in Verbindung stehen. Beispielsweise könnten für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen die 5 Lernenden mit hohen Kompetenzzuwächsen *und* hoch ausgeprägten Eingangsvoraussetzungen bezüglich der fachinhaltlichen Kompetenzen mit den 2 Lernenden verglichen werden, die ebenfalls hohe Kompetenzzuwächse aufweisen, aber niedrig ausgeprägte fachinhaltliche Eingangsvoraussetzungen haben (3 weitere Lernende weisen mittlere Kompetenzzuwächse auf und könnten ggf. ebenfalls, aber in einer anderen Weise berücksichtigt werden – nämlich zur vorläufigen Prüfung der im Kontrast formulierten Hypothesen). Derartige Fallvergleiche für die kleinen Stichproben bieten sich aus Sicht des Autors besonders dann an, wenn neben den Prozessvariablen zu den Vorstellungen auch weitere Informationen über die Bearbeitungsprozesse der jeweiligen Lernenden einbezogen werden (etwa aus Transkripten oder Videosichtungen oder aus anderen vorliegenden Kodierungen). Der vorliegende Abschnitt zur Forschungsfrage F-Vor7 könnte um ein solches Vorgehen ergänzt werden; dies ist allerdings aus Zeitgründen nicht mehr erfolgt.

6.10.1 Gruppierung der Lernenden für die einzelnen Testvariablen

Für jede der *Eingangsvoraussetzungen* (FCI, KFT, SWK, INT sowie FM-Prä; siehe S. 397) wurde eine Gruppierung vorgenommen, die auf der Instruktionsvarianten-Stichprobe basiert. Das heißt, die Gruppierungsgrenzen wurden anhand der Werte aller Personen festgelegt, die die explizit-fachmethodische Instruktionsvariante bearbeitet haben – unabhängig davon, ob diese für die ausgewählten Instruktionsextrakte videografiert oder analysiert wurden oder nicht. Als hoch ausgeprägte Eingangsvoraussetzung wird dabei angesehen, wenn der Testwert einer Person höher ist als die Summe aus dem Mittelwert des Testwertes für alle Personen, die die explizit-fachmethodische Instruktionsvariante bearbeitet haben, und dem 0.43-Fachen¹¹⁰ der zugehörigen Standardabweichung. Als niedrig ausgeprägte Eingangsvoraussetzung wird analog angesehen, wenn der Testwert der Person niedriger ist als der Mittelwert über alle Personen abzüglich des 0.43-Fachen der zugehörigen Standardabweichung. Für den Testwert zum eingesetzten Auszug aus dem Force Concept Inventory (FCI), der als Maß für die Eingangsvoraussetzungen der Lernenden hinsichtlich ihrer fachinhaltlichen Kompetenzen genutzt wird, ergibt sich beispielsweise ein Mittelwert von 507.07 über alle Personen, die die explizit-fachmethodische Instruktionsvariante bearbeitet haben.¹¹¹ Da die Standardabweichung des Testwertes über alle Lernenden, die die explizit-fachmethodische Instruktionsvariante bearbeitet haben, 92.85 beträgt, ergibt sich als Gruppierung: Lernende, deren Testwert unter $507.07 - 0.43 \cdot 92.85 = 467.14$ liegt, weisen niedrig ausgeprägte Eingangsvoraussetzungen zu fachinhaltlichen Kompetenzen auf; Lernende, deren Testwert über 546.99 liegt, weisen hoch ausgeprägte Eingangsvoraussetzungen auf. Wie viele Lernende jeweils hohe und niedrige Testwerte aufweisen, ist in Tabelle 6.36 zusammengefasst; dort sind auch Anzahl dafür angegeben, wie viele Personen sich innerhalb der 0.43-Standardabweichungen großen Umgebung um den Mittelwert befinden (mittlere Testwerte) und für wie viele Personen keine Testwerte vorliegen.

Eine Betrachtung der Anzahlen der Personen, die bezüglich der verschiedenen Eingangsvoraussetzungen deutlich unter beziehungsweise deutlich über dem Mittelwert liegen (mithilfe von Tabelle 6.36), zeigt zwar eine gewisse Schiefeile, da meist etwas mehr Personen in die Gruppe über dem Mittelwert gruppiert sind, aber kein schwerwie-

¹¹⁰Als Größe für die Umgebung um den Mittelwert wird das 0.43-Fache gewählt, weil sich bei einer idealen Normalverteilung so eine Drittelung der Stichprobe ergeben würde.

¹¹¹Der hier berichtete Wert weicht von dem bei Vorholzer (2016) berichteten Wert von 512.99 ab, weil unterschiedliche Bereinigungen der Ursprungsdatensätze vorgenommen wurden.

Erinnerung: In der vorliegenden Arbeit werden für alle auf Erhebungen von Vorholzer (2016) basierenden Variablen die dort berechneten Rasch-Measures (in logits) verwendet. Zur Normierung (Mittelwert 500, Standardabweichung 100) der entsprechenden Skalen und für weitere Informationen siehe Vorholzer (2016).

Tabelle 6.36: Übersicht über die Gruppierung der Personen nach ihren Testwerten für die verschiedenen Testvariablen. Aufgeführt ist jeweils die Anzahl der Personen, die der jeweiligen Gruppe angehören.

Ausprägung des Testwertes	Testvariablen						
	FM-Prä	Zuwachs (KnF)	Zuwachs (BuD)	FCI	KFT	SWK	INT
<i>Kennzeichen von naturwissenschaftlichen Fragen</i>							
niedrig	12	8	X	11	8	7	7
mittel	10	10	X	14	10	11	12
hoch	10	8	X	5	12	12	11
negativ	0	3	X	0	0	0	0
kein Testwert	1	4	X	3	3	3	3
<i>Unterscheidung von Beobachtung und Deutung</i>							
niedrig	11	X	7	10	6	6	6
mittel	7	X	7	10	10	9	11
hoch	8	X	7	5	9	10	8
negativ	0	X	4	0	0	0	0
kein Testwert	1	X	2	2	2	2	2

Anmerkungen. FM-Prä = Testwert für Kompetenzen zum experimentbezogenen Denken und Arbeiten vor dem Einsatz der Instruktion; Zuwachs (KnF/BuD) = FM-Post (nach dem Einsatz der Instruktion) minus FM-Prä (für die jeweilige Instruktionsvariante); FCI/KFT/SWK/INT = Testwerte für die fachinhaltlichen Kompetenzen / die allgemeinen kognitiven Fähigkeiten / das naturwissenschaftsbezogene Selbstwirksamkeitskonzept / die physikbezogenen Interessen jeweils vor dem Einsatz der Instruktion. Ausgegraut sind nicht vorhandene Gruppierungen (X) und nicht vorhandene Gruppen (0).

gendes Ungleichgewicht. Es sind immer auch einige Personen in der Gruppe deutlich unter dem Mittelwert vorhanden. Einzig problematisch könnten die Gruppierungen für die fachinhaltlichen Kompetenzen sein, bei denen – entgegen dem grundsätzlichen Trend, dass die Teilstichproben besser sind als die Teststichprobe – jeweils nur 5 Personen deutlich über dem Mittelwert eingruppiert sind. Zusammengenommen sprechen die Werte aus Tabelle 6.36 für eine hinreichende Datenbasis, um Personen anhand ihrer Testwerte zu Eingangsvoraussetzungen zu untersuchen und zu gruppieren.

Für die *Kompetenzzuwächse* der Lernenden (nachfolgend z. T. abgekürzt: Zuwächse) wird die Gruppierung relativ zur analysierten Stichprobe vorgenommen. Es werden jeweils das Drittel der Lernenden mit den höchsten und das Drittel der Lernenden mit den niedrigsten Kompetenzzuwächsen ausgewählt (genauer: es werden jeweils kaufmännisch gerundete 34 % der analysierten Personen ausgewählt). Dadurch ergeben sich symmetrische Gruppengrößen, was auch von Vorteil für die statistischen Vergleiche ist. Allerdings ist der Vergleich mit den Ergebnissen aus Kapitel 5 zu den Aktivitäten der Lernenden nur eingeschränkt möglich, weil dort analog zum vorigen Absatz basierend auf Mittelwerten und Standardabweichungen gruppiert wurde. Eine Änderung oder Ergänzung der detaillierten Auswertungen war aufgrund der parallel bearbeiteten Forschungszugänge allerdings nachträglich aus Zeitgründen nicht mehr möglich. Es kann angemerkt werden, dass sich allerdings im Fall der zu

den beiden Instruktionsextrakten bezüglich ihrer Vorstellungen analysierten Personen nur geringe Abweichungen in den Zuordnungen der Personen zu hohen bzw. niedrigen Kompetenzzuwächsen zeigen.¹¹²

Auch bei den Gruppierungen für die Zuwächse zeigen sich in beiden Teilstichproben keine problematischen methodischen Schwierigkeiten, wie aus Abbildung 6.41 hervorgeht. Für Kennzeichen von naturwissenschaftlichen Fragen hätte durch eine Wahl kleinerer Gruppengrößen durchaus für einen deutlicheren Kontrast der Gruppe mit niedrigem Zuwachs zur Gruppe mit hohem Zuwachs gesorgt werden können, allerdings scheint nach wie vor eine ausreichende Entfernung dieser beiden Gruppen vom Mittelwert sowie eine ausreichende Abgrenzung zur Gruppe mit mittlerem Zuwachs vorzuliegen, so dass die größere Gruppengröße bevorzugt wurde. Für die Unterscheidung von Beobachtung und Deutung hätte eine Gruppierung in zwei Gruppen mit dem Trennwert 100 ebenfalls eine hinreichende Trennung erzeugt, da nur eine Person nahe an diesem Wert liegt. Eine Dreiteilung wurde aber aus Konsistenzgründen vorgenommen. Bei dieser Dreiteilung für die Unterscheidung von Beobachtung und Deutung ist zu beachten, dass die mittlere Gruppe in zwei Lager geteilt zu sein scheint, so dass Analysen zur mittleren Gruppe mit Vorsicht zu interpretieren sind, sofern sie überhaupt vorgenommen werden.

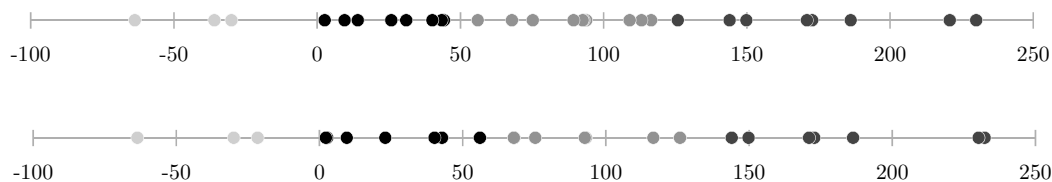


Abbildung 6.41: Verteilung der Personen auf der Skala des Kompetenzzuwachses von Prä- zu Posttest für die Teilstichprobe zu Kennzeichen von naturwissenschaftlichen Fragen (oben, 33 Personen) und die Teilstichprobe zur Unterscheidung von Beobachtung und Deutung (unten, 28 Personen). Von l. n. r. in unterschiedlichen Grautönen: Negativer, niedriger, mittlerer und zuletzt hoher Zuwachs.

6.10.2 Untersuchte Prozessmerkmale aus den vorigen Forschungsfragen

Eine Übersicht der ausgewählten, zu den Forschungsfragen F-Vor1 bis F-Vor6 gehörigen, Prozessvariablen findet sich in Tabelle 6.37 auf Seite 402. Darin ist jeweils durch eine Frage erläutert, was die entsprechende Prozessvariable abbildet.

¹¹²Unklar bleibt in der vorliegenden Arbeit allerdings, ob die wenigen anders zugeordneten Personen aufgrund der insgesamt nicht allzu großen Stichprobengrößen zu statistisch bedeutenden Unterschieden hinsichtlich der nachfolgend berichteten Ergebnisse führen würden, würden alle Analysen mit der abweichenden Gruppierung erneut vorgenommen.

Tabelle 6.37: Übersicht über die Prozessvariablen, die für die Analysen zu F-Vor7 betrachtet werden.

Prozessvariable	Beschreibung bzw. analyseleitende Frage
<i>F-Vor1: Vorliegende Vorstellungen</i>	
Anzahl der Instanzen mit OKs	Bei wie vielen Teilaufgaben sind für die Person Ideen-Oberkategorien zugewiesen?
Anzahl verschiedener OKs	Wie viele verschiedene Ideen-Oberkategorien sind der Person insgesamt zugewiesen?
<i>F-Vor2: Allgemeinheit der Äußerungen, aus denen die Vorstellungen rekonstruiert sind</i>	
Anzahl der Generalisierungen	Wie viele Äußerungen der Person sind mit der Kategorie AllgemeineRegel kodiert?
Anteil der Generalisierungen	Wie groß ist der Anteil der Anzahl der Generalisierungen an der Anzahl aller mit der Kategorie EinschätzungMitHinweis kodierten Äußerungen für die Person?
<i>F-Vor3: Angemessenheit der Vorstellungen</i>	
Angemessenheit verschiedener OKs	Wie groß ist der Wert des Maßes für die Angemessenheit bei Summation über die verschiedenen für die Person vorliegenden Ideen-Oberkategorien?
Angemessenheit gewichtet für OKs	Wie groß ist der Wert des Maßes für die Angemessenheit, wenn die Ideen-Oberkategorien bei der Summation gemäß der Anzahl der Teilaufgaben gewichtet werden, bei denen sie der Person zugewiesen sind?
Anzahl angemessener OKs	Wie viele angemessene Ideen-Oberkategorien sind der Person insgesamt zugewiesen?
Anzahl unangemessener OKs	Wie viele unangemessene Ideen-Oberkategorien sind der Person insgesamt zugewiesen?
<i>F-Vor5: Stabilität und Variabilität der Vorstellungen über Situationen hinweg</i>	
Anzahl stabiler OKs	Wie viele Ideen-Oberkategorien liegen für die Person stabil vor?
Anz. der Instanzen in stabilen OKs	Zähle für jede stabile Ideen-Oberkategorie für die Person: Bei wie vielen Teilaufgaben tritt diese auf?
Anzahl weicher Wechsel	Wie viele weiche Wechsel (jeweils aufeinanderfolgende Teilaufgaben mit sich mindestens in einer Ideen-Oberkategorie unterscheidender Kodierung) treten für die Person auf?
Anzahl harter Wechsel	Wie viele harte Wechsel (jeweils aufeinanderfolgende Teilaufgaben mit in allen zugewiesenen Ideen-Oberkategorien unterschiedlicher Kodierung) treten für die Person auf?
Anzahl der Ketten	Wie viele Ketten (ununterbrochene Teilaufgaben-Reihen mit einer durchgehenden Ideen-Oberkategorie) liegen für die Person vor?
Anzahl freier OKs	Wie viele Ideen-Oberkategorien, die nicht Teil irgendeiner Kette sind, liegen für die Person vor?
<i>F-Vor6: Vorstellungen vor und nach instruktionalen Informationen</i>	
Explizierung: Änd. Angemessenheit	Punkte der Person für Angemessenheit der Ideen-Oberkategorien vor der Explizierung abzüglich der Punkte der Person für Angemessenheit d. I-OKs nach der Explizierung.
Explizierung: Rel. Änd. Angemess.	Relative Punkte (sonst wie vorige Zeile).
Explizierung: Änd. Übereinstimmung	Punkte der Person für Übereinstimmung (der Ideen-Oberkategorien der Person mit den I-OKs der Explizierung) vor der Explizierung abzüglich der Punkte der Person für Übereinstimmung nach der Explizierung.
Explizierung: Rel. Änd. Übereinst.	Relative Punkte (sonst wie vorige Zeile).
<i>F-Vor4: Korrektheit der Entscheidungen</i>	
Korrektheitsmaß für Einschätzungen	Summe der Korrektheiten aller Einschätzungen (mit oder ohne Hinweis) der Person geteilt durch die Anzahl aller Einschätzungen der Person.

Anmerkungen. OKs = Ideen-Oberkategorien. F-Vor4 ist am Ende der Tabelle einsortiert, weil das Korrektheitsmaß keinen direkten Bezug zu den Vorstellungen der Lernenden hat, sondern deren Fähigkeiten abbildet; dennoch entspringt es den Analysen des vorliegenden Kapitels und wird daher nachfolgend betrachtet.

6.10.3 Übersicht über Gemeinsamkeiten und Unterschiede von Personen mit niedrig und hoch ausgeprägten Testwerten

Die Gruppierungen der Personen für jede der Testvariablen (s. Unterabschnitt 6.10.1) und die aus den anderen Forschungsfragen extrahierten Prozessvariablen (s. Unterabschnitt 6.10.2) werden genutzt, um statistische Analysen durchzuführen: Lernende mit niedriger Ausprägung eines jeweiligen Testwerts werden hinsichtlich einer der Prozessvariablen mit Lernenden mit hoher Ausprägung des Testwerts verglichen. Dazu können (ähnlich wie in Unterabschnitt 5.3.2 für die Aktivitätenkodierungen diskutiert) für gewöhnlich parametrische Verfahren genutzt werden, sofern das zu vergleichende Merkmal mindestens intervallskaliert ist. Das zu vergleichende Merkmal ist im vorliegenden Fall die Prozessvariable (Auflistung in Tabelle 6.37); für diese Variablen ist allerdings nicht immer klar, ob die zugehörigen Werte intervallskaliert zu interpretieren sind oder eher ordinal aufgefasst werden sollten (es ist beispielsweise nicht klar, ob der Abstand zwischen einer und elf verschiedenen Ideen-Oberkategorien pro Person als zehnmal so groß anzusehen ist wie der Abstand zwischen einer und zwei verschiedenen Ideen-Oberkategorien pro Person). Nachfolgend werden daher sowohl t-Tests (parametrisch) als auch Mann-Whitney-U-Tests (non-parametrisch) berichtet, wobei der Fokus auf den parametrischen Tests liegt, weil diese grundsätzlich als robust anzusehen sind.

In Tabelle 6.38 sind Ergebnisse der statistischen Vergleiche für alle Prozessvariablen bei Gruppierung der Lernenden hinsichtlich aller Testvariablen berichtet. Dabei sind grundsätzlich die Ergebnisse des t-Tests nach Student berichtet; abgewichen wird davon, wenn die Normalverteilungsannahme für mindestens eine der beiden verglichenen Personengruppen (Testwert niedrig vs. hoch ausgeprägt) nicht erfüllt ist oder sich ein signifikantes Ergebnis für den Mann-Whitney-U-Test findet, während sich ein nicht-signifikantes Ergebnis für den t-Test findet. Anhand von Farbgebung bei Schrift und Hinterlegung wird dies differenziert – die genauen Kombinationen sind in der Legende in Tabelle 6.38 dokumentiert.

Die meisten der berichteten statistischen Zusammenhänge sind positiv. Siehe Tabelle 6.38. Das heißt, dass Personen mit höherem Zuwachs oder höherem Testwert auch bezüglich der Prozessvariable höhere Werte aufweisen, wenn statistisch gesehen Zusammenhänge vorliegen. Negative Zusammenhänge liegen nur für die Anzahlen der weichen und harten Wechsel vor, wo sie auch zu erwarten sind, wie weiter unten diskutiert. Nach der Tabelle wird allerdings zunächst dem Fokus der Arbeit entsprechend eine Gesamtschau zu allen Zusammenhängen von Prozessvariablen mit den Kompetenzzuwächsen der Lernenden vorgenommen.

Tabelle 6.38: Zusammenhänge von Test- und Prozessvariablen jeweils für Vergleiche der Gruppen mit niedrigen vs. hohen Testwerten.

Test- variable	Anzahl der Instanzen mit OKs				Anzahl verschiedener OKs				Angemessenheit gewichtet für OKs				Anzahl angemessener OKs				Anzahl unangemessener OKs			
	stabiler OKs		in stabilen OKs		weicher Wechsel		harter Wechsel		der Ketten		freier OKs									
	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD						
Zuwachs	p*** g	x	p** m	x	x	x	p*** g	x	p* m	x	x	x	p* m	x	x	x	x			
Prätest	p*** m	p* m	x	p* m	x	x	p** m	p** m	x	p* m	x	x	p* m	x	x	x	x			
FCI	p*** g	x	x	x	x	x	p*** m	p*** m	x	x	x	x	p* m	x	x	x	x			
KFT	x	p* m	x	p* m	p* k	x	x	x	x	x	x	p* m	x	x	x	x	x			
SWK	p* m	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x			
INT	x	x	x	x	x	x	p*** m	x	x	x	x	x	x	x	x	x	x			

Test- variable	Anzahl stabiler OKs				Anzahl in stabilen OKs				Anzahl weicher Wechsel				Anzahl harter Wechsel				Anzahl der Ketten				Anzahl freier OKs			
	stabiler OKs		in stabilen OKs		weicher Wechsel		harter Wechsel		der Ketten		freier OKs													
	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD								
Zuwachs	p* m	x	p** m	x	x	x	n*** g	x	p* m	x	p* m	x	p* m	x	p* m	x								
Prätest	x	p* m	x	p* m	x	x	x	x	p*** m	x	x	x	p*** m	x	x	x								
FCI	p* m	x	x	x	n* m	x	x	x	x	x	x	x	x	x	x	x								
KFT	x	x	x	x	x	x	n* m	x	x	x	x	x	x	x	x	x								
SWK	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x								
INT	x	x	x	x	n* m	x	x	x	x	x	x	x	x	x	x	x								

Test- variable	Anzahl der Generalisierungen				Anteil der Generalisierungen				Explizierung: Änd. Angemessenheit				Explizierung: Rel. Änd. Übereinst.				Korrektkeitsmaß für Einschätzungen			
	Generalisierungen		Anteil der Generalisierungen		Angemessenheit		Änd. Angemess.		Explizierung: Rel.		Rel. Änd. Übereinst.		Einschätzungen							
	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD						
Zuwachs	p* m	x	p** m	x	x	x	x	x	x	x	x	p*** g	x	p*** g						
Prätest	x	x	x	x	x	x	x	x	x	x	x	p*** g	x	p*** g						
FCI	p*** g	x	p*** g	x	x	x	x	x	x	x	x	p*** g	x	p*** g						
KFT	x	x	x	x	x	x	x	x	x	x	x	p*** g	x	p*** g						
SWK	p** m	x	p** m	x	x	x	x	x	x	x	x	p*** m	x	p*** m						
INT	x	x	x	x	p** m	x	p** m	x	x	x	x	p* m	x	p* m						

Tabelle 6.38: Zusammenhänge von Test- und Prozessvariablen (Fortsetzung: Legende).

Legende				
KnF / BuD = Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen / zur Unterscheidung von Beobachtung und Deutung. Durchgeführt wurden t-Tests und Mann-Whitney-U-Tests für Gruppen von Personen mit niedrigem vs. hohem Testwert. Testergebnisse sind in Abhängigkeit von den Ergebnissen beider Tests und der Gegebenheit der Normalverteilungsannahme der jeweiligen Stichproben berichtet:				
	t-Test	U-Test	Normalverteilungs- annahme erfüllt	Berichtet
p n x	signifikant	signifikant	egal	Ergebnisse des t-Tests in weiß auf dunkelgrauem Hintergrund
	signifikant	nicht-signifikant	ja	Ergebnisse des t-Tests in schwarz auf dunkelgrauem Hintergrund
	signifikant	nicht-signifikant	nein	Ergebnisse des t-Tests in grau auf hellgrauem Hintergrund
	nicht-signifikant	signifikant	ja	Ergebnisse des U-Tests in hellgrau auf dunkelgrauem Hintergrund
	nicht-signifikant	signifikant	nein	Ergebnisse des U-Tests in schwarz auf hellgrauem Hintergrund
	nicht-signifikant	nicht-signifikant	egal	Ein x
	(vor)signifikant positiver Zusammenhang			k kleine Effektstärke ($0.1 < r < 0.3$)
	(vor)signifikant negativer Zusammenhang			m mittlere Effektstärke ($0.3 < r < 0.5$)
	kein (signifikanter) Zusammenhang			g große Effektstärke ($0.5 < r$)
Signifikanzniveaus:		* $p < .2$	** $p < .1$	*** $p < .05$

6.10.4 Auf die instruktionale Information bezogene Prozessvariablen

Tabelle 6.39: Auszug zu instruktionalen Informationen (=Explizierungen) aus Tabelle 6.38: *Zusammenhänge von Test- und Prozessvariablen jeweils für Vergleiche der Gruppen mit niedrigen vs. hohen Testwerten* (Legende siehe dort).

Test- variable	Explizierung: Änd.		Explizierung: Rel.		Explizierung: Änd.		Explizierung: Rel.	
	Angemessenheit		Änd. Angemess.		Übereinstimmung		Änd. Übereinst.	
	KnF	BuD	KnF	BuD	KnF	BuD	KnF	BuD
Zuwachs	x	x	x	x	x	x	x	x
Prätest	x	x	x	x	x	x	x	x
FCI	x	x	x	x	x	x	x	x
KFT	x	x	x	x	x	x	x	x
SWK	x	x	x	x	x	x	x	x
INT	p** m	x	p** m	x	x	x	x	x

Keine Zusammenhänge von Personenmerkmalen und auf die instruktionale Information bezogenen Prozessvariablen. Ein aus Tabelle 6.38 (Seiten 404 und 405; relevanter Auszug in Tabelle 6.39) direkt ersichtliches Ergebnis besteht darin, dass sich für die im Umfeld der instruktionalen Information (=Explizierung) gemessenen Änderungen von Prozessvariablen keine statistisch bedeutsamen Unterschiede für Personen mit unterschiedlich stark ausgeprägten Personenmerkmalen (abgebildet durch die Testvariablen) finden.¹¹³ Dies lässt sich aufgrund der Deutlichkeit über alle untersuchten potentiellen Zusammenhangspaare hinweg so deuten, dass in der Tat keine Zusammenhänge vorliegen. Der Einwand, dass die Teststärke in der vorliegenden explorativen Arbeit (u. a. aufgrund kleiner Stichproben) nicht gut genug ist, um das Nicht-Vorliegen von Zusammenhängen hinreichend abzusichern, scheint durch die Eindeutigkeit des Ergebnisses über die vielen Tabellenzellen hinweg befriedigend beantwortet. Es könnte allerdings sein, dass die konkret gewählte Operationalisierung der auf die instruktionale Information bezogenen Prozessmerkmale nicht optimal gelungen ist; allerdings ist auch dies durch den Rückgriff auf sowohl absolute als auch relative Änderungsmaße ansatzweise angegangen.

Das Ergebnis scheint auf den ersten Blick irritierend zu sein, weil erwartbar scheint, dass ein Zuwachs an Angemessenheit oder an Übereinstimmung mit explizierten Konzepten (wie bezüglich F-Vor6 untersucht) sich auch im Zuwachs an Kompetenzen zum experimentbezogenen Denken und Arbeiten widerspiegelt (wie durch die

¹¹³Es finden sich bei je 24 pro Instruktionsextrakt untersuchten Zusammenhängen nur zwei mittlere Effekte. Diese finden sich beim Interesse, für das die Validität des Tests für die von Vorholzer (2016) insgesamt untersuchte Stichprobe (von der die hier untersuchte Stichprobe eine Teilmenge ist) nicht optimal ist (siehe Tabelle A.1 im Anhang sowie Vorholzer, 2016, S. 88), weshalb auf eine detaillierte Diskussion dieser beiden Effekte verzichtet wird.

Zuwachs-Variable abgebildet). Wird allerdings berücksichtigt, dass der Zuwachs an Kompetenzen zum experimentbezogenen Denken und Arbeiten sich auf das gesamte bearbeitete Lernmaterial bezieht, während die Prozessvariablen sich auf kurze Instruktionsextrakte aus dem Material zu einem spezifischen Inhaltsbereich beziehen, ist zumindest ansatzweise plausibel, dass keine signifikanten Zusammenhänge gefunden werden. Es könnte sein, dass die kurzen Instruktionsextrakte für die realen Kompetenzzuwächse unbedeutend sind. Vor dem Hintergrund, dass ansonsten (zumindest zum Instruktionsextrakt für Kennzeichen von naturwissenschaftlichen Fragen) die gefundenen und nicht-gefundenen Zusammenhänge gut zu den Ergebnissen zur gesamten Instruktion passen (siehe auch nachfolgende Unterabschnitte), scheint es allerdings eher unplausibel, dass die untersuchten Instruktionsextrakte in keinem Zusammenhang mit den Kompetenzzuwächsen über die gesamte Instruktion hinweg stehen. Etwas nuancierter könnten mögliche Erklärungen folgendermaßen lauten: Auch wenn Kennzeichen von naturwissenschaftlichen Fragen (und analog auch die Unterscheidung von Beobachtung und Deutung) als ein wichtiger Inhaltsbereich zu Kompetenzen zum experimentbezogenen Denken und Arbeiten angesehen werden kann, könnte dessen Stellung – zumindest im eingesetzten Testinstrument von Vorholzer (2016) – nicht zentral genug dafür sein, dass Veränderungseffekte, die sich im Rahmen einer kurzen Sequenz zeigen, sich auch noch für das gesamte Lernmaterial zeigen.¹¹⁴ Es könnte abweichend davon (oder ergänzend dazu) auch sein, dass beobachtbare Änderungen sich nicht während der Bearbeitung des jeweiligen Instruktionsextrakts analytisch feststellen lassen, sondern erst über dem Instruktionsextrakt nachgeschaltete Lerngelegenheiten gefestigt und sichtbar werden (wofür u. a. spräche, dass das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung, für das sich insgesamt nur geringe Zusammenhänge zu Kompetenzzuwächsen finden, am Ende der letzten Einheit der explizit-fachmethodischen Instruktionsvariante verortet ist).

Der nicht-berichtbare Zusammenhang von Kompetenzzuwächsen mit den Prozessvariablen der Zunahme von Angemessenheit oder Übereinstimmung der Vorstellungen könnte zusätzlich zu den Ausführungen des vorigen Absatzes auch als Hinweis darauf gedeutet werden, dass die (absolute oder relative) Änderung der Angemessenheit der Vorstellungen oder ihrer Übereinstimmung mit der Explizierung kein angemessenes

¹¹⁴Weiter unten, auf S. 408, wird erwähnt, dass plausibilisiert werden kann, dass *über* die gesamten (kurzen) Sequenzen zu den Instruktionsextrakten *hinweg erzeugte* Prozessvariablen in verschiedenartigen, theoretisch nachvollziehbaren und erwartbaren Zusammenhängen mit den Personenmerkmalen stehen, obwohl Veränderungseffekte *innerhalb* der kurzen Sequenzen zu den Instruktionsextrakten dies *nicht* tun. Eine mögliche Erklärung lautet, dass die über die zu den Instruktionsextrakten gehörigen Sequenzen generierten Prozessvariablen abbilden, wie die Personen während großer Teile der gesamten Instruktion arbeiten.

Maß für den Zuwachs von Kompetenzen zum experimentbezogenen Denken und Arbeiten ist. Dies ist unter anderem deshalb plausibel, weil nur wenige Situationen betrachtet werden (und die Kompetenzmessung sich auf Stabilität bezieht, indem mehrere Situationen berücksichtigt werden). Anzumerken ist, dass allerdings nicht grundsätzlich davon ausgegangen werden muss, dass die Erhebung der Prozessvariablen an kurzen Sequenzen dazu führt, dass keine Beziehung zum gesamten Lernprozess hergestellt werden kann. Vielmehr sind die Prozessvariablen zur Explizierung in einer ganz spezifischen Weise auf den kurzen Zeitraum bezogen; die anderen untersuchten Prozessvariablen lassen sich eher als zeitlich überdauernde lernendenspezifische Charakteristika der Auseinandersetzung mit Konzepten zum experimentbezogenen Denken und Arbeiten interpretieren (siehe bereits Fußnote 114) und stehen eventuell aus diesem Grund stärker mit den Kompetenzen zum experimentbezogenes Denken und Arbeiten in Zusammenhang (dokumentiert in Tabelle 6.38). Dass dieser Zusammenhang mit den Kompetenzen zum experimentbezogenen Denken und Arbeiten einmal vorrangig auf den Lernzuwachs (nämlich für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen) und einmal vorrangig auf die entsprechenden Eingangsvoraussetzungen (nämlich für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung) beschränkt ist, geht aus den Diskussionen im nächsten Unterabschnitt hervor.

6.10.5 Untersuchte Prozessmerkmale und Kompetenzen zum experimentbezogenen Denken und Arbeiten

Zentral für die vorliegende Arbeit sind die Kompetenzen der Lernenden zum experimentbezogenen Denken und Arbeiten. Diese werden durch die Testvariablen zu Prä- und zu Posttest und zu Kompetenzzuwachs in den Blick genommen. Der Einfachheit halber werden nur die Eingangsvoraussetzungen bezüglich dieser Kompetenzen (Prätest-Variable) und die Veränderung dieser Kompetenzen (Zuwachs-Variable) betrachtet. Die absolut nach der Instruktion erreichten Kompetenzen (Posttest-Variable) werden nicht betrachtet.

Werden die zu den Kompetenzen zum experimentbezogenen Denken und Arbeiten gehörigen Zeilen aus Tabelle 6.38 (die ersten beiden Zeilen des Tabellenkörpers) betrachtet, ergeben sich zwei eindruckliche Ergebnisse, zu denen anschließend vier zusammenführende Deutungen vorgestellt werden. Abschließend wird kurz darauf eingegangen, welche Abweichungen von den Ergebnissen innerhalb der Daten vorliegen und wie diese gedeutet werden können.

Werden die Prozessmerkmale einzeln betrachtet, findet sich zum Instruktionsextrakt für Kennzeichen von naturwissenschaftlichen Fragen oftmals ein Zusammenhang des jeweiligen Prozessmerkmals mit dem Kompetenzzuwachs der Lernenden und zum Instruktionsextrakt für die Unterscheidung von Beobachtung und Deutung oftmals ein Zusammenhang mit den fachmethodischen Eingangsvoraussetzungen. Andere auf fachmethodische Kompetenzen bezogene Zusammenhänge finden sich in sehr wenigen Fällen. In Tabelle 6.40 ist schematisch dargestellt, wie die Zusammenhänge sich für viele der Prozessmerkmale darstellen, wenn jedes für sich genommen wird.

Tabelle 6.40: Schematische Darstellung des Zusammenhangs der einzeln betrachteten Prozessmerkmale mit Kompetenzzuwachs und fachmethodischer Eingangsvoraussetzung. Realisiert als fiktiver Auszug aus Tabelle 6.38: *Zusammenhänge von Test- und Prozessvariablen jeweils für Vergleiche der Gruppen mit niedrigen vs. hohen Testwerten* (Legende siehe dort).

Test- variable	Prozessvariable			
	KnF		BuD	
	p**	g	p*	m
Zuwachs			x	
Prätest	x			
FCI	x		x	
KFT	x		x	
SWK	x		x	
INT	x		x	

Für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen findet sich für gewöhnlich ein statistisch bedeutsamer Unterschied zwischen Lernenden mit unterschiedlich hohen Kompetenzzuwächsen hinsichtlich der Prozessvariable, aber kein statistisch bedeutsamer Unterschied zwischen Lernenden mit unterschiedlichen fachmethodischen Eingangsvoraussetzungen. Für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung findet sich hingegen für gewöhnlich ein statistisch bedeutsamer Unterschied zwischen Lernenden mit unterschiedlichen fachmethodischen Eingangsvoraussetzungen hinsichtlich der Prozessvariable, aber kein statistisch bedeutsamer Unterschied zwischen Lernenden mit unterschiedlich hohen Kompetenzzuwächsen.

In manchen Fällen taucht einer der beiden hinterlegten Zusammenhänge nicht auf (d. h. es findet sich keine grau hinterlegte Zelle in Tabelle 6.38 auf den Seiten 404 und 405). Dies ist beispielsweise für *Anzahl der Generalisierungen* der Fall, weil sich kein statistisch nachweisbarer Unterschied zwischen Lernenden mit unterschiedlichen fachmethodischen Eingangsvoraussetzungen findet, wenn diese hinsichtlich

der Anzahl der allgemeinen, wenig-fallbezogenen Äußerungen verglichen werden, aus denen sich Vorstellungen rekonstruieren lassen. In manchen anderen Fällen taucht keiner der beiden in Tabelle 6.40 grau hinterlegten statistischen Zusammenhänge auf. Dies ist beispielsweise für *Angemessenheit verschiedener OKs* der Fall; dass in diesem Fall kein statistischer Zusammenhang vorliegt, bedeutet, dass sich für die untersuchte Stichprobe weder für Lernende mit unterschiedlichen fachmethodischen Eingangsvoraussetzungen noch für Lernende mit unterschiedlichen Kompetenzzuwächsen systematische Unterschiede hinsichtlich desjenigen Angemessenheitsmaßes finden lassen, welches nur darauf basiert, *ob und wie viele verschiedene* deutlich angemessene / eher angemessene / eher unangemessene / deutlich angemessene Ideen-Oberkategorien mindestens einer der Äußerungen der Person zugewiesen wurden (aber *nicht, wie vielen* Äußerungen diese zugewiesen wurden).¹¹⁵

Nur in zwei Fällen finden sich statistisch bedeutsame Zusammenhänge an entsprechenden Stellen von Tabelle 6.38 (Seiten 404 und 405), an denen in Tabelle 6.40 keine Zusammenhänge eingetragen sind. In beiden Fällen betrifft dies das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen, für das sich statistisch bedeutsame Unterschiede zwischen Personen mit unterschiedlichen fachmethodischen Eingangsvoraussetzungen finden. Die beiden betroffenen Prozessvariablen sind *Anzahl der Instanzen mit OKs* und *Korrektheitsmaß für Einschätzungen*. Diese werden ab Seite 416 einzeln diskutiert.

Werden die Prozessmerkmale in Paketen betrachtet, die sich auf die jeweiligen Forschungsfragen beziehen, finden für nahezu all diese Pakete (vgl. Tabelle 6.37 auf S. 402) – bei qualitativer Zusammenfassung der jeweils zugehörigen Prozessmerkmale – für Kennzeichen von naturwissenschaftlichen Fragen vorrangig Zusammenhänge mit den Kompetenzzuwächsen und für die Unterscheidung von Beobachtung und Deutung vorrangig Zusammenhänge mit den fachmethodischen Eingangsvoraussetzungen. Dieses Ergebnis schließt direkt an das vorige Ergebnis an; es ist auf einer weniger detaillierten Ebene, weil mehrere Prozessmerkmale qualitativ gebündelt werden. Gleichzeitig kann durch das Herauszoomen und zusammenfassende Betrachten eine mit den spezifischen Operationalisierungen der einzelnen Prozessvariablen verbundene Validitätsproblematik zumindest ansatzweise angegangen werden.

Beispielsweise finden sich (in Tabelle 6.38 auf den Seiten 404 und 405) für zwei der Prozessvariablen zur Angemessenheit keine statistisch bedeutsamen Zusammenhänge;

¹¹⁵Für das nach Häufigkeit der Ideen-Oberkategorien gewichtete Angemessenheitsmaß *Angemessenheit gewichtet für OKs* findet sich im Übrigen auch das typische Muster an Zusammenhängen entsprechend der Tabelle 6.40.

für die anderen beiden zur Angemessenheit gehörigen Prozessvariablen finden sich allerdings genau die in Tabelle 6.40 (auf S. 409) schematisch dargestellten Zusammenhänge. Weil statistische Nicht-Signifikanz weniger sicher ist als statistische Signifikanz (u. a. aufgrund der Teststärke), deutet dies darauf hin, dass die in Tabelle 6.40 dargestellten Zusammenhänge auch dann gelten, wenn die Bedeutung der Angemessenheit der Vorstellungen ganz allgemein in den Blick genommen wird.

Für das Paket der sechs Prozessvariablen, die zu der Frage nach der zeitlichen Stabilität der Vorstellungen (F-Vor5) gehören, treten die Zusammenhänge für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen über all die Prozessvariablen hinweg zusammengefasst wesentlich deutlicher hervor als die Zusammenhänge für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung.¹¹⁶ Da die Anzahlen weicher und harter Wechsel im Sinne nicht-vorhandener Stabilität zu interpretieren sind, sind die berichteten Zusammenhänge für einen Abgleich mit den anderen Prozessvariablen zur Stabilität zu invertieren. Dann passen auch sie in das schematisch in Tabelle 6.40 (auf S. 409) zusammengefasste Bild: Für Kennzeichen von naturwissenschaftlichen Fragen finden sich Unterschiede zwischen Lernenden mit unterschiedlichen Kompetenzzuwächsen hinsichtlich der Stabilität ihrer Vorstellungen; für Unterscheidung von Beobachtung und Deutung finden sich Unterschiede zwischen Lernenden mit unterschiedlichen fachmethodischen Eingangsvoraussetzungen hinsichtlich der Stabilität ihrer Vorstellungen; andere statistische Zusammenhänge (mit Bezug zu fachmethodischen Kompetenzen) finden sich nicht.

Für das Paket zur Allgemeinheit der Vorstellungen findet sich ein zu Tabelle 6.40 (auf S. 409) passendes Muster an Zusammenhängen. Aus den Ausführungen auf den folgenden Seiten ergibt sich, dass der sich – als einziges abweichend vom typischen Muster – nicht einstellende statistische Zusammenhang mit dem Prätestwert für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung sowohl durch methodische Probleme bei dem Instruktionsextrakt als auch inhaltlich plausibilisiert werden kann.

Nachfolgend werden die beiden formulierten Ergebnisse ausführlich diskutiert. Es werden zunächst *zwei inhaltliche* Erklärungen bzw. Plausibilisierungen dafür gegeben, dass sich die Ergebnisse wie beschrieben einstellen; anschließend wird alternativ sowie ergänzend dazu anhand von *zwei methodischen* Ausführungen darauf eingegangen, was dazu beigetragen haben könnte, dass die berichteten Ergebnisse entstanden sind.

¹¹⁶Dies passt dazu, dass die Beantwortung der Forschungsfrage zur Stabilität in Bezug auf das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung mit methodischen Schwierigkeiten behaftet war; es wird weiter unten in dem vorliegenden Unterabschnitt bei der Diskussion der Ergebnisse aufgegriffen.

Das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen hat eine größere Bedeutung für die erreichten Kompetenzzuwächse, weil es zu einem früheren Zeitpunkt der Instruktion eingesetzt wird. Das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen befindet sich in etwa zu Beginn des letzten Drittels der ersten Einheit der explizit-fachmethodischen Instruktionsvariante. Das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung befindet sich in etwa zu Beginn des letzten Drittels der dritten und letzten Einheit der explizit-fachmethodischen Instruktionsvariante. Damit wird mehr als die Hälfte des Instruktionsmaterials (und deutlich mehr als die Hälfte der Zeit der Bearbeitung, siehe Unterabschnitt 5.4.1 beginnend auf S. 106) zwischen diesen beiden Instruktionsextrakten bearbeitet. Es könnte sein, dass es einen Einfluss auf das spätere Lernen hat, wie die Lernenden das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen bearbeiten; für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung ist dies weniger plausibel.

Beispielsweise ist es denkbar, dass ein aufgebautes Verständnis zu den im Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen adressierten Inhalten das Erlernen weiterer Inhalte unterstützt oder für dieses sogar notwendig ist. Wer nicht verstanden hat, was eine naturwissenschaftliche Frage von einer nicht-naturwissenschaftlichen Frage unterscheidet, hat es später eventuell auch schwerer, für unterschiedliche naturwissenschaftliche Fragen einzuschätzen, ob diese sich mit einem gegebenen Versuchsaufbau beantworten lassen, weil bereits die Unterscheidung auf der allgemeineren Ebene schwer gefallen ist.¹¹⁷ Es wäre dann allerdings verwunderlich, dass Lernende, für die innerhalb der zu Kennzeichen von naturwissenschaftlichen Fragen untersuchten Sequenz deutliche Zunahmen der Angemessenheit der Vorstellungen gefunden werden, nicht auch zu insgesamt höheren Kompetenzzuwächsen gelangen.¹¹⁸

Abweichend von der Erklärung im vorigen Absatz wäre es auch denkbar, dass Lernende, für die bei den zu Kennzeichen von naturwissenschaftlichen Fragen untersuchten Sequenzen bereits stärker angemessene oder eher zeitlich stabile Vorstellungen

¹¹⁷Das im Text angeführte Argument ist von geringerer Überzeugungskraft als eines, in welchem beispielsweise von einer zu Beginn fehlenden Unterscheidungsfähigkeit zwischen präzise und allgemein formulierten Fragen (d. h. Fragen, in denen die zu untersuchenden Merkmale bzw. Variablen benannte beziehungsweise nicht benannt sind; siehe Tabelle 4.2 auf S. 54) auf eine später schwerer fallendes Unterscheiden von zu einem Versuchsaufbau passenden und unpassenden Fragestellungen geschlossen wird. Dies zeigt, dass der angeführte Erklärungsansatz mit Vorsicht betrachtet werden sollte, wie es auch im nachfolgenden Text geschieht.

¹¹⁸Es wäre denkbar, dass nur Lernende, die bereits für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen alles können, diejenigen sind, die in späteren Teilen der Instruktion viel dazulernen.

rekonstruiert wurden, auch im weiteren Verlauf der Instruktionsbearbeitung immer wieder angemessenere bzw. stabilere Vorstellungen vorweisen (was im Rahmen der vorliegenden Arbeit nicht untersucht ist). Diese alternative Erklärung kann unter anderem so gedeutet werden, dass Lernende, die mit dem Instruktionsmaterial bereits im Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen in intendierter Weise umgehen (können), dies auch im weiteren Verlauf der Instruktionsbearbeitung tun (können) und daher im Mittel höhere Kompetenzzuwächse vorweisen als Lernende, denen der Umgang mit der Instruktion (z. B. dem Format, den affektiven Aspekten oder der gewählten Sprache oder auch dem fachmethodischen Fokus) grundsätzlich oder zumindest zu Beginn der Instruktionsbearbeitung schwerer fällt. Es kann allerdings auch sein, dass die Analysen zum Instruktionsextrakt für Kennzeichen von naturwissenschaftlichen Fragen grundlegende Hinweise darauf geben, wie die Auseinandersetzungen der unterschiedlichen Lernenden mit dem Material vonstatten gehen. Personen, die das Material im Allgemeinen oder kontinuierlich in einer Art und Weise bearbeiten, bei der die untersuchten Prozessmerkmale hoch ausgeprägt sind, würden demnach eher hohe Kompetenzzuwächse erreichen als andere Personen.

Die für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen aufgeführten und diskutierten Einflüsse sind aufgrund des späten Auftauchens des Instruktionsextrakts zur Unterscheidung von Beobachtung und Deutung für letzteres weniger plausibel. Dies könnte erklären, dass sich nur für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen Zusammenhänge von Prozessmerkmalen und erreichten Kompetenzzuwächsen finden. Offen bleibt allerdings, warum die fachmethodischen Eingangsvoraussetzungen für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung in systematischer Weise von Relevanz sind. Die nachfolgend gegebene zweite Erklärung setzt dort an und ist eventuell mit der ersten Erklärung zu kombinieren.

Die Lernenden haben unterschiedlich gute Eingangsvoraussetzungen (Vorwissen etc.) bezüglich der Kennzeichen von naturwissenschaftlichen Fragen und der Unterscheidung von Beobachtung und Deutung. Das Ergebnis, dass sich für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen vorrangig Zusammenhänge von Prozessmerkmalen mit den erreichten fachmethodischen Kompetenzzuwächsen und für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung vorrangig Zusammenhänge von Prozessmerkmalen mit den fachmethodischen Eingangsvoraussetzungen finden, könnte auch dadurch plausibilisiert oder erklärt werden, dass die untersuchten Lernenden bessere Eingangsvoraussetzungen bezüglich der Unterscheidung von Beobachtung und Deutung

mitbringen als bezüglich der Kennzeichen von naturwissenschaftlichen Fragen. Dies ist selbst dann möglich, wenn die Lernenden vergleichbare Testwerte im Prätest aufweisen, weil die Prätestwerte über mehrere Aspekte von Kompetenzen zum experimentbezogenen Denken und Arbeiten gemittelt sind (siehe Abschnitt 4.5).¹¹⁹ Basierend auf den in Abschnitt 2.4, ab S. 25 berichteten Befunden kann davon ausgegangen werden, dass die Lernenden zu Kennzeichen von naturwissenschaftlichen Fragen eher gering ausgeprägte Vorerfahrungen und geringes Vorwissen haben, während sie für die Unterscheidung von Beobachtung und Deutung vermutlich vergleichsweise gute Eingangsvoraussetzungen mitbringen. Während sich für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen aufgrund der Neuheit der Inhalte vor allem Zusammenhänge mit dem Kompetenzzuwachs finden lassen, finden sich dementsprechend für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung keine solchen Zusammenhänge, sondern – wenn überhaupt – Zusammenhänge mit dem personenspezifischen Prätestwert.

Der Umfang der untersuchten Instruktionsextrakte (Anzahl der Aufgaben, zeitlicher Umfang, ...) ist unterschiedlich groß. Zu dem Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen gehören 16 Teilaufgaben, die auf drei zeitlich voneinander getrennte Segmente verteilt sind; die Bearbeitungsdauer aller zugehörigen Aufgaben beträgt im Mittel ca. 12 Minuten. Zu dem Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung gehören 18 Teilaufgaben, die ein zusammenhängendes Segment bilden; die Bearbeitungsdauer aller zugehörigen Aufgaben beträgt im Mittel ca. 4.5 Minuten (Details in Tabelle 6.9 auf S. 281).

Es ist denkbar, dass sich im Verlauf der Bearbeitung des Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen die Ausprägungen der Prozessvariablen im Mittel verändern (z. B. aufgrund einer Zunahme angemessener Vorstellungen). Die Prozessvariablen stellen allerdings über die gesamte Bearbeitungsdauer zusammengefasste Betrachtungen dar, in denen derartige Veränderungen nicht sichtbar werden; vielmehr könnte es sogar sein, dass eine Prozessvariable für zwei Personen mit sehr unterschiedlichen Eingangsvoraussetzungen denselben Wert annimmt, während die eine Person konstant denselben Prozessvariablen-Wert aufweist und für die andere Person der Prozessvariablen-Wert stark zunimmt (denkbar bspw. für die Angemessenheit). Ein Zusammenhang der Prozessvariable mit dem Prä-

¹¹⁹Ein Auflösen von Teilkompetenzen oder gar Testitems, die sich den Instruktionsextrakten zu Kennzeichen von naturwissenschaftlichen Fragen und zur Unterscheidung von Beobachtung und Deutung zuordnen lassen, wäre möglich, wurde aber im Rahmen der vorliegenden Arbeit nicht vorgenommen und bleibt zunächst ein Desiderat.

testwert (also ein Zusammenhang von bspw. Angemessenheit der rekonstruierten Vorstellungen mit den fachmethodischen Eingangsvoraussetzungen) würde sich dann nicht ergeben, weil der Prozessvariablen-Wert für beide Personen derselbe wäre. Für die deutlich kürzere Sequenz zur Unterscheidung von Beobachtung und Deutung fänden sich weniger derartige Mittellungseffekte, sodass Zusammenhänge mit den fachmethodischen Eingangsvoraussetzungen statistisch auffindbar wären. Es wäre dann allerdings verwunderlich, warum für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen gewisse Ausprägungen von Prozessvariablen in einem Zusammenhang mit den erreichten Kompetenzzuwächsen vorliegen.

Es ist auch denkbar, dass durch die unterschiedlich großen zeitlichen Umfänge die Güte der Operationalisierung der Prozessvariablen von Instruktionsextrakt zu Instruktionsextrakt unterschiedlich ist. Beispielsweise kann die zeitliche Stabilität der Vorstellungen der Lernenden über das zeitlich stärker ausgedehnte Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen vermutlich besser untersucht werden; entsprechend finden sich für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung im Paket der Prozessvariablen, die auf die zeitliche Stabilität der Vorstellungen bezogen sind, auch nur sehr vereinzelt Zusammenhänge.

Das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung ist weniger gut zur (durchgeführten) Untersuchung von Vorstellungen geeignet als das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen. Dass sich für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen viele statistische Zusammenhänge ergeben, die zudem nachvollziehbar und einleuchtend sind, während dies für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung nicht der Fall ist, könnte auch so gedeutet werden, dass sich die Instruktionsextrakte unterschiedlich gut dafür eignen, die Vorstellungen der Lernenden zu untersuchen. Es könnte sein, dass bei der instruktionsbasierten Wahl der Instruktionsextrakte weitere Auswahlkriterien relevant gewesen wären wie etwa Hinzunahme der Redezeiten und -anteile der Lernenden (dann wäre beispielsweise aufgefallen, dass diese für die beiden Instruktionsextrakte unterschiedlich auf die Lernenden verteilt sind). Es könnte sein, dass Vorstellungen zur Unterscheidung von Beobachtung und Deutung sich allgemein schlechter im Bearbeitungsprozess von Instruktion rekonstruieren lassen und stattdessen besser in Interviewformaten analysiert werden sollten, beispielsweise aus dem Grund. Damit zusammenhängend könnte es sein, dass das gewählte Kodierverfahren zur qualitativen Inhaltsanalyse aufgrund der situativ erfolgenden Schlagwortvergabe zur Rekonstruktion von Vorstellungen zu Beobachtungen und Deutungen nicht optimal geeignet

ist, während es für Kennzeichen von naturwissenschaftlichen Fragen gut funktioniert (was u. a. deshalb plausibel wäre, weil mit der Entwicklung des Verfahrens bei dem Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen begonnen wurde). Nur eine Hinzunahme weiterer Instruktionsextrakte zu weiteren Konzepten des experimentbezogenen Denkens und Arbeitens kann hierüber wirklich Aufschluss geben.

Abschließend wird noch auf die beiden vom typischen Muster (siehe Tabelle 6.40 auf S. 409) signifikant abweichenden Prozessvariablen *Anzahl der Instanzen mit OKs* und *Korrektheitsmaß für Einschätzungen* eingegangen. In beiden Fällen finden sich statistisch bedeutsame Unterschiede zwischen Personen mit unterschiedlichen fachmethodischen Eingangsvoraussetzungen beim Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen; für andere Prozessvariablen treten derartige Unterschiede nicht signifikant auf – ob sie eventuell für manche Prozessvariablen in der Tendenz vorliegen, wurde allerdings nicht im Detail untersucht; es kann also, insbesondere aufgrund der kleinen Stichproben, zum Teil der Fall sein. In jedem Fall ist das Auftreten der Zusammenhänge für die beiden genannten Prozessvariablen gut nachvollziehbar, wie nachfolgend dargestellt wird.

Anzahl der Instanzen mit Ideen-Oberkategorien. Diese Prozessvariable bildet ab, bei wie vielen (Teil-)Aufgaben einer/einem Lernenden Ideen-Oberkategorien zugewiesen wurden. Sie ist somit gewissermaßen ein Maß dafür, wie sehr sich aus dem Gesprochenen eines einzelnen Lernenden Vorstellungen rekonstruieren lassen. Der Zusammenhang, der sich mit den fachmethodischen Eingangsvoraussetzungen ergibt, ist von kleinerer Effektstärke als der Zusammenhang, der sich mit den erreichten Kompetenzzuwächsen ergibt. Ohne weitere Zusammenhangsanalysen (bspw. mit Moderationsmodellen) ist es allerdings müßig, über spezifische Deutungen zu spekulieren. An dieser Stelle muss es genügen, dass es bei allen betrachteten Prozessvariablen für die Prozessvariable zur *Anzahl der Instanzen mit OKs* am plausibelsten ist, Zusammenhänge zu den fachmethodischen Eingangsvoraussetzungen zu finden, selbst unter Berücksichtigung nur minimaler Vorwissensunterschiede (weil eventuell nahezu alle Lernenden ohne Vorerfahrungen zu dem Konzept des experimentbezogenen Denkens und Arbeitens in die Bearbeitung des Instruktionsextrakts starten; siehe weiter oben im Text). Es ist plausibel, dass die fachmethodischen Eingangsvoraussetzungen beeinflussen, wie häufig sich die Vorstellungen der Lernenden aus ihren Äußerungen rekonstruieren lassen, weil bessere Vertrautheit mit dem Inhaltsbereich vermutlich bedeutet, dass mehr Vorstellungen artikuliert werden können oder zumindest die Arti-

kulation sprachlich ausgereifter ist (und daher eine Rekonstruktion von Vorstellungen vom Beobachter besser erfolgen kann).

Korrektheit der Einschätzungen. Die Prozessvariable bildet ab, wie viele der getroffenen fachmethodischen Entscheidungen (spezifischer: Klassifikationsentscheidungen zum betrachteten fachmethodischen Konzept) korrekt waren. Insofern ist diese Prozessvariable auf die Fähigkeiten der Lernenden bezogen. Einerseits ist es plausibel, dass die Fähigkeiten deutlich stärker mit den Vorerfahrungen und damit auch mit den fachmethodischen Eingangsvoraussetzungen zusammenhängen als die in den anderen Prozessvariablen abgebildeten spezifischen Merkmale der Vorstellungen. Andererseits finden sich für diese Prozessvariable beim Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen Zusammenhänge für alle untersuchten Testvariablen, weshalb spezifischere Deutungen ohne weitere Analysen nicht vorgenommen werden sollten.

6.10.6 Untersuchte Prozessmerkmale und nicht-fachmethodische Eingangsvoraussetzungen

An dieser Stelle werden nur einige besonders relevante und eindruckliche Ergebnisse benannt.

Erstens findet sich für fachinhaltliche Eingangsvoraussetzungen (Testvariable FCI) der plausible Zusammenhang mit den beiden Prozessvariablen, die auf die Allgemeinheit der Vorstellungen bezogen sind. Lernende, die fachinhaltlich höhere Kompetenzen aufweisen, leisten also auch fachmethodisch stärker Abstraktionsarbeit. Interessanterweise findet sich allerdings kein Zusammenhang mit den allgemeinen kognitiven Fähigkeiten, obwohl dies ebenfalls erwartbar wäre.

Zweitens sind genauere Analysen erforderlich, um Zusammenhänge von fachinhaltlichen Eingangsvoraussetzungen und anderen Prozessvariablen zu klären. Es bleibt zunächst offen, warum beispielsweise Lernende mit höheren fachinhaltlichen Eingangsvoraussetzungen beim Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen auch mehr Äußerungen aufweisen sollten, aus denen sich Vorstellungen rekonstruieren lassen. Eventuell bleiben diese Lernenden eher bei der Aufgabe, die Beispielfragen zu klassifizieren, und weichen weniger als andere Lernende dahin ab, die Beispielfragen beantworten zu wollen. Es bleibt ferner offen, warum die fachmethodische Angemessenheit der Vorstellungen oder die Stabilität der fachmethodischen Vorstellungen für Lernende mit unterschiedlichen fachinhaltlichen Eingangsvoraussetzungen unterschiedlich sein sollten.

Drittens finden sich für die Testvariablen zu allgemeinen kognitiven Fähigkeiten, zum Selbstwirksamkeitskonzept und zum Interesse nur sehr wenige Zusammenhänge. Grundlegend ist dies plausibel, wenngleich manche der Zusammenhänge verwunderlich sind und sich eventuell aufgrund der kleinen Stichproben als Artefakte erklären lassen. Ad hoc plausibel scheint hierbei allerdings der Zusammenhang von Selbstwirksamkeitskonzept und den Prozessvariablen zu Generalisierungen zu sein: Wird berücksichtigt, dass über 45 % der Generalisierungen sich bei den Teilaufgaben 1-22-ii (für Kennzeichen von naturwissenschaftlichen Fragen) und 3-22-ii (für die Unterscheidung von Beobachtung und Deutung) finden, scheint es plausibel, dass Lernende, die ein höheres Selbstwirksamkeitskonzept haben auch verbal aktiver sind und daher selbstsicherer Verallgemeinerungen verbalisieren, wenn sie dazu aufgefordert werden, als Lernende mit einem niedrigen Selbstwirksamkeitskonzept.

7 Zusammenführungen, Diskussionen und Ausblicke

Der Aufbau und die Förderung von Kompetenzen zum experimentbezogenen Denken und Arbeiten sind zentrale Ziele des naturwissenschaftlichen Unterrichts (siehe Kapitel 1). Ein zentrales Ziel der explorativ angelegten Arbeit ist, basierend auf Ergebnissen zum Denken und Handeln der Lernenden bei der Bearbeitung von (kompetenzförderlicher) Instruktion zum experimentbezogenen Denken und Arbeiten Hinweise zu Wirkmechanismen von Instruktion zu erlangen sowie erste Erkenntnisse zur Prozessen des Kompetenzaufbaus zu generieren. Hierzu wurden die Bearbeitungsprozesse von Lernenden bei zwei verschiedenen fachmethodischen Instruktionsvarianten (dargestellt in Kapitel 4) untersucht. Die beiden Instruktionsvarianten bestehen jeweils aus drei zeitlich separierten Einheiten (siehe Abschnitt 4.2, ab S. 55) und unterscheiden sich insbesondere im Grad des ausdrücklichen Adressierens von konzeptuellen fachmethodischen Informationen (z. B. reines eigenständiges Entdecken vs. ausdrückliche instruktionale Explizierung von Regeln zur Variablenkontrollstrategie; ausführlicher in Abschnitt 4.3, ab S. 57). Im Gegensatz zu vielen vorangegangenen Forschungsarbeiten wurden bei der Untersuchung neben den sonst häufig mittels Prä-Post-Daten fokussierten Kompetenzzuwächsen auch die videographierten Bearbeitungsprozesse während der Instruktion betrachtet und mit den Kompetenzzuwächsen der einzelnen Lernenden in Beziehung gesetzt.

Aus Zusammenführungen und kritischen Diskussionen von in Kapitel 5 und 6 dargestellten Ergebnissen zu Aktivitäten und Vorstellungen der Lernenden bei der Bearbeitung der Instruktion werden im vorliegenden Kapitel Hypothesen zum Kompetenzaufbau der Lernenden, zu Wirkmechanismen von Instruktionen und zu Prozessen des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten abgeleitet. Jeder der folgenden Abschnitte beinhaltet eine themenspezifische Zusammenführung mit angeschlossener Diskussion.

Kurzübersicht

Abschnitt 7.1 (ab S. 420): Vielfalt in Aktivitäten und Vorstellungen

Abschnitt 7.2 (ab S. 431): Äußerungen mit fachmeth. Gehalt und Kompetenzaufbau

Abschnitt 7.3 (ab S. 463): Diverses zu Unterschieden zwischen Instruktionsvarianten

Abschnitt 7.4 (ab S. 475): Instruktionale Informationen und Praxis

Abschnitt 7.5 (ab S. 478): Prozesse des Kompetenzaufbaus

7.1 Vielfalt in den Aktivitäten und Vorstellungen von Lernenden bei der Bearbeitung von Instruktion zum experimentbezogenen Denken und Arbeiten

In Fazit 1 (unten auf der vorliegenden Seite) werden die Ergebnisse der Arbeit unter dem Gesichtspunkt der Vielfalt zusammengefasst. Dieser Gesichtspunkt hat sich bei der Betrachtung der Daten und Kategorisierungen als strukturierend herausgestellt.

Da die beiden untersuchten Instruktionsvarianten in Teams zu je zwei bis drei Personen bearbeitet wurden, wurden die Lernenden sowohl auf Ebene der einzelnen Personen als auch auf Ebene von Teams analysiert, um Ergebnisse nicht nur vor dem Hintergrund der materiellen Umgebung des Lernens (d. h. der Instruktion), sondern auch der personellen Umgebung des Lernens zu generieren, sofern dies möglich und sinnvoll erscheint. Die Ergebnisse lassen sich hinsichtlich dreier Analysefokusse klassifizieren:

- Analysen *in der Gesamtschau* sind auf die komplette Bearbeitung der Instruktion oder auf eine der drei Instruktionseinheiten gerichtet und basieren auf Mittelwert- oder Summenbildungen oder qualitativen Charakterisierungen der Bearbeitung als Gesamtes (z. B. Häufigkeit des Auftretens einer ausgewählten Aktivität oder Vorstellungen innerhalb einer Instruktionseinheit).
- Analysen *für spezifische Situationen* (z. B. Bearbeitung bestimmter Aufgabenkarten innerhalb der Instruktion) sind so angelegt, dass einzelne Situationen aufgelöst und Vergleiche für jede einzelne Situation vorgenommen werden.
- Analysen *im zeitlichen Verlauf* lösen für die einzelnen Lernenden (d. h. die einzelnen Personen oder die einzelnen Teams) auf, welche Aktivitäten oder Vorstellungen in miteinander vergleichbaren zeitlich aufeinanderfolgenden Situationen auftreten.

Fazit 1

Das Denken und Handeln der Lernenden ist bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante von Vielfalt gekennzeichnet:

- (a) Die Aktivitäten unterschiedlicher Lernender unterscheiden sich sowohl in der Gesamtschau als auch für je spezifische Situationen deutlich voneinander (z. B. gibt es Personen, für die der zeitliche Anteil nonverbaler Experimentierhandlungen sehr groß ist, und Personen für die dieser Anteil sehr klein ist, siehe Unterabschnitt 5.4.2, ab S. 123).

- (b) Aktivitäten von Lernenden mit *vergleichbaren* Kompetenzzuwächsen *unterscheiden* sich in der Gesamtschau und für je spezifische Situationen zum Teil deutlich voneinander, während Aktivitäten von Lernenden mit *unterschiedlichen* Kompetenzzuwächsen zum Teil *große Ähnlichkeit* in der Gesamtschau und für je spezifische Situationen aufweisen (siehe insbesondere Abschnitte 5.4.3–5.4.5, ab S. 163).
- (c) Die Äußerungen und Vorstellungen der Lernenden zu spezifischen Konzepten des experimentbezogenen Denkens und Arbeitens sind inhaltlich (abgebildet durch Ideen-Kategorien und Ideen-Oberkategorien; Details in Abschnitt 6.1, ab S. 242) sowohl in der Gesamtschau als auch für je spezifische Situationen sehr vielfältig (z. B. treten 14 unterschiedliche Ideen-Oberkategorien bei der Aufgabe auf, in der einem fiktiven Gegenüber der Unterschied von naturwissenschaftlichen und nicht-naturwissenschaftlichen Fragestellungen erklärt werden soll; Unterabschnitt 6.5.1, ab S. 295; siehe auch insbesondere Abschnitte 6.4, ab S. 283, und 6.8, ab S. 326). Für viele der einzelnen Personen findet sich ferner im zeitlichen Verlauf inhaltliche Vielfalt der Äußerungen bzw. Vorstellungen zu den ausgewählten Konzepten (z. B. häufiger Wechsel zwischen den Ideen-Oberkategorien statt stabile über Aufgaben hinweg zugewiesene Ideen-Oberkategorien; siehe insbesondere Abschnitt 6.8).

Die in diesem Fazit bezüglich der Aktivitäten aufgeführten Ergebnisse gelten im Wesentlichen auch für die etwas weniger detailliert untersuchte implizit-fachmethodische Instruktionsvariante (siehe dafür Abschnitt 5.5, ab S. 207). Wie die im Fazit für die explizit-fachmethodische Instruktionsvariante benannte Vielfalt sich real manifestiert, ist ausführlicher in den nachfolgenden Unterabschnitten 7.1.1 und 7.1.2 dargestellt (als Synthese von Ergebnissen aus den Kapiteln 5 und 6 zu den Aktivitäten und Vorstellungen der Lernenden); angeschlossen sind jeweils ausgewählte Diskussionen von Methoden und Ergebnissen der Arbeit vor dem Hintergrund der Forschung. In Unterabschnitt 7.1.3 wird abschließend eine Hypothese aus der Zusammenführung der Ergebnisse abgeleitet.

7.1.1 Aktivitäten sowie Verknüpfung von Aktivitäten und Kompetenzzuwächsen

Die Ergebnisse der Arbeit zu den *Aktivitäten* der Lernenden bei der Bearbeitung der untersuchten Instruktionsvarianten sind insofern im Einklang mit Befunden aus Instruktions- und Unterrichtsforschung aus dem Bereich der Naturwissenschaftsdidaktik, als dass in ihnen von hoher Varianz bezüglich der Handlungen von Lernenden innerhalb derselben Lernumgebung berichtet wird (vgl. u. a. Klug, 2017; Knobloch, 2011; Rogge, 2010; Steckenmesser-Sander, 2015; allgemeiner auch einige Beiträge in Rauin et al., 2016). Die Ergebnisse dieser Arbeit sind u. a. auf Aktivitätsprofile in der Gesamtschau bezogen, d. h., darauf, in welchem Verhältnis die verschiedenen zeitlichen Anteile der Aktivitäten für einzelne Personen zueinander stehen. Die Aktivitätsprofile wurden in Typen geclustert und es lassen sich viele sehr verschiedene Aktivitätsprofiltypen finden (für die explizit-fachmethodische Instruktionsvariante beispielsweise je nach Instruktionseinheit 5–7 Typen für 19–33 Lernende; genauere Darstellung und Interpretation in den Abschnitten 5.4.2, ab S. 123, und 5.5.2, ab S. 214; Vielfalt in der Gesamtschau). Auch eine Betrachtung der einzelnen Karten der Instruktion, die jeweils als Situationen aufgefasst werden, zeigt, dass große Streuung in der Bearbeitung durch verschiedene Personen besteht (insbesondere Abschnitte 5.4.1, ab S. 106, 5.4.5, ab S. 184, und 5.4.6, ab S. 190; Vielfalt für spezifische Situationen). Vor dem Hintergrund, dass die einzelnen Karten jeder Instruktionseinheit zwar grob vergleichbar sind, aber verschiedene Fähigkeiten und Verständnisse adressieren, ist es ein naheliegendes Ergebnis, dass die Anteile der Aktivitäten ferner für verschiedene Situationen (unterschiedliche Aufgaben, Versuchskontexte, Fragestellungen, ...) deutlich voneinander abweichen (insbesondere Abschnitte 5.4.1, ab S. 106, 5.4.5, ab S. 184, und 5.4.6, ab S. 190; Vielfalt im zeitlichen Verlauf).¹²⁰

Die gefundene Varianz in den Aktivitäten bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante wurde als Grundlage für weitere Analysen zum Kompetenzaufbau genutzt, indem sie mit dem Kompetenzaufbau der Lernenden

¹²⁰Die Aussage ist sehr generell formuliert und demnach auch auf miteinander nicht vergleichbare Situationen bezogen, für die Abweichungen in den Aktivitäten grundsätzlich erwartbar sind; sie gilt spezifischer auch dann, wenn die für *Aktivitäten* in der Arbeit verglichenen verschiedenen Experimentiersituationen betrachtet werden. Allerdings ist auch für diese anzumerken, dass sie zwar besser miteinander vergleichbar sind als beliebigen Karten derselben Einheit (oder auch desselben inhaltlichen Instruktionsabschnitts), sich allerdings auch für diese eine Einschränkung für Aussagen zum zeitlichen Verlauf ergibt, weil die Situationen in verschiedener Hinsicht (z. B. von Lernenden genutztes Material, Messmethode) zum Teil deutlich voneinander abweichen. Analysen zu den *Vorstellungen* von Lernenden nehmen inhaltlich stärker vergleichbare Karten der Instruktion in den Blick und vergleichen diese hinsichtlich der Inhalte der Vorstellungen von Lernenden; daher sind die weiter unten zu Vorstellungen berichteten Ergebnisse *nicht* mit denselben Einschränkungen verbunden.

abgeglichen wurde. Dabei wurde in der Arbeit insbesondere der Frage nachgegangen, ob die Varianz in den *Aktivitäten in einem Zusammenhang mit den Kompetenzzuwächsen* der Lernenden steht. Hierbei lassen sich zwar Unterschiede in den Aktivitäten von Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen finden; systematische Unterschiede beschränken sich allerdings im Wesentlichen auf die Aktivität der fachmethodischen Äußerungen (sowohl hinsichtlich des zeitlichen Anteils als isoliert betrachtete Aktivität, Unterabschnitt 5.4.4, ab S. 173, als auch hinsichtlich ihrer Stellung in Aktivitätsprofilen, Unterabschnitt 5.4.3, ab S. 163) und weisen auf die Bedeutung des zeitlichen Umfangs fachmethodischer Überlegungen hin (mehr in Abschnitt 7.2, ab S. 431). Im Einklang mit Befunden aus Prozess-Produkt-Forschungen zum Lehren und Lernen, in denen die Bearbeitungsprozesse die erzielten Lernergebnisse meist nicht erklären (allgemein z. B. Ditton, 2002; Gräsel & Göbel, 2015; spezifisch fachdidaktisch und für das Schülerhandeln Knobloch, 2011), finden sich insgesamt jedoch keine eindeutigen Zusammenhänge von Aktivitäten und Kompetenzzuwächsen. So umfassen die Ergebnisse zur explizit-fachmethodischen Instruktionsvariante beispielsweise, dass sich für ähnliche Kompetenzzuwächse (sowohl hohe als auch niedrige) mehrere jeweils verschiedene Aktivitätsprofile (d. h. Vielfalt in der Gesamtschau) finden lassen (Unterabschnitt 5.4.3, ab S. 163).

7.1.2 Äußerungen und Vorstellungen zu ausgewählten Konzepten des experimentbezogenen Denkens und Arbeitens

Wie variantenreich das Denken und Handeln der Lernenden bei der Bearbeitung der untersuchten Instruktion ist, spiegelt sich nicht nur in den Aktivitäten der Lernenden, sondern in ersten Asnätzen auch in den für die explizit-fachmethodische Instruktionsvariante *inhaltlich* untersuchten *Äußerungen* der Lernenden wider, die zu einzelnen ausgewählten Konzepten des experimentbezogenen Denkens und Arbeitens identifiziert wurden. Diese Äußerungen sind in der Gesamtschau *inhaltlich vielfältig*. Vor dem Hintergrund, dass sich die untersuchten Konzepte zu Kennzeichen von naturwissenschaftlichen Fragen und zur Unterscheidung von Beobachtungen und Deutungen inhaltlich jeweils eng umreißen lassen (siehe Abbildung C.2, auf S. 542, und Abbildung D.2, auf S. 546), wurden beispielsweise trotz dieser vergleichsweise engen Bezugsrahmen jeweils sehr viele inhaltlich verschiedene Ideen-Kategorien zur Klassifikation der Äußerungen induktiv generiert (53 Kategorien zu Kennzeichen von naturwissenschaftlichen Fragen, 28 Kategorien zur Unterscheidung von Beobachtung und Deutung; Abschnitt 6.4, ab S. 283: Vielfalt in der Gesamtschau). Auch legen Analysen der Ideen-Kategorien sowie mehrere angeführte kurze Transkriptstücke nahe, dass für verschiedene Lernende inhaltlich sehr unterschiedliche Äußerungen

zu denselben Situationen (z. B. zu jeweils hinsichtlich der Naturwissenschaftlichkeit einzuschätzenden Beispielfragen) vorliegen (z. B. Abschnitt 6.8, ab S. 326: Vielfalt für spezifische Situationen).

Dass die Äußerungen der Lernenden insgesamt und auch dann, wenn einzelne Situationen für sich genommen betrachtet werden, anhand der generierten Ideen-Kategorien als *inhaltlich* vielfältig anzusehen sind, könnte unter anderem dadurch bedingt sein, dass die Ideen-Kategorien mit starker Bindung an die je konkreten aufgefundenen Formulierungen differenziert wurden (siehe Unterabschnitt 6.1.1, ab S. 245). Daher wurden die Ideen-Kategorien in sogenannten Ideen-Oberkategorien zusammengefasst, wenn sie inhaltliche Ähnlichkeit hinsichtlich der Klassifikation von naturwissenschaftlichen Fragen bzw. von Beobachtung und Deutung aufwiesen (siehe Unterabschnitt 6.1.7, ab S. 264). Ein wesentliches Ergebnis der Arbeit ist, dass sich auch auf dieser Ebene große Vielfalt einstellt. Es wird so gedeutet, dass die Vorstellungen der Lernenden zu den untersuchten Konzepten vielfältig sind:

Für die Ideen-Oberkategorien findet sich Vielfalt in der Gesamtschau (u. a. Abschnitt 6.4, ab S. 283) und Vielfalt für spezifische Situationen (u. a. Abschnitt 6.8, ab S. 326). Ferner liegen umfangreiche Ergebnisse vor, die die Vielfalt der Ideen-Oberkategorien im zeitlichen Verlauf für je spezifische Lernende belegen. So wechseln beispielsweise im Verlauf der Bearbeitung die den einzelnen Personen – und mit leichten Einschränkungen auch den Teams als einzelnen Entitäten – zugewiesenen Ideen-Oberkategorien häufig (siehe Abschnitte 6.8, ab S. 326, und 6.9, ab S. 353), obwohl die Aufgaben keine Wechsel von Ideen-Oberkategorien erfordern (was daran deutlich wird, dass auf Kohortenebene für jede Aufgabe breite Spektren an Ideen-Oberkategorien vorliegen; siehe Unterabschnitt 6.8.1, ab S. 326). Nur wenigen Personen werden (wenige) Ideen-Oberkategorien stabil über die Situationen hinweg immer wieder zugewiesen (siehe Abschnitt Unterabschnitt 6.8.2, ab S. 330). Einschränkend ist anzumerken, dass innerhalb von Teams durchaus über die Situationen hinweg eine geringe Zahl (meist zwei oder drei) Ideen-Oberkategorien immer wieder zugewiesen werden (siehe Unterabschnitt 6.8.2). Diese unterscheiden sich jedoch für verschiedene Teams. Insgesamt findet sich also eine große Vielfalt der Vorstellungen, wenn die Ideen-Oberkategorien als Vorstellungen der Lernenden interpretiert werden.

Die berichteten Vorstellungen passen inhaltlich gut zu den Befunden anderer Forschung, scheinen allerdings zum Teil feingliedriger zu sein als andernorts berichtet (ausführlicher in Abschnitt 6.4 beginnend auf den Seiten 288 und 292). Die Vielfalt der Vorstellungen im zeitlichen Verlauf kann allerdings als diskrepant dazu angesehen werden, dass den Vorstellungen in der Literatur meist – zum Beispiel, wenn

Konsequenzen für den Unterricht abgeleitet werden – (zumindest implizit) zeitliche und kognitive Stabilität zugeschrieben wird (vgl. z. B. von Aufschnaiter & Rogge, 2015). Nachfolgend wird daher diskutiert, inwiefern das Ergebnis der Arbeit, dass große inhaltliche Vielfalt und kaum Stabilität im zeitlichen Verlauf gefunden werden kann, in dem gewählten methodischen Vorgehen begründet liegen könnte:

- *Dem angewendeten methodischen Vorgehen liegt ein spezifischer Vorstellungsbegriff zugrunde.* In der vorliegenden Arbeit wird nicht davon ausgegangen, dass Vorstellungen bereits stabile Konstrukte im Denken der Lernenden sind. Für eine Übersicht über Kritik an begrifflichen Fassungen von Vorstellungen, die stabile Wissens Elemente oder Wissensstrukturen betonen, auf denen konkrete Äußerungen von Lernenden beruhen (z. B. p-prims oder mentale Modelle) siehe Rogge (2010). Es wird vielmehr grundsätzlich von Vorstellungen gesprochen (im Sinne davon, was Lernende durch ihre Äußerungen den Zuhörenden vorstellen), auch dann, wenn beispielsweise Ideen-Oberkategorien nur einmal für eine Person zugewiesen sind. Damit wird allerdings nicht gesagt, dass alle berichteten Vorstellungen für alle Personen, denen sie zugewiesen sind, auch in derselben Weise (zeitlich oder gar kognitiv) stabil sind. Wenn also eine große Vielfalt an Vorstellungen berichtet wird, darf dies nicht so gelesen werden, dass diese Vorstellungen bei vielen untersuchten Personen auch in großer (zeitlicher oder kognitiver) Stabilität vorliegen. Vielmehr zeigt sich die große Vielfalt bezüglich der Vorstellungen unter anderem gerade darin, dass große Varianz für einzelne Situationen und über die Situationen hinweg vorliegt. Weitere Forschung könnte die in der vorliegenden Arbeit berichteten Vorstellungen ausführlicher darauf hin untersuchen, welche dieser Vorstellungen zum experimentbezogenen Denken und Arbeiten Lernende tatsächlich mit großer Stabilität generieren.¹²¹ Es ist möglich, dass sich bereits dadurch die Ergebnisse der Arbeit in den sonstigen Forschungsstand harmonisch einfügen lassen (wenn dort jeweils utilisierte Vorstellungsbegriffe hinsichtlich der mit ihnen verbundenen kognitiven Stabilität berücksichtigt werden).
- *Dem angewendeten methodischen Vorgehen liegt die Annahme zugrunde, dass Bedeutungskonstruktion situativ erfolgt und somit die von Lernenden konstruierte Bedeutung nicht notwendigerweise über größere Zeitfenster hinweg konstant ist.* Die an einzelnen Äußerungen vorgenommene Kodierung, die ohne

¹²¹Anschließend an Forschung dazu, welche Vorstellungen mit welcher Stabilität vorliegen, könnten Konsequenzen für Unterricht abgeleitet werden, die sowohl die Stabilität als auch die Variabilität unterschiedlicher Vorstellungen berücksichtigen.

Einbeziehen weiterer Äußerungen derselben Person erfolgte, gründet in dem Ziel, zeitliche Veränderung identifizieren zu können, aber auch in der Annahme, dass Lernende Bedeutung situativ konstruieren (vgl. Rogge, 2010; von Aufschnaiter, 1999). Es wurde daher darauf verzichtet, aus Beobachtersicht scheinbar mehrere Äußerungen erklärende latente Konstrukte aus solchen Äußerungen abzuleiten, die in aus Sicht der Lernenden (aller Wahrscheinlichkeit nach) verschiedenen Kontexten entstanden sind (vgl. Zeitspanne für die Konstruktion von Bedeutung; z. B. von Aufschnaiter, 1999; Rogge, 2010).¹²² Dieses Vorgehen berücksichtigt ferner, dass zwar aus Beobachtersicht Gemeinsamkeiten bestehen können, dies aus Lernendensicht allerdings noch lange nicht der Fall sein muss (vgl. Perspektiven erster und zweiter Ordnung; für diese Lesart von Marton & Booth, 1997, siehe beispielsweise von Aufschnaiter, 2003, und folgend Rogge, 2010; ähnlich auch Beobachtungen erster und zweiter Ordnung von Luhmann, 1997, wie beispielsweise rezipiert bei Schäfer, 2016).

- *Die Ergebnisse der Arbeit könnten darauf hindeuten, dass das methodische Vorgehen zu stark an den beiden vorigen Annahmen ausgerichtet ist und zu wenig gegenstandsbezogen ist.* Neben der Abweichung der Ergebnisse der vorliegenden Arbeit von anderer Forschung (siehe weiter oben) könnte das folgende Ergebnis der Arbeit als ein Hinweis darauf gedeutet werden, dass die Vorstellungen der Lernenden nicht optimal durch die Ideen-Oberkategorien beschrieben werden: Werden die Informationstexte kodiert, die den Lernenden in Textform auf Karten dargereicht werden, werden viele Ideen-Kategorien zugewiesen (siehe Unterabschnitt 6.9.1 beginnend auf den Seiten 360 und 376). Diese Ideen-Kategorien sind verschiedenen Ideen-Oberkategorien zugehörig (in der Zahl nur leicht weniger als Ideen-Kategorien, weil fast jede Ideen-Kategorie zum Informationstext zu einer anderen Ideen-Oberkategorie gehört). Hinter den verschiedenen Ideen-Kategorien steht jedoch ein inhaltlich zusammenhängender und auf ein einziges Konzept bezogener (instrukionaler) Text. Dieser könnte als latentes Vorstellungs-Konstrukt zu den Ideen-Kategorien bezeichnet werden. Demnach dürften nicht die verschiedenen Ideen-Oberkategorien als jeweils einzelne Vorstellungen gedeutet werden, sondern die Vorstellung bestünde aus *einem* latenten Konstrukt, welches hinter den entsprechenden Ideen-Kategorien und damit auch den Ideen-Oberkategorien stünde und dem jeweils in einem

¹²²Der Verzicht erfolgte, indem Ideen-Kategorien niemals aus dem Zusammenziehen mehrerer Äußerungen erfolgten. Die Ideen-Kategorien wurden vielmehr anhand einer einzelnen Äußerung generiert und dann weiteren Äußerungen in anderen Kontexten, auch von anderen Personen, zugewiesen, sofern sie passten.

Satz formulierbaren Konzept entspricht (inhaltlich ausgeführt in Unterabschnitt Unterabschnitt 6.9.1, ab S.354).

Analog zum Vorgehen, einem Informationstext nur eine passende Vorstellung (oder nur sehr wenige passende Vorstellungen) zuzuschreiben, könnte auch mit den Äußerungen einer Person (oder den ihr zugewiesenen Ideen-Kategorien) verfahren werden: Es wäre möglich, qualitativ nach wenigen latenten Vorstellungen zu suchen, aus denen jeweils möglichst viele der Äußerungen (oder Ideen-Kategorien) einer Person entspringen könnten. In der quantitativen Forschung entspräche ein derartiges Vorgehen in etwa ein latenten Klassen- oder Prozessanalyse (z. B. Geiser, 2009; Lanza et al., 2006). Ein solches Vorgehen wurde aufgrund der oben erwähnten Annahmen allerdings nicht gewählt. Es ist wahrscheinlich, dass bei solch einem Vorgehen eine deutlich geringere Vielfalt der Vorstellungen (mindestens im zeitlichen Verlauf, aber vermutlich auch in der Gesamtschau und für einzelne Situationen) berichtet würde.

7.1.3 Vielfalt im Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten

Basierend auf den Ergebnissen aus den vorigen Unterabschnitten wird die folgende Hypothese formuliert:¹²³

Hypothese 1

Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten ist von Vielfalt gekennzeichnet.

Die formulierte Hypothese stellt im Wesentlichen eine Generalisierung der berichteten Ergebnisse dar. Sie geht dabei in mindestens zweifacher Hinsicht über die Ergebnisse hinaus:

¹²³Sowohl Hypothese 1 als auch weitere generierte Hypothesen sind nicht im Sinne der in Kapitel 4 aufgeführten Definition von Hypothesen (im Kontext naturwissenschaftlicher Denk- und Arbeitsweisen) als Hypothesen im engeren Sinne anzusehen. Sie basieren auf empirischen Ergebnissen und theoretischen Überlegungen aus der vorliegenden Arbeit sowie empirischen und theoretischen Befunden anderer Arbeiten und sind als Aussagen über vermutete Gesetzmäßigkeiten und Zusammenhänge für weitere Forschung nutzbar; allerdings sind sie nicht dezidiert prüfbar formuliert, sondern bedürfen der Ausschärfung in je spezifischen Studien. Diese Form wurde basierend auf dem explorativen Charakter der Arbeit gewählt und dient ferner dazu, die formulierten Hypothesen möglichst einfach zugänglich zu halten und eine grobe Richtung für weitere Arbeiten anzugeben.

Die Hypothese ist auf Kompetenzaufbau (zum experimentbezogenen Denken und Arbeiten) im Allgemeinen bezogen. Das in der Arbeit untersuchten Denken und Handeln sowie die damit in Verbindung gesetzten Kompetenzzuwächse stellen verschiedene und vermutlich wesentliche Facetten von Kompetenzaufbau dar (siehe Kapitel 2); es ist zunächst davon auszugehen, dass sich auch bezüglich anderer Aspekte des Denkens und Handelns sowie hinsichtlich anderer Maße für Kompetenzveränderung Vielfalt einstellt. Zusätzlich zu den Befunden aus der vorliegenden Arbeit sprechen dafür unter anderem Befunde von Dean und Kuhn (2007), die so zusammengefasst werden, dass »both interindividual and intraindividual variability are the norm« (S. 396). Unterstützend kann auch auf Resultate aus anderen Studien verwiesen werden, in denen beispielsweise die Fehler, die Lernende beim Experimentieren machen, vielfältig sind und bleiben (Kechel, 2016; Telser, 2019); auch erste Ergebnisse von Kirstein et al. (2019), die dezidiert Zusammenhänge zwischen Kompetenzzuwächsen und Aktivitäten im Lernprozess herstellen, zeigen sehr ähnliche Ergebnisse zur vorliegenden Arbeit, weil dort die Aktivitäten der Lernenden und die zeitlichen Anteile verschiedener Experimentierphasen in keinem Zusammenhang mit den Kompetenzzuwächsen zu stehen scheinen.¹²⁴

Die Validität der Generalisierung ist u. a. durch folgende Faktoren beschränkt:

- Die betrachteten Aktivitäten und Äußerungen stellen nur einen kleinen Ausschnitt von dem Denken und Handeln der Lernenden dar, welches für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten möglicherweise von Relevanz ist. Beispielsweise könnten nicht nur *weitere Aktivitäten* untersucht werden, sondern es könnte auch danach gefragt werden, welche Überlegungen der Lernenden sich nicht nur in (flüchtigen) mündlichen Äußerungen, sondern auch in (weniger flüchtigen) *schriftlichen Notizen* der Lernenden (wieder)finden. Außerdem sind die analysierten Sequenzen auf zwei ausgewählte Konzepte bezogen – Kennzeichen von naturwissenschaftlichen Fragen und die Unterscheidung von Beobachtung und Deutung –, die experimentbezogenes Denken und Arbeiten nicht umfassend abbilden. Untersuchungen zu *weiteren Konzepten* des experimentbezogenen Denkens und Arbeitens (z. B. zur häufig untersuchten Variablenkontrolle) bleiben demnach ein, vermutlich mit einer der in der Arbeit genutzten vergleichbaren Methode (siehe insbesondere Abschnitt 6.1, ab S. 242) bearbeitbares, Desiderat.

¹²⁴ Ähnlich finden sich auch fachinhaltliche Varianz in den Bearbeitungen z. B. bei Rogge (2010), und fehlende Unterschiede in auf Fachinhalten bezogenen verbalen Aktivitäten trotz deutlich unterschiedlicher fachinhaltlicher Kompetenzzuwächse z. B. bei Knobloch (2011).

- Aussagen zum Kompetenzaufbau umfassen im Allgemeinen auch Aussagen über die vorliegenden und sich zeitlich verändernden Fähigkeiten, Fertigkeiten, Verständnisse und Bereitschaften von Lernenden (basierend auf dem Kompetenzbegriff aus Kapitel 2). *Fertigkeiten* und *Bereitschaften* bleiben im Rahmen der Arbeit ununtersucht und damit Desiderate. Die Vielfalt in den *Verständnissen* der Lernenden wird durch die berichtete Vielfalt in den Vorstellungen der Lernenden eindrücklich illustriert. Gleichwohl gelten die bereits im vorigen Aufzählungspunkt erwähnten Einschränkungen durch die eng umgrenzten untersuchten Konzepte des experimentbezogenen Denkens und Arbeitens (Kennzeichen naturwissenschaftlicher Fragen, Unterscheidung von Beobachtung und Deutung). Für Vielfalt in den *Fähigkeiten* der Lernenden finden sich erste Hinweise in der Arbeit, weil Lernende mit ähnlichen Kompetenzzuwächsen in vergleichbaren Situationen deutlich verschiedene Aktivitäten zeigen (z. B. Unterschiede in den fachmethodischen Aktivitäten bei den Experimentiersituationen, siehe Unterabschnitt 5.4.6, ab S. 190). Wenngleich die durchgeführten Analysen der Aktivitäten keine direkten Aussagen über die Fähigkeiten ermöglichen, erlauben sie unter anderem, zwischen fachinhaltlichen und fachmethodischen Fokussierungen sowie der Orientierung auf andere Inhaltsbereiche aus der Physik oder dem Alltag zu unterscheiden. In den Aktivitäten wird damit auch – zumindest in ersten Ansätzen – deutlich, welche Fähigkeiten bzw. Strategien der Aufgabenbearbeitung eingesetzt werden, ob also z. B. überhaupt fachmethodisches Denken und Handeln in den Blick genommen wird oder nicht. Aus einer Vielfalt in den Aktivitäten kann also mit gewisser Plausibilität eine hypothetische Aussage über Vielfalt in den Fähigkeiten abgeleitet werden.

Die Hypothese ist nicht nur auf die beiden in der Arbeit untersuchten Instruktionsvarianten bezogen. Die in der Arbeit untersuchten Instruktionsvarianten bestehen aus strukturiert angelegten Arbeitsaufträgen und Informationstexten (basierend auf Annahmen zum Kompetenzaufbau, siehe Kapitel 4), die so beschaffen sind, dass sie die Bearbeitung stark strukturieren – wenngleich innerhalb der Arbeitsaufträge für selbstständiges Arbeiten und Autonomie der Lernenden gesorgt ist. Für andere Instruktionen ist zunächst davon auszugehen, dass die Vielfalt im Denken und Handeln eher zu als abnimmt:

- Für Instruktionsansätze, die den Lernenden im Vergleich zu den beiden untersuchten Instruktionen deutlich *größere Autonomie* zugestehen (bezogen auf experimentbezogenes Denken und Arbeiten zum Beispiel im Sinne von Vorhol-

zer & von Aufschnaiter, 2019b, S. 4: »the extent in which research questions, methods of investigation etc. are ‘prescribed’ by the teacher«), ist zunächst davon auszugehen, dass die Vielfalt in den Aktivitäten und Vorstellungen bei der Bearbeitung eher zu als abnimmt.

- Andere Instruktionen können sich auch insofern im Grad der Strukturierung von den untersuchten Instruktionsvarianten unterscheiden, als dass sie *weniger Informationen* zum experimentbezogenen Denken und Arbeiten bereitstellen als die untersuchte explizit-fachmethodische Instruktionsvariante (z. B. *minimal-guidance*-Ansätze wie bei Kirschner et al., 2006): »assistance that explicitly or implicitly provides specific rules, strategies, explanations etc. (conceptual knowledge) before and/or during an inquiry in order to help students perform (a prescribed or self-chosen) inquiry activity correctly and successfully« (Vorholzer & von Aufschnaiter, 2019b, S. 5).¹²⁵ Für derartige Instruktionen ist hinsichtlich der betrachteten Fähigkeiten und Konzepte zu erwarten, dass die Vielfalt in den Aktivitäten eher größer ist als bei den untersuchten Instruktionsvarianten, da die Anregung von Aktivitäten weniger spezifisch erfolgt. Ferner dürfte die Vielfalt der Vorstellungen aufgrund einer geringeren Anzahl fachmethodischer Beiträge (denn bereits für die in der vorliegenden Arbeit untersuchte implizit-fachmethodische Instruktionsvariante ist der Anteil fachmethodischer Beiträge deutlich geringer) nicht auszumachen sein. Dies ist allerdings vermutlich vorrangig aufgrund methodischer Einschränkungen der Fall, denn die fachmethodischen Beiträge bei der in der vorliegenden Arbeit untersuchten implizit-fachmethodischen Instruktionsvariante sind trotz geringer Zahl auf vielfältige Inhalte bezogen.

¹²⁵Der Umfang bereitgestellter Informationen ist vermutlich zum Teil mit dem Grad der Autonomie (voriger Aufzählungspunkt) verschränkt (siehe Vorholzer & von Aufschnaiter, 2019b).

7.2 Bedeutung von Äußerungen mit fachmethodischem Gehalt für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten

Die in der Forschung im Zusammenhang mit dem naturwissenschaftlichen Denken und Arbeiten untersuchten instruktionalen Informationen (z. B. in Informationstexten oder Prompts) und Reflexion(sanregung)en zielen oft darauf ab, das Denken der Lernenden auf spezifische Aspekte von Fachmethoden zu lenken und sie zum Sprechen darüber anzuregen (siehe Kapitel 2). Es wird als naheliegend angesehen, dass Überlegungen zu verschiedenen Aspekten zu Fachmethoden eine Bedeutung für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten haben. Daher wurde in der Arbeit das Sprechen über Fachmethoden in zweierlei Hinsicht analysiert: (a) Die in der Arbeit vorgenommenen Analysen der Aktivitäten von Lernenden (Kapitel 5) beinhalten Kategorien, mithilfe derer sich auf die *Bedeutung des zeitlichen Umfangs von Äußerungen mit fachmethodischem Gehalt* schließen lässt. Nachfolgend wird zunächst auf die Bedeutung des zeitlichen Umfangs solcher verbaler Aktivitäten eingegangen, die in Kapitel 5 als Operationalisierung von Äußerungen mit fachmethodischem Gehalt genutzt werden, sogenannte *fachmethodische Beiträge*. (b) Die Analysen zu den Vorstellungen von Lernenden zu ausgewählten Konzepten des experimentbezogenen Denkens und Arbeitens (Kapitel 6) nehmen die *spezifischen Inhalte von Äußerungen mit fachmethodischem Gehalt* in den Blick. Im weiteren Verlauf des vorliegenden Abschnitts wird auch auf Ergebnisse eingegangen, die sich daraus in Bezug auf fachmethodische Beiträge ableiten lassen (u. a. in Unterabschnitt 7.2.1).

Vor dem Hintergrund der großen Vielfalt in den Aktivitäten und Vorstellungen sowie in den jeweiligen Verknüpfungen dieser mit den Kompetenzzuwächsen (siehe Abschnitt 7.1 für eine Zusammenfassung) ist es plausibel, davon auszugehen, dass *für Kompetenzaufbau relevante Prozessmerkmale* eher schwer auszumachen sind. Die Ergebnisse der Arbeit zeigen in der Tat, dass die Identifikation relevanter Prozessmerkmale nur bedingt gelingt. Beispielsweise finden sich einerseits verschiedenste Aktivitätsprofile bei sehr ähnlichen Kompetenzzuwächsen sowie andererseits sehr ähnliche Aktivitätsprofile bei verschiedensten Kompetenzzuwächsen; nur für sehr wenige Aktivitätsprofile lässt sich ein systematisches gemeinsames Auftreten mit hohen bzw. niedrigen Kompetenzzuwächsen finden (insb. Unterabschnitt 5.4.3, ab S. 163). Bezüglich der detailliert untersuchten fachmethodischen Beiträge finden sich dennoch einige Hinweise zur Bedeutung des zeitlichen Umfangs für den Kompetenzaufbau, die im Folgenden (stark) aggregiert berichtet und, insbesondere mit Blick auf Wirkmechanismen von fachmethodischen Instruktionen, diskutiert werden.

Fazit 2A

Die zeitlichen Anteile von fachmethodischen Beiträgen der Lernenden stehen in einem nicht ganz eindeutigen, eher positiven Zusammenhang mit Kompetenzzuwächsen der Lernenden zum experimentbezogenen Denken und Arbeiten:

- (a) Bei Bearbeitung der *explizit-fachmethodischen Instruktionsvariante* gehen größere zeitliche Anteile fachmethodischer Beiträge mit höheren Kompetenzzuwächsen einher; andersherum gehen höhere Kompetenzzuwächse allerdings nicht notwendigerweise mit größeren zeitlichen Anteilen fachmethodischer Beiträge einher (ergibt sich aus Intra-Instruktionsvarianten-Lernendenvergleichen; siehe insbesondere Tabelle 5.18 auf S. 179).
- (b) Bei Bearbeitung der *implizit-fachmethodischen Instruktionsvariante* gehen größere zeitliche Anteile fachmethodischer Beiträge mit höheren Kompetenzzuwächsen einher; auch andersherum gehen höhere Kompetenzzuwächse mit größeren zeitlichen Anteilen fachmethodischer Beiträge einher (ergibt sich aus Intra-Instruktionsvarianten-Lernendenvergleichen; siehe insbesondere Unterabschnitt 5.5.3, ab S. 228).
- (c) Sowohl die zeitlichen Anteile fachmethodischer Beiträge der Lernenden als auch die mittleren Kompetenzzuwächse der Lernenden sind bei der explizit-fachmethodischen Instruktionsvariante bedeutsam größer/höher als für die implizit-fachmethodische Instruktionsvariante (ergibt sich aus dem groben Vergleich der Varianten; siehe insbesondere Unterabschnitt 5.5.2.1, ab S. 214).
- (d) Bei ähnlich hohen Kompetenzzuwächsen sind die zeitlichen Anteile fachmethodischer Beiträge für die explizit-fachmethodische Instruktionsvariante deutlich größer als für die implizit-fachmethodische Instruktionsvariante; selbst im Vergleich von Lernenden mit niedrigen Kompetenzzuwächsen bei der explizit-fachmethodischen Instruktionsvariante vs. Lernenden mit hohen Kompetenzzuwächsen bei der implizit-fachmethodischen Instruktionsvariante sind die zeitlichen Anteile fachmethodischer Beiträge für die Lernenden aus der explizit-fachmethodischen Instruktionsvariante deutlich höher (ergibt sich aus dem Vergleich einzelner Personen aus beiden Instruktionsvarianten; siehe insbesondere Unterabschnitt 5.5.3.3, ab S. 232).

Anzumerken ist insbesondere, dass trotz paralleler Formulierung der Aufzählungspunkte (a) und (b) die zugehörigen Ergebnisse nicht in genau derselben Weise generiert wurden; außerdem sind die Ergebnisse zur explizit-fachmethodischen Instruktionsvariante statistisch deutlich besser abgesichert als die Ergebnisse zur implizit-fachmethodischen Instruktionsvariante.¹²⁶ Ferner ist das Fazit auch insofern stark aggregiert, als dass die Unterscheidung verschiedener fachmethodischer Beiträge (fachmethodisch-indiziert und fachmethodisch-vermutet sowie zu den Teilprozessen des experimentbezogenen Denkens und Arbeitens; siehe u. a. Abschnitt 5.1, ab S. 68) nicht aufgegriffen wird, sondern über alle zugehörigen Ergebnisse hinweg zusammengefasst wurde. In den nachfolgenden Diskussionen wird allerdings darauf eingegangen.

Das Ergebnis, dass – trotz grundsätzlich vergleichbarer Gesamt- und Einheitenbearbeitungsdauern der beiden Instruktionsvarianten (siehe Unterabschnitt 5.5.1, ab S. 208)¹²⁷ – der zeitliche Anteil fachmethodischer Beiträge bei der explizit-fachmethodischen Instruktionsvariante deutlich größer ist als bei der implizit-fachmethodischen Instruktionsvariante (siehe Unterabschnitt 5.5.2, ab S. 214), ist erwartungsgemäß, weil in der explizit-fachmethodischen Instruktionsvariante deutlich stärker fachmethodische Beiträge angeregt werden und zu diesen aufgefordert wird (siehe Abschnitt 4.3, ab S. 57). In einem punktuell mit der der Arbeit zugrundeliegenden Studie vergleichbaren Forschungsprojekt, kontrastiert Kalthoff (2019) unter anderem eine explizit-fachmethodische Instruktion mit Informationstexten und Reflexionsaufgaben (die sogenannte Variante *explizit II*) mit einer implizit-fachmethodischen Instruktion mit größerem Anteil an Messwiederholungen (bezeichnet als Variante *implizit*). Wie bereits im von Vorholzer (2016) durchgeführten Vergleich der in der vorliegenden Arbeit genutzten Instruktionsvarianten findet sich auch in dem angeführten Forschungsprojekt, dass bei der explizit-fachmethodischen Instruktionsvariante der über alle Personen der Gesamtstichprobe gemittelte Kompetenzzuwachs größer ist als bei der implizit-fachmethodischen Instruktionsvariante (Kalthoff, 2019; Kalthoff et al., 2018). Vergleichbar mit den Ergebnissen der vorliegenden Arbeit findet Kalthoff (2019) zudem, dass die zeitlichen Anteile fachmethodischer Aktivitäten innerhalb der explizit-fachmethodischen Variante explizit II größer ist als innerhalb der implizit-fachmethodischen Instruktionsvariante.

¹²⁶Methodische Limitationen wurden bei den Ergebnisdarstellungen in den vorherigen Kapiteln diskutiert; auf einige Limitationen wird in den folgenden Unterabschnitten zusammenfassend eingegangen, sofern dies für die jeweils abgeleiteten Schlussfolgerungen oder Hypothesen relevant ist.

¹²⁷Auf die in Kapitel 5 detaillierter untersuchte Vergleichbarkeit der beiden Instruktionsvarianten wird in dem vorliegenden Kapitel der Zusammenführungen aus Gründen der Übersichtlichkeit nicht weiter eingegangen. Eine unter anderen Gesichtspunkten erfolgende Zusammenführung zur Bedeutung der Bearbeitungsdauern findet sich in Unterabschnitt 7.3.1 (ab S. 463; basierend auf der ausführlichen Diskussion in Unterabschnitt 5.5.1.2, ab S. 210).

Wenn, wie durch den Vergleich der Instruktionsvarianten nahegelegt, tatsächlich höhere Kompetenzzuwächse mit größerem zeitlichen Anteil fachmethodischer Beiträge einhergehen sollten, werfen die Ergebnisse der vorliegenden Arbeit im Hinblick auf die Zusammenhänge zu den Kompetenzzuwächsen mindestens die folgenden zwei Fragen auf:

- (1) Weshalb finden sich bei der explizit-fachmethodischen Instruktionsvariante Lernende, die *trotz* ihrer (im Vergleich zur implizit-fachmethodischen Instruktionsvariante) *großen zeitlichen Anteile* fachmethodischer Beiträge *niedrige Kompetenzzuwächse* aufweisen?¹²⁸
- (2) Weshalb finden sich bei der implizit-fachmethodischen Instruktionsvariante Lernende, die *trotz* ihrer (im Vergleich zur explizit-fachmethodischen Instruktionsvariante) *geringen zeitlichen Anteile* fachmethodischer Beiträge *hohe Kompetenzzuwächse* aufweisen?¹²⁹

In den nachfolgenden Unterabschnitten werden anhand der Erörterung dieser beiden Fragen Einordnungen und generelle Deutungen der Ergebnisse der Arbeit zur Bedeutung fachmethodischer Beiträge für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten vorgenommen. Dabei wird unter anderem darauf Bezug genommen, dass vermutlich nicht der zeitlich Umfang fachmethodischer Beiträge alleine, sondern auch weitere Charakteristika – d. h. beispielsweise in der Situation, dem Kontext, der Formulierung oder dem spezifischen Inhalt liegende Merkmale – dieser Beiträge für den Kompetenzaufbau relevant sind (Fazit 2b). Außerdem werden Hypothesen zum Kompetenzaufbau und zu Wirkmechanismen der beiden Instruktionsvarianten formuliert.

¹²⁸Die zeitlichen Anteile fachmethodischer Beiträge sind zwar nicht mehrere Größenordnungen größer als die zeitlichen Anteile anderer Aktivitäten, sie sind aber – vor dem Hintergrund anderer Studien, in denen zeitliche Anteile von auf spezifische Konzepte eines Lernmaterials bezogenen fachlichen Beiträgen untersucht werden (insbesondere Rogge, 2010) – als eher größer als erwartbar einzuschätzen. Daher sind die zeitlichen Anteile auch grundsätzlich (und nicht nur im Vergleich zur implizit-fachmethodischen Instruktionsvariante) als groß einzuschätzen.

¹²⁹Da für nahezu alle Lernenden, die die implizit-fachmethodische Instruktionsvariante bearbeitet haben, die zeitlichen Anteile fachmethodischer Beiträge sehr gering sind, lässt sich die Frage verschärft auch so formulieren: Weshalb finden sich bei der implizit-fachmethodischen Instruktionsvariante überhaupt Lernende mit hohen Kompetenzzuwächsen, obwohl die zeitlichen Anteile fachmethodischer Beiträge so gering sind?

7.2.1 Charakteristika von Äußerungen mit fachmethodischem Gehalt

Die nicht ganz systematischen Befundlagen zum Zusammenhang zwischen den zeitlichen Anteilen fachmethodischer Beiträge und den Kompetenzzuwächsen, die sich innerhalb der explizit-fachmethodischen Instruktionsvariante zeigen, sind möglicherweise auf Inhalte der fachmethodischen Beiträge zurückführbar. Dies kann – wie im Anschluss an den graue Kasten erläutert – als eine mögliche Antwort auf die in Abschnitt 7.2 aufgeworfenen Fragen angesehen werden; in etwas allgemeinerer Form wird es daher als Hypothese, der in weiteren Studien nachgegangen werden kann, postuliert:

Hypothese 2.1

(Vorrangig) Äußerungen, die gewisse Charakteristika (z. B. spezifische Inhalte) haben, sind relevant für den Kompetenzaufbau.

Spezifisch auf die Äußerungen mit fachmethodischem Gehalt bezogen¹³⁰ ist es also der Hypothese folgend relevant, worauf die fachmethodischen Beiträge der Lernenden im Spezifischen gerichtet sind. Es ist beispielsweise denkbar, dass einige Lernende trotz geringem zeitlichen Anteil fachmethodischer Beiträge vor allem solche Beiträge tätigen, die eine hohe Relevanz für die aufzubauende Kompetenz haben, während andere Lernende zwar häufig fachmethodisch beitrugen, die Äußerungen aber beispielsweise wenig zu den anzustrebenden Konzepten passen. Lernende, die *niedrige* Kompetenzzuwächse erreichen und *mäßig große* zeitliche Anteile fachmethodischer Beiträge aufweisen, würden demnach nur fachmethodische Beiträge machen, die für den Kompetenzaufbau *wenig zielführend oder gar hinderlich* sind; Lernende mit *hohen* Kompetenzzuwächsen und *mäßig großen* zeitlichen Anteilen fachmethodischer Beiträge würden nur für den Kompetenzaufbau *relevante* Beiträge machen. Diese Annahme könnte auch plausibilisieren, dass Lernende, die *sehr große* zeitliche Anteile fachmethodischer Beiträge aufweisen, grundsätzlich *hohe* Kompetenzzuwächse erreichen,

¹³⁰Schon die Entscheidung, in der vorliegenden Arbeit vorrangig fachmethodische Beiträge im Kontext von fachinhaltlichen Beiträgen und von Beiträgen zur Vorbereitung und Durchführung von Versuchen zu betrachten (allgemeiner formuliert: die Wahl der betrachteten Kategorien, siehe Abschnitt 5.1), basiert im Prinzip auf einer allgemeinen Version der hier formulierten Hypothese. Es wird beispielsweise davon ausgegangen, dass Beiträge, die auf Gruppenarbeitsorganisation gerichtet sind, zur spezifischen auf fachmethodischen Kompetenzaufbau gerichteten Untersuchung nur einen geringen Beitrag leisten. An dem Beispiel wird deutlich, dass es durchaus Fälle geben kann und wird, in denen genau das Gegenteil der Hypothese zutrifft – aggregiert über viele Äußerungen scheint sich die Hypothese aus den Ergebnissen der vorliegenden Arbeit allerdings gut ableiten zu lassen.

weil in ihren sehr großen zeitlichen Anteilen immer auch für den Kompetenzaufbau *relevante* fachmethodische Beiträge enthalten sind.

Für die explizit-fachmethodische Instruktionsvariante wurden fachmethodische Beiträge, als eine Operationalisierung von Beiträgen mit fachmethodischem Gehalt (siehe Abschnitt 5.1, ab S. 68), hinsichtlich verschiedener Aspekte differenziert. Dazu gehörte das zeitliche Umfeld und der Aufgabenkontext der Beiträge (anhand der Zeitmarken) sowie die Deutlichkeit des fachmethodischen Bezugs (anhand der Subkategorien fachmethodisch-indiziert und fachmethodisch-vermutet, für die Unterscheidung siehe insbesondere Abschnitt 5.1). Auch bei der Untersuchung der Vorstellungen der Lernenden wurden Äußerungen mit fachmethodischem Gehalt differenziert kategorisiert (siehe insbesondere Abschnitt 6.1 und noch spezifischer Unterabschnitt 6.1.3, ab S. 254). Anhand der so aufgespannten Dimensionen wird anhand von Vergleichen von Lernenden mit unterschiedlichen Kompetenzzuwächsen auf solche potentiellen Charakteristika von Äußerungen mit fachmethodischem Gehalt geschlossen, die besonders relevant für den Kompetenzaufbau sein könnten. Das folgende Fazit fasst die zugehörigen Ergebnisse zusammen, die sodann in Bezug auf die ableitbaren Charakteristika diskutiert werden.¹³¹

Fazit 2B

Für die explizit-fachmethodische Instruktionsvariante finden sich im Vergleich von Lernenden mit unterschiedlichen Kompetenzzuwächsen Unterschiede hinsichtlich verschiedener Merkmale der Äußerungen mit fachmethodischem Gehalt:

- (a) Der Zusammenhang des zeitlichen Anteils fachmethodischer Beiträge mit dem Kompetenzzuwachs von Lernenden unterscheidet sich zwischen den Subkategorien zur Deutlichkeit des fachmethodischen Bezugs (fachmethodisch-indiziert vs. -vermutet; siehe Unterabschnitt 5.4.4, ab S. 173). Der Zusammenhang fällt zudem für die Einheiten der Instruktion, die unterschiedliche Teilprozesse des experimentbezogenen Denkens und Arbeitens adressieren, unterschiedlich aus (siehe Unterabschnitt 5.4.4).

¹³¹Für die implizit-fachmethodische Instruktionsvariante wurden ebenfalls die fachmethodischen Beiträge analysiert (während die ausführlichen vorstellungsbezogenen Analysen nicht vorgenommen wurden), sie werden an dieser Stelle allerdings nicht berichtet, da die Ergebnisse (evtl. aufgrund der Beschaffenheit der Stichprobe) unklar bleiben.

- (b) In den zeitlichen Anteilen fachmethodischer Beiträge, die mit Variablenkontrolle in Verbindung gebracht werden können, zeigen sich deutliche positive Unterschiede für Lernende mit hohen vs. niedrigen Kompetenzzuwächsen (siehe Unterabschnitt 5.4.5, ab S. 184). Für andere Inhalte oder Situationen finden sich keine systematischen Zusammenhänge.
- (c) In den Aktivitätsprofilen, die vorrangig Lernenden mit hohen Kompetenzzuwächsen zugeordnet sind, finden sich zumeist auch höhere Anteile spezifischer anderer Aktivitäten als in den Aktivitätsprofilen, die Lernenden mit niedrigen oder durchmischten Kompetenzzuwächsen zugeordnet sind; die jeweils betroffene spezifische Aktivität unterscheidet sich für die analysierten Einheiten der Instruktion (Unterabschnitt 5.4.3, ab S. 163).
- (d) Kompetenzzuwächse von Personen hängen eher mit den Aktivitäten des betrachteten Individuums als den Aktivitäten der zugehörigen Teammitglieder zusammen. Dies gilt insbesondere für den zeitlichen Umfang fachmethodischer Beiträge (siehe Abschnitte 5.4.4, ab S. 173, und 5.4.5, ab S. 184).
- (e) Für Lernende mit hohen Kompetenzzuwächsen ist die Anzahl der für die einzelnen Lernenden stabilen rekonstruierten Vorstellungen größer als für Lernende mit niedrigen Kompetenzzuwächsen (Unterabschnitt 6.10.5, ab S. 408).
- (f) Für Lernende mit hohen Kompetenzzuwächsen ist die Angemessenheit der für die einzelnen Lernenden rekonstruierten Vorstellungen größer als für Lernende mit niedrigen Kompetenzzuwächsen (Unterabschnitt 6.10.5).
- (g) Für Lernende mit hohen Kompetenzzuwächsen ist die Allgemeinheit/Generalisiertheit der für die einzelnen Lernenden rekonstruierten Vorstellungen größer als für Lernende mit niedrigen Kompetenzzuwächsen (Unterabschnitt 6.10.5).

Generell ist zu den aufgeführten Ergebnissen anzumerken, dass sie weniger gut abgesichert sind als die in Fazit 1 und Fazit 2A aufgeführten Ergebnisse. Insbesondere basieren sie nur auf Teilmengen der dort jeweils berücksichtigen Äußerungen. Dennoch können die Ergebnisse andeuten, dass eine genauere Untersuchung von Äußerungen mit fachmethodischem Gehalt hinsichtlich der benannten Merkmale zielführend für das Identifizieren von Charakteristika der für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten in besonderer Weise relevanten Äußerungen mit fachmethodischem Gehalt ist:

Bedeutung des zeitlichen Anteils fachmethodischer Beiträge differenziert nach Deutlichkeit des Bezug zu Fachmethoden und nach bearbeiteten Instruktionseinheiten (Fazit 2B,a). Im Rahmen der Arbeit wurden fachmethodische Beiträge entlang der Deutlichkeit des Bezugs zu fachmethodischen Konzepten differenziert: fachmethodisch-indizierte Beiträge beziehen sich auf Äußerungen, in denen durch Worte oder Begründungselemente vergleichsweise deutlich erkennbar ist, dass ein Bezug zu Fachmethoden (bspw. zu fachmethodischen Regeln oder Zuweisungen von unabhängigen/abhängigen Variablen) vorliegt; fachmethodisch-vermutete Beiträge beziehen sich auf Äußerungen, bei denen aus der Art oder dem unmittelbaren Kontext der Äußerung geschlossen werden kann, dass fachmethodische Überlegungen angestellt wurden (ausführlich in Abschnitt 5.1, ab S.68). Als Ergebnisse finden sich mit Bezug zu dieser Unterscheidung (siehe Unterabschnitt 5.4.4, ab S.173):

- In der ersten Einheit, die auf den Teilprozess des Formulierens von Fragen und Hypothesen bezogen ist, findet sich ein stärkerer Zusammenhang von Kompetenzzuwächsen mit dem zeitlichen Anteil fachmethodisch-indizierter Beiträge als mit dem zeitlichen Anteil fachmethodisch-vermuteter Beiträge.
- Für die dritte Einheit, die auf den Teilprozess des Auswertens und Interpretierens bezogen ist, findet sich ein stärkerer Zusammenhang von Kompetenzzuwächsen mit dem zeitlichen Anteil fachmethodisch-vermuteter Beiträge als mit dem zeitlichen Anteil fachmethodisch-indizierter Beiträge.
- Alle Zusammenhänge (fachmethodisch-indiziert, -vermutet; alle drei Einheiten) sind (zumindest in der Tendenz) positiv.

Dass sich in Einheit 1 vorrangig für fachmethodisch-indizierte, in Einheit 3 vorrangig für fachmethodisch-vermutete Beiträge Zusammenhänge mit den Kompetenzzuwächsen der Lernenden zeigen, lässt sich auf verschiedene Weise deuten:

- Die Lernenden haben bereits hoch entwickelte Kompetenzen zu den in Einheit 1 adressierten Teilprozessen des Formulierens von Fragen und Hypothesen und können daher die Regeln explizieren (fachmethodisch-indizierte Beiträge); sie haben aber nur gering ausgeprägte Kompetenzen zu den Teilprozessen des Auswertens und Interpretierens und müssen daher sehr stark intuitiv vorgehen (vgl. Studien zum fachinhaltlichen Lernen bspw. Rogge, 2010; Saniter, 2003), was sich in fachmethodisch-vermuteten Beiträgen widerspiegelt. Dass sich bei Aufteilung der Einheiten nur für die dritte Einheit (zu den Teilprozessen des

Auswertens und Interpretierens) signifikante Unterschiede in den Kompetenzzuwächsen für die explizit-fachmethodische vs. die implizit-fachmethodische Instruktionsvariante ergeben (siehe Abschnitt 4.5, ab S. 63; Details in Vorholzer, 2016), könnte als ein Hinweis auf große Potentiale für Kompetenzzuwachs bezüglich dieses Teilprozesses gedeutet werden. Befunde zu großen Schwierigkeiten von Lernenden insbesondere bei den Teilprozessen des Formulierens von Fragen und Hypothesen (z. B. Grube, 2011; teilweise auch für höhere Klassenstufen wie die hier untersuchte elfte Klassenstufe) sowie niedrige Werte im Prä-Test zum experimentbezogenen Denken und Arbeiten (auch wenn die Betrachtung auf die zu den Teilprozessen des Formulierens von Fragen und Hypothesen gehörigen Aufgaben eingeschränkt wird; Vorholzer, 2016, Abschnitt 10.2, Ab S. 122) deuten allerdings darauf hin, dass die untersuchten Lernende nicht mit bereits hoch entwickelten Kompetenzen in die Bearbeitung der Instruktion starten.

- Während erfolgreicher Kompetenzaufbau bei Konzepten zu den Teilprozessen des Formulierens von Fragen und Hypothesen stärker von Verbalisierungen der Regeln und Entscheidungen begleitet wird, ist für erfolgreichen Kompetenzaufbau bei Konzepten zu den Teilprozessen des Auswertens und Interpretierens von geringerer Relevanz, ob Lernende Regeln explizieren. Ein Grund dafür könnte sein, dass in den Teilprozessen des Auswertens und Interpretierens intuitives Handeln ausreicht, weil auch fachinhaltliche Kompetenzen genutzt werden, beispielsweise indem zum Einschätzen der Plausibilität von Deutungen erzielte Ergebnisse mit der Realität abgeglichen werden (für ähnliche Überlegungen bspw. Zimmerman, 2007, S. 187f.). Ein anderer Grund dafür, das für die Einheiten unterschiedlich große zeitliche Anteile auf das Verbalisieren von Regeln vorliegen, könnte in der Anlage der explizit-fachmethodischen Instruktion liegen (Betrachtungen des Instruktionsmaterials legen nahe, dass fachmethodische Beiträge in Einheit 1 stärker bzw. häufiger angeregt werden als in Einheit 3); allerdings bleibt bei dem aus instruktionaler Sicht genannten Grund offen, warum das Explizieren von Konzepten durch die Lernenden für Einheit 1 in einem größeren Zusammenhang mit den Kompetenzzuwächsen steht als für Einheit 3.

Vor dem Hintergrund der Ergebnisse der Arbeit lässt sich nur schwer entscheiden, welche der beiden Deutungen die größere Plausibilität aufweist. Dies bleibt im Rahmen der Arbeit ein Desiderat. Mit der Entscheidung oder Formulierung alternativer

Erklärungen geht allerdings einher, ob, inwiefern und in welchen Fällen die Deutlichkeit des fachmethodischen Bezugs als ein für den Kompetenzaufbau relevantes Charakteristikum angesehen werden kann und sollte.

Bedeutung des zeitlichen Anteils fachmethodischer Beiträge mit Bezug zur Variablenkontrolle (Fazit 2B,b). Das Ergebnis, dass sich für Lernende mit höheren Kompetenzzuwächsen größere zeitlichen Anteile von fachmethodischen Beiträgen, die auf Variablenkontrolle bezogen sind, vorfinden als für Lernende mit niedrigen Kompetenzzuwächsen, basiert auf in den Abschnitten 4.5 (ab S. 63) und 5.4.6 (ab S. 190) berichteten Analysen, in denen unter anderem Inhalte von Karten der Instruktion mit bedeutsamen Unterschieden zwischen Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen untersucht werden. Dieses Ergebnis könnte beispielsweise so gedeutet werden, dass die Beschäftigung mit der Variablenkontrolle den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten insgesamt fördert. Die Variablenkontrollstrategie hat eine zentrale Stellung für das experimentbezogene Denken und Arbeiten. Beispielsweise ist es – wodurch zur Deutung passende Befunde von Hofstein et al., 2005 ggf. erklärt werden könnten – möglicherweise auch für das Formulieren von naturwissenschaftlichen Fragestellungen relevant, abhängige und unabhängige Variablen zu identifizieren (u. a., weil bei »Hat X einen Einfluss auf Y?« die Rollen von X als unabhängige und Y als abhängige Variable zumindest intuitiv für die Formulierung herangezogen werden können). In der untersuchten explizit-fachmethodischen Instruktionsvariante wird sogar ein Zusammenhang zwischen Begriffen der Variablenkontrolle und dem Formulieren von Fragen erfragt (Karte 2-52, im Anhang in Abbildung G.2 auf S. 594) und erläutert (Karte 2-52, im Anhang in Abbildung G.2 auf S. 594). Auch für das Auswerten und Interpretieren von Daten lässt sich eine Bedeutung der Variablenkontrolle identifizieren. Kompetenzen zum experimentbezogenen Denken und Arbeiten würden demnach besonders stark dadurch gefördert, dass vermehrt und länger über Variablenkontrolle gesprochen und nachgedacht wird, weil dies auch positiven Einfluss auf andere Facetten von Kompetenzen zum experimentbezogenen Denken und Arbeiten hat. Allerdings sind vielfältige alternative Deutungen denkbar (etwa, dass die Aufgaben des EDAW-Tests, der in den Prä- und Post-Erhebungen eingesetzt wird, sich gut durch Variablenkontrolle lösen lassen, obwohl dies eventuell in anderen Testsituationen, wie beispielsweise Realsituationen, nicht grundsätzlich der Fall wäre).

Bei der Deutung des Ergebnisses ist ferner zu bedenken, dass das methodische Vorgehen zur Identifikation der inhaltlichen Bezüge der fachmethodischen Beiträge nur eine grobe Näherung ermöglicht, weil nicht der tatsächliche Inhalt der Äußerung

untersucht wurde, sondern nur vermerkt wurde, bei welcher Karte die Äußerung auftrat. Aus der Kartenummer lässt sich entnehmen, welche Aufgaben bearbeitet wurden; anhand der Aufgaben kann wiederum darauf geschlossen werden, welche fachmethodischen Inhalte bei der explizit-fachmethodischen Instruktion im Fokus stehen. Es ist durchaus möglich, dass Lernende bei der Aufforderung, zwischen präzisen und allgemeinen naturwissenschaftlichen Fragen zu unterscheiden (siehe Karten aus der Instruktion im Anhang in Abbildung G.1 auf S. 591), über Variablenkontrolle sprechen. Umgekehrt ist es auch möglich, dass Lernende bei einer Aufforderung zur Zuweisung von abhängigen und unabhängigen Variablen andere fachmethodische Beiträge machen. Zumindest die Feststellung, dass Lernende mit hohen Kompetenzzuwächsen in Kontexten, in denen von der Instruktion auf Variablenkontrolle fokussiert wird, zeitlich umfangreicher fachmethodisch beitragen, bleibt allerdings bestehen. Es ist demnach zwar naheliegend, aber ohne zusätzliche detailliert auf Inhalte ausgerichtete Analysen nicht eindeutig, dass Äußerungen mit fachmethodischem Gehalt, die sich auf Variablenkontrolle beziehen von besonderer Bedeutung für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten sind.

Ergänzend zu den Ergebnissen für die explizit-fachmethodische Instruktionsvariante lässt sich anmerken, dass bei der implizit-fachmethodischen Instruktionsvariante für Beiträge, die sich auf die Variablenkontrolle beziehen, keine Unterschiede zwischen Lernenden mit unterschiedlichen Kompetenzzuwächsen gefunden werden (Unterabschnitt 5.5.4, ab S. 237). Konträr zur obigen Argumentation, gemäß welcher die Äußerungen mit Bezug zur Variablenkontrolle besonders relevant für den Kompetenzaufbau sein dürften, könnte es daher auch sein, dass derartige Beiträge in keinem direkten Zusammenhang mit dem Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten stehen: Bei der implizit-fachmethodischen Instruktionsvariante finden sich eventuell deshalb keine unterschiedlichen Umfänge fachmethodischer Beiträge mit Bezug zur Variablenkontrolle für Lernende mit unterschiedlichen Kompetenzzuwächsen, weil die Instruktion nicht zu fachmethodischen Beiträgen anregt; bei der explizit-fachmethodischen Instruktionsvariante finden sich hingegen unterschiedliche Umfänge, weil Lernende, die bei Anregung zu fachmethodischen Beiträgen besonders umfangreich antworten, zu größeren Kompetenzzuwächsen gelangen als Lernende, die weniger umfangreich antworten, und weil dies nur für die zeitlich dominanten Beiträge mit Bezug zur Variablenkontrolle deutlich wird (etwa, weil hierzu die Lernenden besonders gut etwas sagen können). Demnach wäre es nicht zuerst ein Charakteristikum von für den Kompetenzaufbau besonders relevanten Äußerungen, dass sie auf Variablenkontrolle bezogen sind, sondern es ginge vielmehr

darum, wie stark und umfangreich Lernende auf die Anregung durch die Instruktion reagieren.

Ausblicksartig lässt sich darauf verweisen, dass in der vorliegenden Arbeit nur der Gesamtkompetenzzuwachs betrachtet wurde, der die Teilprozesse des experimentbezogenen Denkens und Arbeitens zusammenfasst. Das Testinstrument erlaubt eine gesonderte Auflösung der Teilprozesse (etwa im Sinne von Teilkompetenzen), die genauere Hinweise darauf liefern könnte, inwiefern Äußerungen, die auf Variablenkontrolle bezogen sind, mit den Kompetenzen zum experimentbezogenen Denken und Arbeiten (und ggf. den einzelnen Teilkompetenzen) in Verbindung stehen. Eventuell sind fachmethodische Beiträge mit Bezug zur Variablenkontrolle für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten gerade deshalb besonders relevant, weil ihr zeitlicher Umfang im Zusammenhang mit dem erreichten Fähigkeitszuwachs zum Teilprozess des Planens und Durchführens von Untersuchungen in Zusammenhang steht. Das wäre sehr naheliegend, könnte allerdings nicht erklären, dass die Unterschiede nur für die explizit-fachmethodische Instruktionsvariante deutlich werden, nicht aber für die implizit-fachmethodische Instruktionsvariante, weil sich die beiden Instruktionsvarianten bezüglich der Subskala zum Teilprozess des Planens von Untersuchungen nur geringfügig (und nicht signifikant) unterscheiden (siehe Vorholzer, 2016, Abschnitt 10.2, ab S. 122). Eventuell sind fachmethodische Beiträge mit Bezug zur Variablenkontrolle für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten gerade deshalb besonders relevant, weil ihr zeitlicher Umfang im Zusammenhang mit dem erreichten Fähigkeitszuwachs zum Teilprozess des Auswertens und Interpretierens in Zusammenhang steht, bei dem durch mangelndes Verständnis der Variablenkontrolle leicht Fehlentscheidungen getroffen werden können. Das ist zwar weniger naheliegend, könnte aber erklären, dass die Unterschiede nur für die explizit-fachmethodische Instruktionsvariante deutlich werden, da der mittlere Kompetenzzuwachs der Stichprobe bezüglich dieses Teilprozesses bei Bearbeitung der implizit-fachmethodischen Instruktionsvariante geringer als für die anderen Teilprozesse ist und der mittlere Kompetenzzuwachs der vergleichbaren Stichprobe zur explizit-fachmethodischen Instruktionsvariante deutlich größer ist (siehe Vorholzer, 2016, S. 123–124).¹³²

¹³² Allerdings wäre der geschilderte potentielle Zusammenhang zwischen fachmethodischen Beiträgen mit Bezug zur Variablenkontrolle und den verschiedenen Zuwächsen für die Kompetenzteilskalen nicht die wahrscheinlichste Erklärung für die Unterschiede. Auf eine plausiblere Erklärung, die Vorholzer (2016) vermutet hat, wird in Unterabschnitt 7.3.3, ab S. 467 eingegangen.

Bedeutung des zeitlichen Anteils fachmethodischer Beiträge vor dem Hintergrund des zeitlichen Anteils anderer Aktivitäten (Fazit 2B,c). In Unterabschnitt 5.4.3 (ab S. 163) wird für die drei Einheiten der explizit-fachmethodischen Instruktionsvariante diskutiert, inwiefern die Kompetenzzuwächse der Lernenden sich als in Zusammenhang mit den gefundenen Aktivitätsprofilen stehend sehen lassen. Ein Aktivitätsprofil stellt dabei die Gesamtheit der zeitlichen Anteile der untersuchten Aktivitäten dar (symbolisch zusammengefasst in einem Balkendiagramm). Während beim Suchen nach Charakteristika fachmethodischer Beiträge zum einen darauf geschaut werden kann, wodurch ein einzelner fachmethodischer Beitrag gekennzeichnet ist (wie direkt der Bezug zu Fachmethoden ist, ob er einen Bezug zur Variablenkontrolle hat usw.), kann zum anderen auch anhand generell aufdeckbarer Kovariationen mit anderen Aktivitäten auf mögliche Charakteristika fachmethodischer Beiträge geschlossen werden, wie nachfolgend skizziert.

Als Ergebnis für die erste Einheit findet sich, dass für die zeitlichen Anteile des nonverbalen Vorbereitens und Durchführens von Versuchen keine systematischen Unterschiede zwischen Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen vorliegen. Allerdings stellen sich zwei interessante Befunde ein: Erstens wird für einen Aktivitätsprofiltyp (E1-P-P1-C3) ein insgesamt recht geringer zeitlicher Anteil fachmethodischer Beiträge und ein sehr großer zeitlicher Anteil nonverbaler Aktivitäten des Vorbereitens und Durchführens festgestellt. Dieser Aktivitätsprofiltyp geht mit niedrigen Kompetenzzuwächsen einher. Zweitens wird für einen Aktivitätsprofiltyp (E1-P-P1-C1) ein großer zeitlicher Anteil fachmethodisch-indizierter Beiträge festgestellt. Auf den großen Anteil fachmethodischer Beiträge scheint es allerdings nicht zurückzuführen sein, dass mit diesem Aktivitätsprofiltyp hohe Kompetenzzuwächse einhergehen, da auch für andere Aktivitätsprofiltypen, bei denen die Kompetenzzuwächse durchmischt oder niedrig ausfallen, vergleichbare Anteile fachmethodisch-indizierter Beiträge auffindbar sind. Allerdings findet sich für den zweiten hier diskutierten Aktivitätsprofiltyp zusätzlich zum großen Anteil fachmethodisch-indizierter Beiträge ein zeitlich großer Anteil nonverbaler Aktivitäten des nonverbalen Vorbereitens und Durchführens von Versuchen.

Das Ergebnis, dass erst das gemeinsame Auftreten großer zeitlicher Anteile sowohl fachmethodischer Beiträge als auch nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen mit hohen Kompetenzzuwächsen in Zusammenhang zu stehen scheint, könnte darauf hindeuten, dass es zumindest innerhalb der ersten Einheit der explizit-fachmethodischen Instruktionsvariante von Bedeutung für den Kompetenzaufbau ist, die fachmethodischen Inhalte mit den durchgeführten

Versuchen zu verknüpfen. Die durchgeführten Analysen geben zwar keine Auskunft darüber, ob Lernende real Verknüpfungen hergestellt haben, aber es scheint so, dass gerade solche Lernenden, die sowohl genug praktische Erfahrungen machen als auch hinreichend über Fachmethoden sprechen, zu hohen Kompetenzzuwächsen gelangen. Zukünftige Forschung könnte zeigen, ob und inwiefern Verknüpfungen von experimentellem Handeln und fachmethodischen Inhalten hergestellt werden – was in Bezug auf Einheit 1 insbesondere deshalb interessant wäre, weil für das Formulieren von Fragen und Hypothesen zwar Forschungsbefunde existieren, die nahelegen, dass Versuchsdurchführungen förderlich sind (Hofstein et al., 2005), die Beiträge, die Erfahrungen des Vorbereitens und Durchführens zur Entwicklung der Fähigkeit des Formulierens von Fragen und Hypothesen leisten, allerdings aus theoretischer Sicht nicht optimal nachvollziehbar sind (geschweige denn bisher empirisch nachvollzogen wurden).

Als eines der Ergebnisse für die zweite Einheit scheint ein relativ zu zeitlichen Anteilen anderer Aktivitäten großer zeitlicher Anteil fachmethodischer Beiträge nur dann positiv mit Kompetenzzuwächsen zusammenzuhängen, wenn der zeitliche Anteil fachinhaltlicher Beiträge relativ zu anderen Aktivitätsprofiltypen groß ist. Es scheint also ein Mindestmaß an zeitlichem Anteil fachinhaltlicher Beiträge zu geben, ab dem der zeitliche Umfang fachmethodischer Beiträge relevant für den Kompetenzzuwachs ist.

Dieses Ergebnis ist in mindestens zweierlei Hinsicht interessant. Zum einen könnte es ein Charakteristikum fachmethodischer Beiträge sein, wie sie in fachinhaltliche Äußerungen eingebettet sind. Es könnte etwa einen Unterschied machen, ob eine fachinhaltliche Beobachtung zu den Versuchsgegenständen gemacht wird und ein fachmethodischer Beitrag zur Variablenkontrolle oder ob nur ein fachmethodischer Beitrag zur Variablenkontrolle gemacht wird.¹³³ Zum anderen könnte es sein, dass sich im zeitlichen Anteil fachinhaltlicher Beiträge vorhandene fachinhaltliche Kompetenzen widerspiegeln¹³⁴; obwohl also fachinhaltliche Kompetenzen grundsätzlich nur einen geringen Einfluss auf die erreichten Kompetenzzuwächse haben (wie bei Vorholzer, 2016 für die hier genutzten Instruktionsvarianten gezeigt), könnten daher im Prozess – zumindest für die zweite Einheit der explizit-fachmethodischen

¹³³Es ist nicht Gegenstand dieses Abschnitts, aber es wäre ferner auch spannend, ob es für den Kompetenzaufbau einen Unterschied macht oder im Prozess des Kompetenzaufbaus einen Unterschied darin gibt, ob *zuerst* die fachinhaltliche Beobachtung zu den Versuchsgegenständen gemacht wird und *dann* ein fachmethodischer Beitrag zur Variablenkontrolle oder ob dies andersherum erfolgt. Diesen und ähnlichen Fragen muss allerdings in einem anderem Rahmen als der vorliegenden Arbeit nachgegangen werden.

¹³⁴Es könnten sich in den größeren zeitlichen Anteilen fachinhaltlicher Beiträge prinzipiell auch fehlende fachinhaltliche Kompetenzen widerspiegeln, z. B., weil fachinhaltlicher Klärungsbedarf besteht, für den mehr Zeit eingesetzt werden muss. Aufgrund dessen, dass die Instruktion so konzipiert ist, dass die fachinhaltlichen Gegenstände und Konzepte den Lernenden grundsätzlich vertraut sein sollten (siehe Kapitel 4), ist allerdings eher die Vermutung im Text naheliegend.

Instruktionsvariante, die auf das Planen von naturwissenschaftlichen Untersuchungen bezogen ist – Einflüsse vorliegen. Hieran zeigt sich u. a., dass ein Ausweiten der vorgenommenen Analysen zu den Aktivitäten durch ein Differenzieren von Lernenden nach Eingangsvoraussetzungen zwar grundsätzlich spannend sein könnte, allerdings genauere inhaltliche Analysen zu den fachmethodischen *und* den fachinhaltlichen Beiträgen *und* deren inhaltlichen Zusammenhängen im Prozess der Instruktionsbearbeitung von großem Mehrwert für die Forschung zum Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten wären.

Bedeutung der Eigenständigkeit der Äußerungen mit fachmethodischem Gehalt (Fazit 2B,d). Ein stärkerer Zusammenhang der Kompetenzzuwächse von Personen mit den Aktivitäten des Individuums im Vergleich zu den Aktivitäten des gesamten Teams zeigt sich in verschiedenen Analysen: Zum einen finden sich bei den Analysen von Aktivitätsprofilen für Personen deutlichere Zusammenhänge mit Kompetenzzuwächsen als für Teams (Unterabschnitt 5.4.3, ab S. 163). Zum anderen finden sich in den zeitlichen Anteilen fachmethodischer Beiträge für Personen mit hohen vs. niedrigen Kompetenzzuwächsen deutlichere Unterschiede als für Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen (für eine Diskussion dieser Ergebnisse siehe Unterabschnitt 5.4.4, ab S. 173). Auch in Bezug auf die in Kapitel 6 untersuchten Vorstellungen der Lernenden finden sich vorrangig Hinweise darauf, dass die Ergebnisse auf Teamebene weniger deutlich sind als die auf Ebene der Individuen. In Übereinstimmung mit theoretischen Annahmen zum *individuellen* Lernen von Lernenden innerhalb der materialen und sozialen Lernumgebung (z. B. von Aufschnaiter, 2003, contra *shared cognition*) dürften also für den Kompetenzaufbau besonders relevante Äußerungen solche sein, die von dem Individuum, dessen Kompetenzaufbau untersucht wird, selbst geäußert wurden.

Bedeutung der Stabilität, Angemessenheit und Allgemeinheit der Vorstellungen der Lernenden (Fazit 2B,e+f+g). Die Analyse der Vorstellungen der Lernenden wurde in einer überblicksartigen Weise auf die Kompetenzzuwächse der Lernenden bezogen (siehe Abschnitt 6.10, ab S. 397). Die zugehörigen Ergebnisse basieren anders als die zuvor dargestellten Ergebnisse nur auf den Videoausschnitten zu zwei vergleichsweise kurzen Instruktionsextrakten und nicht auf dem gesamten Instruktionsmaterial. Dennoch ergeben sich erste Hinweise darauf, dass es für den Kompetenzaufbau von Prä- zu Post-Erhebung relevant ist, dass die Lernenden ihre Vorstellungen stabil anbringen, grundsätzlich eher angemessene Vorstellungen anbringen und – falls vom Material gefordert oder benötigt – ihre Vorstellungen verallgemeinert, statt nur fallspezifisch anbringen können.

In Unterabschnitt 6.10.4 (ab S. 406) wird dafür argumentiert, dass sich die Zusammenhänge der Kompetenzzuwächse und der untersuchten Merkmale der Äußerungen mit fachmethodischem Gehalt, aus denen auf die Vorstellungen der Lernenden geschlossen wurde, vermutlich gerade aus dem Grund ergeben, dass sie generelle Charakteristika des verbalen Beitrags der entsprechenden Lernenden abbilden, weil aufgrund der Kürze der untersuchten Instruktionsextrakte andere Zusammenhänge nicht gut denkbar sind. Allerdings müssten weitere Analysen zeigen, ob sich die Äußerungen mit fachmethodischem Gehalt, die einzelne Lernende in weiteren verschiedenen Instruktionsextrakten zu verschiedenen fachmethodischen Konzepten tatsächlich in ähnlicher Weise charakterisieren lassen, da hierzu basierend auf den beiden untersuchten Instruktionsextrakten nicht genügend Hinweise generiert werden können.

Die Ausführungen zu Fazit 2B legen in ihrer Gesamtheit nahe, dass einige Charakteristika der fachmethodischen Beiträge, die für Lernende mit hohen Kompetenzzuwächsen zeitlich umfangreicher vorliegen, zur Wirksamkeit der Instruktion und zum Kompetenzaufbau der Lernenden beitragen. In weiteren Studien könnten einzelne Charakteristika spezifisch untersucht werden, indem beispielsweise der Kompetenzaufbau von Lernenden detailliert verglichen wird, die sich hinsichtlich eines in Fazit 2B genannten Merkmals (z. B. des zeitlichen Anteils von Beiträgen mit Bezug zur Variablenkontrolle) unterscheiden. Auch ein instruktionales Anregen von fachmethodischen Beiträgen mit spezifischen Charakteristika (etwa durch häufiges instruktionales Fragen nach Generalisierungen/Verallgemeinerungen) könnte in weiteren Studien genauer untersucht werden. Es ist nicht davon auszugehen, dass für alle aufgefundenen Charakteristika ein größerer zeitlicher Anteil zugehöriger fachmethodischer Beiträge grundsätzlich mit höheren Kompetenzzuwächsen einhergeht – beispielsweise dürfte es auch relevant sein, wann und in welchem instruktionalem Kontext generalisierende fachmethodische Beiträge auftreten. Wenn Lernende an Stellen der Instruktion keine fachmethodischen Beiträge machen, obwohl diese klar erforderlich sind, kann beispielsweise davon ausgegangen werden, dass dies abträglich für den Kompetenzzuwachs sein dürfte, egal wie viel sie dies durch fachmethodische Beiträge an anderen Stellen zeitlich ausgleichen. Hieran schließen sich auch potentielle Fragen an, die darauf bezogen sind, wie gut die Aktivitäten der Lernenden zu den in der Instruktion intendierten Aktivitäten passen (im Sinne einer Kohärenz der Aufgabenbearbeitung, siehe Neumann, 2020) und zwar nicht nur gemittelt über die gesamte Bearbeitungsdauer, sondern auch je spezifisch für die einzelnen Aufgaben.

7.2.2 Zusammenhang von Fähigkeiten und Verständnissen

Die, in den Fragen auf Seite 434 kondensierten, nicht ganz systematischen Befundlagen zum Zusammenhang zwischen den zeitlichen Anteilen fachmethodischer Beiträge und den Kompetenzzuwächsen könnten – verbunden mit den sehr geringen Anteilen fachmethodischer Beiträge bei der implizit-fachmethodischen Instruktionsvariante – als ein Hinweis darauf verstanden werden, dass fachmethodische Beiträge in anderem und geringerem Zusammenhang mit Fähigkeitsaufbau stehen als mit Vorstellungsaufbau:

Hypothese 2.2

Äußerungen mit fachmethodischem Gehalt sind besonders relevant für den Vorstellungsaufbau, aber nicht in gleicher Weise für den in den Prä- und Post-Erhebungen gemessenen Fähigkeitsaufbau.

Aufbauend auf eine von Zohar und David (2008) vorgeschlagene Unterscheidung von »the ›linguistic component‹ and the ›practical experiences‹« (S. 79) der dort adressierten Variablenkontrolle vermuten Lorch et al. (2010), dass sowohl explizites Adressieren als auch experimentelle Exploration »may be fundamental« (S. 91) für den Kompetenzaufbau zur Variablenkontrolle. Sie fassen die Hypothesen von Zohar und David (2008) zusammen

They hypothesized that the linguistic component of their intervention (i.e., explicit instruction based on verbal discussion) is critical to the development of metastrategic knowledge associated with CVS understanding. In addition, they hypothesized that opportunities to apply CVS in experiments are critical to the acquisition of the thinking strategy because the experience makes concrete what would otherwise remain abstract knowledge. (Lorch et al., 2010, S. 99)

und schließen mit zusätzlichen Stützungen aus anderen Studien. In der Diskussion schreiben sie, dass ihre Ergebnisse »demonstrate that explicit instruction and experimentation make distinguishable contributions to students' understanding of CVS« (S. 99). Lorch et al. (2010) führen zwar aus, dass die unterscheidbaren Beiträge einerseits das Konstanthalten der Kontrollvariablen (explizit-fachmethodische Instruktionsvariante), andererseits das Variieren der unabhängigen Variable (implizit-fachmethodische Instruktionsvariante) betreffen; allerdings scheint diese konkrete Zuschreibung mit dem genutzten Studiendesign zusammenzuhängen. Für die vorliegende Arbeit könnten die unterscheidbaren Beiträge der impliziten Adressierung

durch fachmethodisch korrekt angelegte Versuchsanleitungen und Anweisungen und Hinweise, die beiden Instruktionsvariante gemein ist (die Lernenden explorieren und experimentieren), und der expliziten Adressierung durch zeitlich strukturierte auf fachmethodische Konzepte fokussierende Aufgabenserien, die nur für die explizit-fachmethodische Instruktionsvariante vorgenommen wird (die Lernenden werden explizit instruiert), darin bestehen, dass in beiden Fällen die in den Testinstrumenten geprüften Fähigkeiten gefördert werden, aber nur für die explizit-fachmethodische Instruktionsvariante auch die Verständnisse, vergleichbar mit der »linguistic component« aus dem obigen Zitat, gefördert werden.

Nachfolgend wird zuerst erläutert, dass es plausibel ist, davon auszugehen, dass der Aufbau von Verständnissen zum experimentbezogenen Denken und Arbeiten sich wesentlich zwischen den beiden Instruktionsvarianten unterscheidet und nur beschränkt mit dem gemessenen Kompetenzaufbau in Zusammenhang steht (Unterunterabschnitt 7.2.2.1). Anschließend wird ein Ergebnis zur explizit-fachmethodischen Instruktionsvariante berichtet, welches nahelegt, dass Vorstellungen und Fähigkeiten bereits während der Bearbeitung der Instruktion nur in einem geringen Zusammenhang stehen (Unterunterabschnitt 7.2.2.2, ab S.449); dies lässt sich als Hinweis darauf verstehen, dass die in der Hypothese 2.2 vorgenommene Unterscheidung von Fähigkeits- und Verständnisaufbau angemessen ist. Den potentiell unterschiedlichen Zusammenhängen von fachmethodischen Beiträgen mit Fähigkeits- und Vorstellungsaufbau wurde allerdings in der vorliegenden Arbeit nicht nachgegangen; die Hypothese 2.2 ist also, wie alle in der vorliegenden Arbeit abgeleiteten Hypothesen, mit einem Desiderat verbunden.

7.2.2.1 Aufbau von Verständnissen bei den beiden Instruktionsvarianten

Die eingesetzten Instruktionsvarianten unterscheiden sich wesentlich durch die Explizitheit der Adressierung fachmethodischer Konzepte. Weil also die Konzepte zum experimentbezogenen Denken und Arbeiten, zu denen die Lernenden Verständnisse aufbauen sollen, in der explizit-fachmethodischen Instruktionsvariante deutlich stärker im Fokus stehen, liegt es nahe, davon auszugehen, dass die Verständnisse auch eher in der intendierten Weise aufgebaut werden. In der Forschungsliteratur wird zwar zum Teil darauf hingewiesen, dass bei explizit-fachmethodische Instruktion die Möglichkeit zum intensiven und umfangreichen Aufbau von Verständnissen durch die klare Kommunikation der angestrebten Konzepte beschränkt wird (z. B. Kuhn & Dean, 2005); für die Frage danach, bei welcher Instruktionsvariante der Verständnisaufbau stärker in instruktional intendierter Weise erfolgt, ergibt sich jedoch

7.2 Bedeutung von Äußerungen mit fachmethod. Gehalt für den Kompetenzaufbau

grundsätzlich die Antwort, dass dies plausiblerweise bei der explizit-fachmethodischen Instruktionsvariante geschehen dürfte (und zumeist auch geschieht, vgl. Kapitel 2, z. B. Lazonder & Egberink, 2014).

Das genutzte Testinstrument für Kompetenzen zum experimentbezogenen Denken und Arbeiten (Vorholzer et al., 2016, EDAW-Test) setzt Aufgaben ein, die die *Fähigkeiten* der Lernenden zum experimentbezogenen Denken und Arbeiten testen; es ist allerdings kein Vorstellungstest, mit dem sinnvoll auf real existierenden *Vorstellungen* der Lernenden zu Konzepten des experimentbezogenen Denkens und Arbeitens geschlossen werden kann. Es könnte daher sein, dass Lernende in der implizit-fachmethodischen Instruktionsvariante höchstens Fähigkeiten aufbauen (und daher manche Lernende hohe Kompetenzzuwächse von Prä- zu Posttest erreichen), in der explizit-fachmethodischen Instruktionsvariante hingegen sowohl Fähigkeiten als auch Vorstellungen. Dies würde sich dann darin widerspiegeln, dass sich bei der explizit-fachmethodischen Instruktionsvariante deutlich größere zeitliche Anteile fachmethodischer Beiträge finden, die entlang der Argumentationslinie vor allem für den Vorstellungsaufbau relevant sein dürften, als bei der implizit-fachmethodischen Instruktionsvariante. Die im Mittel über alle Lernenden höheren Kompetenzzuwächse für die explizit-fachmethodische Instruktionsvariante (sowie die für einige einzelne Lernende aus der explizit-fachmethodischen Instruktionsvariante auftretenden Kompetenzzuwächse, die sehr viel größer sind als alle Kompetenzzuwächse von Lernenden bei der implizit-fachmethodischen Instruktionsvariante) könnten dann damit zusammenhängen, dass aufgebaute Vorstellungen letztlich positive Auswirkung auf Fähigkeiten haben. Insofern läge für den gemessenen Fähigkeitsaufbau auch ein indirekter Zusammenhang mit dem zeitlichen Umfang fachmethodischer Beiträge vor, obwohl dieser zeitliche Umfang direkt nur in Zusammenhang mit dem Vorstellungsaufbau steht. Offen bliebe bei einem derartigen Plausibilisierungsversuch allerdings, warum für die explizit-fachmethodische Instruktionsvariante einige der gefundenen Zusammenhänge zwischen dem zeitlichen Anteil ausgewählter fachmethodischer Beiträge und den gemessenen Kompetenzzuwächsen sehr deutlich sind.

7.2.2.2 Fähigkeiten und Verständnisse während der Bearbeitung von Instruktion

Zu einer Betrachtung des Kompetenzaufbaus kann neben der von einem Zeitpunkt zu einem späteren erreichten Kompetenz*veränderung* (also der Veränderung von Fähigkeiten, Fertigkeiten, Verständnissen und Bereitschaften) auch eine Betrachtung der in den Bearbeitungsprozessen offenbar werdenden Fähigkeiten, Fertigkeiten,

Verständnisse und Bereitschaften gezählt werden (vgl. Abschnitt 2.2, ab S. 13). Wenn eine Hypothese zum Verhältnis von Fähigkeiten und Verständnissen (spezifischer: Vorstellungen) formuliert wird, wie in Hypothese 2.2 (auf S. 447) geschehen, ist es daher auch bedeutsam, in welcher Beziehung die beiden Bestandteile von Kompetenz zueinander in den Prozessen der Bearbeitung von Instruktion stehen. Wenngleich dies nicht dezidiert in der Arbeit untersucht wurde, findet sich ein zugleich auf Fähigkeiten und Verständnisse von Lernenden bezogenes Ergebnis für die untersuchte explizit-fachmethodische Instruktionsvariante (siehe ausführlicher in Abschnitt 6.7, ab S. 312):

Die Lernenden *handeln* in Entscheidungsprozessen dazu, ob spezifische Fragen naturwissenschaftliche Fragestellungen sind oder nicht, *überwiegend sachangemessen*. Äußerungen von Lernenden, in denen sie ihr eigenes Handeln und Entscheiden erläutern, lassen aber häufig auf *unangemessene Vorstellungen* schließen. Für Entscheidungen zur Klassifizierung von naturwissenschaftlichen Fragen zeigt sich, dass inkorrekte Entscheidungen fast ausschließlich mit unangemessenen Vorstellungen einhergehen, wenn innerhalb der Entscheidungsäußerung Hinweise auf zugehörige Vorstellungen vorhanden sind. Wenn alle Äußerungen betrachtet werden, aus denen sich Vorstellungen rekonstruieren lassen, lassen sich unangemessene Vorstellungen aber eben auch sehr häufig bei korrekten Entscheidungen auffinden.

Weil dieses Ergebnis auf eine inhaltlich sehr eng umrissene Sequenz zu naturwissenschaftlichen Fragestellungen bezogen ist, ist unklar, ob es sich auch für andere Handlungsprozesse zeigt. Für andere Ausschnitte der Instruktion scheinen gemäß einiger Eindrücke aus den Kodierprozessen zu den Aktivitäten und weiterer Videosichtungen zumindest die Handlungen ebenfalls eher angemessen zu sein; nur für die Variablenkontrolle bei einzelnen Versuchen ist es anders. Die Vorstellungen wurden nur zu zwei Konzepten dezidiert untersucht (siehe Abschnitt 6.2, ab S. 269); Eindrücke aus probenhalber für andere Konzepte durchgeführten Analysen sowie den Kodierprozessen zu den Aktivitäten legen nahe, dass – wie für die dezidiert untersuchten Instruktionsextrakte – sowohl eher angemessene als auch eher unangemessene Vorstellungen aufgefunden werden können.

Das Ergebnis, dass sachangemessenes Handeln und Entscheiden nicht grundsätzlich mit aus fachmethodischer Sicht als angemessen einzuschätzenden Vorstellungen einhergeht, ist in der vorliegenden Deutlichkeit unerwartet. Das Ergebnis zeigt, dass es hilfreich ist, bei Untersuchungen zum Kompetenzaufbau zwischen Fähigkeiten und Verständnissen (hier: Vorstellungen) zu unterscheiden und nicht automatisch vom einen auf das jeweils andere zu schließen. Wie sich sachangemessenes Handeln/Ent-

scheiden und Vorstellungen beim experimentbezogenen Denken und Arbeiten zueinander verhalten, benötigt weitere Untersuchungen, mindestens aus dem Grund, dass die betrachteten Konzepte nur einen sehr kleinen Teil des experimentbezogenen Denkens und Arbeitens darstellen und die untersuchte explizit-fachmethodische Instruktion ein sehr spezifischer instruktionaler Ansatz ist, der unter Umständen Artefakte generiert. Dabei könnte beispielsweise darauf eingegangen werden, wie sich sachangemessenes Handeln/Entscheiden und Vorstellungen sowie deren ggf. voneinander unabhängigen Entwicklungen zueinander verhalten. In den hier betrachteten Fällen sind verschiedene Erklärungen dafür, dass trotz problematischer Vorstellungen korrekte Entscheidungen getroffen werden (können), vor dem Hintergrund der Videoaufzeichnungen und Analysen plausibel, etwa, dass die Entscheidung intuitiv korrekt vorgenommen wird und die Vorstellung losgelöst von der Entscheidung (z. B. erst in einer nachfolgend geäußerten Begründung) aktiviert wird, jedoch gar nicht handlungsleitend ist.

Abschließend lässt sich anmerken, dass die Angemessenheit von Äußerungen mit fachmethodischem Gehalt *nicht situationsspezifisch* untersucht wurde. Die Einschätzungen der Angemessenheit beschränken sich auf die Videosequenzen, die zur Rekonstruktion von Vorstellungen genutzt wurden, und beschreiben, ob die einer Vielzahl von Äußerungen zugeschriebene Vorstellung (genauer: Ideen-Oberkategorie) grundsätzlich eher angemessen oder unangemessen ist. Es ist jedoch möglich, dass einzelne fachmethodische Beiträge, die als generell eher unangemessen einzuschätzende Vorstellungen nahelegen, in den je spezifischen Situationen als sachangemessen einzuschätzen sind (beispielhaft in Abschnitt 6.6, ab S. 306, erläutert); dies gilt vermutlich erst recht für Teile des Materials, die nicht untersucht wurden (wie beispielsweise das Anlegen von Tabellen und Diagrammen). Ein Desiderat der Arbeit besteht also darin, die situationsspezifische Sachangemessenheit von Äußerungen zu untersuchen und diese ggf. in Beziehung zu den erreichten Kompetenzzuwächsen zu setzen, weil es naheliegend ist, dass die Sachangemessenheit fachmethodischer Beiträge von Bedeutung für den Kompetenzaufbau ist.

7.2.3 Allgemeine oder fachmethodische kognitive Aktivität

Die in der Forschung untersuchten Funktionen von instruktionalen fachmethodischen Informationen und Reflexion(sanregung)en zielen oft darauf ab, das Denken der Lernenden auf spezifische Aspekte von Fachmethoden zu lenken (siehe Kapitel 2). Solch fachmethodische kognitive Aktivität der Lernenden wird auch in der der Arbeit zugrundeliegende Studie als von großer Bedeutung in den Prozessen

des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten angesehen. Da fachmethodische kognitive Aktivität als Nachdenken über Fachmethoden empirisch nicht direkt erfassbar ist, werden an dieser Stelle verbale Aktivitäten mit fachmethodischem Gehalt (im Kontext aller anderen Aktivitäten der jeweils sprechenden Person) als Indikatoren für fachmethodische kognitive Aktivität angesehen. Auf das Denken der Lernenden mithilfe ihrer Sprechhandlungen zu schließen, ist ein häufig angewendetes und plausibles Vorgehen (theoretische Argumente bspw. bei von Aufschnaiter, S., 19xx). Gleichzeitig ist einschränkend anzumerken, dass fachmethodische Beiträge als eine mögliche Operationalisierung fachmethodischer kognitiver Aktivität keinesfalls mit fachmethodischer kognitiver Aktivität an sich gleichzusetzen sind. Beispielsweise können Lernende fachmethodisch kognitiv aktiv sein, ohne dass dies Eingang in fachmethodische Äußerungen finden (vgl. z. B. van der Graaf et al., 2015, S. 395); fachmethodische Beiträge können auch von Forschenden zugeschrieben werden, obwohl den Äußerungen keine fachmethodischen Überlegungen seitens der Lernenden zugrunde liegen (z. B. könnte die Aussage »Lass das mal immer von da fallen« von Forschenden als ein Hinweis auf ein intuitives Verständnis der Variablenkontrolle gedeutet werden, obwohl von Lernende damit eventuell einfach nur sicherstellen wollen, dass sie nicht von dem fallenden Gegenstand getroffen werden). Dass die in der Arbeit untersuchten Lernenden in Teams von (zwei bis) drei Personen arbeiten, soll dazu beitragen, dass möglichst viele fachmethodische kognitive Aktivitäten in Verbalisierungen münden. Um möglichst selten fälschlicherweise auf das Vorhandensein fachmethodischer kognitiver Aktivitäten zu schließen, werden unter anderem eine sorgfältige Trennung verschiedener Kategorien verbaler Aktivitäten im Kodiermanual vorgenommen und Kurzantworten wie Ja oder Nein nicht bei den Aktivitätskodierungen mitkodiert (ausführlicher in Abschnitt 5.1, ab S. 68).

Eine Erklärung für die gefundenen Disparitäten zwischen den beiden untersuchten Instruktionsvarianten könnte auch in dem Anteil *allgemeiner* kognitiver Aktivität liegen. Es ist wäre dann im Vergleich der beiden untersuchten Instruktionsvarianten für den erreichten Kompetenzzuwachs nicht vor allem relevant, wie groß der Anteil *fachmethodischer* kognitiver Aktivität ist, sondern wie kognitiv aktiv die Lernenden ganz allgemein sind. Vor dem Hintergrund der Methoden und Ergebnisse der Arbeit wird nachfolgend zuerst diskutiert, dass es durchaus sein kann, dass die Operationalisierung fachmethodischer kognitiver Aktivität durch die fachmethodischen Beiträge (insbesondere für die implizit-fachmethodische Instruktionsvariante) problembehaftet ist (Unterunterabschnitt 7.2.3.1). Anschließend wird eine Hypothese zur Bedeutung der allgemeinen kognitiven Aktivität der Lernenden für den

7.2 Bedeutung von Äußerungen mit fachmethod. Gehalt für den Kompetenzaufbau

Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten eingeführt und formuliert (Unterunterabschnitt 7.2.3.2).

Eine Variation der Erklärung des vorigen Absatzes für die gefundenen Unterschiede zwischen den Instruktionsvarianten könnte darin liegen, dass nicht alle Beiträge, die als fachmethodische Beiträge kodiert sind, auf eine kognitive Aktivität hinweisen, die für den Kompetenzaufbau relevant ist. Basierend auf einem für Fachinhalte vorliegenden Kodiermanual von Rogge (2010) wird ausblicksartig ein Kodiermanual entwickelt, mit dem das Konzeptualisierungsniveau fachmethodischer Beiträge untersucht werden kann, also »inwiefern in einer [fachlichen, d. h. im originalen Kontext vor allem: fachinhaltlichen – in der vorliegenden Arbeit eben: fachmethodischen] Äußerung selbst etwas Konzeptuelles, d. h. Generalisierendes[,] formuliert wird oder nicht« (Rogge, 2010, S. 107). Das Kodiermanual wird in Unterunterabschnitt 7.2.3.3 vorgestellt und erste Einblicke in potentielle Ergebnisse werden gegeben.

7.2.3.1 Operationalisierung fachmethodischer kognitiver Aktivität

Wenn davon ausgegangen wird, dass fachmethodische kognitive Aktivität eine zentrale Rolle im Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten spielt, könnte der geringe zeitliche Anteil fachmethodischer Beiträge bei der Bearbeitung der implizit-fachmethodischen Instruktionsvariante auf eine unzureichende Operationalisierung fachmethodischer kognitiver Aktivität durch verbale fachmethodische Beiträge hindeuten:

Es könnte sein, dass sich bei implizit-fachmethodischer Instruktion fachmethodische kognitive Aktivität der Lernenden *nur zu sehr geringen Teilen in verbalen Beiträgen widerspiegelt*. Es ist klar, dass Denken sich nicht vollständig in Verbalisierungen zeigt. Wenn jedoch nur ein sehr geringer Anteil fachmethodischer Überlegungen Eingang in verbale Aktivitäten findet, ist der Bedeutung fachmethodischer kognitiver Aktivität bei implizit-fachmethodischer Instruktion nur schwer empirisch nachzugehen. Eventuell böten sich hier ergänzende Analysen nonverbaler Handlungen oder Befragungen von Lernenden an.

Es könnte sein, dass *die in der Arbeit genutzten Kategorien für fachmethodische Beiträge nicht ausreichend abbilden*, welche Äußerungen von Lernenden Rückschlüsse auf fachmethodische kognitive Aktivität erlauben:

- Zum einen finden sich im Kodiermanual auf drei Teilprozesse bezogene Kategorien, die um Kategorien zu weiteren Teilprozessen des experimentellen Denkens und Arbeitens ergänzt werden könnten. (Bei der Entwicklung und dem Einsatz des Kodiermanuals für die untersuchten Instruktionsvarianten

fanden sich allerdings nur sehr vereinzelt Äußerungen, die fachmethodischen Gehalt hatten, aber nicht in die Kategorien zu den Teilprozessen passten.)

- Zum anderen wird bei den in Kapitel 5 berichteten Aktivitätskodierungen zwar zwischen fachmethodisch-indizierten und fachmethodisch-vermuteten Beiträgen unterschieden, allerdings weisen die detaillierten Analysen, die in Kapitel 6 berichtet sind, darauf hin, dass auch aus einigen anderen Äußerungen Hinweise auf Vorstellungen von Lernenden zu Konzepten des experimentbezogenen Denkens und Arbeitens abgeleitet werden können. (Dass die Kodierungen zu den Aktivitäten und zu den Vorstellungen, die anhand zweier grundsätzlich verschiedener Verfahren generiert wurden, nicht vollständig deckungsgleich sind, könnte allerdings auch daran liegen, dass bei den Vorstellungskodierungen zu vielen Äußerungen fachmethodischer Gehalt zugeschrieben wurde.) Für Analysen von fachlichen (Sprech-)Handlungen von Lernenden findet sich bei von Aufschnaiter und Rogge (2010) beispielsweise ein neustufiges Kategoriensystem, bei dem alleine drei Kategorien für exploratives (Sprech-)Handeln angegeben sind, also für solche (Sprech-)Handlungen, die »sich auf konkrete Objekte bzw. Sachverhalte [beziehen] und [...] aus Beobachtersicht wenig zielorientiert [erfolgen]« (Rogge, 2010, S. 108), aber zugleich im Zusammenhang mit dem fachlichen Lernen stehen. Ob derartige (Sprech-)Handlungen als Explorationen zu Fachmethoden oder zu Fachinhalten anzusehen sind – was für die vorliegende Arbeit von großer Relevanz wäre –, ist basierend auf Explorationen des Autors am analysierten Videomaterial aus Sicht des Autors in vielen Fällen nur schwer zu entscheiden. Weitere Untersuchungen könnten hier genauere Aufschlüsse geben.

7.2.3.2 Allgemeine relevanter als fachmethodische kognitive Aktivität

Die der Arbeit zugrundeliegende Studie ist so angelegt, dass Lernende in beiden Instruktionsvarianten implizit zu fachmethodischer kognitiver Aktivität angeregt werden (implizite Adressierung), aber nur in der explizit-fachmethodischen Instruktionsvariante zu Konzepten des experimentbezogenen Denkens und Arbeitens informiert werden und zeitlich strukturierte Explorationsphasen und Übungsphasen durchlaufen (explizite Adressierung). Insbesondere werden Lernende in der implizit-fachmethodischen Instruktionsvariante zu fachinhaltlicher kognitiver Aktivität angeregt, um gleiche Bearbeitungsdauern zu ermöglichen, anstatt zum erneuten Durchführen von Versuchen. Es ist daher zunächst aufgrund der Anlage der Instruktion nicht plausibel, davon auszugehen, dass Lernende bei der implizit-fachmethodischen Instruktionsva-

riante aufgrund anderer Instruktionselemente (wie etwa wiederholtes Durchführen von Experimenten) zu fachmethodischen Überlegungen angeregt werden als bei der explizit-fachmethodischen Instruktionsvariante.

Wird davon ausgegangen, dass fachmethodische kognitive Aktivität durch die Kategorien der fachmethodischen Beiträge zumindest ansatzweise angemessen abgebildet wird, weisen die Ergebnisse der Arbeit darauf hin, dass für die implizit-fachmethodische Instruktionsvariante unerwarteter Weise nur kognitive Aktivität, die auf andere Aspekte gerichtet ist, mit den Kompetenzzuwächsen zum experimentbezogenen Denken und Arbeiten zusammenhängt: Es gibt keinen Zusammenhang von Kompetenzzuwächsen und dem zeitlichen Umfang fachmethodischer Beiträge bei der implizit-fachmethodischen Instruktionsvariante, aber es gibt Zusammenhänge von Kompetenzzuwächsen und fachinhaltlichen Beiträgen und verbalen technischen Beiträgen zur Vorbereitung und Durchführung von Versuchen (siehe Unterabschnitt 5.5.2, ab S. 214). Demnach wäre insgesamt für Kompetenzzuwächse zum experimentbezogenen Denken und Arbeiten nur von Bedeutung, ob die Lernenden grundsätzlich kognitiv aktiv sind (was sich in dem utilisierten Untersuchungssetting für viele der Lernende in verbalen Beiträgen äußern dürfte; Betrachtungen relativ zu den Redeanteilen ermöglichen weitere Aufklärung, siehe Unterabschnitt 5.5.3.3, ab S. 232). Dazu passt, dass bei der explizit-fachmethodischen Instruktionsvariante die kognitive Aktivität zu großen Teilen auf Fachmethoden, bei der implizit-fachmethodischen Instruktionsvariante zu großen Teilen auf Fachinhalte gerichtet ist – je höher die kognitive Aktivität also ganz allgemein ist, desto wahrscheinlicher ist in beiden Fällen der Kompetenzzuwachs zum experimentbezogenen Denken und Arbeiten. Dies wird in der folgenden Hypothese zusammengefasst:

Hypothese 2.3

Allgemeine kognitive Aktivität der Lernenden ist (zumindest bei der implizit-fachmethodischen Instruktionsvariante) relevanter für den Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten als fachmethodische kognitive Aktivität.

Zukünftige genauere Analysen könnten beispielsweise Äußerungen mit fachmethodischem Gehalt für implizit-fachmethodische Instruktionsvarianten spezifischer in den Blick nehmen (beispielsweise durch Rekonstruktion von Vorstellungen analog zum hier nur auf die explizit-fachmethodische Instruktionsvariante gerichteten Kapitel 6). Ferner könnten weitere Analysen darauf gerichtet sein, ob auch innerhalb explizit-

fachmethodischer Instruktionsvarianten der Anteil fachlicher Beiträge insgesamt (unabhängig davon, ob sie fachmethodisch oder fachinhaltlich oder technisch sind), also kognitive Aktivität im Allgemeinen, die Kompetenzzuwächse besser erklärt als der Anteil fachmethodischer Beiträge.

7.2.3.3 Höhere Konzeptualisierungsniveaus als spezifische fachmethodische kognitive Aktivität

Eine Betrachtung davon, wie generalisierend die fachmethodischen Äußerungen der Lernenden formuliert sind, könnte – wie bereits im letzten Aufzählungspunkt in Unterunterabschnitt 7.2.3.1 erwähnt – spezifischere Hinweise darauf geben, inwiefern fachmethodische kognitive Aktivität im Zusammenhang mit Kompetenzaufbau steht. Dort wurde erläutert, dass die in der Arbeit vorgenommene Operationalisierung von Äußerungen mit fachmethodischem Gehalt durch die Kategorien der fachmethodisch-indizierten und -vermuteten Beiträge eventuell zu wenig umfassend ist, weil insbesondere explorative Äußerungen mit fachmethodischem Gehalt eventuell nicht berücksichtigt werden. Es könnte allerdings andersherum auch sein, dass die Kategorien fachmethodisch-indizierter und -vermuteter Beiträge zu breit gefasst sind; eventuell stehen nur solche fachmethodischen Beiträge in Zusammenhang mit den erreichten Kompetenzzuwächsen, die eher generalisierend formuliert sind. Beispielsweise könnte es sein, dass Lernende, die die Variablenkontrolle in dem Satz »Von allen Variablen, die einen Einfluss auf die abhängige Variable haben könnten, sollte möglichst nur die unabhängige Variable verändert werden«¹³⁵ zusammenfassen können, höhere Kompetenzzuwächse erreichen als Lernende, die – an denselben Stellen der Instruktion bzw. als höchste Stufe all ihrer Generalisierungen – nur Sätze wie »Du darfst hier die Masse nicht verändern, sie könnte ja einen Einfluss auf die Falldauer haben; wir wollen ja nur den Einfluss des Kegeldurchmessers testen« formulieren. Es könnte sein, dass vor allem Lernende, die im Bearbeitungsprozess viele generalisierende kognitive fachmethodische Aktivitäten ausführen oder häufiger bzw. überhaupt bei diesen ankommen (ggf. auf Basis von vorangehenden weniger stark generalisierenden Äußerungen wie für Fachinhalte bspw. von Rogge, 2010, beschrieben), hohe Kompetenzzuwächse aufweisen. Dies könnte dann ggf. wiederum als ein Hinweis darauf gedeutet werden, dass das (auch für die in der Einleitung der Arbeit angesprochenen Selbsterklärungen relevante) Abstrahieren einen wesentlichen Beitrag zum Kompetenzaufbau leistet.

¹³⁵ Alternativ könnte der Satz auch ohne die entsprechenden Fachbegriffe aus der untersuchten Instruktion mit etwas mehr Wortumfang formuliert werden: »Von allen Optionen im Versuch sollte möglichst nur die Option verändert werden, für die wir wissen wollen, ob sie einen Einfluss auf die betrachtete Größe hat; natürlich können Optionen geändert werden, bei denen wir sicher sind, dass sie keinen Einfluss haben.«

Ausführlicher Ausblick (Teil I von II)

Ein Maß für den Grad der Generalisiertheit von fachlichen Äußerungen können die Konzeptualisierungsniveaus von von Aufschnaiter und Rogge (2010a) darstellen; diese sind allerdings in bisherigen Publikationen auf fachinhaltliche Konzepte bezogen ausformuliert. Ein erster Übertrag auf fachmethodische Konzepte wird nachfolgend innerhalb des grauen Kastens vorgenommen; außerdem werden anschließend an den grauen Kasten Ergebnisse aus der Erprobung an einer sehr kleinen Stichprobe vorgestellt und an diesen Möglichkeiten der Weiterarbeit aufgezeigt.

Für den Übertrag der Konzeptualisierungsniveaus auf die in der vorliegenden Arbeit betrachteten fachmethodischen Beiträge wird nachfolgend davon ausgegangen, dass diese sich grundsätzlich auf alle fachlichen Beiträge anwenden lassen (sie wurden schließlich nicht vor dem Hintergrund einer expliziten Unterscheidung von Fachmethoden und Fachinhalten entwickelt) und als auf Strukturmerkmale von Äußerungen gerichtete Kategorien quer zu der auf Inhalte gerichteten Unterscheidung zwischen Fachinhalten und Fachmethoden usw. kodiert werden können.¹³⁶ Der Übertrag auf Fachmethoden besteht daher vor allem darin, zu prüfen, welche der Kategorienbeschreibungen allgemeiner formuliert werden müssen, um fachmethodische Beiträge nicht aufgrund einer fachinhaltliche Bezüge implizierenden Formulierung per se auszuschließen, sowie darin, angemessene Beispiele zu ergänzen. Beim Übertragen stellte sich heraus, dass einige Subkategorien als grundsätzlich fachmethodisch und eben nicht als fachinhaltlich anzusehen sind (etwa 4b: Ein bestimmtes Verhalten wird vorgeschlagen), während sich andere Subkategorien nicht auf Fachmethoden übertragen lassen und grundsätzlich fachinhaltlich sind. Obwohl also keine explizite Unterscheidung von Fachinhalten und Fachmethoden im Kategoriensystem von von Aufschnaiter und Rogge (2010a) vorgenommen wird, finden sich in den Beispielen erste Anknüpfungspunkte für diese Unterscheidung.

In Tabelle 7.1 (auf den Seiten 458 und 459) ist die Struktur der Darstellung des Kategoriensystems von Rogge (2010, S. 108–109) so gut es ging abgebildet; sofern möglich wurde der Wortlaut beibehalten oder nur minimal abgewandelt, um die Übereinstimmungen zu verdeutlichen. Allerdings werden aufgrund der Fokussierung

¹³⁶Es wäre – abweichend von der im Text geschilderten orthogonalen Kategorienrelation – sicherlich auch plausibel, inhärent fachmethodische Konzeptualisierungsniveaus zu formulieren, die auf allen Niveaus von bereits existierenden Bezügen zu Fachmethoden ausgehen und nur spezifizieren, was einen fachmethodischen Beitrag in Abgrenzung von anderen fachmethodischen Beiträgen kennzeichnen muss, um ihn mit einem spezifischen Niveau zu kodieren (hierarchische Kategorienrelation). Darauf kann allerdings an dieser Stelle nicht detaillierter eingegangen werden; es könnte sogar sein, dass die auf die beiden Weisen entstehenden Kategorien (einmal orthogonal, einmal hierarchisch) letztlich in der Anwendung kaum differieren.

Tabelle 7.1: Konzeptualisierungsniveaus basierend auf Rogge (2010, S.108–109) adaptiert mit Beispielen zu fachmethodischen Beiträgen (Teil 1).

(Sprech)Handlung	Kategorie mit Kurzbeschreibung	Indikatoren & Beispiele
Explorativ: Die Handlungen und Äußerungen des Probanden beziehen sich auf konkrete Objekte bzw. Sachverhalte und erfolgen aus Beobachtersicht wenig zielorientiert. Für Fachmethoden bspw. zu ergänzen: Sprechweisen erkunden, probeweise Versuchsaufbau ändern; konkrete Objekte (z.B. Tabelle oder Diagramme), Situationen und Phänomene (unter fach-methodischen Gesichtspunkten) beschreiben. (Umfasst die hier nicht weiter ausdifferenzierten Niveaus 1 <i>handeln (und beschreiben)</i> , 2 <i>beschreiben mit Vorlage</i> und 3 <i>beschreiben ohne Vorlage</i> .)		
Intuitiv	4 <i>vermuten:</i>	
regelbasiert: Die Äußerungen des Probanden beziehen sich auf konkrete Objekte bzw. Sachverhalte und beruhen aus Beobachtersicht auf einer (implizit) erfassten Regelmäßigkeit, die jedoch nicht expliziert wird, z. B.: Versuchsausgänge vermuten, experimentelles Verhalten vorschlagen, relevante Sachverhalte betonen, physikalische oder fachmethodische Bezeichnungen systematisch zur Beschreibung von Objekten bzw. Phänomenen nutzen, Ursache-Wirkungs-Zusammenhänge an konkreten Objekten bzw. Phänomenen erläutern, experimentelles Verhalten in konkreten Situationen (durch fachinhaltliche oder fachmethodische Auswirkungen innerhalb der konkreten Situation) begründen.	a) Der Ausgang eines Phänomens oder Experimentes wird sicher vermutet, aber nicht geraten. b) Ein bestimmtes Verhalten wird vorgeschlagen. c) Ein relevanter Sachverhalt (in Bezug auf eine konkrete Situation) wird betont. 5 <i>zuschreiben:</i> a ₁) <i>Einfache Zuschreibung.</i> Sortierung von Gegenständen zu Kategorien (fachliche und nicht-fachliche Einordnungen). Nach dem Ausgang eines Experimentes wird dessen Ergebnis auf ähnliche Fälle bezogen. Die Merkmale einer konkreten fachmethodischen Entität (Fragestellung, Hypothese, Versuchsplanung, Tabelle, Diagramm usw.) mit einer anderen konkreten fachmethodischen Entität in Verbindung bringen. a ₂) <i>Zuschreib. mit fachmeth. Vokabular.</i> Zuordnung von Worten zu Situationen, Sortierung von Gegenständen zu Kategorien (fachliche und nicht-fachliche Einordnungen). Eine konkrete fachmethod. Entität (Fragestellung, Hypothese, Versuchsplanung, Tabelle, Diagramm usw.) ohne weitere Erläuterung mit einem Fachbegriff in Verbindung bringen. b) <i>Ausdifferenzierte Zuschreibung.</i> Abläufe werden unter systematischer Nutzung von bestimmten Wörtern (vor allem Fachbegriffe) beschrieben.	→ <i>nur Fachinhaltlich</i> - „Müssen die Spalten der Tabelle nicht getauscht werden?“ - Es wird expliziert, dass eine vorgegebene Unterscheidung nicht den Erwartungen entspricht (Bsp.: „Für mich ist das das gleiche [egal, in welche Spalte einer Tabelle die UV bzw. die AV gehört]“). - Einer physikalischen Größe eine fachmethodische Funktion zuschreiben (Bsp.: „Die Spannung ist das X, für das wir prüfen wollen, ob es einen Einfluss hat“). → <i>nur Fachinhaltlich</i> - „Das ist das Gleiche wie bei der anderen Frage, für die wir einschätzen sollten, ob sie naturwissenschaftlich ist.“ - Der Verlauf eines experimentellen Arbeitsgangs wird auf ähnliche Fälle bezogen (z.B. hinsichtlich der Abfolge von Frage, Hypothese, Planung und Auswertung). - Einer physikalischen Größe eine fachmethodische Funktion zuschreiben (Bsp.: „Die Spannung ist die unabhängige Variable“). - „Wahrscheinlich müssen wir wieder Variablenkontrolle prüfen.“ → <i>unklar (evtl. nur Fachinhaltlich; ausdifferenz. fachmeth. Zuschreib. scheinen häufig bereits 6 erläutern zuzugehören)</i>

Tabelle 7.1: Konzeptualisierungsniveaus basierend auf Rogge (2010, S.108–109)
adaptiert mit Beispielen zu fachmethodischen Beiträgen (Teil 2).

(Sprech)Handlung	Kategorie mit Kurzbeschreibung	Indikatoren & Beispiele
	<i>6 erläutern</i>	
	FI Es werden inhaltliche Zusammenhänge zwischen zwei (oder mehr) Aspekten expliziert (z.B. Angabe einer Anweisung zur Durchführung zusammen mit einer darauf bezogenen fachinhaltlichen/fachmethodischen Begründung). Dabei erfolgt keine Angabe einer Generalisierung.	- „Das ist eine Beobachtung, weil das ist, welche schneller war.“ - „Wir haben die Variablenkontrolle eingehalten, weil wir das richtige Ergebnis bekommen haben.“
	FM	- „Das ist eine Beobachtung, weil du sehen kannst, welche schneller war.“ - „Man muss die Masse konstant halten, weil man ja sonst nicht weiß, was [welche der Variablen] die Falldauer beeinflusst.“
Explizit	<i>7 verallgemeinern</i>	
regelbasiert:	a) Für eine bestimmte Klasse von Objekten bzw. Situationen erfassen/benennen, dass diese eine gemeinsame Eigenschaft aufweisen, und diesen Zusammenhang explizieren.	- „Man muss die Einheiten in jeder Zeile der Tabelle angeben.“ - „Wie beim s(t)-Diagramm kommt auch hier beim I(U)-Diagramm zuerst die unabhängige Variable.“
Die Äußerungen des Probanden beziehen sich auf Klassen von Situationen und Objekten, z. B.: Fachmethodische Gemeinsamkeiten von (experimentellen) Handlungen und Objekten, Allgemeine Aussagen über Versuchsdurchführungen oder über Merkmale von fachmethodischen Entitäten, Generalisierungen zur Erklärung von (konkreten oder allgemeinen) Fällen oder zur Entscheidung zwischen (konkreten oder allgemeinen) fachmethodischen Möglichkeiten nutzen.	b) Es wird ein regelbasierter Zusammenhang zwischen zwei oder mehr Aspekten beschrieben oder eine Vorschrift/Definition formuliert.	- „Die unabhängige Variable ist die Variable, für die man wissen will, ob sie einen Einfluss hat.“ - „Die Variable, für die wir wissen wollen, welchen Einfluss sie hat, gehört auf die zweite, die untere Achse im Diagramm.“
	<i>8 erklären</i>	
	(Kausale) Verknüpfung von einem bekannten bzw. beobachteten (konkreten oder allgemeinen) Fall mit einer Verallgemeinerung (Wert 7).	- Bereits getroffene fachmethodische Zuweisungen begründen (Bsp.: „Das hier ist eine präzise Frage, weil alle präzise Fragen laut Definition des Materials die unabhängige und die abhängige Variable benennen“).
	<i>9 theoriebasiert vorhersagen</i>	
	Auf der Basis einer Verallgemeinerung (Wert 7) wird der Ausgang von einem oder das Vorgehen in einem bisher nicht beobachteten bzw. nicht bekannten (konkreten oder allgemeinen) Fall beschrieben.	- „Du musst die Masse gleich lassen, weil sie nicht die unabhängige, sondern eine Kontrollvariable ist.“ - „Weil die Variable, für die wir wissen wollen, welchen Einfluss sie hat, auf die zweite, die untere Achse gehört, musst Du die Masse da hinschreiben.“

Anmerkungen. Die Unterteilung von Niveau 6 in eine fachmethodische und eine fachinhaltliche Erläuterungskomponente kann vermutlich auch auf die Niveaus 7 bis 9 übertragen werden; im bisher untersuchten Material lassen sich allerdings keine eindeutig auf Fachinhalte bezogenen fachmethodischen Beiträge für diese Niveaus finden.

auf fachmethodische Beiträge nur für Kategorien, die sich für fachmethodische Beiträge zuweisen lassen, Indikatoren und Beispiele angegeben. Auch wurden einzelne Subkategorien ergänzt, weil sich bei einer Unterscheidung von Fachinhalten und Fachmethoden auch innerhalb solcher Konzeptualisierungen, die Verknüpfungen mehrerer Entitäten herstellen, unterscheiden lässt, ob bspw. eine fachmethodische Entscheidung durch einen Fachinhalt erläutert wird (Niveau 6-FI, Beispiel in der Tabelle) oder durch einen fachmethodischen Sachverhalt erläutert wird (Niveau 6-FM).

Zuletzt sei noch angemerkt, dass die Kategorieentwicklung anhand zweier betrachteter Videos natürlich nicht abgeschlossen werden konnte. Erst ein weiteres Anwenden der Kategorien auf zusätzliche Lernende kann zeigen, ob sich anhand der formulierten Indikatoren und Beispiele grundsätzlich passende und aussagekräftige Kodierungen erstellen lassen. Die im Anschluss dargestellten ausblicksartigen Ergebnisse sind daher mit Vorsicht zu lesen und zu deuten.

Die in Tabelle 7.1 vorgestellten Konzeptualisierungsniveaus wurden zur Erstellung des Ausblicks für zwei Teams à 3 Personen allen Äußerungen zugewiesen, die von den Hilfskräften bereits als fachmethodisch-indizierte oder -vermutete Beiträge kodiert waren. Die beiden Teams wurden anhand der Kompetenzzuwächse der Personen so ausgewählt, dass ein Kontrast von Lernenden mit hohen vs. niedrigen Kompetenzzuwächsen möglichst gut möglich ist (Kennwerte in Tabelle 7.2). Fachmethodische Beiträge, die zwei verschiedene Konzeptualisierungsniveaus beinhalten, wurden mit dem höheren der beiden Niveaus rekodiert. Das Rekodieren der fachmethodischen

Tabelle 7.2: Übersicht über die ausblicksartig mithilfe der Konzeptualisierungsniveaus analysierten Lernenden, insbesondere ihre Kompetenzzuwächse.

Person	Anwesend/Analysiert					Test-Measures							Gruppierung	
	E1	E2	E3	KnF	BuD	Zuwachs	Pre	Post	FCI	KFT	SWK	INT	Zuwachs	
G19	✓	✓	✓	✓	✓	40,2	374,8	415,0	436,9	475,9	437,0	322,6	N	
G20	✓	✓	✓	✓	✗	43,5	448,2	491,6	538,5	528,0	445,1	468,4	N	
G21	✓	✓	✓	✓	✓	-29,9	479,9	450,0	524,4	529,1	579,7	583,1	N	
H22	✓	(✓)	✓	✓	✓	186,4	515,1	701,5	327,5	590,7	519,2	520,1	H	
H23	✓	(✓)	✓	✓	✓	149,9	759,9	909,8	631,5	685,8	559,7	556,0	H	
H24	✓	(✓)	✓	✓	✓	56,1	364,7	420,8	495,7	416,5	494,9	542,5	N	

Anmerkungen. Anwesend/Analysiert: ✓ = Anwesend und in der vorliegenden Arbeit analysiert; (✓) = Anwesend und nicht analysiert; ✗ = Nicht anwesend (G20 verlässt den Raum vorzeitig). E1/E2/E3 = Einheit der explizit-fachmethodischen Instruktionsvariante; KnF/BuD = Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen bzw. zur Unterscheidung von Beobachtung und Deutung. Test-Measures: Pre/Post = Prä-/Post-Erhebung zu Kompetenzen zum experimentbezogenen Denken und Arbeiten; FCI = Fachinhaltliche Kompetenzen; KFT = Allgemeine kognitive Fähigkeiten; SWK = physikbezogenes Fähigkeitsselbstkonzept; INT = physikbezogenes Interesse. Gruppierung des Zuwachs entstammt Kapitel 5 der Arbeit: H = Hoch (Werte ab minimal 109,5), N = Niedrig (Werte bis maximal 56,1); NEG = stark negative Zuwächse (Werte unter -37,7).

7.2 Bedeutung von Äußerungen mit fachmethod. Gehalt für den Kompetenzaufbau

Beiträge erfolgte aus Zeitgründen nur für die dritte Einheit (in der sowohl fachmethodische Beiträge zum Planen und Durchführen von Untersuchungen als auch vorrangig zum Auswerten und Interpretieren von Daten vorhanden sind). Fachmethodische Beiträge, für die aus akustischen oder anderen Gründen (wie bspw. Unvollständigkeit des Satzes, die sich nicht durch den unmittelbaren Kontext der Äußerung ausdeuten lässt) zwar von den Hilfskräften ein fachmethodischer Beitrag kodiert wurde, vom Autor der Arbeit allerdings kein Konzeptualisierungsniveau zugewiesen werden konnte, wurden mit einem Niveau 0 (unklar) kodiert. Fachmethodische Beiträge, denen eine der explorativ-fachmethodischen Kategorien zugewiesen werden konnte, fanden sich nur sehr wenige: meistens aufgrund von Herrumraten der Lernenden.

In Tabelle 7.3 ist dargestellt, wie die im Laufe der gesamten dritten Einheit auftretenden Konzeptualisierungsniveaus für die untersuchten Lernenden verteilt sind. Zunächst ist deskriptiv sofort ersichtlich, dass die beiden Personen H22 und H23, die hohe Kompetenzzuwächse zum experimentbezogenen Denken und Arbeiten erreichen (Zuwachs Gesamt), sich hinsichtlich der auftretenden Konzeptualisierungsniveaus deutlich von den anderen Personen unterscheiden: insbesondere ist die gewichtete Summe als Maß des im Mittel erreichten Konzeptualisierungsniveaus für diese Personen deutlich höher als für die anderen Personen. Um auszuschließen, dass dies nur aufgrund insgesamt größerer Häufigkeit fachmethodischer Beiträge der Fall ist, wurde ein gewichtetes Verhältnis berechnet, indem durch die Anzahl aller mit Konzeptualisierungsniveaus kodierten fachmethodischen Beiträge geteilt wurde. Auch hierfür sind die beiden Personen mit den hohen Kompetenzzuwächsen zum experimentbezogenen Denken und Arbeiten deutlich von den anderen Personen abgrenzbar. Dass diese

Tabelle 7.3: Übersicht über die Konzeptualisierungsniveaus, die die ausblicksartig untersuchten Lernenden erreicht haben.

Person	Konzeptualisierungsniveau							Gewicht. Summe	Gewicht. Verhältnis	FM-Zeit (in Sek.)	FM-Zeit (relativ)	Zuwachs Gesamt	Zuwachs AI
	0	4	5	6	7	8	9						
G19	2	1	4	3	0	0	0	42	4,20	39,7	1,4%	40,24	40,12
G20	3	1	4	3	1	0	0	49	4,08	53,2	1,9%	43,49	29,38
G21	5	0	6	6	2	0	0	80	4,21	83,7	3,0%	-29,89	-4,69
H22	7	0	13	14	7	2	1	223	5,07	245,8	5,3%	186,35	239,23
H23	6	2	15	12	11	3	0	256	5,22	230	4,9%	149,86	180,06
H24	4	1	2	2	2	0	0	40	3,64	21,8	0,5%	56,08	126,09

Anmerkungen. Gewicht. Summe = Anzahl der Kodierungen multipliziert mit Niveaubezeichnung. Gewicht. Verhältnis = Gewichtete Summe geteilt durch Anzahl der insgesamt vorgenommenen Kodierungen zu Konzeptualisierungsniveaus. FM-Zeit = Absoluter Anteil fachmethodischer Beiträge. Zuwachs Gesamt = Erreichter Kompetenzzuwachs (über alle drei Einheiten hinweg). Zuwachs AI = Erreichter Kompetenzzuwachs für den Teilprozess des Auswertens und Interpretierens von Daten (Subskala).

beiden Personen auch insgesamt und relativ deutlich mehr Zeit mit fachmethodischen Beiträgen zubringen, könnte also eine bedeutsame Grundlage für das Erreichen höherer Konzeptualisierungsniveaus sein. Allerdings lässt sich mit der ausblicksartig untersuchten Stichprobe nicht sagen, ob dies tatsächlich der Fall ist. Dafür müssten beispielsweise Lernende gefunden werden, die hohe zeitliche Anteile fachmethodischer Beiträge haben und niedrige Kompetenzzuwächse erreichen; sofern diese nur niedrige Konzeptualisierungsniveaus erreichen, könnte dies als Bestätigung dafür angesehen werden, dass die im Mittel höheren Konzeptualisierungsniveaus in einem Zusammenhang mit den höheren Kompetenzzuwächsen stehen. Dazu würde passen, dass die Konzeptualisierungsniveaus 7, 8 und 9 fast nur von den beiden Lernenden mit hohen Kompetenzzuwächsen erreicht werden.

Die im vorigen Absatz vorgenommene Deutung, dass im Mittel höhere Konzeptualisierungsniveaus mit den hohen erreichten Kompetenzzuwächsen in Zusammenhang stehen könnten, wird für die untersuchte Stichprobe allerdings mehrfach einschränkt. In anschließenden Studien müsste beispielsweise darauf eingegangen werden, ob die höheren Konzeptualisierungsniveaus aufgrund von Eingangsvoraussetzungen erreicht werden; dies scheint für die vorliegende Stichprobe sehr plausibel zu sein, weil insbesondere die fachmethodischen Eingangsvoraussetzungen und die allgemeinen kognitiven Fähigkeiten der beiden Personen H22 und H23 (die eben gerade die beiden Lernenden sind, die hohe Kompetenzzuwächse erreichen) deutlich stärker ausgeprägt sind als bei den anderen Personen.

Ein interessantes Ergebnis, welches aufgrund der kleinen Stichprobe nicht gut gedeutet werden kann, besteht darin, dass für H24 zwar die Werte zu den Konzeptualisierungsniveaus im Sinne der obigen Deutung zu dem insgesamt erreichten Kompetenzzuwachs zum experimentbezogenen Denken und Arbeiten (Zuwachs Gesamt) passen, aber der auf der Subskala zum Teilprozess des Auswertens und Interpretierens erreichte Zuwachs (Zuwachs AI) nicht passt. Vielmehr erreicht die Schülerin H24 trotz geringer Werte zu den Konzeptualisierungsniveaus einen sehr hohen Zuwachs zum Auswerten und Interpretieren. Es stellt sich hier unter anderem die Frage nach Effekten der Zusammenarbeit von Lernenden, unter denen sich einige befinden, die sehr stark dazulernen; eventuell hat H24 von den fachmethodischen Beiträgen der anderen Personen gerade deshalb so stark profitiert, weil diese anderen Lernenden so hohe Konzeptualisierungsniveaus erreicht haben. Dies wäre plausibel, weil H24 eventuell die von den anderen Personen formulierten Konzepte besser verstehen konnte als die instruktional formulierten Informationen zu den Konzepten. Weitere Studien könnten derartige Prozesse des Arbeitens in Teams in den Blick nehmen.

7.3 Weitere Ansätze zur Erklärung von Unterschieden zwischen Instruktionsvarianten

Hypothesen zu den Wirkmechanismen fachmethodischer Instruktion lassen sich nicht nur, wie in Abschnitt 7.2 vorgenommen, ausgehend von Betrachtungen der fachmethodischen Beiträge der Lernenden bei der Bearbeitung der untersuchten Instruktionsvarianten generieren, sondern können auch aus anderen Aspekten des Kompetenzaufbaus abgeleitet werden. Nachfolgend wird die Bearbeitung der Instruktionsvarianten auf einer allgemeineren Ebene zur Diskussion potentieller Faktoren der Wirksamkeit von fachmethodischer Instruktion hinzugezogen, indem kurz auf die generierten Daten zu den *Bearbeitungsdauern* eingegangen wird (Unterabschnitt 7.3.1). Zudem wird auf einige Äußerungen von Lernenden eingegangen, die beim Kodieren aufgefunden wurden und auf die *Anlage von Instruktions- und Testungsmaterial* bezogen sind (Unterabschnitt 7.3.2). Aus einer anderen Perspektive werden darauffolgend zwei in der Forschungslandschaft diskutierte Faktoren – *Unterschiedlichkeit des Kompetenzaufbaus für verschiedene Teilkompetenzen des experimentbezogenen Denkens und Arbeitens*, Unterabschnitt 7.3.3, und *Unterschiedlichkeit des Kompetenzaufbaus und dessen Förderung für Lernende mit unterschiedlichen Eingangsvoraussetzungen*, Unterabschnitt 7.3.4 – auf die Bearbeitungsprozesse bezogen; dies geschieht insbesondere anhand von Diskussionen von konkret für die untersuchten Instruktionsvarianten formulierten Vermutungen von Vorholzer (2016).

7.3.1 Bearbeitungsdauern

Die untersuchte explizit-fachmethodische und die untersuchte implizit-fachmethodische Instruktionsvariante wurden von Vorholzer (2016) so konzipiert, dass im wesentlichen ähnliche Bearbeitungsdauern erreicht werden sollten. Grundsätzlich wurde dieses Ziel erreicht: Die Bearbeitungsdauern, die jeweils als Summen von Kartenbearbeitungsdauern berechnet wurden, liegen insgesamt und für jede der drei Einheiten in vergleichbaren Größenordnungen (siehe Unterabschnitt 5.5.1.1, ab S. 208). Allerdings stellen sich systematische Unterschiede zwischen den untersuchten Instruktionsvarianten und innerhalb der explizit-fachmethodischen Instruktionsvariante auch für die untersuchten Lernenden ein: Die Bearbeitungsdauern sind für die implizit-fachmethodische Instruktionsvariante systematisch größer als für die explizit-fachmethodische Instruktionsvariante (siehe Unterabschnitt 5.5.1.1); und innerhalb der explizit-fachmethodischen Instruktionsvariante sind die Bearbeitungsdauern für Lernende mit hohen Kompetenzzuwächsen systematisch größer als für Lernende mit niedrigen Kompetenzzuwächsen (siehe Unterabschnitt 5.4.1, ab S. 106).

Wird das Ergebnis, dass die Bearbeitungsdauer im Mittel für die implizit-fachmethodische Instruktionsvariante größer ist als für die explizit-fachmethodische Instruktionsvariante, mit dem Befund abgeglichen, dass die Kompetenzzuwächse für die explizit-fachmethodische Instruktionsvariante im Mittel höher sind als für die implizit-fachmethodische Instruktionsvariante, lässt sich anmerken: Eine höhere Effektivität der untersuchten explizit-fachmethodischen Instruktionsvariante beruht nicht auf einer umfangreicheren Bearbeitungsdauer. Für die in der vorliegenden Arbeit vorgenommenen Auswertungen und Deutungen wäre es problematisch, wenn die Bearbeitungsdauern für die explizit-fachmethodische Instruktionsvariante größer wären als für die implizit-fachmethodische Instruktionsvariante, da dann Rückschlüsse auf die positiveren Effekte der explizit-fachmethodischen Instruktionsvariante durch deren Anlage *oder* die höheren Bearbeitungsdauern bedingt sein könnten – so wie das Ergebnis aber real ausfällt, bleibt diese Einschränkung erspart.

Das Ergebnis, dass innerhalb der explizit-fachmethodischen Instruktionsvariante ein positiver Zusammenhang zwischen Bearbeitungsdauer und Kompetenzzuwachs vorliegt (siehe oben, basierend auf Unterabschnitt 5.4.1), während innerhalb der implizit-fachmethodischen Instruktionsvariante die Bearbeitungsdauern in keinem systematischen Zusammenhang mit dem Kompetenzzuwächsen zu stehen scheinen (siehe Unterabschnitt 5.5.1, ab S. 208), könnte so gedeutet werden, dass es bei der explizit-fachmethodischen Instruktionsvariante stärker als bei der implizit-fachmethodischen Instruktionsvariante darauf ankommt, wie die Lernenden ihre Bearbeitungsprozesse (in dem ihnen durch die Instruktion gesteckten Rahmen) organisieren. Das könnte heißen, dass es innerhalb der explizit-fachmethodischen Instruktionsvariante grundsätzlich relevant für den erreichten Kompetenzaufbau ist, wie gut die Lernende ihre Zeit im Blick haben, wie tiefgehend die Lernenden die Aufgaben bearbeiten, ob die Lernenden alle Aufgaben bearbeiten usw. (ergänzend lässt sich hierzu anführen, dass auch generelle Eindrücke aus den Kodierprozessen bestätigen, dass innerhalb der explizit-fachmethodischen Instruktionsvariante Lernende mit hohen Kompetenzzuwächsen konzentrierter arbeiten; spezifisch darauf ausgerichtete Analysen, wie bspw. Ratings, wurden allerdings nicht vorgenommen). Innerhalb der implizit-fachmethodischen Instruktionsvariante scheinen hingegen eher singuläre Ereignisse (die aufgrund kurzer Dauer kaum Einfluss auf die insgesamt erreichte Bearbeitungsdauer haben, wie etwa spezifische fachmethodische Beiträge, beispielsweise die eigenständige Erläuterung des Konzepts der Variablenkontrolle) bedeutsam für den Kompetenzaufbau zu sein. Im Rückschluss könnten diese Deutungen auch heißen, dass die allermeisten Elemente der explizit-fachmethodischen Instruktionsvariante relevant sind und direkt zum

Kompetenzzuwachs beitragen (weil es vor allem darauf ankommt, dass diese mit großem Zeiteinsatz – vermutlich umfassend und tiefgehend – bearbeitet werden).¹³⁷

7.3.2 Kompetenzaufbau und Testsituationen

Anhand von Hinweisen, die sich bei der Durchführung der Analysen ergaben, könnte der hohe Zuwachs für einige Lernende bei der implizit-fachmethodischen Instruktionsvariante auch dadurch plausibilisiert werden, dass es diesen besonders gut gelingt, die Parallelitäten zwischen Lern- und Testsituationen zu identifizieren, wodurch sie die Posttestung deutlich besser bearbeiten als die Prätestung:

Aus einigen Äußerungen von Lernenden, die bei der Bearbeitung der implizit-fachmethodischen Instruktionsvariante getroffen werden, geht hervor, dass die Lernenden bei der Bearbeitung der Instruktion die Fragen aus dem Prätest zum Teil noch in Erinnerung haben. Beispielsweise: »Hier musst Du wahrscheinlich auch darauf achten, wie das mit der Höhe ist; ist so ähnlich wie im Test.« Derartige Äußerungen zeigen, dass die Bearbeitung des Prätests einen Einfluss auf die nachfolgende Instruktionsbearbeitung haben kann: Sie können eine Sensibilisierung für experimentbezogenes Denken und Arbeiten herbeiführen (weil sie eine Fokussierung erzeugen, die Lernende in nachfolgenden Situationen eigenständig ausprobieren) oder sogar direkt zum Kompetenzaufbau beitragen (obwohl im Test keine Konzepte des experimentbezogenen Denkens und Arbeitens expliziert werden, könnten die Aufgaben doch als Explorations- oder Übungsaufgaben fungiert haben; dass Testungen positive Effekte auf gemessene Kompetenzen haben können, findet sich auch bei Tempel et al., 2018; vgl. auch Dean & Kuhn, 2007). Demnach würde der Prätest (zumindest für einige Lernende) die Funktion eines orientierenden Prompts erfüllen, aufgrund dessen die nachfolgende Bearbeitung der Instruktion stärker von fachmethodischer Fokussierung geprägt ist. Es bleibt aber unklar, ob dies nur für einige Lernende der Fall ist und ggf. wodurch diese Lernenden gekennzeichnet sind. Außerdem scheint es nicht vollständig plausibel zu sein, dass die fachmethodischen Fokussierungen sich nicht für alle Lernenden mit hohen Kompetenzzuwächsen bei der implizit-fachmethodischen Instruktionsvariante an zumindest einzelnen Stellen auch in verbalen Aktivitäten niederschlagen. Ferner wurde nicht detailliert untersucht, ob derartige Äußerungen auch innerhalb der explizit-fachmethodischen Instruktionsvariante vorkommen; dann könnten die Unterschiede zwischen den Instruktionsvarianten nur bedingt dadurch

¹³⁷ Auch bei der implizit-fachmethodischen Instruktionsvariante könnte es sein, dass singuläre Ereignisse erst basierend auf dem Zusammenwirken von mehreren instruktionalen Elementen und einer lernendenspezifischen Bearbeitung dieser entstehen. Dieser Möglichkeit analytisch nachzugehen, dürfte jedoch ein schwieriges Unterfangen sein.

plausibilisiert werden. Zuletzt lässt sich anmerken, dass mangels systematischer Analyse von Äußerungen der Lernenden, aus denen sich auf ein Rückbeziehen auf die Testumgebung schließen lässt, keine Aussage darüber getroffen werden kann, ob derartige Äußerungen vor allem innerhalb der ersten Einheit der Instruktion vorkommen (was durchaus plausibel wäre, weil diese direkt nach der Prä-Erhebung bearbeitet wurde).

Ausgehend von den oben zitierten Äußerungen könnte auch vermutet werden, dass es im Posttest einige Lernende gibt, die auf die Erfahrungen aus der Bearbeitung der Instruktion (und ggf. aus dem Prätest) zurückgreifen. Es könnte also sein, dass einige Lernende während der Bearbeitung der Instruktion keine substantiellen Fortschritte in ihren Fähigkeiten und Verständnissen machen, sondern erst in der Bearbeitung des Posttests die Erfahrungen reflektieren und dadurch dazu in der Lage sind, den Test angemessen zu bearbeiten. Der eigentliche Fähigkeitsaufbau fände demnach erst bei der Posttest-Bearbeitung statt. Aufgrund der strukturellen Ähnlichkeit von Testinstrument und explizit-fachmethodischer Instruktionsvariante nahe, dass insbesondere Lernende, die diese Instruktionsvariante bearbeitet haben, im Posttest Rückbezüge herstellen.¹³⁸ Dies würde zu den im Mittel deutlich größeren Kompetenzzuwächsen bei der explizit-fachmethodischen Instruktionsvariante passen. Es könnte demnach sein, dass die bei der Instruktionsbearbeitung gemachten experimentellen Erfahrungen bei beiden Instruktionsvarianten sowie für die explizit-fachmethodische Instruktionsvariante zusätzlich die Erfahrung im Umgang mit Aufgaben, die strukturell aufgrund der expliziten Ausrichtung auf Fachmethoden den Testaufgaben sehr vergleichbar sind, verbunden mit einem Gewöhnungseffekt von Prä- zu Post-Erhebung zu den gemessenen Kompetenzzuwächsen führen. Dies könnte Bedeutung dafür haben, wie die Befunde zu den untersuchten Instruktionsvarianten kommuniziert werden, weil zwar reale Kompetenzzuwächse zu verzeichnen sind, diese aber ggf. auf eine sehr eng umrissene Klasse von Aufgaben beschränkt sein könnten; Studien zum Transfer der mit den eingesetzten Instruktionsvarianten aufgebauten Kompetenzen in andere, eventuell komplexere Kontexte des experimentbezogenen Denkens und Arbeitens könnten hierzu Aufschluss geben.

Als Hypothese, der in weiteren Studien nachgegangen werden kann, wird dieser Abschnitt so zusammengefasst:

¹³⁸Es wäre dabei unter anderem spannend, zu untersuchen, ob sich die bereits in Unterabschnitt 7.2.2 (ab S. 447) unterschiedenen Dimensionen von Fähigkeiten und Verständnissen auch in den Rückbezügen der Lernenden, die unterschiedliche Instruktionsvarianten bearbeitet haben, finden lassen: Greifen Lernende, die die explizit-fachmethodische Instruktionsvariante bearbeitet haben, bei der Testbearbeitung tatsächlich stärker auf in der Instruktion adressierte Konzepte des experimentbezogenen Denkens und Arbeitens zurück? Oder bestehen die Fähigkeiten bei der Testbearbeitung weitestgehend unabhängig von den Verständnissen bzw. deren expliziter Verwendung?

Hypothese 3.1

Der Kompetenzaufbau einer Person ist wesentlich davon abhängig, wie diese die administrierten Testinstrumente und die Erlebnisse aus der Instruktionsbearbeitung zusammenbringt.

7.3.3 Teilkompetenzen

In der vorliegenden Arbeit wird eine in zahlreichen Studien vorgenommene Unterscheidung in drei Teilprozesse genutzt (siehe Abschnitt 2.2, ab S. 13): Formulieren von Fragen und Hypothesen (FH), Planen und Durchführen von Untersuchungen (PU) und Auswerten und Interpretieren von Daten (AI). Es wird häufig davon ausgegangen, dass die zugehörigen Fähigkeiten, Fertigkeiten und Verständnisse voneinander abgrenzbare Teilkompetenzen darstellen (vgl. Abschnitt 2.2). In der Forschungslandschaft wird auch die Frage aufgeworfen, inwiefern die Wirksamkeit explizit-fachmethodischer Instruktion sich in Bezug auf Teilprozesse des experimentbezogenen Denkens und Arbeitens bzw. zugehörige Teilkompetenzen unterscheidet.

In der der Arbeit zugrundeliegenden Studie findet sich zwar bei Betrachtung der Gesamtskala zum experimentbezogenen Denken und Arbeiten, dass die explizit-fachmethodische Instruktionsvariante der implizit-fachmethodischen Instruktionsvariante signifikant überlegen ist. Allerdings findet sich bei einer Differenzierung der zu den Teilprozessen gehörigen Subskalen FH, PU und AI, dass die explizit-fachmethodische Instruktionsvariante der implizit-fachmethodischen Instruktionsvariante nur in Bezug auf den Teilprozess des Auswertens und Interpretierens signifikant überlegen ist. Für die beiden anderen Teilprozesse des experimentbezogenen Denkens und Arbeitens finden sich keine oder nur sehr kleine Effekte, die zudem nicht signifikant sind. Diese Ergebnisse entstammen der Dissertation von Vorholzer (2016; insb. Abschnitt 10.2 und Abbildung 10.2). In der Diskussion der Ergebnisse stellt Vorholzer (2016), der zudem die Konstruktion der Instruktion vorgenommen hat, eine ausführliche Vermutung dazu auf, wie diese Ergebnisse erklärt werden könnten:

Eine mögliche Ursache für die unterschiedliche Wirkung des expliziten und des impliziten Instruktionsansatzes auf die einzelnen Teilkompetenzen ist die Strukturierung der Lerneinheiten. Obwohl jede der durchgeführten Lerneinheiten auf eine bestimmte Teilkompetenz fokussiert (vgl. Tabelle 9.3, S. 95), gibt es einige Aspekte der Teilkompetenzen FH und PU, welche in allen Lerneinheiten berücksichtigt werden mussten. So werden z. B. in jeder der drei Lerneinheiten Experimente durchgeführt, bei

deren Durchführung immer auch die Einhaltung der Variablenkontrolle zu berücksichtigen ist (z. B. Karte 4 und 11 aus Lerneinheit I der Treatmentgruppe, vgl. Anhang B.1 oder Karte 4, 5 und 9 aus Lerneinheit I der Kontrollgruppe, vgl. Anhang B.2). Die übergeordnete Teilkompetenz »Planung von Untersuchungen« wird jedoch lediglich in Lerneinheit II fokussiert. In gleicher Weise sind z. B. auch Aspekte der Teilkompetenz »Formulierung von Fragen und Hypothesen« außerhalb der zugehörigen Lerneinheit I von Relevanz, da auch die im Rahmen späterer Einheiten durchgeführten Experimente häufig einer vorgegeben oder von den Schülerinnen entwickelten Fragestellung nachgehen. Im Gegensatz dazu werden die Aspekte der Teilkompetenzen AI [sic] fast ausschließlich in der zugehörigen dritten Lerneinheit berücksichtigt. In den vorhergehenden Lerneinheiten finden Auswertung und Interpretation dagegen nur vereinzelt und auf einem basalen Niveau statt (z. B. Vervollständigen von Sätzen oder Ankreuzen von Ergebnissen; z. B. Karte 6 und 7 aus Lerneinheit I der Kontrollgruppe, vgl. Anhang B.2). Folglich bieten die Lerneinheiten der Treatment- und der Kontrollgruppe insgesamt deutlich mehr Gelegenheiten, etwas über das Formulieren von Fragen und Hypothesen oder das Planen von Untersuchungen zu lernen, als dies für das Auswerten und Interpretieren von Daten der Fall ist. (Vorholzer, 2016, S. 132–133)

Demnach könnten die Unterschiede zwischen den Instruktionsvarianten sich dadurch erklären lassen, dass die (aufgrund der in beiden Instruktionsvarianten vorhandenen Strukturierung entlang fachmethodischer Konzepte existierenden) Möglichkeiten der Anwendung und (Wieder-)Entdeckung fachmethodischer Konzepte für FH und PU stärker genutzt werden als für AI. Vorholzer (2016) formuliert dies in einer zusammenfassenden Deutung so:

Implizite Instruktion scheint besonders dann eine ähnliche Wirkung wie explizite Instruktion zu entfalten, wenn sie systematisch entlang prozessbezogener Konzepte strukturiert ist und zahlreiche Möglichkeit zur (Wieder-)Entdeckung dieser Konzepte bietet. Umgekehrt scheint sich die Überlegenheit einer expliziten Strukturierung umso deutlicher zu zeigen, je weniger Zeit (bzw. Gelegenheiten) und je weniger Struktur die Instruktion anbietet. (Vorholzer, 2016, S. 134–135)

Die für die vorliegende Arbeit durchgeführten Analysen der Bearbeitungsprozesse können genauere Aufschlüsse darüber geben, inwiefern sich die vermutete Erklärung von Vorholzer (2016) als plausibel erweist. Aus der nach Teilprozessen differenzierten

7.3 Weitere Ansätze zur Erklärung von Unterschieden zwischen Instruktionsvarianten

Betrachtung der fachmethodischen Beiträge für die explizit-fachmethodische Instruktionsvariante (Unterabschnitt 5.4.4, ab S. 173) und für die implizit-fachmethodische Instruktionsvariante (Unterabschnitt 5.5.3, ab S. 228) lässt sich ableiten, dass sich innerhalb der explizit-fachmethodischen Instruktionsvariante die Schwerpunkte der Einheiten durchaus in den fachmethodischen Beiträgen widerspiegeln:¹³⁹

Insbesondere finden sich faktisch keine fachmethodischen Beiträge zum Auswerten und Interpretieren in den ersten beiden Einheiten der explizit-fachmethodischen Instruktionsvariante. Dies könnte durch das Vorgehen beim Kodiertraining bedingt sein: Zwei der Hilfskräfte haben die Kategorie der fachmethodischen Beiträge zum Auswerten und Interpretieren von Daten zwar zu Beginn – vor dem Kodieren von Einheit 1 – kennengelernt, aber erst beim zeitlich auf das Kodieren von Einheit 1 folgende Kodieren von Einheit 3 intensiv genutzt. Da allerdings auch bei Einheit 2, die von dem Autor der Arbeit und einer anderen Hilfskraft kodiert wurde, kaum fachmethodische Beiträge mit Bezug zum Auswerten und Interpretieren von Daten aufgefunden wurden, liegt es sehr nahe, dass Lernende in den ersten beiden Einheiten der explizit-fachmethodischen Instruktionsvariante keine fachmethodischen Überlegungen zum Auswerten und Interpretieren von Daten verbalisieren. Damit ist nicht gesagt, dass sie keine Auswertungen oder Interpretationen vornehmen (wiewohl diese basal sein dürften, wie aus dem ersten Zitat von Vorholzer oben hervorgeht), sondern dass sie beim (basalen) Auswerten und Interpretieren keine Äußerungen machen, die fachmethodischen Gehalt haben – vielmehr scheint es den Lernenden nur um die fachinhaltlichen Ergebnisse der durchgeführten Experimente zu gehen.

Die Vermutung, dass Lernende zu den beiden anderen Teilprozessen in allen Einheiten fachmethodisch beitragen, scheint sich nur zum Teil zu bestätigen. Die in Abbildung 5.34 (auf S. 174) dargestellten Ergebnisse zu den Anteilen fachmethodische Beiträge scheinen sich besser dadurch deuten zu lassen, dass jeweils der in der vorangegangenen Einheit fokussierte Teilprozess erneut von den Lernenden fachmethodisch beleuchtet wird. Diese Deutung würde auch mit den im Material angelegten Wiederholungen übereinstimmen, die bereits vorher adressierte fachmethodische Konzepte auch in späteren Einheiten erneut kurz und bündig explizit in den Fokus der Lernenden rücken, weil sich diese vor allem in Einheit 2 für Fragen und Hypothesen sowie in Einheit 3 für das Planen und Durchführen von Untersuchungen finden. Vor dem Hintergrund, dass die erste Einheit (zu Fragen und Hypothesen) nur 45 Minuten

¹³⁹Dass in der vorliegenden Arbeit nur diese recht allgemeinen – für die sich im Text anschließende ausführliche Diskussion allerdings ausreichenden – Analysen durchgeführt wurden, hatte vor allem zeitliche Gründe. Im Sinne von Priorisierung wurden detaillierte Analysen wie bspw. zu Aktivitätsprofilen oder zu Karten der Instruktion allerdings aus inhaltlichen Gründen zurückgestellt: Psychometrisch lassen sich die Teilkompetenzen nicht optimal trennen (siehe Vorholzer, 2016) und die Wahl der analysierten Lernenden erlaubt für manche der Einheiten keine sinnvollen Vergleiche (es müssten also weitere Teams kodiert werden).

angesetzte Bearbeitungsdauer umfasst, während die anderen beiden Einheiten mit 90 Minuten Bearbeitungsdauer angesetzt sind, könnte es daher sein, dass es nicht so sehr auf die insgesamt einem Teilprozess zugewandte Zeit ankommt, sondern vielmehr darauf, dass eine zeitlich separierte und/oder in einem anderen inhaltlichen Kontext stattfindende instruktional angeleitete Wiederholung vorgenommen wird (beides ist nur für die untersuchte explizit-fachmethodische Instruktionsvariante der Fall, siehe Abschnitt 4.2, ab S. 55, daher ist nicht klar, ob beides auch relevant ist).

Während sich also für die explizit-fachmethodische Instruktionsvariante die Schwerpunkte der fachmethodischen Beiträge deutlich aus den schwerpunktmäßig einheitenweise instruktional adressierten Teilprozessen des experimentbezogenen Denkens und Arbeitens zu ergeben scheinen, stellt sich für die implizit-fachmethodische Instruktionsvariante ein anderes Ergebnis ein (siehe insb. Abbildung 5.52 auf S. 230): Fachmethodische Beiträge zu den Teilprozessen des Formulierens von Fragen und Hypothesen und des Auswertens und Interpretierens von Daten finden sich nur zu geringen zeitlichen Anteilen; auch der Anteil fachmethodischer Beiträge zum Planen und Durchführen von Untersuchungen ist im Vergleich zur explizit-fachmethodischen Instruktionsvariante sehr gering, allerdings ist er deutlich größer als die Anteile zu den anderen beiden Teilprozessen. Für die im nächsten Absatz folgenden Deutungen ist zudem relevant, dass fachmethodische Beiträge mit Bezügen zu Fragen und Hypothesen nur in der ersten Einheit aufgefunden wurden, während Bezüge zu den anderen beiden Teilprozessen in allen drei Einheiten gefunden wurden.

Dass fachmethodische Beiträge mit Bezügen zum Planen und Durchführen von Untersuchungen in der implizit-fachmethodischen Instruktionsvariante den Großteil der fachmethodischen Beiträge ausmachen, kann als Stützung einer im weiteren Kontext der obigen Zitate von Vorholzer (2016) formulierten Vermutung angesehen werden, dass nämlich die untersuchte explizit-fachmethodische Instruktionsvariante anders als in vielen Studien, in denen explizite und implizite Varianten zur Variablenkontrolle miteinander verglichen werden, nur deshalb nicht deutlich effektiver im Bereich der als Bestandteil des Teilprozesses zum Planen und Durchführen von Untersuchungen aufgefassten Variablenkontrolle ist als die untersuchte implizit-fachmethodische, weil die implizit-fachmethodische Instruktionsvariante bereits substantiell Gelegenheiten zum (Wieder-)Entdecken zugehöriger fachmethodischer Konzepte beinhaltet.¹⁴⁰ Aller-

¹⁴⁰Vorholzer (2016) schreibt u. a.: »Im Gegensatz zu den in dieser Arbeit verwendeten Lerneinheiten [...], enthält eine solche Umsetzung von impliziter Instruktion beispielsweise kein nach dem Prinzip der VKS [Variablenkontrollstrategie] strukturierten Versuchsanleitungen und bietet daher vermutlich auch kaum Gelegenheiten für Schülerinnen, Konzepte zur VKS bei der Bearbeitung der Aufgaben zu „entdecken“. [...] Obwohl sowohl der Instruktionsansatz von Chen und Klahr als auch der in dieser Arbeit verwendete Instruktionsansatz als „implizit“ bezeichnet werden können, unterscheiden sich die Ansätze deutlich in ihrer Umsetzung (vgl. Kapitel 7) und haben daher, vermutlich, eine unterschiedliche Wirkung auf den Kompetenzaufbau.«

dings ist einschränkend zu bedenken, dass die zeitlichen Anteile fachmethodischer Beiträge (unabhängig von der Zugehörigkeit zu den Teilprozessen und damit eben auch für den Teilprozess des Planens und Durchführens von Untersuchungen) für die implizit-fachmethodische Instruktionsvariante deutlich geringer sind als für die explizit-fachmethodische Instruktionsvariante (siehe auch die verschiedenen Diskussionen in Abschnitt 7.2). Weiterhin passen die Ergebnisse zu den anderen Teilprozessen des experimentbezogenen Denkens und Arbeitens nicht optimal zu der Vermutung von Vorholzer, weil sich für den Teilprozess des Auswertens und Interpretierens in allen Einheiten der implizit-fachmethodischen Instruktionsvariante fachmethodische Beiträge finden, obwohl sich für diesen die deutlichsten Unterschiede zwischen den Instruktionsvarianten einstellen, während für den Teilprozess des Formulierens von Fragen und Hypothesen, zu dem sich keine Unterschiede zwischen den Instruktionsvarianten finden, nur bei Einheit 1 der implizit-fachmethodischen Instruktionsvariante fachmethodische Beiträge finden. Ob also tatsächlich die Nutzung der Gelegenheiten zum (Wieder-)Entdecken der fachmethodischen Konzepte den Ausschlag für die geringen Unterschiede geben, scheint sich anhand der untersuchten fachmethodischen Beiträge nicht gut entscheiden zu lassen. Vielmehr dürften erst Analysen dazu, ob und wie umfangreich Lernende bei den beiden Instruktionsvarianten eigenständig fachmethodische Konzepte formulieren (und welche dies sind), erste Aufschlüsse dazu liefern (für erste Ansätze dazu siehe Ausführlicher Ausblick – Teil I, ab Seite 457).

In einer Rückschau auf die vorhergehenden Absätze lässt sich zusammenfassen, dass die von Vorholzer (2016, S. 132–135) formulierte Vermutung, dass eine systematische Strukturierung entlang von fachmethodischen Konzepten und ein Beinhalt von zahlreichen Möglichkeiten zu (Wieder-)Entdeckung dieser Konzepte Merkmale effektiver implizit-fachmethodischer Instruktion sind, durch die in der vorliegenden Arbeit berichteten Ergebnisse nicht optimal gestützt werden kann; gleichwohl stehen die Ergebnisse nicht in starkem Kontrast zu dieser Vermutung, so dass sie aufgrund der theoretischen Plausibilität auch weiterhin aufrechterhalten werden kann. Erst zukünftige Untersuchungen – die stärker als in der vorliegenden Arbeit auf implizit-fachmethodische Instruktion eingehen – werden Klärung bringen können. Dies wäre wünschenswert sowohl für Forschung als auch für Unterrichtspraxis, denn aus zeitlichen Gründen wird eine kleinschrittige explizite Adressierung im Regelunterricht nicht für alle fachmethodischen Konzepte möglich sein und auch eine lange erprobte, intensiv durchdachte implizite Adressierung dürfte Lehrkräfte vor große Herausforderungen stellen. Forschungsergebnisse zur Relevanz der Merkmale könnten deutlicher herausstellen, welche Aspekte bei der Unterrichtsplanung bedacht werden müssen und welche Aspekte von geringerer Relevanz für den Aufbau von Kompetenzen zum experimentbezogenen Denken und Arbeiten sind.

Mit Blick auf die explizit-fachmethodische Instruktionsvariante lässt sich zusammenfassend bemerken, dass sich die jeweils vorgenommene Adressierung spezifischer Teilprozesse deutlich in den Bezügen der fachmethodischen Beiträge widerspiegelt. Daraus lässt sich zwar folgern, dass Lernende mit einer Fokussierung auf einen einzelnen Teilprozess gut zurechtkommen, ohne Ständig Bezug zu den anderen Teilprozessen herzustellen; allerdings könnte es dennoch sein, dass die Reihenfolge der Einheiten der explizit-fachmethodischen Instruktionsvariante Relevant für den Kompetenzaufbau ist, weil Verständnisse und/oder Fähigkeiten zu den Teilkompetenzen zum Planen und Durchführen oder insbesondere zum Auswerten und Interpretieren erst vor dem Hintergrund von bereits (zumindest ansatzweise) vorhandenen Verständnissen und/oder Fähigkeiten zu den anderen hier im Vorhinein adressierten Teilprozessen aufgebaut werden können. Eine Durchsicht des Instruktionsmaterials ergibt beispielsweise, dass insbesondere bei der Adressierung des Auswertens und Interpretierens Rückbezüge zu Konzepten des Planens und Durchführens hergestellt werden. In weiteren Studien könnte sowohl geprüft werden, inwiefern die Teilprozesse sich unabhängig voneinander instruktionale realisieren lassen als auch anschließend daran, inwiefern Reihenfolgeeffekte existieren.

7.3.4 Eingangsvoraussetzungen

Wie in Unterunterabschnitt 2.4.2.4 (ab S. 37) beschrieben, finden sich im Vergleich explizit-fachmethodischer und implizit-fachmethodischer Instruktionsansätze zum Teil unterschiedliche Effekte auf Lernende mit unterschiedlichen Eingangsvoraussetzungen. Die Darstellung der Arbeit bezieht sich vor allem *fachmethodische* Eingangsvoraussetzungen, bezüglich derer sich Unterschiede in der Wirkung der Instruktionsansätze ausmachen lassen: Lernende mit *guten* Eingangsvoraussetzungen profitieren in den meisten Studien sowohl von explizit- als auch implizit-fachmethodischen Instruktionsansätzen, während Lernende mit *schlechten* fachmethodischen Eingangsvoraussetzungen vor allem von den stärker-informierenden explizit-fachmethodischen Instruktionsansätzen profitieren (meist etwas weniger als Lernende mit besseren Eingangsvoraussetzungen) – von weniger-informierenden implizit-fachmethodischen Instruktionsansätzen oder nicht-informierenden Minimal-guidance-Ansätzen allerdings nur wenig oder zum Teil gar nicht.

Für die in der vorliegenden Arbeit betrachteten Instruktionsvarianten hat Vorholzer (2016) keine der erwartbaren Einflüsse von Eingangsvoraussetzungen aufgefunden. Insbesondere finden sich keine Unterschiede in den erreichten Kompetenzzuwächsen zwischen Lernenden mit unterschiedlichen Eingangsvoraussetzungen für die implizit-fach-

7.3 Weitere Ansätze zur Erklärung von Unterschieden zwischen Instruktionsvarianten

methodische Instruktionsvariante, für die vor dem Hintergrund des vorigen Absatzes die Einflüsse der Eingangsvoraussetzungen am deutlichsten werden müssten. Obwohl das Hauptaugenmerk der vorliegenden Arbeit nicht auf der implizit-fachmethodischen Instruktionsvariante liegt, lassen sich aus den Ergebnissen der Arbeit erste Hinweise dazu ableiten, warum der Einfluss der Eingangsvoraussetzungen für diese und die explizit-fachmethodische Instruktionsvariante sowie im Vergleich der Instruktionsvarianten erwartungskonform nicht vorhanden ist. Diese Hinweise werden ausgehend von der von Vorholzer (2016) auf den Seiten 136 bis 137 seiner Dissertation formulierten potentiellen Erklärung abgeleitet; anschließend werden weitere Aspekte diskutiert.

Vorholzer (2016) kontrastiert die von ihm entwickelten und auch in der vorliegenden Arbeit untersuchten Instruktionsvarianten mit typischem Physikunterricht. Weil die Instruktion aufgrund der Kleingruppenarbeit an verschriftlichter Instruktion stärker zu aktivem Explorieren und Diskutieren anregen als dies im gewöhnlich zu großen Teilen aus Klassengesprächen bestehenden Physikunterricht der Fall sei, würden die Eingangsvoraussetzungen nicht zur Kompensation fehlender Lerngelegenheiten herangezogen. Ferner unterstütze die auf das (Wieder-)Entdecken hin angelegte »Redundanz innerhalb einer Lerneinheit und über mehrere Lerneinheiten hinweg« den Kompetenzaufbau aller Lernenden gleichermaßen, weil sie alle zielgerichtet (intelligent) üben würden (S. 136). Aus den Ergebnissen der Arbeit geht nun, wie zum Teil in Unterabschnitt 7.3.3 diskutiert, hervor, dass die Gelegenheiten zum (Wieder-)Entdecken fachmethodischer Konzepte *vor allem innerhalb der dem jeweiligen Teilprozess zugehörigen Einheit* der *explizit-fachmethodischen* Instruktionsvariante vorzufinden sind; über die Verteilung der Lerngelegenheiten innerhalb der einzelnen Lerneinheiten kann jedoch aufgrund fehlender inhaltlicher Auflösung der fachmethodischen Beiträge keine Aussage gemacht werden.¹⁴¹ Außerdem findet sich für die *implizit-fachmethodische* Instruktionsvariante nur ein sehr geringer Anteil fachmethodischer Beiträge für jede der untersuchten Personen, so dass davon auszugehen ist, dass bei der Bearbeitung der implizit-fachmethodischen Instruktionsvariante von keiner Person alle in den Tests adressierten fachmethodischen Konzepte entdeckt werden (zumindest wenn dies so operationalisiert wird, dass diese Entdeckung sprachlich irgendwie angezeigt wird).

¹⁴¹Die Analysen zu den Vorstellungen (Kapitel 6) beschränken sich auf zeitlich eng umrissene Instruktionsextrakte und können daher ebenfalls keine Hinweise dazu geben, inwiefern innerhalb derselben Lerneinheit der explizit-fachmethodischen Instruktionsvariante die Lernenden (bereits zuvor adressierte) Konzepte (wieder-)entdecken, die zu dem in der Einheit fokussierten Teilprozess des experimentbezogenen Denkens und Arbeitens gehören. Die zur Wahl der Instruktionsextrakte notwendigen Analysen des Instruktionsmaterials geben Hinweise darauf, wo im Material spezifische Übungen zu den untersuchten Konzepten angelegt sind und die zugehörigen Analysen geben Aufschluss darüber, dass diese grundsätzlich intensiv genutzt wurden; aber es dürfte vermutlich zusätzlich relevant sein, wann Lernende ohne direkt instruktional angelegte Übungen die Konzepte (wieder-)entdecken.

Passend zur von Vorholzer (2016) formulierte Erklärung legen die stichprobenartig ausgewerteten sehr geringen Anteile der Zeit, die die Lernenden *off task* sind (siehe Unterabschnitt 5.4.1, ab S. 106), nahe, dass die abweichend von typischem Physikunterricht (siehe dazu, von den in der vorliegenden Arbeit zitierten Quellen vermutlich am ehesten Zohar & David, 2008) geringen Einflüsse der Eingangsvoraussetzungen auf die hohe kognitive Aktivität und die ausdauernde Beteiligung der Lernenden zurückgeführt werden können. Allerdings finden sich auch für Studien, in denen die Intervention stärker auf Individuen oder Kleingruppen ausgerichtet ist, Einflüsse der Eingangsvoraussetzungen im Vergleich von explizit-fachmethodischen und implizit-fachmethodischen Instruktionsvarianten (z. B. Lorch et al., 2010; Zohar & Peled, 2008).

Aufschluss darüber, warum den Eingangsvoraussetzungen keine große Bedeutung zukommt, könnten zum einen detailliertere Untersuchungen der implizit-fachmethodischen Instruktionsvariante geben. Dabei wären auch Vergleiche mit anderen implizit-fachmethodischen Instruktionsansätzen und Vergleiche der jeweils zugehörigen Bearbeitungsprozesse relevant; dies lässt sich unter anderem daran illustrieren, dass Vorholzer, 2016 Unterschiede zwischen der untersuchten Instruktionsvariante und der impliziten Instruktion von Chen & Klahr, 1999 diskutiert und diese auf die bereits weiter oben diskutierten Gelegenheiten zum (Wieder-)Entdecken zurückführt, aber vor dem Hintergrund der vorliegenden Arbeit offen bleibt, ob diese Erklärung zutreffen ist. Zum anderen bietet es sich an, der bereits in Unterabschnitt 7.2.3 in Ansätzen diskutierten Frage nach der Stellung von allgemeiner und spezifischer fachmethodischer kognitiver Aktivität nachzugehen. Diese können eventuell nicht nur die Unterschiede zwischen den Instruktionsvarianten erklären, sondern auch Hinweise darauf geben, warum die Wirkung beider Instruktionsvariante unabhängig von den Eingangsvoraussetzungen ist. Aufschluss darüber, warum den Eingangsvoraussetzungen keine große Bedeutung zukommt, könnten auch Untersuchungen geben, in denen die Aktivitäten und Vorstellungen – und ggf. Fähigkeiten im Prozess der Bearbeitung, wie bspw. dazu, korrekte fachmethodische Entscheidungen zu treffen – nicht (nur) für Lernende mit unterschiedlichen Kompetenzzuwächsen, sondern (zudem auch) für Lernende mit unterschiedlichen Eingangsvoraussetzungen differenziert werden. Aufgrund der geringen Bedeutung der Eingangsvoraussetzungen für die erreichten Kompetenzzuwächse, wurde diese Differenzierung in der vorliegenden Arbeit nicht detailliert vorgenommen; es könnte allerdings sein, dass sich die *Bearbeitungsprozesse* für Lernende mit unterschiedlich guten Eingangsvoraussetzungen *deutlich voneinander unterscheiden* und diese Unterschiede Hinweise darauf enthalten, warum es in den letztlich *erreichten Kompetenzzuwächsen keine* Unterschiede gibt.

7.4 Instruktionale Informationen und Praxis

Von großer Wichtigkeit für den Kompetenzaufbau bei explizit-fachmethodischer Instruktion ist der Umgang mit in der Instruktion gegebenen Informationen. Für das Umfeld von zwei Informationstexten zu ausgewählten Konzepten des experimentbezogenen Denkens und Arbeitens – Kennzeichen von naturwissenschaftlichen Fragen sowie Unterscheidung von Beobachtungen und Deutungen – liegen detaillierte Analysen der Äußerungen von Lernenden vor (berichtet in Kapitel 6). Grundsätzlich ist zu betonen, dass während der Bearbeitung der Informationskarten nur sehr vereinzelt Äußerungen von Lernenden auftreten (abgesehen von lautem Vorlesen) und noch seltener Äußerungen mit fachmethodischem Gehalt, aus denen sich beispielsweise auf Vorstellungen der Lernenden schließen ließe (geht sowohl aus den in Kapitel 5 berichteten als auch den in Kapitel 6 berichteten Analysen hervor). In der Arbeit wird nicht nur das der Information nachlaufende zeitliche Umfeld betrachtet, um die Nutzung der Information zu erfassen; auch die vorlaufende Bearbeitung der Instruktion wird analysiert, unter anderem, weil durch den Vergleich vorlaufender und nachlaufender Vorstellungen auf den Einfluss der Information geschlossen wird.

Fazit 4.1

Durch Informationstexte werden typischerweise bereits vorhandene (eventuell nicht zu den Informationen passende) Vorstellungen nicht abgelöst, sondern eher weitere, zusätzliche Vorstellungen etabliert.

Die Profile von Vorstellungen im zeitlichen Verlauf (z. B. Abbildung 6.20 auf S. 345) zeigen, dass viele der Vorstellungen, die vor den Informationstexten geäußert wurden, auch danach weiterhin vorliegen (selbst wenn sie nicht zu den auf der Information angeführten Konzepten passen; insbesondere Abschnitt 6.8, ab S. 326, und auch Abschnitt 6.9, ab S. 353). Dieses Ergebnis passt gut zu Befunden aus der Lern- und Vorstellungsforschung, in denen grundsätzlich die Beständigkeit von Vorwissen/Präkonzepten hervorgehoben wird (z. B. Méheut, 2012; Schecker & Duit, 2018). Anzumerken ist allerdings, dass nur für etwa die Hälfte der untersuchten Personen Profile des zeitlichen Verlaufs ausgemacht werden konnten; für die anderen Lernenden ließen sich nicht genug Vorstellungen zuweisen. Das Ergebnis zeigt, dass eine wesentliche Herausforderung von Studien, die Kompetenzaufbau mithilfe von Prozessaufzeichnungen untersuchen, darin besteht, genügend Hinweise auf Zwischenstände der Lernenden (verbale Äußerungen oder auch nonverbale Hinweise auf Kompetenzen) zu erhalten.

Fazit 4.2

Direkt nach den Informationstexten sind die Vorstellungen der Lernenden angemessener und stärker in Übereinstimmung mit den explizierten Konzepten als vor den Informationstexten; wenige Minuten sowie eine Woche später (nächste Unterrichtsstunde) scheint dieser Effekt jedoch nur noch sehr gering zu sein (Abschnitt 6.9, ab S. 353).

Vor dem Hintergrund der oft nur geringen Wirkung von Erklärungen (zusammenfassend Wittwer & Renkl, 2008) findet sich eine kleine, aber nicht zu vernachlässigende Wirkung der Informationstexte. Die Befunde im direkten Umfeld der Informationstexte deuten darauf hin, dass Gelegenheiten zum Explorieren (vorlaufend zum Informationstext) sowie vor allem zum Üben des Anwendens der mitgeteilten Informationen (nachlaufend) hilfreich sind, um angestrebte Vorstellungen aufzubauen. Hier deutet sich an, dass vermutlich deutlich mehr (verschieden eingekleidete) Informationstexte mit zugehörigen Übungsphasen zu ein und demselben Konzept erforderlich wären, damit angestrebte Vorstellungen hinreichend stabil aufgebaut werden. Diese Deutung steht zwar im Einklang mit der Bedeutung, die dem (intelligenten) Üben für den Unterricht zugeschrieben wird (z. B. Hattie, 2009; Meyer, 2018), es ist jedoch in Bezug auf die eng gefassten adressierten Konzepte, die teilweise fast schon trivial anmuten, intuitiv eher verwunderlich. Für Forschung zur Wirkung von instruktionalen Informationen bzw. Erklärungen insbesondere im größeren instruktionalen/unterrichtlichen Kontext ergibt sich somit unter anderem, dass Wiederholungsgelegenheiten zentrale Analysestellen darstellen sollten. Ein derartiger Fokus würde auch Untersuchungen dazu erlauben, ob Wiederholungen tatsächlich in der aus der vorliegenden Arbeit ableitbaren Weise bedeutsam für die letztlich erzielte Wirkung von Erklärungen sind. Vor dem Hintergrund der scheinbar teilweise vorliegenden Notwendigkeit von immer wiederkehrenden Lernunterstützungen wie Prompts oder Scaffolds in Studien zum experimentbezogenen Denken und Arbeiten (z. B. Koenen, 2014; Arnold et al., 2014; siehe auch Unterabschnitt 2.4.2) wäre beispielsweise auch denkbar, dass Lernende Wiederholungen von Konzepten über sehr lange Zeiträume benötigen oder dass Lernende zu späteren Zeitpunkten keine vollständigen Wiederholungen benötigen, sondern ihnen Gedankenstützen (etwa »Denk daran, was eine naturwissenschaftliche Frage kennzeichnet« oder »Achte auf die Variablenkontrolle«) ausreichen. Vor dem Hintergrund der zum Teil großen bei der implizit-fachmethodischen Instruktionsvariante erreichten Kompetenzzuwächse könnte beispielsweise untersucht werden, ob die entsprechenden Lernenden eigenständig Wiederholungen (von zuvor eigenständig explizierten Konzepten) vornehmen.

Die Bedeutung von Explorations- und Übungsphasen lässt sich auch anhand eines weiteren Ergebnisses aus den Videoanalysen illustrieren, welches in den vorherigen Kapiteln der Arbeit nicht berichtet ist, weil es nicht-berichteten Analysen zu Instruktionsextrakten sowie dem Eindruck des Autors bei der Betrachtung der vollständigen Videos und Transkripte (bspw. bei Schulungen von Hilfskräften oder Materialsichtungen für Beispiele) entstammt. Über die beiden hereingegebenen Informationstexte, zu denen in Kapitel 6 Analysen berichtet sind, hinaus, finden sich in der explizit-fachmethodischen Instruktionsvariante weitere Informationstexte. Es scheint – etwas anders als für Kennzeichen von naturwissenschaftlichen Fragen, wozu in Kapitel 6 detaillierte Untersuchungen berichtet sind – so zu sein, dass Lernende bei explizit-fachmethodischer Instruktion die Informationstexte zum Teil nicht verstehen (beispielsweise, weil sie die adressierten Konzepte nicht sachangemessen wiedergeben bzw. nicht wiedergeben können) oder zumindest im direkten Nachgang nicht anwenden (beispielsweise bearbeiten sie nachfolgende Aufgaben zum Teil nicht korrekt, die durch angemessene Nutzung der Information korrekt bearbeitet werden könnten). Erst im weiteren Verlauf der Instruktion äußern die Lernenden anhand verschiedener Gelegenheiten zur Anwendung der Konzepte nach und nach angemessene Vorstellungen und handeln angemessen. Es ist plausibel, dass die Lernenden nach Informationstexten zunächst erneut Explorationen anstellen (Rogge, 2010; von Aufschnaiter & von Aufschnaiter, 2003) und sie ihre Kompetenz erst durch die gegebene Information *und* die nachlaufenden Gelegenheiten zur Anwendung entwickeln.

Die Analysen der Arbeit legen nahe, dass Sequenzen aus den Bearbeitungsprozessen untersucht wurden, die zentral für Prozesse des Kompetenzaufbaus zu den ausgewählten Konzepten des experimentbezogenen Denkens und Arbeitens sind. Allerdings wurde die spezifische Wirksamkeit der jeweils in sich geschlossenen Sequenzen aus zeitlich direkt aufeinanderfolgenden Explorationen, Informationen und Übungsaufgaben für den Kompetenzaufbau nicht untersucht. Ein erster Hinweis darauf, dass die Zusammenhänge nicht so eindeutig sind, wie zu erwarten wäre, lässt sich aus dem unerwarteten Befund ableiten, dass die Vorstellungsänderungen im Umfeld der hereingegebenen Information nur in einem nicht-signifikanten Zusammenhang von kleiner Effektstärke mit dem insgesamt erreichten Kompetenzzuwachs steht (ausführlicher diskutiert in Unterabschnitt 6.10.4, ab S. 406). Welche Bedeutung die innerhalb der Sequenzen aufzufindenden Vorstellungsänderungen für den insgesamt erreichten Kompetenzzuwachs hat, sollte daher in weiteren Studien detaillierter untersucht werden. Im Kontext der vorliegenden Arbeit könnten beispielsweise die Aufgaben aus Prä- und Posttest, die zu den adressierten Konzepten gehörige Fähigkeiten betreffen, aus den Testinstrumenten ausgelöst und einzeln betrachtet werden.

7.5 Prozesse des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten

Aus den Ergebnissen der Arbeit und den Diskussionen im vorliegenden Kapitel lassen sich auch Erkenntnisse und Hypothesen zu Prozessen des Kompetenzaufbaus ableiten, also dazu, wie sich das Denken und Handeln der Lernenden im zeitlichen Verlauf verändert. In Unterabschnitt 7.5.1 wird ausgeführt, wie sich Prozesse des Kompetenzaufbaus anhand der zeitlichen Veränderungen von Äußerungen mit fachmethodischem Gehalt beschreiben lassen. Ergebnisse der Arbeit werden aus diesem Blickwinkel diskutiert und ein ausführlicher Ausblick zum Einsatz von Konzeptualisierungsniveaus wird gegeben (als Fortführung von Ausführlicher Ausblick – Teil I, beginnend auf Seite 457).

In ersten Ansätzen finden sich Ergebnisse zur zeitlichen Veränderung von Vorstellungen bereits in Abschnitt 7.4. Diese werden in Unterabschnitt 7.5.2 rekapituliert und als Grundlage für einen kurzen Ausblick dazu herangezogen, wie eine Untersuchung von Vorstellungen im zeitlichen Verlauf aussehen kann und inwiefern diese für die Forschung relevant sein kann.

7.5.1 Zeitliche Veränderungen in den Charakteristika von Äußerungen mit fachmethodischem Gehalt

In Unterabschnitt 7.5.1.1 wird diskutiert, inwiefern es plausibel ist, dass erfolgreicher Kompetenzaufbau sich in den Prozessen mit einer Zunahme von Äußerungen einhergeht, die zwar als fachmethodisch angesehen werden können, aber keine expliziten Bezüge zu fachmethodischen Konzepten herstellen (in der vorliegenden Arbeit als fachmethodisch-vermutete Beiträge kategorisiert). In Unterabschnitt 7.5.1.2 werden ansatzweise in der Arbeit vorhandene Kategorien als Ausgangspunkt für einen Ausblick dazu genutzt, wie untersucht werden könnte, ob die Generalität der Äußerungen mit fachmethodischem Gehalt (d. h. deren Konzeptualisierungsniveau) zeitliche Veränderungen sinnvoll beschreibt.

7.5.1.1 Zeitliche Veränderung bezüglich der untersuchten Kategorien fachmethodischer Beiträge

Das in Unterabschnitt 5.4.5.1 (ab S. 184) berichtete Ergebnis, dass Lernende mit hohen Kompetenzzuwächsen für die drei Abschnitte der ersten Instruktionseinheit zunehmend höhere zeitliche Anteile fachmethodisch-vermuteter Beiträge aufweisen, die sich zudem mit zunehmend größeren Effekten von den geringen zeitlichen Anteilen

fachmethodisch-vermuteter Beiträge der Lernenden mit niedrigen Kompetenzzuwächsen unterscheiden, lässt sich eventuell im Sinne einer Veränderung im Denken und Handeln der Lernenden im zeitlichen Verlauf deuten: Erfolgreiche Prozesse des Kompetenzaufbaus gehen mit einer relativen Zunahme fachmethodisch-vermuteter Beiträge einher. Anders als fachmethodisch-indizierte Beiträge, bei denen durch Worte oder Begründungselemente vergleichsweise deutlich erkennbar ist, dass ein Bezug zu Fachmethoden (bspw. zu fachmethodischen Regeln oder Zuweisungen von unabhängigen/abhängigen Variablen) vorliegt, werden fachmethodisch-vermutete Beiträge solche Äußerungen kodiert, bei denen nur aus der Art oder dem unmittelbaren Kontext der Äußerung geschlossen werden kann, dass fachmethodische Überlegungen angestellt wurden. Damit ist verbunden, dass Lernende in fachmethodisch-vermuteten Beiträgen die fachmethodischen Regeln und Entscheidungen nicht in gleicher Deutlichkeit verbalisieren. Demnach könnte es sein, dass es bei weniger fortgeschrittenem Kompetenzaufbau bedeutsam ist, fachmethodische Regeln und Entscheidungen explizit zu verbalisieren (fachmethodisch-indizierte Beiträge), während es im späteren Verlauf genügt, verbal weniger sprachlich eindeutige Verknüpfungen von Fachmethoden und konkreten Situationen vorzunehmen (fachmethodisch-vermutete Beiträge).

Eine Einschränkung der Deutung zur Entwicklung über die Abschnitte der ersten Einheit der Instruktion hinweg besteht allerdings darin, dass bei einer Veränderung im Denken und Handeln der Lernenden im Sinne einer Entwicklung im Prozess des Kompetenzaufbaus zunächst nur davon ausgegangen werden sollte, dass *in vergleichbaren Situationen* unterschiedliche Aktivitäten ausgeführt werden. In gewisser Hinsicht sind die Situationen, die bearbeiteten Aufgaben, vergleichbar, weil sie demselben Instruktionsmaterial entstammen; allerdings gibt es deutliche Unterschiede zwischen den Aufgaben, bspw. darin, wie klar zur Explizierung eines Konzeptes aufgefordert wird (etwa durch »begründe«), die sich vermutlich auch darin widerspiegeln dürften, welche Kategorie fachmethodischer Beiträge überwiegt. Da allerdings die Deutung auf großen Abschnitten der Instruktion beruht, die viele Situationen umfassen, wird zunächst davon ausgegangen, dass sich die Unterschiede, die zwischen einzelnen Aufgaben existieren, herausmitteln, d. h. ausgleichen. Inwiefern dies korrekt ist, müssten detailliertere Analysen zeigen.

Ein weiteres Ergebnis der Arbeit lässt sich – allerdings mit sehr starken Einschränkungen – in vergleichbarer Weise deuten, weil es ebenfalls nahelegen könnte, dass es bei erfolgreichem Kompetenzaufbau im späteren Verlauf genügt, verbal weniger eindeutige Verknüpfungen von Fachmethoden und konkreten Situationen vorzunehmen:

In Einheit 3 hängen fachmethodisch-vermutete Beiträge signifikant und deutlich stärker als in Einheit 1 mit den Kompetenzzuwächsen zusammen (genauer in Unterunterabschnitt 5.4.4.1, ab S.175); ferner finden sich für Einheit 3 nur schwach ausgeprägte Zusammenhänge fachmethodisch-indizierter Beiträge mit den Kompetenzzuwächsen. Wenn davon ausgegangen wird, dass über alle drei Einheiten hinweg eine zusammenhängende Kompetenz untersucht wird und eine Aufgliederung in Teilprozesse oder gar Teilkompetenzen zu den Teilprozessen wenig zielführend ist, könnten die Einheiten 1 und 3 auch als Ausschnitte aus einem integrierten Kompetenzaufbau aufgefasst werden: ein Ausschnitt früher und einer später im Kompetenzaufbau. Solch ein Integrieren scheint aus psychometrischer Perspektive eine Möglichkeit darzustellen, aus fachdidaktischer Sicht vermutlich erscheint es allerdings eher nicht sinnvoll (vgl. Abschnitt 2.2, ab S.13). Zudem können die Videoaufzeichnungen zu Einheit 1 und zu Einheit 3 der Instruktion zwar durchaus als zwei zu analysierende zeitliche Ausschnitte aus einem Kompetenzaufbau aufgefasst werden, die Vergleichbarkeit der Einheiten ist allerdings durch die unterschiedlichen adressierten Fähigkeiten und Konzepte deutlich eingeschränkt. Wird trotz all dieser Einschränkungen ein Vergleich der Einheiten als Vergleich zweier Analysen zur selben Kompetenz angesehen, dann unterstützt das genannte Ergebnis, die Deutung, dass fachmethodisch-vermutete Beiträge in erfolgreichen Prozessen des Kompetenzaufbaus später eine Funktion haben, während fachmethodisch-indizierte Beiträge vor allem zu Beginn von Bedeutung sind.

Mit den vorgestellten Interpretationen vergleichbare Befunde fänden sich für fachinhaltliches Lernen beispielsweise bei Saniter (2003) oder Roth (2003); verwunderlich wäre allerdings, dass die Ergebnisse in der vorliegenden Arbeit für einen Zeitraum einiger Minuten (oder in Bezug auf den Vergleich von Einheiten 1 und 3: weniger Unterrichtsstunden) berichtet werden können, während es sich in den zitierten Studien um Unterschiede zwischen fortgeschrittenen Physikstudierenden und Diplomphysikern handelt (also einen deutlich größeren zeitlichen Unterschied). Bei den untersuchten kurzen Zeitspannen dürfte wahrscheinlich vielmehr für die zu Beginn vermutlich wenig fortgeschrittenen Lernenden analog zu den Ergebnissen bei Rogge (2010; dort zu Fachinhalten) davon auszugehen sein, dass – für jedes spezifische Konzept einzeln betrachtet – die Anteile der Explizierungen zu den Konzepten durch die Lernenden im zeitlichen Verlauf im Mittel zunehmen. Eine Rekodierung der fachmethodischen Beiträge anhand von spezifisch auf Explizierungen angelegten Kategorien (etwa in Anlehnung an die Niveaus aus den Studien von Saniter, 2003, Roth, 2003, oder Rogge, 2010) könnten dazu genaueren Aufschluss geben, unter anderem, weil die Kategorien fachmethodisch-indizierter und -vermuteter Beiträge sich zwar mit Blick

7.5 Prozesse des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten

auf Explizierungen von Regeln deuten lassen, allerdings nicht spezifisch zu diesem Zweck konzipiert wurden. Dabei könnte dann auch die relevante Frage untersucht werden, ob sich Veränderungen hinsichtlich der Explizitheit der fachmethodischen Beiträge nur gemittelt auf Kohortenebene oder auch für (manche der) Individuen (in ggf. unterschiedlicher Weise) zeigen (auswertungsmethodisch z. B. ähnlich dazu, wie es für manche Kategorien zu Vorstellungen in Kapitel 6 vorgenommen wurde).

Prozesse des Kompetenzaufbaus könnten sich auch darin widerspiegeln, dass sich die Reihenfolgen von Aktivitäten ändern. So wäre es beispielsweise denkbar, dass fachmethodische Begründungen von den Lernenden zunächst nur auf Aufforderung und nach dem Ausführen einer fachmethodischen Tätigkeit (z. B. dem Formulieren von Fragen oder dem Durchführen eines Versuchs) gegeben werden, später im weiter fortgeschrittenen Kompetenzaufbau allerdings mehr und mehr eine steuernde Funktion einnehmen, die vermutlich auch daran ersichtlich sein dürfte, dass die Beiträge zeitlich vor dem Ausführen der Tätigkeiten auftreten. In ersten Ansätzen zeigt sich in den Analysen der Arbeit, dass fachmethodische Beiträge und insbesondere Begründungen vorrangig auf Aufforderung und erst nachgeschaltet auftreten (VERWEIS). Die in der Arbeit genutzten Kodierungen könnten für zukünftige Analysen zu Reihenfolgen von Aktivitäten genutzt werden; zur Beantwortung der Frage, ob fachmethodische Beiträge begründende Funktion haben oder lediglich die getroffene Entscheidung beschreiben, könnten die Kodierungen zu den Vorstellungen ansatzweise genutzt werden. Derartige erste Einblicke wären allerdings durch spezifischere Analysen der inhaltlichen Zusammenhänge der jeweiligen (verbalen und nonverbalen) Aktivitäten zu ergänzen, weil die eingesetzten Kodierungen nicht abbilden, ob zwei aufeinander folgende Aktivitäten inhaltlich aufeinander bezogen sind.

7.5.1.2 Zeitliche Veränderung in der Generalität der Äußerungen mit fachmethodischem Gehalt

Die in Unterunterabschnitt 7.5.1.1 geschilderten Ergebnisse könnten so interpretiert werden, dass Nennungen fachmethodischer Konzepte im zeitlichen Verlauf tendenziell abnehmen. Dies wurde in der Breite allerdings nicht spezifisch untersucht. Einzig für die Konzepte zu Kennzeichen von naturwissenschaftlichen Fragen und Beobachtungen und Deutungen finden sich Ergebnisse, aus denen allerdings eine derartige Abnahme nicht gefolgert werden kann. Vielmehr finden sich dort Generalisierungen insbesondere bei den Aufgaben, die Lernende gezielt dazu auffordern. Weitere Untersuchungen sind also notwendig, um spezifischere Bezüge herzustellen und Aussagen

dazu abzuleiten, welche Funktion das (eigenständig initiierte oder instruktional aufgeforderte) Verbalisieren von fachmethodischen Konzepten im Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten hat.

Ausführlicher Ausblick (Teil II von II)

Entwickeltes Kategoriensystem zu Konzeptualisierungsniveaus zur Analyse von zeitlicher Veränderung nutzen

Wie bereits in Unterunterabschnitt 7.5.1.1 ausgeführt, kann eine Untersuchung von Explizierungen von Konzepten zeigen, ob sich in den Prozessen des Kompetenzaufbaus nach und nach mehr Explizierungen von Konzepten finden und somit dadurch eine zeitliche Entwicklung im Kompetenzaufbau beschrieben werden kann. Anhand der bereits in dem ersten Teil des ausführlichen Ausblicks (beginnend auf Seite 457) vorgestellten Kategorien zu den Konzeptualisierungsniveaus kann nicht nur dieser Frage nachgegangen werden, sondern sogar detaillierter beantwortet werden, ob die Äußerungen mit fachmethodischem Gehalt für die einzelnen Lernenden nach und nach immer stärker von den konkreten Objekten und fachmethodischen Entitäten losgelöst formuliert werden. Für die sechs probenhalber untersuchten Personen aus den beiden ausblickshalber betrachteten Teams sind im Anschluss an die grauen Kästen erste Ergebnisse skizziert und Anmerkungen notiert, wie weitergearbeitet werden könnte, um Entwicklungen des Denkens und Handelns im zeitlichen Verlauf des Kompetenzaufbaus zum experimentbezogenen Denken und Arbeiten aufzulösen.

Mithilfe der Ausgabefunktion von MAXQDA wurden für den Ausblick Verlaufsdiagramme erstellt, in denen die Kodierungen zu den Konzeptualisierungsniveaus im zeitlichen Verlauf visualisiert sind. Diese finden sich in Tabelle 7.4 (auf S. 484). Der Abstand zwischen zwei vertikalen langen hellgrauen Strichen entspricht einem Zeitintervall von 10 Minuten Dauer. Da Team H insgesamt länger an der dritten Einheit der explizit-fachmethodischen Instruktionsvariante gearbeitet hat als Team G, finden sich hier auch mehr vertikale lange hellgraue Striche. Die kurzen vertikalen Striche symbolisieren die Kodierungen der fachmethodischen Beiträge auf dem jeweiligen Konzeptualisierungsniveau. Die breite der kurzen Striche steht für den zeitlichen Umfang der jeweils zugehörigen Äußerung; es kann allerdings in seltenen Fällen vorkommen, dass zwei Kodierungen so nah beieinander liegen, dass in der Visualisierung daraus ein dicker Strich wird. Um die Visualisierung nicht mit Details zu überfrachten, wurden die eigentlichen Konzeptualisierungsniveaus berichtet, nicht

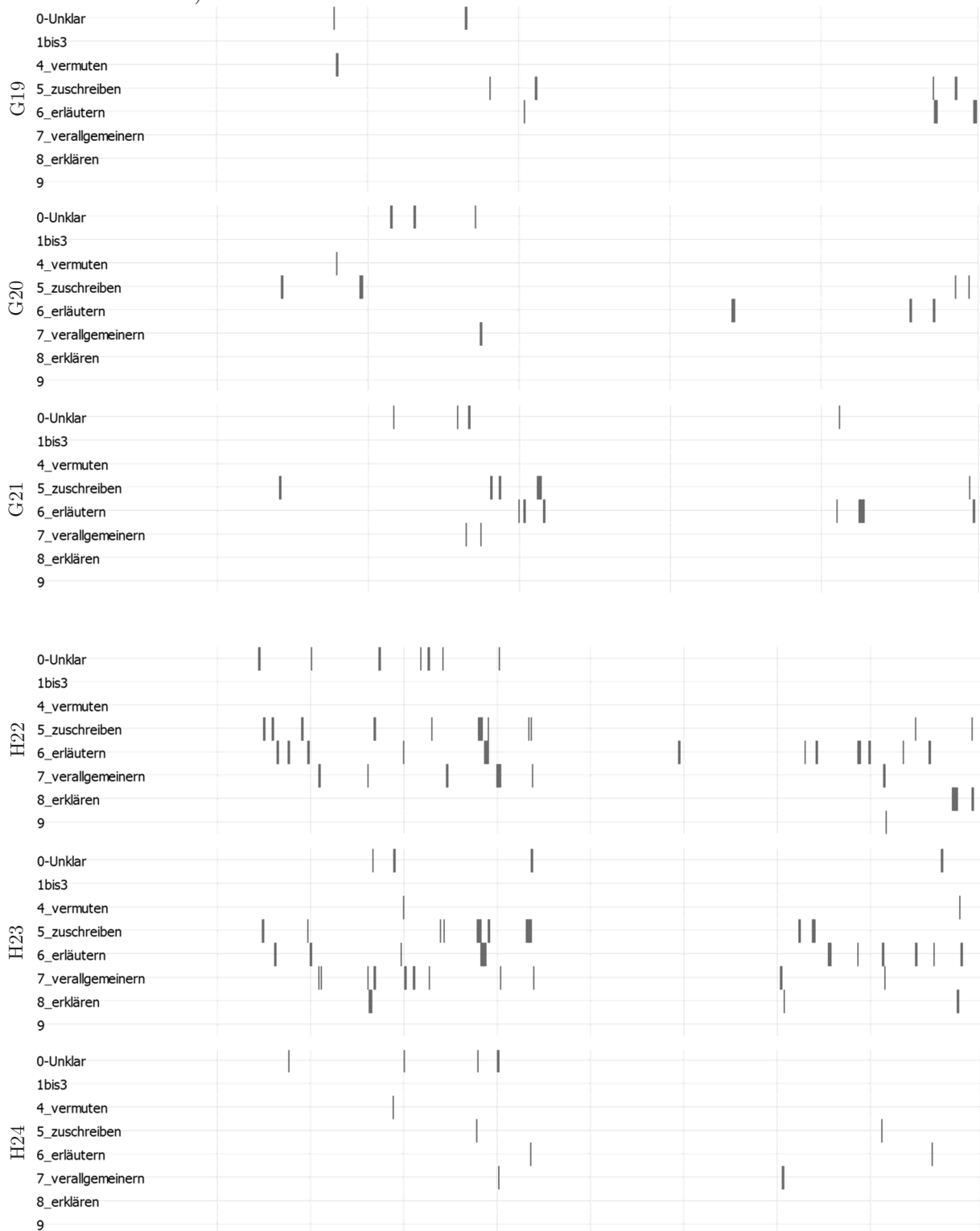
jedoch die Subkategorien, aus denen diese Zusammengesetzt sind (z. B. 7 *verallgemeinern* statt 7a und 7b; vgl. Tabelle 7.1 auf den Seiten 458 und 459). Weitere Aufbereitungen könnten Trennlinien enthalten, anhand derer auf die im jeweiligen Zeitabschnitt bearbeiteten Karten geschlossen werden könnte; dadurch könnte die Vergleichbarkeit der Darstellungen über mehrere, unterschiedlich schnell arbeitende Teams, gesichert werden.

Aus Tabelle 7.4 (auf S. 484) geht zunächst hervor, was bereits in Unterunterabschnitt 7.2.3.3 (ab S. 456) benannt wurde: Die beiden Lernenden mit den hohen Kompetenzzuwächsen weisen deutlich mehr fachmethodische Beiträge auf (deutlich mehr Striche) und nur für die Lernenden mit den hohen Kompetenzzuwächsen finden sich substantielle Anzahlen bei den höheren Konzeptualisierungsniveaus (in der Grafik horizontal gespiegelt: weiter unten liegende Bereiche sind für Lernende mit niedrigen Kompetenzzuwächsen sehr wenig gefüllt). Vergleichbar ist hingegen für alle Lernenden der grobe Zeitraum, in dem sich Häufungen von fachmethodischen Beiträgen finden, sowie auch – recht grob betrachtet – das jeweils zunächst erreichte Konzeptualisierungsniveau. Dies legt nahe, dass der Ausgangspunkt fachmethodischer Beiträge bis auf wenige Ausnahmen, die in der Grafik auch als optisch alleinstehend auffallen, die Instruktion ist. Die höheren Konzeptualisierungsniveaus, die sich für einige der Lernenden finden, scheinen grundsätzlich auf einige in zeitlicher Nähe davor liegende fachmethodische Beiträge zu folgen, die ein oder zwei Konzeptualisierungsniveaus darunter liegen (Ausnahme: ca. Minute 70 für H23 und H24). Diese Ergebnisse passen sehr gut zu den Ergebnissen von Rogge (2010) für fachinhaltlichen Konzeptaufbau.

An die geschilderten Ergebnisse anschließend lässt sich zukünftig nun zum einen auf die Ausnahmen fokussieren. Was führt zu fachmethodischen Beiträgen, die nicht innerhalb der großen (vermutlich instruktionsangeregten) Ansammlungen, vorgefunden werden können? Wodurch sind fachmethodische Beiträge (in sich selbst oder in ihrem Umfeld) gekennzeichnet, die hohe Konzeptualisierungsniveaus aufweisen, aber nicht am Ende, sondern eher am Anfang von Kaskaden fachmethodischer Beiträge stehen? »Steuern« diese beispielsweise die nachfolgenden (Sprech-)Handlungen? Zum anderen lässt sich dem aufgefundenen grundsätzlichen Muster der Konzeptualisierungsniveaus nachgehen, indem u. a. geprüft wird, ob die aufsteigenden Kaskaden auch inhaltlich zusammenhängen (d. h. zum selben fachmethodischen Konzept), oder indem für größere Stichproben geprüft wird, inwiefern Lernende mit zu hohen Konzeptualisierungsniveaus aufsteigenden Kaskaden fachmethodischer Beiträge tatsächlich höhere Kompetenzzuwächse erreichen. Auch eine Analyse der Passung von Anfor-

7 Zusammenführungen, Diskussionen und Ausblicke

Tabelle 7.4: Konzeptualisierungsniveaus der sechs ausblicksartig untersuchten Personen zeitlich aufgelöst ((ca. 10 Minuten Bearbeitungsdauer pro vertikalem grauem durchgezogenem Strich).



derung der Instruktion (Fordert die Aufgabe eine Begründung?) und erreichtem Konzeptualisierungsniveau (Wird von der Person eine Begründung – Niveau 6 – gegeben?) dürfte ertragreich sein, u. a. zur Beantwortung der Frage, inwiefern eine anforderungskonforme Bearbeitung bei Lernenden mit niedrigen sowie Lernenden mit hohen Kompetenzzuwächsen vorliegt, und ob es nötig ist, die auf Ebene der Konzeptualisierungsniveaus durch die Instruktion gestellten Anforderungen zu übertreffen, um zu hohen Kompetenzzuwächsen zu gelangen.

7.5.2 Zeitliche Veränderung von Vorstellungen

Die in Kapitel 6 berichteten Analysen liefern u. a. Ergebnisse dazu, welche Vorstellungen sich für die Lernenden im zeitlichen Verlauf rekonstruieren lassen. In Abschnitt 7.4 (ab S. 475) wurde diskutiert, welche Rolle den instruktionalen Informationstexten im Vorstellungsaufbau zukommt. Nachfolgend wird darauf eingegangen, inwiefern die Analysen und Ergebnisse der Arbeit erste Ansätze für generelle Beschreibungen von Prozesse des Vorstellungsaufbaus darstellen (können).

In Unterabschnitt 6.9.2 (ab S. 380) wurden verschiedene Profile untersucht, die sich im zeitlichen Umfeld des zu Kennzeichen von naturwissenschaftlichen Fragen eingesetzten instruktionalen Informationstextes vorfinden lassen. Dabei wurden die vor und nach der Information stabil auftretenden Vorstellungen betrachtet. Für etwa die Hälfte der Lernenden finden sich dann Profile im zeitlichen Verlauf, aus denen sich bei einer inhaltlichen Betrachtung der jeweils stabil auftretenden Vorstellungen ableiten lässt, dass in den meisten Fällen nach der Information angemessene Vorstellungen stabil auftreten – solche, die bereits vorher vorlagen, oder neu hinzukommende. Nur im Falle des Profils, in dem alle stabilen Vorstellungen sich wohl vor als auch nach der Information vorfinden (d. h. keine der Vorstellungen wird abgelegt oder neu aufgenommen; Profil B), finden sich vorrangig Vorstellungen, die aus fachmethodischer Sicht eher unangemessen sind. Hieraus lässt sich die Vermutung ableiten, dass unterschiedliche Lernende voneinander abweichende Verläufe des Vorstellungsaufbaus aufweisen und dass dies eventuell darin begründet liegt, dass sie unterschiedlich stark an bereits Gelerntem festhalten. Um zu untersuchen, inwiefern tatsächlich abweichende Verläufe ein Charakteristikum von Vorstellungsaufbau zum experimentbezogenen Denken und Arbeiten sind, sind allerdings sowohl hinsichtlich der untersuchten Konzepte als auch der untersuchten Personen größere Anzahlen erforderlich, da die Ergebnisse Artefakte des spezifischen Umfelds der untersuchten instruktionalen Informationen darstellen könnten. Ferner zeigt die Tatsache, dass sich nur für etwa die Hälfte der Lernenden Profile im zeitlichen Verlauf ausmachen lassen, weil sich für die andere Hälfte keine stabilen Vorstellungen finden lassen, dass weitere

Analysen auch Wege finden müssten, um stärker fluide Äußerungen der Lernenden zu berücksichtigen; ein Ignorieren solcher Äußerungen scheint zwar mit Blick auf eine Rekonstruktion von besonders prominenten Vorstellungen sowie aus methodischer Sicht zur Reduktion von Fehlschlüssen bei der Rekonstruktion der Vorstellungen angemessen zu sein, für Beschreibungen von Verläufen der Vorstellungsentwicklung dürften aber gerade diese z. T. singulären Äußerungen bedeutsam sein.

Die Analysen der Arbeit sind vor allem auf sehr kurze Zeitskalen von wenigen Minuten bezogen. Innerhalb dieser kurzen Zeit werden zwar relativ viele Aufgaben bearbeitet, in deren Kontext Äußerungen getroffen werden, aus denen sich auf die Vorstellungen der Lernenden zum experimentbezogenen Denken und Arbeiten schließen lässt; allerdings weisen die nur ansatzweise in den Blick genommenen eine Unterrichtsstunde später auftretenden Äußerungen der Lernenden darauf hin, dass sich auf einer größeren Zeitskala abweichende Verläufe einstellen könnten (bspw. erneute Rückschritte im Verständnis). Eine Herausforderung zukünftiger Studien wird darin bestehen, sowohl die kurzen Zeitskalen als auch die etwas größeren Zeitskalen zu untersuchen und miteinander in Verbindung zu bringen. Es ist denkbar, dass Verständnisse zu einzelnen Konzepten des experimentbezogenen Denkens und Arbeitens mehrfach in ähnlicher Weise aufgebaut werden müssen, damit sie langfristig vorhanden sind; dann würden sich auf kurzen Zeitskalen sehr vergleichbare Verläufe zeigen, obwohl die kurzen Zeitabschnitte an unterschiedlichen Stellen im größeren zeitlichen Bild platziert wären, und es würde sich die Frage stellen, woran Entwicklung hin zu stabilen angemessenen Vorstellungen sichtbar wird.

Die Bereite der zu den Vorstellungen vorgenommenen Analysen zeigt, dass nicht nur eine Betrachtung der Inhalte der Vorstellungen vorgenommen werden kann, wenn zeitliche Änderung von Vorstellungen untersucht werden soll. Beispielsweise kann auch die Stabilität der Vorstellungen oder die Generalität der Äußerungen, aus denen die Vorstellungen rekonstruiert werden, in den Blick genommen werden. Die in Abschnitt 6.7 (ab S. 312) vorgenommene Verknüpfung von Vorstellungen und Entscheidungen zeigt, dass ferner für eine Untersuchung des Kompetenzaufbaus ein Aufeinanderbeziehen von Vorstellungen und Fähigkeiten relevant sein dürfte. Es konnten beispielsweise aufgrund der Anlage der untersuchten Ausschnitte der Instruktion bisher keine Aussagen über zeitliche Verläufe der Zusammenhänge von Vorstellungen und Fähigkeiten getroffen werden. Anschließend an das Ergebnis, dass für angemessene Entscheidungen nicht notwendigerweise angemessene Vorstellungen notwendig sind, lässt sich beispielsweise fragen, ob Fähigkeitsaufbau dem Vorstellungsaufbau zeitlich vorausgeht.

8 Schluss

Ausgehend von einer typischerweise auf die Messung, Modellierung und Förderung von Kompetenzen gerichteten Forschungslandschaft zum Experimentieren wurde in der vorliegenden explorativen Arbeit der Aufbau von Kompetenzen zum experimentbezogenen Denken und Arbeiten untersucht. Vor dem Hintergrund der Bedeutung, die Informationen und expliziter Fokussierung auf Fachmethoden in der Forschungslandschaft zukommt, wurde für eine explizit-fachmethodische Instruktion, deren Kompetenzförderlichkeit bereits für die betrachtete Population nachgewiesen war, das Denken und Handeln der Lernenden bei der Bearbeitung in den Blick genommen.

Anhand von Vergleichen von Lernenden mit hohen und niedrigen Kompetenzzuwächsen und Vergleichen mit einer parallel angelegten implizit-fachmethodischen Instruktion wurde die Relevanz davon, wie viel Lernende tatsächlich über Fachmethoden sprechen, mithilfe kategorienbasierter Videoanalysen untersucht. Es zeigt sich unter anderem, dass die zeitlichen Anteile von Äußerungen mit fachmethodischem Gehalt in einem (nicht ganz systematischen) Zusammenhang mit den erreichten Kompetenzzuwächsen stehen und dass sich einige Charakteristika fachmethodischer Beiträge ausmachen lassen, die besonders relevant für den Kompetenzaufbau sein dürften. Indem Instruktion und Aktivitäten der Lernenden aufeinander bezogen wurden, konnten einige Annahmen zu Wirkmechanismen von fachmethodischer Instruktion genauer beleuchtet werden.

Die Ergebnisse sind jeweils mit Einschränkungen verbunden, insbesondere dadurch, dass nur zwei vergleichsweise stark vorstrukturierende und zudem strukturell ähnliche Instruktionen betrachtet wurden und dass die Zahl der analysierten Lernenden angesichts der Heterogenität der Population und angesichts der Arbeit in Teams eher klein ist. Die eingesetzten Methoden der Kodierung der Videos und der Auswertung der Kodierungen (u. a. durch Clusteranalysen kompletter Aktivitätsprofile und durch instruktionsabschnittsspezifische Betrachtungen, die zum Teil zeitliche Auflösung ermöglichen) konnten als Grundlage für weitere Forschung erprobt werden; angesichts der generierten Ergebnisse dürfte eine Ergänzung um spezifisch inhaltsbezogene und strukturmerkmalbezogene Analysen der Äußerungen mit fachmethodischem Gehalt angebracht sein, wenn nach spezifischeren Beschreibungen von kompetenzförderlichen Aktivitäten der Lernenden und (zugehörigen) kompetenzförderlichen instruktionalen Elementen gefragt wird.

Die in der Arbeit vorgenommenen Analysen zu den Vorstellungen der Lernenden, die sich aus Äußerungen mit fachmethodischen Gehalt zu zwei betrachteten Konzepten rekonstruieren lassen, geben erste Einblicke in Inhalte und Strukturmerkmale (wie bspw. Fallbezogenheit der Äußerungen, aus denen Vorstellungen rekonstruiert werden, oder zeitliche Stabilität der Vorstellungen). Die vorgenommenen Verknüpfungen der detaillierten Vorstellungsanalysen mit den erreichten Kompetenzzuwächsen sind allerdings unter anderem dadurch eingeschränkt, dass die kurzen Ausschnitte der Instruktion, für die die Vorstellungen analysiert wurden, nur einzelne Aspekte des Konstrukts betreffen, welches der Messung der Kompetenzzuwächse zugrunde liegt. Es wurde allerdings diskutiert, dass und wie weitere Arbeiten diesem Umstand abhelfen können.

Die zu den Vorstellungen der Lernenden generierten Ergebnisse sind weniger durch die Brille der Bezüge zu den Kompetenzzuwächsen zu betrachten, sondern vielmehr als detaillierte Auflösungen von im Kompetenzaufbau rekonstruierten Vorstellungen zu Konzepten des experimentbezogenen Denkens und Arbeitens zu sehen, die Ausgangspunkte für weitere Forschung zu solchen Vorstellungen darstellen können. Beispielsweise geben die in der Arbeit berichteten Analysen nicht nur Hinweise darauf, welche Vorstellungen bei Lernenden auftreten, sondern auch, wie stabil sie im zeitlichen Verlauf aktiviert werden. Auch der hergestellte Bezug von instruktionalen Informationen mit den bei Lernenden rekonstruierten Vorstellungen liefert Hinweise wie z. B., dass Wiederholungsgelegenheiten sowohl mit Blick auf Förderung als auch auf Forschung zentrale Stellen von Instruktion sind. Das zur Generierung der Ergebnisse eingesetzte und in der Arbeit ausführlich geschilderte qualitativ-inhaltsanalytische Verfahren wird vor dem Hintergrund theoretischer Annahmen diskutiert; es lässt sich auch auf weitere Konzepte des experimentbezogenen Denken und Arbeiten übertragen und ist somit ebenfalls als ein Ertrag der Arbeit anzusehen.

Ein zentraler Beitrag der explorativen Arbeit besteht darin, einen Kontext und eine Motivation für weitere Studien zu liefern, die spezifischer auf einzelne Aspekte des fachmethodischen Kompetenzaufbaus eingehen. Insbesondere mit Blick auf entlang von Konzepten des experimentbezogenen Denkens und Arbeitens angelegte explizit-fachmethodische Instruktion motiviert die vorliegende Arbeit eine Reihe anschließender Studien. Beispielsweise stellen die in der Einleitung thematisierten Selbsterklärungen von Lernenden gewissermaßen Spezialfälle von den untersuchten Äußerungen mit fachmethodischem Gehalt dar, deren Beziehung zu instruktionalen Elementen geklärt werden könnte. Und auch die Bedeutung von fachmethodischen und fachinhaltlichen Eingangsvoraussetzungen, auf die in der vorliegenden Arbeit

nur randständig eingegangen wurde, lässt sich z. B. durch Differenzierung ausgewählter Ergebnisse der Arbeit für unterschiedliche Lernende untersuchen. Zudem stellen die zum Teil ausführlich diskutierten Hypothesen und Desiderata in Kapitel 7 Ausgangspunkte für Studien dar.

Zusammenfassend können die in der vorliegenden Arbeit eingesetzten Methoden und generierten sowie ausführlich diskutierten Ergebnisse Ausgangspunkte für weitere Forschung zum bisher nur in Ansätzen beleuchteten Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten darstellen. Damit kann die Arbeit mittelbar auch einen Beitrag dazu leisten, Unterricht zu verbessern, der Kompetenzen zum experimentbezogenen Denken und Arbeiten fördert.

Literaturverzeichnis

- Abd-El-Khalick, F., BouJaoude, S., Duschl, R., Lederman, N. G., Mamlok Naaman, R., Hofstein, A., Niaz, M., Treagust, D. & Tuan, H. L. (2004). Inquiry in Science Education: International Perspectives. *Science Education*, 88(3), 397–419. <https://doi.org/10.1002/sce.10118>
- Abrahams, I. & Millar, R. (2008). Does Practical Work Really Work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education*, 30(14), 1945–1969. <https://doi.org/10.1080/09500690701749305>
- Adusah, A. K. & Brooks, G. P. (2011). Type I Error Inflation of the Separate-Variances Welch t test with Very Small Sample Sizes when Assumptions Are Met. *Journal of Modern Applied Statistical Methods*, 10(1), 362–372. <https://doi.org/10.22237/jmasm/1304224320>
- Alfieri, L., Brooks, P. J., Aldrich, N. J. & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, 103(1), 1–18. <https://doi.org/10.1037/a0021017>
- Allie, S., Buffler, A., Campbell, B. & Lubben, F. (1998). First-year physics students' perceptions of the quality of experimental measurements. *International Journal of Science Education*, 20(4), 447–459. <https://doi.org/10.1080/0950069980200405>
- Amsel, E. & Brock, S. (1996). The development of evidence evaluation skills. *Cognitive Development*, 11(4), 523–550. [https://doi.org/10.1016/S0885-2014\(96\)90016-7](https://doi.org/10.1016/S0885-2014(96)90016-7)
- Arndt, K. (2016). *Experimentierkompetenz erfassen*. Logos-Verl.
- Arnold, J., Kremer, K. & Mayer, J. (2016). Concept Cartoons als diskursiv-reflexive Szenarien zur Aktivierung des Methodenwissens beim Forschenden Lernen. *Zeitschrift für Didaktik der Biologie*, 20, 33–43. <https://doi.org/10.4119/zdb-1636>
- Arnold, J., Kremer, K. & Mayer, J. (2017). Scaffolding beim Forschenden Lernen. *Zeitschrift für Didaktik der Naturwissenschaften*, 23(1), 21–37. <https://doi.org/10.1007/s40573-016-0053-0>
- Arnold, J. C., Kremer, K. H. & Mayer, J. (2014). Understanding Students' Experiments—What kind of support do they need in inquiry tasks? *International*

- Journal of Science Education*, 36(16), 2719–2749. <https://doi.org/10.1080/09500693.2014.930209>
- Arnold, J. C., Wellnitz, N. & Mayer, J. (2010). Beschreibung und Messung von Beobachtungskompetenz bei Schülerinnen und Schülern der Sekundarstufe I. *Erkenntnisweg Biologiedidaktik*, 10, 7–22. Zugriff am 1. Februar 2021 unter https://www.bcp.fu-berlin.de/biologie/arbeitsgruppen/didaktik/Erkenntnisweg/2010/2010_2_arnold.pdf
- Australian Curriculum, Assessment and Reporting Authority. (2015). The Australian Curriculum: Science – Version 7.5 (15.05.2015). Zugriff am 10. September 2015 unter <http://www.australiancurriculum.edu.au/download/f10>
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. Holt Rinehart and Winston.
- Beretz, A.-K. (2013). *Wie verstehen und bearbeiten Schüler/innen Aufgaben zum Kompetenzbereich Erkenntnisgewinnung? Videogestützte Analyse des Einsatzes eines Testinstrumentes* (Wissenschaftliche Hausarbeit im Rahmen der Ersten Staatsprüfung für das Lehramt an Gymnasien). Justus-Liebig-Universität. Gießen, Physik. Zugriff am 12. September 2015 unter <https://www.uni-giessen.de/cms/fbz/fb07/fachgebiete/physik/einrichtungen/didaktik/doku/AnnKathrin-2013>
- Blanchard, M. R., Southerland, S. A., Osborne, J. W., Sampson, V. D., Annetta, L. A. & Granger, E. M. (2010). Is inquiry possible in light of accountability? A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Science Education*, 94(4), 577–616. <https://doi.org/10.1002/sce.20390>
- Börlin, J. (2012). *Das Experiment als Lerngelegenheit: Vom interkulturellen Vergleich des Physikunterrichts zu Merkmalen seiner Qualität*. Logos-Verl.
- Bortz, J. & Weber, R. (2005). *Statistik für Human- und Sozialwissenschaftler: Mit 242 Tabellen* (6., vollst. überarb. und aktualisierte Aufl.). Springer Medizin.
- Boudreaux, A., Shaffer, P. S., Heron, P. R. L. & McDermott, L. C. (2008). Student understanding of control of variables: Deciding whether or not a variable influences the behavior of a system. *American Journal of Physics*, 76(2), 163–170. <https://doi.org/10.1119/1.2805235>
- Brennan, R. L. & Prediger, D. J. (1981). Coefficient kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, 41(3), 687–699.

- Bullock, M. & Ziegler, A. (1999). Scientific reasoning: Developmental and individual differences. In F. E. Weinert & W. Schneider (Hrsg.), *Individual development from 3 to 12* (S. 38–54). Cambridge Univ. Press.
- Bybee, R. W. (1997). *Achieving scientific literacy: From purposes to practices*. Heinemann.
- Čančula, M. P., Planinšič, G. & Etkina, E. (2015). Analyzing patterns in experts' approaches to solving experimental problems. *American Journal of Physics*, 83(4), 366–374. <https://doi.org/10.1119/1.4913528>
- Chalmers, A. F. (2001). *Wege der Wissenschaft: Einführung in die Wissenschaftstheorie* (5., völlig überarb. und erw. Aufl.). Springer.
- Chen, Z. & Klahr, D. (1999). All Other Things Being Equal: Acquisition and Transfer of the Control of Variables Strategy. *Child Development*, 70(5), 1098–1120. <https://doi.org/10.1111/1467-8624.00081>
- Chi, M. T., de Leeuw, N., Chiu, M.-H. & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18(3), 439–477. https://doi.org/10.1207/s15516709cog1803_3
- Chinn, C. A. & Malhotra, B. A. (2002). Children's responses to anomalous scientific data: How is conceptual change impeded? *Journal of Educational Psychology*, 94(2), 327–343. <https://doi.org/10.1037//0022-0663.94.2.327>
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Cohen, J. W. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. Aufl.). Erlbaum.
- Crawford, B. A. (2014). From Inquiry to Scientific Practices in the Science Classroom. In N. G. Lederman & S. K. Abell (Hrsg.), *Handbook of research on science education* (S. 515–541). Routledge.
- de Winter, J. (2013). Using the Student's t-test with extremely small sample sizes. *Practical Assessment, Research, and Evaluation*, 18. <https://doi.org/10.7275/E4R6-DJ05>
- Dean, D. & Kuhn, D. (2007). Direct instruction vs. discovery: The long view. *Science Education*, 91(3), 384–397. <https://doi.org/10.1002/sce.20194>
- Delacre, M., Lakens, D. & Leys, C. (2017). Why Psychologists Should by Default Use Welch's t-test Instead of Student's t-test. *International Review of Social Psychology*, 30(1), 92. <https://doi.org/10.5334/irsp.82>

- Department for Education. (2013). National curriculum in England – Science programmes of study: key stage 3 [September 2013]. Zugriff am 15. März 2021 unter https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335174/SECONDARY_national_curriculum_-_Science_220714.pdf
- Dickmann, M. (2016). *Messung von Experimentierfähigkeiten*. Logos-Verl.
- Ding, L. (2018). Progression trend of scientific reasoning from elementary school to university: A large-scale cross-grade survey among chinese students. *International Journal of Science and Mathematics Education*, 16(8), 1479–1498. <https://doi.org/10.1007/s10763-017-9844-0>
- Ditton, H. (2002). Unterrichtsqualität – Konzeptionen, methodische Überlegungen und Perspektiven. *Unterrichtswissenschaft*, 30(3), 197–212.
- Dobber, M., Zwart, R., Tanis, M. & van Oers, B. (2017). Literature review: The role of the teacher in inquiry-based education. *Educational Research Review*, 22, 194–214. <https://doi.org/10.1016/j.edurev.2017.09.002>
- Dresing, T. & Pehl, T. (2017). *Praxisbuch Interview, Transkription & Analyse: Anleitungen und Regelsysteme für qualitativ Forschende* (7. Auflage). Eigenverlag. <https://books.google.de/books?id=tzQIR8Hl36QC>
- Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp & H. Reinders (Hrsg.), *Implicit and explicit knowledge in second language learning, testing and teaching* (S. 3–25). Multilingual Matters.
- Emden, M. (2011). *Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens: Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I*. Logos-Verl.
- Erpenbeck, J. & von Rosenstiehl, L. (2007a). Einführung. In J. Erpenbeck & L. von Rosenstiehl (Hrsg.), *Handbuch Kompetenzmessung* (S. XVII–XLVI). Schäffer-Poeschel.
- Erpenbeck, J. & von Rosenstiehl, L. (Hrsg.). (2007b). *Handbuch Kompetenzmessung: Erkennen, Verstehen und Bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis* (2., überarb. und erw. Aufl.). Schäffer-Poeschel.
- Ertl, D. (2010). The Nature of Science. Das Wesen / die Natur der Naturwissenschaften. *Plus Lucis*, 1, 5–7.
- Etkina, E., Karelina, A., Ruibal-Villasenor, M., Rosengrant, D., Jordan, R. & Hmelo-Silver, C. E. (2010). Design and Reflection Help Students Develop Scientific

- Abilities: Learning in Introductory Physics Laboratories. *Journal of the Learning Sciences*, 19(1), 54–98. <https://doi.org/10.1080/10508400903452876>
- Europäische Kommission. (2008). Der Europäische Qualifikationsrahmen für lebenslanges Lernen (EQR). Zugriff am 5. August 2020 unter <https://www.fibaa.org/fileadmin/files/folder/Zertifizierung/EQF.pdf>
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4. Aufl.). SAGE Publications.
- Field, A., Miles, J. & Field, Z. (2013). *Discovering statistics using R* (Reprint). Sage.
- Fischer, H. E. (1998). Scientific Literacy und Physiklernen. *Zeitschrift für Didaktik der Naturwissenschaften*, 4(2), 41–52.
- Frey, A. (2006). Strukturierung und Methoden zur Erfassung von Kompetenz. *Bildung und Erziehung*, 59(2). <https://doi.org/10.7788/bue.2006.59.2.125>
- Fritz, C. O., Morris, P. E. & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of experimental psychology: General*, 141(1), 2–18. <https://doi.org/10.1037/a0024338>
- Furtak, E. M., Seidel, T., Iverson, H. & Briggs, D. C. (2012). Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching: A Meta-Analysis. *Review of Educational Research*, 82(3), 300–329. <https://doi.org/10.3102/0034654312457206>
- Gabi, J. (2018). Auswertung von schriftlichen Schülerantworten zu Lernaufgaben zum naturwissenschaftlichen Denken und Arbeiten in der Sekundarstufe II: Dokumentation des methodischen Vorgehens und der Ergebnisse [Interner Bericht].
- Gehlen, C. (2016). *Kompetenzstruktur naturwissenschaftlicher Erkenntnisgewinnung im Fach Chemie*. Logos-Verl.
- Geiser, C. (2009). *Datenanalyse mit Mplus: Eine anwendungsorientierte Einführung* (2. Aufl.). VS Verlag für Sozialwissenschaften.
- Glug, I. (2009). Entwicklung und Validierung eines Multiple-Choice-Tests zur Erfassung prozessbezogener naturwissenschaftlicher Grundbildung.
- Gott, R., Duggan, S., Roberts, R. & Hussain, A. (2015). Research into Understanding Scientific Evidence. Zugriff am 8. Dezember 2016 unter https://community.dur.ac.uk/rosalyn.roberts/Evidence/CofEv_Gott%20et%20al.pdf
- Gräber, W., Nentwig, P., Koballa, T. R. & Evans, R. H. (Hrsg.). (2013). *Scientific Literacy: Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung*. VS Verlag für Sozialwissenschaften.

- Gräsel, C. & Göbel, K. (2015). VII-3 Unterrichtsqualität. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung* (S. 107–119). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-19994-8_8
- Großmann, N. & Wilde, M. (2019). Experimentation in biology lessons: guided discovery through incremental scaffolds. *International Journal of Science Education*, 44(9), 1–23. <https://doi.org/10.1080/09500693.2019.1579392>
- Grube, C. (2011). *Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung*. Universitätsbibliothek Kassel. Zugriff am 15. März 2021 unter <https://kobra.bibliothek.uni-kassel.de/handle/urn:nbn:de:hebis:34-2011041537247>
- Grygier, P. (2008). *Wissenschaftsverständnis von Grundschulern im Sachunterricht*. Klinkhardt.
- Gut, C. (2012). *Modellierung und Messung experimenteller Kompetenz: Analyse eines large-scale Experimentiersets*. Logos-Verl.
- Gut-Glanzmann, C. & Mayer, J. (2018). Experimentelle Kompetenz. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Theorien in der naturwissenschaftsdidaktischen Forschung* (S. 121–140). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-56320-5_8
- Hägele, J. J. (2022). Transkripte und Diagramme zur Dissertation ‚Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten‘. *JLUpub*. <https://doi.org/10.22029/jlupub-542>
- Hammann, M. (2004). Kompetenzentwicklungsmodelle: Merkmale und ihre Bedeutung – dargestellt anhand von Kompetenzen beim Experimentieren. *Der mathematische und naturwissenschaftliche Unterricht*, 57(4), 196–203.
- Hammann, M. & Asshoff, R. (2014). *Schülervorstellungen im Biologieunterricht: Ursachen für Lernschwierigkeiten* (1. Aufl.). Klett/Kallmeyer.
- Hammann, M., Thi, T. H. P. & Ehmer, M. (2006). Fehlerfrei Experimentieren. *Der mathematische und naturwissenschaftliche Unterricht*, 59(5), 292–299.
- Handl, A. (2010). *Multivariate Analysemethoden*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-14987-0>
- Hartig, J. & Klieme, E. (Hrsg.). (2007). *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik: Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung* (Bd. 20). Bundesministerium für Bildung und Forschung (BMBF).
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

- Heeren, T. & D'Agostino, R. (1987). Robustness of the two independent samples t-test when applied to ordinal scaled data. *Statistics in medicine*, 6(1), 79–90. <https://doi.org/10.1002/sim.4780060110>
- Heinicke, S. (2012). *Aus Fehlern wird man klug: Eine genetisch-didaktische Rekonstruktion des „Messfehlers“*. Logos-Verl.
- Heinicke, S. & Heering, P. (2013). Discovering Randomness, Recovering Expertise: The Different Approaches to the Quality in Measurement of Coulomb and Gauss and of Today's Students. *Science & Education*, 22(3), 483–503. <https://doi.org/10.1007/s11191-011-9430-8>
- Heller, K. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision: KFT 4-12+R*. Beltz Test GmbH.
- Helmke, A. & Schrader, F.-W. (2014). Vorwort zur 3. Auflage. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S.6–8). Beltz Verlagsgruppe.
- Hild, P., Buff, A., Gut, C. & Parchmann, I. (2020). Adaptives kompetenzbezogenes Feedback beim selbstständigen praktisch-naturwissenschaftlichen Arbeiten. *Zeitschrift für Didaktik der Naturwissenschaften*, 26(1), 19–35. <https://doi.org/10.1007/s40573-020-00109-8>
- Hofheinz, V. (2008). *Erwerb von Wissen über Nature of Science: Eine Fallstudie zum Potenzial impliziter Aneignungsprozesse in geöffneten Lehr-Lern-Arrangements am Beispiel von Chemieunterricht* (Dissertation). Universität Siegen. Fachbereich 8.
- Hofmann, J. (2015). *Untersuchung des Kompetenzaufbaus von Physiklehrkräften während einer Fortbildungsmaßnahme*. Logos-Verl.
- Hofstein, A. & Kind, P. M. (2012). Learning In and From Science Laboratories. In B. J. Fraser, K. Tobin & C. J. McRobbie (Hrsg.), *Second International Handbook of Science Education* (S.189–207). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9041-7_15
- Hofstein, A. & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, 88(1), 28–54. <https://doi.org/10.1002/sce.10106>
- Hofstein, A., Navon, O., Kipnis, M. & Mamlok-Naaman, R. (2005). Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories. *Journal of Research in Science Teaching*, 42(7), 791–806. <https://doi.org/10.1002/tea.20072>
- Holliday, W. G. (2006). A Balanced Approach To Science Inquiry Teaching. In L. B. Flick & N. G. Lederman (Hrsg.), *Scientific Inquiry and Nature of Science*

- (S. 201–217). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5814-1_10
- Höttecke, D. (2001a). *Die Natur der Naturwissenschaften historisch verstehen: Fachdidaktische und wissenschaftshistorische Untersuchungen*. Logos-Verl.
- Höttecke, D. (2001b). Die Vorstellungen von Schülern und Schülerinnen von der „Natur der Naturwissenschaften“. *Zeitschrift für Didaktik der Naturwissenschaften*, 7, 7–23.
- Höttecke, D. & Hopf, M. (2018). Schülervorstellungen zur Natur der Naturwissenschaften. In H. Schecker, T. Wilhelm, M. Hopf & R. Duit (Hrsg.), *Schülervorstellungen und Physikunterricht* (S. 271–287). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-57270-2_13
- Höttecke, D. & Rieß, F. (2015). Naturwissenschaftliches Experimentieren im Lichte der jüngeren Wissenschaftsforschung – Auf der Suche nach einem authentischen Experimentbegriff der Fachdidaktik. *Zeitschrift für Didaktik der Naturwissenschaften*, 21(1), 127–139. <https://doi.org/10.1007/s40573-015-0030-z>
- Hübinger, R. (2008). *Schüler auf Weltreise: Entwicklung und Evaluation von Lehr-/Lernmaterialien zur Förderung experimentell-naturwissenschaftlicher Kompetenzen für die Jahrgangsstufen 5 und 6*. Logos-Verl.
- Kalthoff, B. (2019). *Explizit oder implizit? Untersuchung der Lernwirksamkeit verschiedener fachmethodischer Instruktionen im Hinblick auf fachmethodische und fachinhaltliche Fähigkeiten von Sachunterrichtsstudierenden*. Logos-Verl.
- Kalthoff, B., Theyssen, H. & Schreiber, N. (2018). Explicit promotion of experimental skills. And what about the content-related skills? *International Journal of Science Education*, 40(11), 1305–1326. <https://doi.org/10.1080/09500693.2018.1477262>
- Kauertz, A. (2014). Entwicklung eines Rasch-skalierten Leistungstests. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 341–353). Springer Spektrum. https://doi.org/10.1007/978-3-642-37827-0_27
- Kechel, J.-H. (2016). *Schülerschwierigkeiten beim eigenständigen Experimentieren*. Logos-Verl.
- Keselman, A. (2003). Supporting inquiry learning by promoting normative understanding of multivariable causality. *Journal of Research in Science Teaching*, 40(9), 898–921. <https://doi.org/10.1002/tea.10115>

- Kind, P. & Osborne, J. (2017). Styles of Scientific Reasoning: A Cultural Rationale for Science Education? *Science Education*, 101(1), 8–31. <https://doi.org/10.1002/sce.21251>
- Kircher, E., Girwidz, R. & Häußler, P. (Hrsg.). (2015). *Physikdidaktik: Theorie und Praxis* (3. Aufl.). Springer Spektrum.
- Kirschner, P. A., Sweller, J. & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, 41(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1
- Kirstein, D., Habig, S. & Walpuski, M. (2019, 25. September). *Individuelles Lernen mit kooperativen Experimentieraufgaben im Chemieunterricht* [Vortrag]. GDCP-Jahrestagung, Wien.
- Klahr, D. (1999). Comment: Scientific Reasoning. In F. E. Weinert & W. Schneider (Hrsg.), *Individual development from 3 to 12* (S. 55–60). Cambridge Univ. Press.
- Klahr, D. & Carver, S. M. (1995). Scientific thinking about scientific thinking. *Monographs of the Society for Research in Child Development*, 60(4), 137–151. <https://doi.org/10.1111/j.1540-5834.1995.tb00233.x>
- Klahr, D. & Dunbar, K. (1988). Dual Space Search During Scientific Reasoning. *Cognitive Science*, 12(1), 1–48. https://doi.org/10.1207/s15516709cog1201_1
- Klahr, D. & Nigam, M. (2004). The equivalence of learning paths in early science instruction: effect of direct instruction and discovery learning. *Psychological science*, 15(10), 661–667. <https://doi.org/10.1111/j.0956-7976.2004.00737.x>
- Klieme, E. (2004). Was sind Kompetenzen und wie lassen sie sich messen. *Pädagogik (Weinheim)*, 6, 10–13.
- Klieme, E., Funke, J., Leutner, D., Reimann, P. & Wirth, J. (2001). Problemlösen als fächerübergreifende Kompetenz. Konzeption und erste Resultate aus einer Schulleistungsstudie. *Zeitschrift für Pädagogik*, 47(2), 179–200.
- Klieme, E. & Hartig, J. (2007). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. *Zeitschrift für Erziehungswissenschaft*, 10(8), 11–29.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. *Zeitschrift für Pädagogik*, 52(6), 876–903.
- Klug, T. (2017). *Wirkung kontextorientierter physikalischer Praktikumsversuche auf Lernprozesse von Studierenden der Medizin*. Logos-Verl.

- Knobloch, R. (2011). *Analyse der fachinhaltlichen Qualität von Schüleräußerungen und deren Einfluss auf den Lernerfolg: Eine Videostudie zu kooperativer Kleingruppenarbeit*. Logos-Verl.
- Koenen, J. (2014). *Entwicklung und Evaluation von experimentunterstützten Lösungsbeispielen zur Förderung naturwissenschaftlich-experimenteller Arbeitsweisen*. Logos-Verl.
- Kohlauf, L., Rutke, U. & Neuhaus, B. J. (2010). Beobachten als wissenschaftliche Erkenntnismethode. Entwicklung eines Kompetenzmodells für Kinder ab dem Vorschulalter. *Erkenntnisweg Biologiedidaktik*, 9, 165–178. Zugriff am 1. Februar 2021 unter https://www.bcp.fu-berlin.de/biologie/arbeitsgruppen/didaktik/Erkenntnisweg/2010/Kohlhauf_2010_2-11.pdf
- Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. MIT Press.
- Krajcik, J. S., Czerniak, C. & Berger, C. (1999). *Teaching children science: A project-based approach*. McGraw-Hill.
- Krell, M. (2018). Schwierigkeitserzeugende Aufgabenmerkmale bei Multiple-Choice-Aufgaben zur Experimentierkompetenz im Biologieunterricht: Eine Replikationsstudie. *Zeitschrift für Didaktik der Naturwissenschaften*, 24(1), 1–15. <https://doi.org/10.1007/s40573-017-0069-0>
- Kruse, A. & Denz, C. (2016). *Philosophie und Physik am außerschulischen Lernort: Konzepte zur Natur der Naturwissenschaften an Schule und Hochschule*. Springer Spektrum. <https://doi.org/10.1007/978-3-658-11851-8>
- Kuhn, D. (1995). Microgenetic Study of Change: What Has It Told Us? *Psychological science*, 6(3), 133–139. <https://doi.org/10.1111/j.1467-9280.1995.tb00322.x>
- Kuhn, D., Arvidsson, T. S., Lesperance, R. & Corprew, R. (2017). Can engaging in science practices promote deep understanding of them? *Science Education*, 101(2), 232–250. <https://doi.org/10.1002/sce.21263>
- Kuhn, D., Black, J., Keselman, A. & Kaplan, D. (2000). The Development of Cognitive Skills To Support Inquiry Learning. *Cognition and Instruction*, 18(4), 495–523. https://doi.org/10.1207/S1532690XCI1804_3
- Kuhn, D. & Dean, D. (2005). Is developing scientific thinking all about learning to control variables? *Psychological science*, 16(11), 866–870. <https://doi.org/10.1111/j.1467-9280.2005.01628.x>
- Kuhn, D. & Ho, V. (1980). Self-directed activity and cognitive development. *Journal of Applied Developmental Psychology*, 1(2), 119–133. [https://doi.org/10.1016/0193-3973\(80\)90003-9](https://doi.org/10.1016/0193-3973(80)90003-9)

- Kuhn, D. & Phelps, E. (1982). The Development of Problem-Solving Strategies. *Advances in Child Development and Behavior*, 17, 1–44. [https://doi.org/10.1016/S0065-2407\(08\)60356-0](https://doi.org/10.1016/S0065-2407(08)60356-0)
- Kurtz, T. (2010). Der Kompetenzbegriff in der Soziologie. In T. Kurtz & M. Pfadenhauer (Hrsg.), *Soziologie der Kompetenz* (S. 7–28). VS Verlag für Sozialwissenschaften.
- Lange, S. (2012). *Steigerung selbstregulierten Lernens durch computerbasiertes Feedback beim Erwerb von Experimentierkompetenz im Fach Biologie* (Dissertation). Georg-August-Universität. Göttingen, Georg-August University School of Science (GAUSS).
- Lanza, S. T., Flaherty, B. P. & Collins, L. M. (2006). Latent Class and Latent Transition Analysis. In I. B. Weiner (Hrsg.), *Handbook of psychology*. Wiley. <https://doi.org/10.1002/0471264385.wei0226>
- Lazonder, A. W. & Egberink, A. (2014). Children’s acquisition and use of the control-of-variables strategy: Effects of explicit and implicit instructional guidance. *Instructional Science*, 42(2), 291–304. <https://doi.org/10.1007/s11251-013-9284-3>
- Lazonder, A. W. & Harmsen, R. (2016). Meta-Analysis of Inquiry-Based Learning: Effects of Guidance. *Review of Educational Research*, 86(3), 681–718. <https://doi.org/10.3102/0034654315627366>
- Leach, J., Driver, R., Millar, R. & Scott, P. (1996). Progression in learning about the ‘Nature of Science’: Issues of conceptualisation and methodology. In M. Hughes (Hrsg.), *Progression in learning* (S. 109–139). Multilingual Matters.
- Lederman, J. S., Lederman, N. G., Bartos, S. A., Bartels, S. L., Meyer, A. A. & Schwartz, R. S. (2014). Meaningful assessment of learners’ understandings about scientific inquiry—The views about scientific inquiry (VASI) questionnaire. *Journal of Research in Science Teaching*, 51(1), 65–83. <https://doi.org/10.1002/tea.21125>
- Lederman, N. G., Abd-El-Khalick, F., Bell, R. L. & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners’ conceptions of nature of science. *Journal of Research in Science Teaching*, 39(6), 497–521. <https://doi.org/10.1002/tea.10034>
- Lehrer, R. & Schauble, L. (2015). The Development of Scientific Thinking. In R. M. Lerner (Hrsg.), *Handbook of child psychology and developmental science* (S. 1–44). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118963418.childpsy216>

- Lersch, R. (2007). Kompetenzfördernd unterrichten – 22 Schritte von der Theorie zur Praxis. *Pädagogik (Weinheim)*, (12), 36–43.
- Lin, X. & Lehman, J. D. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching*, 36(7), 837–858. [https://doi.org/10.1002/\(SICI\)1098-2736\(199909\)36:7<837::AID-TEA6>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1098-2736(199909)36:7<837::AID-TEA6>3.0.CO;2-U)
- Litt, T. (1952). Naturwissenschaft und Menschenbildung. *Physikalisch Blätter*, 8(11), 481–492. <https://doi.org/10.1002/phbl.19520081101>
- Lorch, R. F., Lorch, E. P., Calderhead, W. J., Dunlap, E. E., Hodell, E. C. & Freer, B. D. (2010). Learning the control of variables strategy in higher and lower achieving classrooms: Contributions of explicit instruction and experimentation. *Journal of Educational Psychology*, 102(1), 90–101. <https://doi.org/10.1037/a0017972>
- Luhmann, N. (1997). *Die Gesellschaft der Gesellschaft*. Suhrkamp.
- Maiseyenko, V. (2014). *Modellbasiertes Experimentieren im Unterricht: Praxistauglichkeit und Lernwirkungen*. Logos-Verl.
- Marton, F. & Booth, S. A. (1997). *Learning and awareness*. Lawrence Erlbaum Associates.
- Matlen, B. J. & Klahr, D. (2013). Sequential effects of high and low instructional guidance on children’s acquisition of experimentation skills: Is it all in the timing? *Instructional Science*, 41(3), 621–634. <https://doi.org/10.1007/s11251-012-9248-z>
- Mayring, P. (2013). Qualitative Inhaltsanalyse. In U. Flick, E. v. Kardorff & I. Steinke (Hrsg.), *Qualitative Forschung* (10. Aufl.). Rowohlt-Taschenbuch-Verl.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12., überarb. Aufl.). Beltz.
- McComas, W. F. (Hrsg.). (1998). *The nature of science in science education: Rationales and strategies* (Bd. 5). Kluwer Acad. Publ.
- McNeill, K. L. & Berland, L. (2017). What is (or should be) scientific evidence use in k-12 classrooms? *Journal of Research in Science Teaching*, 54(5), 672–689. <https://doi.org/10.1002/tea.21381>
- Méheut, M. (2012). Preconceptions and Learning. In N. M. Seel (Hrsg.), *Encyclopedia of the Sciences of Learning* (S. 2662–2664). Springer US. https://doi.org/10.1007/978-1-4419-1428-6_1850

- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749.
- Methodenberatung Zürich. (2021). Clusteranalyse (J. Schwarz & H. Bruderer Enzler, Hrsg.). Zugriff am 9. März 2021 unter https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/interdependenz/gruppierung/cluster.html
- Meyer, H. (2018). *Was ist guter Unterricht?* (13. Auflage). Cornelsen.
- Millar, R., Gott, R., Lubben, F. & Duggan, S. (1996). Children's performance in investigative tasks in science: A framework for considering progression. In M. Hughes (Hrsg.), *Progression in learning* (S. 82–108). Multilingual Matters.
- Muth, L. (2017). *Einfluss der Auswertephase von Experimenten im Physikunterricht*. Logos-Verl.
- Nehring, A. (2014). *Wissenschaftliche Denk- und Arbeitsweisen im Fach Chemie: Eine kompetenzorientierte Modell- und Testentwicklung für den Bereich der Erkenntnisgewinnung*. Logos-Verl.
- Nehring, A. & Schwichow, M. (2020). Was ist Wissen, was ist Können? Deutungen des Kompetenzbegriffs und deren psychometrische Konsequenzen im Kontext von Fachwissen und Variablenkontrollstrategie. *Zeitschrift für Didaktik der Naturwissenschaften*, 26(1), 73–87. <https://doi.org/10.1007/s40573-020-00113-y>
- Nehring, A., Stiller, J., Nowak, K. H., Upmeyer zu Belzen, A. & Tiemann, R. (2016). Naturwissenschaftliche Denk- und Arbeitsweisen im Chemieunterricht – eine modellbasierte Videostudie zu Lerngelegenheiten für den Kompetenzbereich der Erkenntnisgewinnung. *Zeitschrift für Didaktik der Naturwissenschaften*. <https://doi.org/10.1007/s40573-016-0043-2>
- Neuendorf, K. A. (2017). *The content analysis guidebook* (Second edition). Sage.
- Neumann, K. (2020). Die Bedeutung instruktionaler Kohärenz für eine systematische Kompetenzentwicklung. *Unterrichtswissenschaft*, 48(1), 1–10. <https://doi.org/10.1007/s42010-020-00068-6>
- Neumann, K., Kauertz, A., Lau, A., Notarp, H. & Fischer, H. E. (2007). Die Modellierung physikalischer Kompetenz und ihrer Entwicklung: Modelling structure and development of students' physics competence. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 209–232.
- NGSS Lead States (Hrsg.). (2013). NGSS DCI Combined. Zugriff am 23. November 2016 unter <http://www.nextgenscience.org/sites/default/files/NGSS%20DCI%20Combined%2011.6.13.pdf>

- Oberauer, K. (1993). Prozedurales und deklaratives Wissen und das Paradigma der Informationsverarbeitung. *Sprache & Kognition*, 12(1), 30–43.
- OECD. (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. OECD Publishing. <https://doi.org/10.1787/19963777>
- Osborne, J. (2014). Scientific Practices and Inquiry in the Science Classroom. In N. G. Lederman & S. K. Abell (Hrsg.), *Handbook of research on science education* (S. 579–599). Routledge.
- Oser, F. & Spychiger, M. (2005). *Lernen ist schmerzhaft: Zur Theorie des negativen Wissens und zur Praxis der Fehlerkultur*. Beltz.
- Ostlund, K. (1998). What Research Says about Science Process Skills: How Can Teaching Science Process Skills Improve Student Performance in Reading, Language Arts, and Mathematics? *Electronic Journal of Science Education*, 2(4). Zugriff am 1. Februar 2021 unter <http://www.ejse.southwestern.edu/article/view/7589/5356>
- Pedaste, M., Mäeots, M., Leijen, Ä. & Sarapuu, T. (2012). Improving Students' Inquiry Skills through Reflection and Self-Regulation Scaffolds. *Technology, Instruction, Cognition and Learning*, 9, 81–95.
- Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A., Kamp, E. T., Manoli, C. C., Zacharia, Z. C. & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, 14, 47–61. <https://doi.org/10.1016/j.edurev.2015.02.003>
- Perla, R. J. & Carifio, J. (2008). Can our Conception of the Nature of Science be Tentative Without Qualification? Special Issue: Designing and Assessing Contextual Approaches for the Teaching of Science and Mathematics. *The Journal of Educational Thought / Revue de la Pensée Éducative*, 42(2), 127–150.
- Petermann, V. (2017). *Vorstellungen von Schülerinnen, Schülern und Studierenden zum naturwissenschaftlichen Beobachten und Deuten* (Wissenschaftliche Hausarbeit im Rahmen der Ersten Staatsprüfung für das Lehramt an Gymnasien). Justus-Liebig-Universität. Gießen.
- Piekny, J. & Maehler, C. (2013). Scientific reasoning in early and middle childhood: the development of domain-general evidence evaluation, experimentation, and hypothesis generation skills. *The British journal of developmental psychology*, 31(2), 153–179. <https://doi.org/10.1111/j.2044-835X.2012.02082.x>
- Prediger, S. & von Aufschnaiter, C. (2017). Umgang mit heterogenen Lernvoraussetzungen aus fachdidaktischer Perspektive: Fachspezifische Anforderungs- und

- Lernstufungen berücksichtigen. In T. Bohl, J. Budde & M. Rieger-Ladich (Hrsg.), *Umgang mit Heterogenität in Schule und Unterricht* (S. 291–308). Verlag Julius Klinkhardt.
- Priemer, B. (2006). Deutschsprachige Verfahren der Erfassung von epistemologischen Überzeugungen. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 159–175.
- Rabe, T. & Mikelskis, H. F. (2007). Kohärenzbildungshilfen und Selbsterklärungen: Fördern sie das Physiklernen? *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 33–52.
- Rasch, B., Frieze, M., Hofmann, W. & Naumann, E. (2006). *Quantitative Methoden: Einführung in die Statistik – Band 1* (2., erw. Aufl.). Springer.
- Rauin, U., Herrle, M. & Engartner, T. (Hrsg.). (2016). *Videoanalysen in der Unterrichtsforschung: Methodische Vorgehensweisen und Anwendungsbeispiele*. Beltz Juventa.
- Rehm, M., Bünder, W., Haas, T., Buck, P., Labudde, P., Brovelli, D., Østergaard, E., Rittersbacher, C., Wilhelm, M., Genseberger, R. & Svoboda, G. (2008). Legitimationen und Fundamente eines integrierten Unterrichtsfachs Science. *Zeitschrift für Didaktik der Naturwissenschaften*, 14, 99–124.
- Reinisch, B., Helbig, K. & Krüger, D. (2020). *Biologiedidaktische Vorstellungsfor-schung: Zukunftsweisende Praxis*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-61342-9>
- Renkl, A. (2002). Worked-out examples: instructional explanations support learning by self-explanations. *Learning and Instruction*, 12(5), 529–556. [https://doi.org/10.1016/S0959-4752\(01\)00030-5](https://doi.org/10.1016/S0959-4752(01)00030-5)
- Renkl, A. (2012). Modellierung von Kompetenzen oder von interindividuellen Kompetenzunterschieden. *Psychologische Rundschau*, 63(1), 50–53. <https://doi.org/10.1026/0033-3042/a000110>
- Renkl, A. & Schworm, S. (2002). Lernen, mit Lösungsbeispielen zu lehren. *Zeitschrift für Pädagogik, Beiheft 45*, 259–270.
- Rieß, W. & Robin, N. (2012). Das Experimentieren in der fachdidaktischen empirischen Forschung. In W. Rieß, M. Wirtz, B. Barzel & A. Schulz (Hrsg.), *Experimentieren im mathematisch-naturwissenschaftlichen Unterricht* (S. 129–152). Waxmann.
- Rieß, W., Wirtz, M., Barzel, B. & Schulz, A. (Hrsg.). (2012). *Experimentieren im mathematisch-naturwissenschaftlichen Unterricht: Schüler lernen wissenschaftlich denken und arbeiten*. Waxmann.

- Rizzo, K. L. & Taylor, J. C. (2016). Effects of Inquiry-Based Instruction on Science Achievement for Students with Disabilities: An Analysis of the Literature. *Journal of Science Education for Students with Disabilities*, 19(1), 1–16.
- Roberts, R. (2001). Procedural understanding in biology: The ‘thinking behind the doing’. *Journal of Biological Education*, 35(3), 113–117. <https://doi.org/10.1080/00219266.2001.9655758>
- Rogge, C. (2010). *Entwicklung physikalischer Konzepte in aufgabenbasierten Lernumgebungen*. Logos-Verl.
- Romesburg, H. C. (2004). *Cluster analysis for researchers*. Lulu Press.
- Rönnebeck, S., Bernholt, S. & Ropohl, M. (2016). Searching for a common ground – A literature review of empirical research on scientific inquiry activities. *Studies in Science Education*, 1–37. <https://doi.org/10.1080/03057267.2016.1206351>
- Ropohl, M. & Scheuermann, H. (2018). Welche Rückmeldungen wirken am besten? *Zeitschrift für Didaktik der Naturwissenschaften*, 24(1), 151–165. <https://doi.org/10.1007/s40573-018-0080-0>
- Ross, J. A. (1988). Controlling Variables: A Meta-Analysis of Training Studies. *Review of Educational Research*, 58(4), 405–437. <https://doi.org/10.3102/00346543058004405>
- Ruiz-Primo, M. A. & Shavelson, R. J. (1996). Rhetoric and reality in science performance assessments: An update. *Journal of Research in Science Teaching*, 33(10), 1045–1063. [https://doi.org/10.1002/\(SICI\)1098-2736\(199612\)33:10%3C1045::AID-TEA1%3E3.0.CO;2-S](https://doi.org/10.1002/(SICI)1098-2736(199612)33:10%3C1045::AID-TEA1%3E3.0.CO;2-S)
- Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student’s t-test and the Mann–Whitney U test. *Behavioral Ecology*, 17(4), 688–690. <https://doi.org/10.1093/beheco/ark016>
- Schaake, S. (2011). Die Natur der Naturwissenschaften verstehen lernen: Historische, gesellschaftliche und kulturell relevante Stationen für den Chemieunterricht. In Zentrum für Lehrerbildung der Universität Kassel (Hrsg.), *Reihe Studium und Forschung*. kassel university press. Zugriff am 22. Oktober 2015 unter <https://kobra.bibliothek.uni-kassel.de/bitstream/urn:nbn:de:hebis:34-2011102639442/1/SchaakeNaturwissenschaften.pdf>
- Schäfer, P. (2016). *Bedrohte Identität und Veränderungen im arabischen Sicherheitsdiskurs*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-12207-2>
- Schalk, L., Edelsbrunner, P. A., Deiglmayr, A., Schumacher, R. & Stern, E. (2019). Improved application of the control-of-variables strategy as a collateral be-

- nefit of inquiry-based physics education in elementary school. *Learning and Instruction*, 59, 34–45. <https://doi.org/10.1016/j.learninstruc.2018.09.006>
- Scharfenberg, F.-J. (2005). *Experimenteller Biologieunterricht zu Aspekten der Gentechnik im Lernort Labor: empirische Untersuchung zu Akzeptanz, Wissenserwerb und Interesse (am Beispiel des Demonstrationslabors Bio-/Gentechnik der Universität Bayreuth mit Schülern aus dem Biologie-Leistungskurs des Gymnasiums)* (Dissertation). Universität Bayreuth. Fakultät für Biologie, Chemie und Geowissenschaften.
- Scharmann, L. C., Smith, M. U., James, M. C. & Jensen, M. (2005). Explicit Reflective Nature of Science Instruction: Evolution, Intelligent Design, and Umbrellaology. *Journal of Science Teacher Education*, 16(1), 27–41. <https://doi.org/10.1007/s10972-005-6990-y>
- Schauble, L. (1999). Syntax or Substance? What is Scientific Reasoning? *Human development*, 42(5), 278–282. <https://doi.org/10.1159/000022636>
- Schauble, L., Klopfer, L. E. & Raghavan, K. (1991). Students' transition from an engineering model to a science model of experimentation. *Journal of Research in Science Teaching*, 28(9), 859–882. <https://doi.org/10.1002/tea.3660280910>
- Schecker, H. & Duit, R. (2018). Schülervorstellungen und Physiklernen. In H. Schecker, T. Wilhelm, M. Hopf & R. Duit (Hrsg.), *Schülervorstellungen und Physikunterricht* (S. 1–21). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-57270-2_1
- Schecker, H., Neumann, K., Theyßen, H., Eickhorst, B. & Dickmann, M. (2016). Stufen experimenteller Kompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, 22(1), 197–213. <https://doi.org/10.1007/s40573-016-0050-3>
- Schecker, H. & Parchmann, I. (2006). Modellierung naturwissenschaftlicher Kompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 45–66.
- Schecker, H., Wilhelm, T., Hopf, M. & Duit, R. (Hrsg.). (2018). *Schülervorstellungen und Physikunterricht*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-57270-2>
- Schendera, C. (2011). *Clusteranalyse mit SPSS: Mit Faktorenanalyse*. De Gruyter.
- Schmandt, T. B. (2017). *Umgang von Schülerinnen und Schülern mit fachbezogenem Feedback, Informationen und Hilfen* (Wissenschaftliche Hausarbeit im Rahmen der Ersten Staatsprüfung für das Lehramt an Gymnasien). Justus-Liebig-Universität. Gießen.
- Schmidt, D. (2016). *Modellierung experimenteller Kompetenzen sowie ihre Diagnostik und Förderung im Biologieunterricht*. Logos-Verl.

- Schmidt-Weigand, F., Hänze, M. & Wodzinski, R. (2009). Complex Problem Solving and Worked Examples. *Zeitschrift für Pädagogische Psychologie*, 23(2), 129–138. <https://doi.org/10.1024/1010-0652.23.2.129>
- Schreiber, N. (2012). *Diagnostik experimenteller Kompetenz: Validierung technologiegestützter Testverfahren im Rahmen eines Kompetenzstrukturmodells*. Logos-Verl.
- Schreiber, N., Theyßen, H. & Schecker, H. (2014). Diagnostik experimenteller Kompetenz: Kann man Realexperimente durch Simulationen ersetzen? *Zeitschrift für Didaktik der Naturwissenschaften*, 20(1), 161–173. <https://doi.org/10.1007/s40573-014-0017-1>
- Schreier, M. (2013). *Qualitative content analysis in practice*. Sage.
- Schwichow, M., Christoph, S., Boone, W. J. & Härtig, H. (2016). The impact of sub-skills and item content on students' skills with regard to the control-of-variables strategy. *International Journal of Science Education*, 38(2), 216–237. <https://doi.org/10.1080/09500693.2015.1137651>
- Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik*, 60, 850–866.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004). Einheitliche Prüfungsanforderungen in der Abiturprüfung Physik.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.). (2005). *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss (Jahrgangsstufe 10) [Beschluss der Kultusministerkonferenz vom 16.12.2004]*. Luchterhand (Wolters Kluwer).
- Sonnenschein, I. (2019). *Naturwissenschaftliche Denk- und Arbeitsprozesse Studierender im Labor*. Logos-Verl.
- Steckenmesser-Sander, K. S. (2015). *Gemeinsamkeiten und Unterschiede physikbezogener Handlungs-, Denk- und Lernprozesse von Mädchen und Jungen*. Logos-Verl.
- Stender, A., Geller, C., Neumann, K. & Fischer, H. E. (2013). Der Einfluss der Unterrichtstaktung auf die Strukturiertheit und Abgeschlossenheit von Lernprozessen. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 189–207.
- Stiller, J. (2015). *Scientific Inquiry im Chemieunterricht* (Diss.). Humboldt-Universität zu Berlin. Berlin, Mathematisch-Naturwissenschaftliche Fakultät. <https://doi.org/10.18452/17503>

- Strippel, C. G. (2017). *Naturwissenschaftliche Erkenntnisgewinnung an chemischen Inhalten vermitteln*. Logos-Verl.
- Stulz, N. (2012). Effektstärken: Zur klinischen Relevanz von Studienergebnissen. *Psychotherapie, Psychosomatik, medizinische Psychologie*, 62(7), 288–289. <https://doi.org/10.1055/s-0032-1305059>
- Teig, N. (im Druck für 2021). Inquiry in science education. In T. Nilsen, A. Stancel-Piatak & J. E. Gustafsson (Hrsg.), *International Handbook of Comparative Large-Scale Studies in Education*. Springer Open. <https://doi.org/10.31219/osf.io/r3acn>
- Telser, V. (2019). *Erfassung und Förderung experimenteller Kompetenz von Lehrkräften im Fach Chemie*. Logos-Verl.
- Tempel, T., Kaufmann, K., Kranz, J. & Möller, A. (2018). Retrieval-based skill learning: Testing promotes the acquisition of scientific experimentation skills. *Psychological research*. <https://doi.org/10.1007/s00426-018-1088-2>
- Tesch, M. (2005). *Das Experiment im Physikunterricht: Didaktische Konzepte und Ergebnisse einer Videostudie*. Logos-Verl.
- van der Graaf, J., Segers, E. & Verhoeven, L. (2015). Scientific reasoning abilities in kindergarten: dynamic assessment of the control of variables strategy. *Instructional Science*, 43(3), 381–400. <https://doi.org/10.1007/s11251-015-9344-y>
- van Harmelen, E. (2017). *Combining direct instruction on the Control-of-Variables strategy with task segmentation: Is there a positive synergistic effect?* (MSc. Thesis). University of Twente. Enschede, NL, Faculty of Behavioral, Management; Social Sciences, Department of Instructional Technology. Zugriff am 15. März 2021 unter <https://pdfs.semanticscholar.org/ee31/a28cffab50d4cb95843e867cebd20b2838d6.pdf>
- VERBI Software. (2018). MAXQDA 2018 Manual. Zugriff am 15. März 2021 unter <https://www.maxqda.de/download/manuals/MAX2018-Online-Manual-Complete-DE.pdf>
- VERBI Software. (2019). MAXQDA 2020 Manual. Zugriff am 15. März 2021 unter <https://www.maxqda.de/download/manuals/MAX2020-Online-Manual-Complete-DE.pdf>
- Villanueva, M. G. & Hand, B. (2011). Data versus Evidence. Investigating the Difference. *Science Scope*, 35(1), 42–45.
- von Aufschnaiter, C. (1999). *Bedeutungsentwicklungen, Interaktionen und situatives Erleben beim Bearbeiten physikalischer Aufgaben: Fallstudien zu Bedeutungs-*

entwicklungsprozessen von Studierenden und Schüler(inne)n in einer Feld- und einer Laboruntersuchung zum Themengebiet Elektrostatik und Elektrodynamik. Logos-Verl.

- von Aufschnaiter, C. (2003). Interactive processes between university students: Structures of interactions and related cognitive development. *Research in Science Education*, 33, 341–374.
- von Aufschnaiter, C., Capell, J., Dübbele, G., Ennemoser, M., Mayer, J., Stiensmeier-Pelster, J., Sträßer, R. & Wolgast, A. (2015). Diagnostische Kompetenz: Theoretische Überlegungen zu einem zentralen Konstrukt der Lehrerbildung. *Zeitschrift für Pädagogik*, 61(5), 738–758.
- von Aufschnaiter, C. & Rogge, C. (2010a). Misconceptions or missing conceptions. *EURASIA Journal of Mathematics, Science & Technology Education*, 6(1), 3–18.
- von Aufschnaiter, C. & Rogge, C. (2010b). Wie lassen sich Verläufe der Entwicklung von Kompetenz modellieren? *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 95–114.
- von Aufschnaiter, C. & Rogge, C. (2015). Conceptual Change in Learning. In R. Gunstone (Hrsg.), *Encyclopedia of science education* (S. 209–218). Springer.
- von Aufschnaiter, C. & von Aufschnaiter, S. (2003). Theoretical framework and empirical evidence of students' cognitive processes in three dimensions of content, complexity, and time. *Journal of Research in Science Teaching*, 40(7), 616–648. <https://doi.org/10.1002/tea.10102>
- von Aufschnaiter, C. & Hofmann, J. (2014). Kompetenz und Wissen – Wechselseitige Zusammenhänge und Konsequenzen für die Unterrichtsplanung. *Der mathematische und naturwissenschaftliche Unterricht*, 67(1), 10–16.
- Vorholzer, A. (2016). *Wie lassen sich Kompetenzen des experimentellen Denkens und Arbeitens fördern? Eine empirische Untersuchung der Wirkung eines expliziten und eines impliziten Instruktionsansatzes.* Logos-Verl.
- Vorholzer, A. (2017). Lernaufgaben zu fachmethodischen Kompetenzen. *Der mathematische und naturwissenschaftliche Unterricht*, 70(2), 83–89.
- Vorholzer, A., Hägele, J. J. & von Aufschnaiter, C. (2020). Instruktionen kohärent anlegen und Kompetenzaufbau untersuchen: Zugänge und Herausforderungen am Beispiel experimentbezogener Kompetenz. *Unterrichtswissenschaft*, 48(1), 61–89. <https://doi.org/10.1007/s42010-019-00064-5>

- Vorholzer, A. & von Aufschnaiter, C. (2019a, 12.–16. August). *Impact of Prior Knowledge and Targeted Inquiry Abilites on the Effectivness of Explicit Instruction* [Paper-Präsentation]. EARLI 2019, Aachen.
- Vorholzer, A. & von Aufschnaiter, C. (2019b). Guidance in inquiry-based instruction – an attempt to disentangle a manifold construct. *International Journal of Science Education*, 2(3), 1–16. <https://doi.org/10.1080/09500693.2019.1616124>
- Vorholzer, A., von Aufschnaiter, C. & Boone, W. J. (2018). Fostering upper secondary students' ability to engage in practices of scientific investigation: A comparative analysis of an explicit and an implicit instructional approach. *Research in Science Education*, 103(1), 1. <https://doi.org/10.1007/s11165-018-9691-1>
- Vorholzer, A., von Aufschnaiter, C. & Kirschner, S. (2016). Entwicklung und Erprobung eines Tests zur Erfassung des Verständnisses naturwissenschaftlicher Denk- und Arbeitsweisen // Entwicklung und Erprobung eines Tests zur Erfassung des Verständnisses experimenteller Denk- und Arbeitsweisen. *Zeitschrift für Didaktik der Naturwissenschaften*, 22(1), 25–41. <https://doi.org/10.1007/s40573-015-0039-3>
- Wagenschein, M. (1995). *Die pädagogische Dimension der Physik* (1. Neuaufl.). Hahner Verl.-Ges.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Hrsg.), *Defining and selecting key competencies* (S. 45–66). Hogrefe & Huber.
- Weinert, F. E. (2014). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (3. Aufl., S. 17–32). Beltz Verlagsgruppe. (Original erschienen 2001)
- Wellnitz, N. & Mayer, J. (2013). Erkenntnismethoden in der Biologie – Entwicklung und Evaluation eines Kompetenzmodells. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 315–345.
- Wentura, D. & Pospeschill, M. (2015). *Multivariate Datenanalyse*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-531-93435-8>
- Werner, M. & Kremer, K. (2010). „Ein Experiment ist das, was der Lehrer macht.“ Schülervorstellungen über die Natur der Naturwissenschaften. *Erkenntnisweg Biologiedidaktik*, 9, 135–150.
- White, B. Y., Shimoda, T. A. & Frederiksen, J. R. (1999). Enabling Students to Construct Theories of Collaborative Inquiry and Reflective Learning:

- Computer Support for Metacognitive Development. *International Journal of Artificial Intelligence in Education (IJAIED)*, 10, 151–182.
- Wilhelm, P. & Beishuizen, J. J. (2004). Asking Questions During Self-Directed Inductive Learning: Effects on Learning Outcome and Learning Processes. *Interactive Learning Environments*, 12(3), 251–264. <https://doi.org/10.1080/10494820512331383449>
- Wilson, C. D., Taylor, J. A., Kowalski, S. M. & Carlson, J. (2010). The Relative Effects and Equity of Inquiry-Based and Commonplace Science Teaching on Students' Knowledge, Reasoning, and Argumentation. *Journal of Research in Science Teaching*, 47(3), 276–301. <https://doi.org/10.1002/tea.20329>
- Wirth, J., Thillmann, H., Küsting, J., Fischer, H. E. & Leutner, D. (2008). Das Schülerexperiment im naturwissenschaftlichen Unterricht – Bedingungen der Lernförderlichkeit dieser Lehrmethode. *Zeitschrift für Pädagogik*, (3), 361–375.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Hogrefe.
- Wittwer, J. & Renkl, A. (2008). Why Instructional Explanations Often Do Not Work: A Framework for Understanding the Effectiveness of Instructional Explanations. *Educational Psychologist*, 43(1), 49–64. <https://doi.org/10.1080/00461520701756420>
- Wollenschläger, M. (2012). *Effekte kompetenziellen Feedbacks auf Performanz (wissenschaftliches Denken), Motivation und Metakognition von Lernenden der Sekundarstufe I* (Dissertation). Christian-Albrechts-Universität zu Kiel. Kiel. Zugriff am 15. März 2021 unter https://macau.uni-kiel.de/servlets/MCRFileNodeServlet/dissertation_derivate_00004393/Dissertation_Mareike%20Wollenschlaeger1.pdf
- Zhang, W.-X., Hsu, Y.-S., Wang, C.-Y. & Ho, Y.-T. (2015). Exploring the impacts of cognitive and metacognitive prompting on students' scientific inquiry practices within an e-learning environment. *International Journal of Science Education*, 37(3), 529–553. <https://doi.org/10.1080/09500693.2014.996796>
- Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, 27(2), 172–223. <https://doi.org/10.1016/j.dr.2006.12.001>

- Zimmerman, C. & Klahr, D. (2018). Development of Scientific Thinking. In S. Ghetti (Hrsg.), *Developmental and Social Psychology* (S. 223–248). John Wiley & Sons Inc. <https://doi.org/10.1002/9781119170174.epcn407>
- Zohar, A. & David, A. B. (2008). Explicit teaching of meta-strategic knowledge in authentic classroom situations. *Metacognition and Learning*, 3(1), 59–82. <https://doi.org/10.1007/s11409-007-9019-4>
- Zohar, A. & Peled, B. (2008). The effects of explicit teaching of metastrategic knowledge on low- and high-achieving students. *Learning and Instruction*, 18(4), 337–353. <https://doi.org/10.1016/j.learninstruc.2007.07.001>

Anhang

A Prä- und Post-Erhebungen

A.1 Anlage der Erhebungen

Tabelle A.1: Übersicht über bei Vorholzer eingesetzte Instrumente (basierend auf Vorholzer, 2016, S. 90, Tabelle 9.2: »Zentrale Eigenschaften und statistische Kennwerte der eingesetzten Testinstrumente«).

	EDAWT	FTM	KFT	Interesse an Physik (INT)	Physikbezogene Selbst-einschätzung (SWK)
Messzeitpunkt	Pretest Posttest	Pretest	Pretest	Pretest	Pretest
Aufgabentyp	Multiple-Choice & offene Aufgaben	Multiple-Choice	Multiple-Choice	4 stufige Likert-Skala	4 stufige Likert-Skala
Aufgabenanzahl	24 pro Heft*	17	50 (25 pro Heft*)	8	15
Bearbeitungszeit	35 Minuten	25 Minuten	8 Minuten	5 Minuten	5 Minuten
Item Outfit-MNSQ Bereich (min. – max.)	0.71 – 1.16	0.71 – 1.16	0.53 – 1.48	0.65 – 1.72	0.64 – 1.87
Personen-Reliabilität	.81	.74	.80	.83	.95
Item-Reliabilität	.97	.98	.94	.98	.97

Anmerkungen. *: Ausführliche Anmerkungen zu den Testheften und deren Einsatz finden sich bei Vorholzer (2016).

A.2 Testwerte der untersuchten Personen

In den folgenden vier Tabellen die Testwerte der Personen, die untersucht wurden. Für Personen, deren Aktivitäten für die explizit-fachmethodische bzw. die implizit-fachmethodische Instruktionsvariante untersucht wurden, findet sich jeweils eine Tabelle. Außerdem findet sich für die beiden Konzepte, zu denen die Vorstellungen untersucht wurden, jeweils eine Tabelle. Personen, für die sowohl die Aktivitäten als auch die Vorstellungen untersucht wurden, sind demnach mehrfach aufgeführt.

Tabelle A.2: Übersicht über die zu den Aktivitäten untersuchten Personen (Teil 1: explizit-fachmethodische Instruktionsvariante).

Person	Bearb.	Analysiert für			Werte aus den Prä- und Post-Erhebungen						
	Einh.	E1	E2	E3	Zuwachs	Pre	Post	FCI	KFT	SWK	INT
<i>Explizit-fachmethodische Instruktionsvariante</i>											
A1	3	x	x		230,17	626,81	856,98	631,47	616,84	589,91	528,81
A2	3	x	x		171,03	479,86	650,89				
A3	3	x	x		144,01	419,53	563,54	538,54	509,93	469,59	555,99
B4	3	x	x	x	116,6	386,25	502,85	466,46	459,45	463,13	529,76
B5	3	x	x	x	125,92	593,54	719,46	524,35	529,12	477,9	484,89
B6	2	x	x		113,32	501,11	614,43	451,65	459,45	477,9	404,91
C7	3		x		92,73	515,12	607,85				
C8	2					427,59		510,11	463,22	600,24	611,61
C9	3		x		9,59	515,12	524,71	466,46	511,63	477,9	484,89
D10	2	x		x	118,37	448,15	566,52	481,17	459,45	395,87	322,55
D11	3	x	x	x	133,67	474,18	607,85	552,79	590,8	461,37	484,89
D12	3	x	x	x	163,95	428,39	592,34				
E13	2		x			396,73		538,54	426,53	387,26	468,39
E14	2		x		232,29	462,19	694,48	670,66	547,16	412,57	429,58
E15	2		x		16,03	282,37	298,4	524,35	443,06	436,99	528,81
F16	3	x	x	x	-63,58	417,35	353,77	422,15	568,1	378,4	542,48
F17	2	x	x		44,29	458,56	502,85	422,15	590,8	569,65	468,39
F18	3	x	x	x	172,81	419,53	592,34	451,65	459,45	428,89	450,23
G19	3	x	x	x	40,24	374,8	415,04	436,85	475,91	436,99	322,55
G20	3	x	x	x	43,49	448,15	491,64	538,54	527,98	445,09	468,39
G21	3	x	x	x	-29,89	479,86	449,97	524,35	529,12	579,72	583,14
H22	3	x		x	186,35	515,12	701,47	327,45	590,72	519,22	520,05
H23	3	x		x	149,86	759,91	909,77	631,47	685,8	559,7	555,99
H24	3	x		x	56,08	364,74	420,82	495,73	416,5	494,87	542,48
I25	3	x			31,11	593,13	624,24	597,68	648,27	540,25	450,23
I26	3	x			93,97	458,56	552,53				
I27	3	x			89,56	330,85	420,41	538,54	547,79	512,45	583,14
J28	3	x		x	-18,68	401,81	383,13	582,17	383,67	395,87	468,39
J29	3	x		x	13,58	525,67	539,25	451,65	400,41	494,87	450,23
J30	3	x		x	164,41	443,44	607,85	451,65	529,12	519,22	470,11
K31	3	x			-21,56	645,8	624,24	524,35	547,79	549,89	404,91
K32	3	x			2,28	437,84	440,12	597,68	493,01	644,79	583,14
K33	3	x			42,78	509,75	552,53	466,46	547,79	600,24	528,81
L34	3	x		x	-83,84	407,08	323,24	436,85	426,53	503,57	468,39
L35	3	x		x	22,99	479,86	502,85	567,28	529,12	503,57	429,58
L36	3	x		x	67,88	410,68	478,56	466,46	494,98	412,57	468,39
M37	3	x		x	75,36	437,31	512,67	481,17	447,71	589,91	583,14
M38	3	x		x	68,03	458,56	526,59	466,46	590,8	559,31	468,39
M39	3	x		x	2,71	401,81	404,52	481,17	492,66	428,89	450,23

Anmerkungen. Bearb. Einh. = Anzahl bearbeiteter (aber nicht notwendigerweise auch analysierter Einheiten). Leeres Feld = Nicht analysiert bzw. nicht anwesend bei Erhebung; x = Anwesend bzw. Werte für alle Erhebungen liegen vor; (x) = Werte zu EDAW-Erhebungen liegen vor (insb. also der Kompetenzzuwachs), aber andere Erhebungen sind nicht vollständig. Personenkürzel enthalten einen Buchstaben für das Team und eine fortlaufende Nummer.

Tabelle A.2: Übersicht über die zu den Aktivitäten untersuchten Personen (Teil 2: implizit-fachmethodische Instruktionsvariante).

Person	Bearb.	Analysiert für			Werte aus den Prä- und Post-Erhebungen						
	Einh.	E1	E2	E3	Zuwachs	Pre	Post	FCI	KFT	SWK	INT
Implizit-fachmethodische Instruktionsvariante											
Q50	3	x	x	x	1,67	342,13	343,8	538,54	346,83	369,18	372,83
Q51	2	x	x		0,11	446,31	446,42	379,59	547,79	453,21	500,24
Q52	3	x	x	x	163,54	450,17	613,71	481,17	509,93	540,25	322,55
R53	2	x	x		41,33	551,01	592,34	552,79	616,84	512,45	528,81
R54	3	x	x	x	-11,54	578,06	566,52	538,54	648,27	428,89	322,55
R55	3	x	x	x	-11,5	538,09	526,59	495,73	426,53	445,09	372,83
S56	2	x	x		53,46	374,8	428,26	379,59	492,66	436,99	450,23
S57	3	x	x	x	107,84	342,13	449,97	481,17	509,93	803,3	661,31
S58	3	x	x	x	74,5	445,31	519,81	510,11	494,98	461,37	429,58
T59	3	x	x	x	140,34	386,25	526,59	597,68	432,2	445,09	429,58
T60	3	x	x	x	-34,92	502,61	467,69	510,11	529,12	420,76	468,39
T61	3	x	x	x	73,81	523,18	596,99	495,73	511,63	521,52	484,89
U62	3	x	x	x	-13,9	563,98	550,08	552,79	529,12	540,25	569,49
U63	3	x	x	x	35,06	578,65	613,71	670,66	685,8	486,32	500,24
U64	3	x	x	x	20,39	504,32	524,71	650,25	529,12	503,57	404,91
V65	3	x	x	x	84,28	512,71	596,99	631,47	616,84	672,26	597,11
V66	3	x	x	x	-33,84	407,08	373,24	407,67	347,44	521,52	404,91
V67	3	x	x	x	57,16	316,08	373,24	436,85	65,52	395,87	450,23

Anmerkungen. Bearb. Einh. = Anzahl bearbeiteter (aber nicht notwendigerweise auch analysierter Einheiten). Leeres Feld = Nicht analysiert bzw. nicht anwesend bei Erhebung; x = Anwesend bzw. Werte für alle Erhebungen liegen vor; (x) = Werte zu EDAW-Erhebungen liegen vor (insb. also der Kompetenzzuwachs), aber andere Erhebungen sind nicht vollständig. Personenkürzel enthalten einen Buchstaben für das Team und eine fortlaufende Nummer.

A Prä- und Post-Erhebungen

Tabelle A.3: Übersicht über die Testwerte der zu Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen untersuchten Personen.

Person	Werte aus den Prä- und Post-Erhebungen						
	Zuwachs	Pre	Post	FCI	KFT	SWK	INT
A1	230,17	626,81	856,98	631,47	616,84	589,91	528,81
A2	171,03	479,86	650,89	X	X	X	X
A3	144,01	419,53	563,54	538,54	509,93	469,59	555,99
B4	116,6	386,25	502,85	466,46	459,45	463,13	529,76
B5	125,92	593,54	719,46	524,35	529,12	477,9	484,89
B6	113,32	501,11	614,43	451,65	459,45	477,9	404,91
C7	92,73	515,12	607,85	X	X	X	X
C8	X	427,59		510,11	463,22	600,24	611,61
C9	9,59	515,12	524,71	466,46	511,63	477,9	484,89
F16	-63,58	417,35	353,77	422,15	568,1	378,4	542,48
F17	44,29	458,56	502,85	422,15	590,8	569,65	468,39
F18	172,81	419,53	592,34	451,65	459,45	428,89	450,23
G19	40,24	374,8	415,04	436,85	475,91	436,99	322,55
G20	43,49	448,15	491,64	538,54	527,98	445,09	468,39
G21	-29,89	479,86	449,97	524,35	529,12	579,72	583,14
H22	186,35	515,12	701,47	327,45	590,72	519,22	520,05
H23	149,86	759,91	909,77	631,47	685,8	559,7	555,99
H24	56,08	364,74	420,82	495,73	416,5	494,87	542,48
I25	31,11	593,13	624,24	597,68	648,27	540,25	450,23
I26	93,97	458,56	552,53	X	X	X	X
I27	89,56	330,85	420,41	538,54	547,79	512,45	583,14
M37	75,36	437,31	512,67	481,17	447,71	589,91	583,14
M38	68,03	458,56	526,59	466,46	590,8	559,31	468,39
M39	2,71	401,81	404,52	481,17	492,66	428,89	450,23
N40	109,12	515,12	624,24	745,12	647,54	579,72	661,31
N41	220,88	593,54	814,42	650,25	567,89	672,26	643,29
N42	25,92	452,64	478,56	524,35	234,81	530,78	542,48
O43	X	448,96		524,35	511,63	428,89	372,83
O44	-35,84	440,36	404,52	407,67	492,66	485,13	528,81
O45	14,19	386,25	400,44	436,85	443,06	387,26	450,23
P46	X		701,58	510,11	509,93	503,57	468,39
P47	X	389,28		481,17	590,8	549,89	514,79
P48	93,1	578,65	671,75	481,17	685,8	589,91	528,81

Anmerkungen. Leeres Feld = Nicht analysiert bzw. nicht anwesend bei Erhebung;
 x = Anwesend bzw. Werte für alle Erhebungen liegen vor; (x) = Werte zu EDAW-
 Erhebungen liegen vor, aber andere Erhebungen sind nicht vollständig.

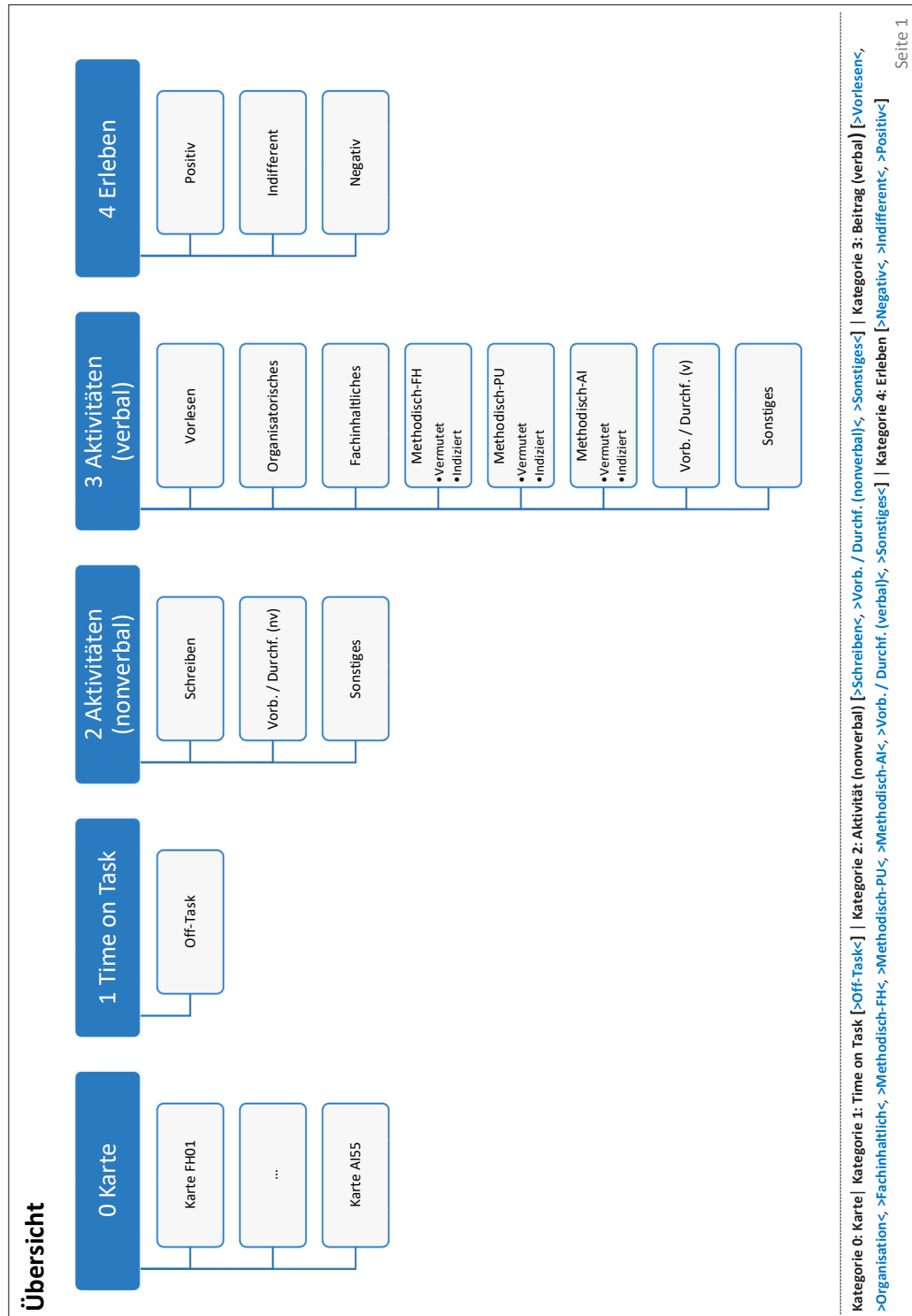
Tabelle A.4: Übersicht über die Testwerte der zu Vorstellungen zur Unterscheidung von Beobachtung und Deutung untersuchten Personen.

Person	Werte aus den Prä- und Post-Erhebungen						
	Zuwachs	Pre	Post	FCI	KFT	SWK	INT
A0 = E14	232,29	462,19	694,48	670,66	547,16	412,57	429,58
A1	230,17	626,81	856,98	631,47	616,84	589,91	528,81
A2	171,03	479,86	650,89	X	X	X	X
A3	144,01	419,53	563,54	538,54	509,93	469,59	555,99
B4	116,6	386,25	502,85	466,46	459,45	463,13	529,76
B5	125,92	593,54	719,46	524,35	529,12	477,9	484,89
C7	92,73	515,12	607,85	X	X	X	X
C8	X	427,59		510,11	463,22	600,24	611,61
C9	9,59	515,12	524,71	466,46	511,63	477,9	484,89
F16	-63,58	417,35	353,77	422,15	568,1	378,4	542,48
F18	172,81	419,53	592,34	451,65	459,45	428,89	450,23
G19	40,24	374,8	415,04	436,85	475,91	436,99	322,55
G21	-29,89	479,86	449,97	524,35	529,12	579,72	583,14
H22	186,35	515,12	701,47	327,45	590,72	519,22	520,05
H23	149,86	759,91	909,77	631,47	685,8	559,7	555,99
H24	56,08	364,74	420,82	495,73	416,5	494,87	542,48
K31	-21,56	645,8	624,24	524,35	547,79	549,89	404,91
K32	2,28	437,84	440,12	597,68	493,01	644,79	583,14
K33	42,78	509,75	552,53	466,46	547,79	600,24	528,81
L34	-83,84	407,08	323,24	436,85	426,53	503,57	468,39
L35	22,99	479,86	502,85	567,28	529,12	503,57	429,58
L36	67,88	410,68	478,56	466,46	494,98	412,57	468,39
M37	75,36	437,31	512,67	481,17	447,71	589,91	583,14
M38	68,03	458,56	526,59	466,46	590,8	559,31	468,39
M39	2,71	401,81	404,52	481,17	492,66	428,89	450,23
P46	X		701,58	510,11	509,93	503,57	468,39
P48	93,1	578,65	671,75	481,17	685,8	589,91	528,81

Anmerkungen. Leeres Feld = Nicht analysiert bzw. nicht anwesend bei Erhebung;
 x = Anwesend bzw. Werte für alle Erhebungen liegen vor; (x) = Werte zu EDAW-
 Erhebungen liegen vor, aber andere Erhebungen sind nicht vollständig.

B Ausführlichere Dokumentation zu Aktivitäten

B.1 Kodiermanual (10 Seiten)



Kategorie 0: Karte

Codes beschreiben, welche Karte von den Schülerinnen und Schülern zu einem bestimmten Zeitpunkt bearbeitet wird. Wird gruppenweise und vorlaufend zu anderen Kategorien kodiert (am besten in einem extra Durchlauf, kann auch HK ohne spez. Vorkenntnisse erledigen). Codes aus dieser Kategorie werden parallel zu Codes aus allen anderen Kategorien vergeben (dort nicht erwähnt).

Codes	Regeln und Beispiele
>Karte FH01< >Karte FH02< ... >Karte FH34< >Karte PU01< ... >Karte PU57< >Karte AI01< ... >Karte AI55<	– Beginn einer Karte ist der Zeitpunkt, ab dem SuS sich inhaltlich mit der Karte auseinandersetzen. Beispiel: SuS beginnen damit, eine Karte still oder laut vorzulesen oder reagieren verbal / nonverbal auf Inhalt der aufgedeckten Karte (z. B. Freude über richtige Lösung...). Achtung: Der Moment in dem die Karte „umgeblättert“ wird, ist nicht immer der Moment in dem die inhaltliche Auseinandersetzung beginnt! – Jeder Stelle des Videos muss genau ein Code dieser Kategorie zugewiesen werden, außer wenn SuS Karten (im Sinne eines stumpfen Weiterblätterns) überspringen, ohne diese inhaltlich zu bearbeiten (-> dann kein Code). Beispiele: a) Wenn SuS Karte 1 schon weglegen, aber Karte 2 noch nicht begonnen haben, wird das Zeitsegment weiterhin als >Karte 1< kodiert b) Wenn SuS Lösung bei der Bearbeitung einer Kontrollkarte (z. B. Karte 2) auf eine vorlaufende Aufgabenkarte (z. B. Karte 1) zurückgreifen, dann wird dem Zeitsegment nur >Karte 2< zugewiesen c) Wird Bearbeitung von Karte 2 abgebrochen und z. B. eine Aufgabe von Karte 1 erneut bearbeitet, dann wird diesem Segment wieder >Karte 1< zugewiesen.

Kategorie 1: Time on Task

Code wird schülerweise vergeben (also unabhängig von allen beliebigen Codes, die einer anderen Person zugeordnet werden könnten) und kann nicht parallel zu anderen Codes für dieselbe Schülerin / denselben Schüler vergeben werden.
Code wird phasenweise vergeben, d. h. nur unterbrochen, wenn zwischendurch die gleiche Person offensichtlich anderen Aktivitäten nachgeht (nicht notwendigerweise Aktivitäten, die kodiert werden – auch länger aufs Material schauen oder auf den Stift beißen zählen oder auch länger ruhig vor sich hinstarren).
Jedes Zeitsegment dem nicht explizit Off-Task zugewiesen wurde, wird in der Auswertung automatisch als On-Task gewertet!

Codes	Beschreibung	Indikatoren Die Schülerin / der Schüler ...	Beispiele
>Off-Task<	Alle nonverbalen Aktivitäten und/oder verbale Beiträge, die mit großer Wahrscheinlichkeit nicht im Zusammenhang mit der Bearbeitung des Lernmaterials stehen.	a) ... redet über private bzw. nicht auf den laufenden Unterricht, das Lernmaterial oder Physik bezogene Dinge. b) ... nutzt Versuchsmaterialien für nicht auf das Lernmaterial bezogenen Aktivitäten („Herumspielen“). c) ... geht offensichtlich nicht auf den Unterricht / das Lernmaterial bezogenen Aktivitäten nach.	a) – „Hast du die Aufgaben für Deutsch gemacht?“ – „Wollen wir nachher Eis essen gehen?“ – „Letzte Woche hat unser Physiklehrer ein kariertes Hemd getragen“ b) – Baut Figuren aus Gliedermaßstab – Spielt mit Stoppuhr c) – Schminkt sich – Verlässt Platz (obwohl nicht in Mat. angelegt)

– Wenn Indikator für >Off-Task< gleichzeitig mit >Aufschreiben< einer beliebigen Person aus der Gruppe auftritt, wird **kein** >Off-Task< kodiert.
– Wenn Indikator b) gleichzeitig mit einem anderen Code der gleichen Person auftritt, wird **kein** >Off-Task< kodiert (z. B. Herumspielen während verbalen Beitrags).

Kategorie 0: Karte | Kategorie 1: Time on Task [>Off-Task<] | Kategorie 2: Aktivität (nonverbal) [>Schreiben<, >Vorb. / Durchf. (nonverbal)<, >Sonsstiges<] | Kategorie 3: Beitrag (verbal) [>Vorlesen<, >Organisation<, >Fachinhaltlich<, >Methodisch-FH<, >Methodisch-PU<, >Methodisch-AI<, >Vorb. / Durchf. (verbal)<, >Sonsstiges<] | Kategorie 4: Erleben [>Negativ<, >Indifferent<, >Positiv<]

Kategorie 2: Aktivität (nonverbal)

Codes aus dieser Kategorie werden schülerweise vergeben und können parallel zu Codes für dieselbe Person aus den Kategorien 3 und 4 vergeben werden. Codes werden phasenweise vergeben, d. h. nur unterbrochen, wenn zwischendurch die gleiche Person offensichtlich anderen Aktivitäten nachgeht (nicht notwendigerweise Aktivitäten, die kodiert werden – auch länger aufs Material schauen oder auf den Stift beißen zählen).

Codes	Beschreibung	Indikatoren Die Schülerin / der Schüler ...	Beispiele
>Schreiben<	Es wird eine Antwort, Rechnung, Zeichnung usw. aufgeschrieben; kann sowohl im Lernmaterial als auch auf einem „Schmierzettel“ o. Ä. erfolgen.	– Wenn Schüler/in beim Aufschreiben den geschriebenen Text laut „mitliest“, wird sowohl >Schreiben< als auch ein verbaler Beitrag (z. B. >Fachinhaltlich< – Code wird entsprechend der vorigen inhaltlichen Kodierung vergeben) kodiert (auch einzelne Wörter, wenn aus dem Kontext klar wird, dass sie Teil eines Sinnzusammenhangs sind – vergleichbar mit dem Fall, dass auf eine vorgelesene Frage nur ein Wort geantwortet wird). – Ankreuzen wird <u>nicht</u> kodiert, weil Zeitfenster i. d. R. nur sehr kurz sind.	
>Vorb. / Durchf. (nonverbal)<	<u>Zielgerichtete</u> Manipulation des Versuchsmaterials, zur Vorbereitung oder Durchführung eines selbst ausgedachten oder im Lernmaterial angelegten Versuchs, der zur Klärung eines Sachverhalts dient.	a) ... baut einen Versuch auf oder bereitet einen Versuch vor. b) ... führt einen Versuch durch bzw. nimmt Messwerte auf* [wird auch kodiert, wenn Schüler/in nur bei Durchführung assistiert, siehe letztes Bsp.].	a) – Wasser wird auf Bechergläser verteilt – Stativmaterial wird zusammengebaut b) – In Trinkhalme pusten und Luftblasen beobachten – Papierkegel fallenlassen und Fallzeiten aufnehmen – Hält einen Zollstock, während andere/r Schüler/in Papierkegel fallen lässt
>Sonstiges<	*Möglichkeit zur Abgrenzung gegenüber „herumspielen“ mit Trinkhalmen, Stoppuhr, Gliedmaßstab usw. (>Off-Task<): Erkenntnisinteresse („Klärung des Sachverhalts“) zeigt sich vermutlich darin, dass gezielt beobachtet wird. Kategorie umfasst <u>nicht</u> entsprechende Sprachhandlungen (z. B. „Du musst in beide Gläser gleich viel Wasser machen!“) Kategorie kann genutzt werden, um möglicherweise relevante auf den Inhalt / das Lernmaterial bezogene nonverbale Aktivitäten zu markieren, die nicht in den zuvor genannten Kategorien abgebildet sind.	– Im Coding-Kommentarfeld von MAXQDA kurz vermerken, welche Überlegung zur Zuweisung geführt hat (evtl. zur Not ein Memo erstellen).	

Kategorie 0: Karte | Kategorie 1: Time on Task [>Off-Task<] | Kategorie 2: Aktivität (nonverbal) [>Schreiben<, >Vorb. / Durchf. (nonverbal)<, >Sonstiges<] | Kategorie 3: Beitrag (verbal) [>Vorlesen<, >Organisation<, >Fachinhaltlich<, >Methodisch-FH<, >Methodisch-FH<, >Methodisch-PU<, >Methodisch-AI<, >Vorb. / Durchf. (verbal)<, >Sonstiges<] | Kategorie 4: Erleben [>Negativ<, >Indifferent<, >Positiv<]

Kategorie 3: Beitrag (verbal)

Codes aus dieser Kategorie können parallel zu Codes für dieselbe Person aus den Kategorien 2 und 4 vergeben werden.

- **NUR EIN CODE:** Grundsätzlich kann **jedem Beitrag nur ein Code** aus dieser Kategorie zugewiesen werden. Ausnahmen bilden einzig die vermuteten methodischen Beiträge, die immer mit einem weiteren Beitrag kodiert werden – diese Ausnahmen für Doppelkodierungen werden unten an den entsprechenden Stellen aufgeführt! Der zugewiesene Code wird gemäß der **einen Hauptfunktion** des Beitrags gewählt. Zum Beispiel können fachmethodische Äußerungen durchaus fachinhaltliche Aspekte beinhalten, die aber nicht mitkodiert werden, weil die Hauptfunktion des Beitrags einer fachmethodischen Äußerung entspricht (so ist der Nebensatz in „Hier wurde die Masse kontrolliert, weil die beiden Gegenstände gleich viel wiegen“ fachinhaltlich, trotzdem wird der gesamte Beitrag nur als fachmethodisch kodiert).
- **FACHLICH UNANGEMESSENE BEITRÄGE:** Fachbezogene Beiträge (inhaltlich und methodisch) können sowohl fachlich angemessen als auch unangemessen sein. Insbesondere werden **unangemessene Beiträge in gleicher Weise kodiert wie angemessene** (z. B. gehört das Nennen unangemessener Regeln für das Formulieren von naturwissenschaftlichen Fragen auch in die Kategorie **>Methodisch-FH|indiziert<**).
- **ÜBERTRAGSREGEL:** Ist aus einer **unmittelbar vorausgegangenen Äußerung** oder aus der gerade bearbeiteten Aufgabe rekonstruierbar, worauf sich eine Äußerung **bezieht** (z. B. verweisen Pronomina deutlich auf Objekte aus dem vorigen Satz oder – sehr eindeutig – aus der Aufgabe; wird ein Halbsatz einer anderen Person oder aus dem Material vervollständigt), so ist die **Äußerung in entsprechender Weise zu kodieren**. Kann allerdings die Bedeutung nicht mit großer Wahrscheinlichkeit eindeutig rekonstruiert werden, ist **>Sonstiges<** zu kodieren.

Beispiel 1: S1: „Ich frage mich irgendwie, ob hier... [bricht Satz ab]“ → Kodierung als **>Sonstiges<**, weil Aussage nicht eindeutig rekonstruiert werden kann.

Beispiel 2: Beendet ein S einen Satz eines anderen S und der erste Teil ist eigenständig nicht rekonstruierbar, ist dem ersten Teil **>Sonstiges<** zuzuweisen, dem zweiten Teil die entsprechende Kategorie des gesamten Satzes.

Beispiel 3: „[Ich würde hier als Antwort aufschreiben:] weil die Wassersäule größer ist“ → mit **>Fachinhaltlich<** zu kodieren.

Beispiel 4: S1 liest „Was sind hier die Kontrollvariablen“, S2 [nach 2 Sekunden]: „Material und Höhe“ → Kodierung als **>Methodisch-PU|indiziert<** (Deutung der Aussage von S2: „Hier sind die Kontrollvariablen Material und Höhe“).

- **3-SEKUNDEN-REGEL:** Haben aufeinanderfolgende Beiträge derselben Person, denen der gleiche Code zugewiesen wird, einen Abstand von weniger als 3 Sekunden, so werden sie als ein Beitrag kodiert. Haben Beiträge, die zum gleichen Satz oder zum gleichen Code gehören, einen Abstand von mehr als 3 Sekunden, so werden sie einzeln (und ggf. jeweils mit dem Code des gesamten Satzes) kodiert.

Achtung: Für diese Regel werden die Doppelzuweisungen „>Methodisch-XY|vermutet< + >Vorb./Durchf. Versuch (V)<“ und „>Methodisch-XY|vermutet< + >Fachinhaltlich<“ als jeweils **ein Code** verstanden (der von allen hier real aufgeführten Codes zu unterscheiden ist).

Beachte: Die 3-Sekunden-Regel **regiert** auch bei inhaltlich verschiedenen Beiträgen des gleichen Codes. Dies gilt insbesondere auch (a) bei Ping-Pong-Unterhaltungen und (b) beim Kartenwechsel (für alle Kategorien inkl. **>Vorlesen<**).

- **FÜLLWÖRTER:** Kurze Füllwörter (z. B. „Ähm“, „Hm“, „Ok“) werden **nicht kodiert**. (Ausnahme: Die 3-Sekunden-Regel sorgt dafür, dass Füllwörter zwischen zwei Beiträgen, die weniger als 3 Sekunden voneinander entfernt sind, vom Code überdeckt werden.)

- **KURZANTWORTEN: Zustimmungde / ablehnende** Antworten und Wiederholungen **ohne konkreten Inhaltsbezug** (z. B. „Ja“, „Okay“, „Ne“, „Stimmt“, ...) werden **nicht** kodiert.

Kategorie 0: Karte | Kategorie 1: Time on Task [**>Off-Task<**] | Kategorie 2: Aktivität (nonverbal) [**>Schreiben<**, **>Vorb./Durchf. (nonverbal)<**, **>Sonstiges<**] | Kategorie 3: Beitrag (verbal) [**>Vorlesen<**, **>Organisation<**, **>Fachinhaltlich<**, **>Methodisch-FH<**, **>Methodisch-PU<**, **>Methodisch-AL<**, **>Vorb./Durchf. (verbal)<**, **>Sonstiges<**] | Kategorie 4: Erleben [**>Negativ<**, **>Indifferent<**, **>Positiv<**]

Codes	Beschreibung	Indikatoren Die Schülerin / der Schüler ...	Beispiele
	Schüler/in liest im Lernmaterial vorgegeben Text (auch einzelnes Wort) laut vor.	a) ... hat den Blick auf das Lernmaterial gerichtet und liest ein Text aus dem Lernmaterial vor. b) ... spricht von Mitschüler/in vorgelesenen Text (nahezu) wortwörtlich.	
>Vorlesen<	- Wird auch zugewiesen, wenn Schüler/in murmelt. - Wird nicht zugewiesen, wenn der vorgelesene Text eindeutig Teil eines Beitrags ist, der einer anderen Kategorie zuzuordnen wäre. Beispiel: „Aber hier steht doch ...“ [liest Text von Karte vor], deshalb muss das hier eine präzise Frage sein“ → Kodierung als >Methodisch-FH<.	a) ... äußert, diskutiert oder fragt danach, wie eine Aufgabe zu verstehen bzw. zu bearbeiten ist.* – „Sollen wir da jetzt was aufschreiben?“ – „Da stand doch, wir sollen zuerst eine Frage aufschreiben“ b) ... äußert, diskutiert oder fragt nach dem Arbeitsablauf (was?), dem Arbeitsstand (wie weit?) oder die Arbeitsteilung (wer?) in der Gruppe.* – „Kann ich umblättern?“ – „Was sollen wir als nächstes machen?“ – „Du kannst schon mit Aufbauen anfangen“	
>Organisation<	Beitrag, der auf die Organisation von Arbeitsabläufen in der Gruppe oder auf die Klärung von Aufgabenstellungen abzielt		
	* Ausnahme: Wenn es sich um eine Äußerung/Frage im Rahmen einer experimentellen Aufgabe handelt, wird immer >Vorbereitung/Durchführung von Versuch (verbal)< (s. u.) kodiert. Abgrenzung zu >Vorb./Durchf.< liegt darin, dass man eine Orga-Äußerung für jedes beliebige Experiment machen könnte („Bau mal auf“ vs. „Hol mal die Schlitten aus der Kiste“ letzteres als >Vorb./Durchf.<).		
	a) ... stellt eine inhaltsbezogene (nicht organisatorische) Frage. b) ... trifft eine Vorhersage über den Ausgang eines Versuchs (mit und ohne Begründung). c) ... äußert / diskutiert über eine Beobachtung (auch Messwerte) zu einem gerade durchgeführten oder erinnerten Versuch (ggf. mit Deutung). d) ... überträgt ein inhaltsbezogenes Konzept auf einen konkreten Fall (z. B. zur Begründung einer Vorhersage oder Deutung einer Beobachtung). e) ... äußert / diskutiert ein inhaltsbezogenes Konzept.	a) – „Das hat doch was mit der Masse des Wassers zu tun, oder?“ – „Hängt die Fallgeschwindigkeit vom Gewicht ab?“ b) – „Ich glaube, im linken Glas steigen zuerst Luftblasen auf“ – „Die Blasen werden größer sein, weil der Durchmesser vom Halm größer ist“ c) – „Im linken Glas sind zuerst Blasen aufgestiegen!“ – „Da sind doch vorhin in beiden Gläsern gleichzeitig Blasen aufgestiegen“ – „Die Stoppuhr zeigt 1,2 Sekunden“ d) – „Der große Kegel wird langsamer fallen, weil die Luftreibung größer ist“ – „In der Unterwasserhöhle ist der Druck kleiner, weil dort die Wassersäule kleiner ist“ e) – „Der Schweredruck hängt immer von der Höhe der Wassersäule ab“ – „Die Wassermenge im Topf hat Einfluss auf die Kochzeit“	
>Fachinhaltlich<	Fachinhaltsbezogene Auseinandersetzung mit einem (Gedanken-)Versuch, z. B. Beobachtungen, Deutungen, ...		
Kategorie 0: Karte Kategorie 1: Time on Task [>Off-Task<] Kategorie 2: Aktivität (nonverbal) [>Schreibens<, >Vorb./ Durchf. (nonverbal)<, >Sontiges<] Kategorie 3: Beitrag (verbal) [>Vorlesens<, >Organisation<, >Fachinhaltlich<, >Methodisch-FH<, >Methodisch-PUS<, >Methodisch-AIS<, >Vorb./ Durchf. (verbal)<, >Sontiges<] Kategorie 4: Erleben [>Negativ<, >Indifferent<, >Positiv<]			

Codes	Beschreibung	Indikatoren Die Schülerin / der Schüler ...	Beispiele
	Äußerungen, die sich auf das Formulieren naturwi. Fragen (F), Vermutungen (V) und Hypothesen (H) beziehen. Hierzu gehören u. a. Frage, Aussage, Beispiel oder Regel dazu, ...was naturwi. F/V/H kennzeichnet und wie F/V/H aus methodischer Sicht beschrieben und/oder bewertet werden können (z. B. präzise, allgemein, naturwissenschaftlich). ...worauf es aus methodischer Sicht beim Formulieren von naturwi. F/V/H ankommt und wie man dabei vorgehen sollte (z. B. Passung von F zu V/H). >Methodisch-FH<	a) ... nutzt fachmethodische Kategorien / Kriterien / Regeln beim Beschreiben / Bewerten naturwissenschaftlicher F/V/H. b) ... nennt / beschreibt / erläutert / erfragt fachmethodische Kategorien / Kriterien / Regeln bezüglich der Beschreibung / Bewertung von naturwi. F/V/H. c) ... nutzt fachmethodische Regeln / Strategien beim Formulieren von naturwi. F/V/H. d) ... nennt / beschreibt / erläutert / erfragt fachmethodische Regeln / Strategien bezüglich des Formulierens von naturwi. F/V/H.	– „Die Frage ... ist auf jeden Fall eine naturwissenschaftliche Frage“ – „Das ist doch eine Vermutung, oder?“ – „... [nennt Frage] wäre z. B. auch eine naturwissenschaftliche Frage, weil man Sie untersuchen kann“ – „Die Frage ist präzise, die legt doch alles genau fest“ – „f) ist ganz sicher eine“ [Aufgabe erfordert, das Fragen bzgl. Krit. eingeordnet werden] – „Ich suche jetzt alle Sätze raus, die eine wissenschaftliche Begründung drin haben“ [vermutet] – „Eine naturwissenschaftliche Frage muss immer was mit Physik zu tun haben“ – „Naturwissenschaftliche Fragen muss man immer eindeutig beantworten können“ – „Du musst jetzt auch noch in die Frage einbauen, welches Merkmal beobachtet werden soll“ [indiziert] – „Die Frage musst du genauer aufschreiben“ [vermutet] – „Wir sollen doch nur eine Vermutung aufschreiben, wenn wir sie auch begründen können“ – „Muss man bei einer Vermutung immer auch eine Begründung schreiben?“
Jeder Beitrag wird genau einer der Kategorien rechts zugeordnet	>vermuteter Konzeptbezug< Art oder Kontext der Äußerung lässt vermuten, dass ein Indikator erfüllt ist, es gibt jedoch keine konkreten (indizierten) Hinweise. >indizierter Konzeptbezug< Einzelnes Wort oder Begründungselement aus der Äußerung zeigt vergleichsweise deutlich, dass ein Indikator erfüllt ist, oder SuS formulieren selbst eine fachmethodische Regel!	– „Die Frage stimmt so glaub ich nicht“ [Wenn sich „stimmen“ im Kontext vmtl. auf Methodisches bezieht (z. B. zu allgemein formulierte Frage)] – „Ich suche jetzt alle Sätze raus, die eine wissenschaftliche Begründung drin haben“ [Vermutung: Unterschied Vermutung/Hypothese] – „Du musst jetzt auch noch in die Frage einbauen, welches Merkmal beobachtet werden soll“ – „Die Frage ist präzise, die legt doch alles genau fest“ – „f) ist ganz sicher eine naturwissenschaftliche Frage“ – „Naturwissenschaftliche Fragen muss man immer eindeutig beantworten können“ [Regel]	– „Die Frage ... ist auf jeden Fall eine naturwissenschaftliche Frage“ – „Das ist doch eine Vermutung, oder?“ – „... [nennt Frage] wäre z. B. auch eine naturwissenschaftliche Frage, weil man Sie untersuchen kann“ – „Die Frage ist präzise, die legt doch alles genau fest“ – „f) ist ganz sicher eine“ [Aufgabe erfordert, das Fragen bzgl. Krit. eingeordnet werden] – „Ich suche jetzt alle Sätze raus, die eine wissenschaftliche Begründung drin haben“ [vermutet] – „Eine naturwissenschaftliche Frage muss immer was mit Physik zu tun haben“ – „Naturwissenschaftliche Fragen muss man immer eindeutig beantworten können“ – „Du musst jetzt auch noch in die Frage einbauen, welches Merkmal beobachtet werden soll“ [indiziert] – „Die Frage musst du genauer aufschreiben“ [vermutet] – „Wir sollen doch nur eine Vermutung aufschreiben, wenn wir sie auch begründen können“ – „Muss man bei einer Vermutung immer auch eine Begründung schreiben?“
	Wird eine F/V/H als Beispiel formuliert, um eine Regel zu verdeutlichen oder ein Merkmal hervorzuheben (siehe u. a. Bsp. all), dann als >Methodisch-FH verm./ind.< kodieren. Kann parallel mit >Fachinhaltlich< kodiert werden, wenn der Konzeptbezug, der die Aussage zu „Methodisch“ werden lässt, nur ein >vermuteter Konzeptbezug< ist.		
Kategorie 0: Karte Kategorie 1: Time on Task [>Off-Task<] Kategorie 2: Aktivität (nonverbal) [>Schreibens<, >Vorb. / Durchf. (nonverbal)<] Kategorie 3: Beitrag (verbal) [>Vorlesens<, >Organisation<, >Fachinhaltlich<, >Methodisch-FH<, >Methodisch-PU<, >Methodisch-AI<, >Vorb. / Durchf. (verbal)<] Kategorie 4: Erleben [>Negativ<, >Indifferent<, >Positiv<]			

Codes	Beschreibung	Indikatoren Die Schülerin / der Schüler...	Beispiele
	Äußerungen, die sich auf das methodische Vorgehen beim Planen einer Untersuchung beziehen. Hierzu gehören u. a. Frage, Aussage, Beispiel oder Regel dazu, ...wie Untersuchungen aus methodischer Sicht beschrieben und/oder bewertet werden können (z. B. UV*, AV*, KV*, fair / unfair, Passung von UV/AV zu F). ...worauf es aus methodischer Sicht beim Planen und Durchführen eines Versuchs ankommt und wie man dabei vorgehen sollte (z. B. Variablenkontrolle). *: Unabhängige Variable (UV), abhängige Variable (AV), Kontrollvariable (KV). >Methodisch-PU<	a) ... nutzt fachmethodische Kategorien / Kriterien / Regeln beim Beschreiben / Bewerten eines Versuchs. b) ... nennt / beschreibt / erläutert / erfragt fachmethodische Kategorien / Kriterien / Regeln bezüglich der Beschreibung / Bewertung eines Versuchs. c) ... nutzt fachmethodische Regeln / Strategien beim Planen von Untersuchungen. d) ... nennt / beschreibt / erläutert / erfragt fachmethodische Regeln / Strategien bezüglich des Planens von Untersuchungen.	– „Die Variable, die in diesem Versuch beobachtet wird, ist ...“ – „Die unabhängige Variable ist bei diesem Versuch...“ – „Ich glaube, die beiden Versuche kann man nicht vergleichen, denn die Masse hat auch einen Einfluss.“ – „Im Versuch wird ja ... verändert, die Frage fragt aber nach ...“ [vermutet] – „Die unabhängige Variable erkennt man daran, dass...“ – „Bei einer fairen Untersuchungen darf immer nur...“ – „Man lässt zwei gleich schwere Kegel aus gleicher Höhe fallen“ [vermutet] – „Wir dürfen immer nur eine Variable gleichzeitig verändern“ [indiziert] – „Du musst die beiden Trinkhalme gleich tief eintauchen, sonst ist das kein fairer Versuch“ [indiziert] – „Achte darauf, dass du beide Kegel <i>gleich hoch</i> hältst!“ [vermutet] – „Bei einer Untersuchung sollte immer nur eine Variable gleichzeitig geändert werden“ – „Bei der Planung einer Untersuchung muss man darauf achten, dass sie zur Fragestellung passt“
Jeder Beitrag wird genau einer der Kategorien rechts zugeordnet	>vermuteter Konzeptbezug< Art oder Kontext der Äußerung lässt vermuten, dass ein Indikator erfüllt ist, es gibt jedoch keine konkreten (indizierten) Hinweise. >indizierter Konzeptbezug< Einzelnes Wort oder Begründungselement aus der Äußerung zeigt vergleichsweise deutlich, dass ein Indikator erfüllt ist, oder SuS formulieren selbst eine fachmethodische Regel!	– „Man lässt zwei gleich große Kegel, zur gleichen Zeit aus gleicher Höhe fallen“ [Vermutung: Variablenkontrolle] – „Im Versuch wird ja ... verändert, die Frage fragt aber nach ...“ [Vermutung: Passung Frage–Versuch] – „Hier wird Variblenkontrolle verletzt.“ – „Bei einer Untersuchung sollte immer nur eine Variable gleichzeitig geändert werden“ [Regel] – „Hier ist doch der Durchmesser die abhängige Variable, weil wir den im Versuch beobachten“	
	– Kann parallel mit >Vorb. / Durchf. (verbal)< kodiert werden, z. B. wenn Anweisung mit Konzept aus dem Bereich PU begründet wird; Indikator/Beispiel c) – Kann parallel mit >Fachinhaltlich< kodiert werden, wenn der Konzeptbezug, der die Aussage zu „Methodisch“ werden lässt, nur ein >vermuteter Konzeptbezug< ist.		
Kategorie 0: Karte Kategorie 1: Time on Task [>Off-Task<] Kategorie 2: Aktivität (nonverbal) [>Schreiben<, >Vorb. / Durchf. (nonverbal)<, >Sonstiges<] Kategorie 3: Beitrag (verbal) [>Vorlesen<, >Organisation<, >Fachinhaltlich<, >Methodisch-FH<, >Methodisch-PU<, >Methodisch-AI<, >Vorb. / Durchf. (verbal)<, >Sonstiges<] Kategorie 4: Erleben [>Negativ<, >Indifferent<, >Positiv<]			

Codes	Beschreibung	Indikatoren Die Schülerin / der Schüler...	Beispiele
Äußerungen, die sich auf das methodische Vorgehen beim Auswerten und Interpretieren von Untersuchungen beziehen. Hierzu gehören u. a. Frage, Aussage, Beispiel oder Regel dazu, ...wie Aspekte der Auswertung (z. B. Tabellen, Diagramme, etc.) oder der Interpretation (z. B. Beobachtung (B), Deutung (D)) aus methodischer Sicht beschrieben und/oder bewertet werden können (z. B. Skalierung / Beschriftung von Tab. / Diag., Klassifizieren als B/D). ...worauf es aus methodischer Sicht beim Auswerten (z. B. Anlegen von Diag. / Tab.) und Interpretieren (z. B. Trennung B/D, Passung von D zu B oder von D zu F) eines Versuchs ankommt und wie man dabei vorgehen sollte. >Methodisch-AI< >Methodisch-AI<		a) ... nutzt fachmethodische Kategorien / Kriterien / Regeln beim Beschreiben / Bewerten einer Auswertung oder einer Interpretation. b) ... nennt / beschreibt / erläutert / erfragt fachmethodische Kategorien / Kriterien / Regeln bezüglich der Beschreibung / Bewertung von Auswertungen und Interpretationen. c) ... nutzt fachmethodische Regeln / Strategien beim Auswerten und Interpretieren von Versuchen. d) ... nennt / beschreibt / erläutert / erfragt fachmethodische Regeln / Strategien bezüglich des Auswertens und Interpretierens von Versuchen.	– „Diese Aussage hier ist sicher eine Beobachtung, weil sie noch keine Interpretation enthält“ – „Das kann man doch gar nicht beobachten, oder?“ – „Sind bei dieser Tabelle die Spalten vertauscht?“ – „Auf diese Achse muss die ..., denn das ist die unabhängige Variable bei diesem Versuch“ – „Die Interpretation stimmt schon, passt aber irgendwie nicht zu der Vermutung oben“ – „In einer Tabelle muss die unabhängige Variable immer in die erste Spalte“ – „Messwerte sind immer Beobachtungen“ – „Eine Schlussfolgerung muss doch immer zur Frage passen, oder?“ – „Wie waren nochmal die Regeln für das Skalieren der Achsen?“ – „Schreib mal die Variable ... auf die x-Achse und ... auf die y-Achse“ – „Kreuz mal die Aussagen a) und b) an, das sind sicher Beobachtungen“ – „Die Achsen müssen so sein, dass die Punkte das gesamte Diagramm ausfüllen“ – „Man darf beim Auswerten halt nicht einzelne Werte weglassen“ – „Muss die Zeit immer auf die x-Achse?“
	>vermuteter Konzeptbezug< Art oder Kontext der Äußerung lässt vermuten, dass ein Indikator erfüllt ist, es gibt jedoch keine konkreten (indizierten) Hinweise.	– „Schreib mal die Variable ... auf die x-Achse und ... auf die y-Achse“ [Vermutung: Achsenbeschriftung] – „Das kann man doch gar nicht beobachten, oder?“ [Vermutung: Beobachtung/Deutung wenn der Kontext es hergibt]	
	Jeder Beitrag wird genau einer der Kategorien rechts zugeordnet >indizierter Konzeptbezug< Einzelnes Wort oder Begründungselement aus der Äußerung zeigt vergleichsweise deutlich, dass ein Indikator erfüllt ist, oder SuS formulieren selbst eine fachmethodische Regel formulieren!	– „Die Interpretation stimmt schon, passt aber irgendwie nicht zu der Vermutung oben“ – „Diese Aussage hier ist sicher eine Beobachtung, weil sie noch keine Interpretation enthält“ – „Messwerte sind immer Beobachtungen“ [Regel]	
– Kann parallel mit >Fachinhaltlich< kodiert werden, wenn der Konzeptbezug, der die Aussage zu „Methodisch“ werden lässt, nur ein >vermuteter Konzeptbezug< ist.			
Kategorie 0: Karte Kategorie 1: Time on Task [>Off-Task<] Kategorie 2: Aktivität (nonverbal) [>Schreibens<, >Vorb. / Durchf. (nonverbal)<] Kategorie 3: Beitrag (verbal) [>Vorlesens<, >Organisation<, >Fachinhaltlich<, >Methodisch-FH<, >Methodisch-PU<, >Methodisch-AI<, >Vorb. / Durchf. (verbal)<] Kategorie 4: Erleben [>Negativ<, >Indifferent<, >Positiv<]			

Codes	Beschreibung	Indikatoren Die Schülerin / der Schüler...	Beispiele
>Vorb. / Durchf. (verbal)<	Fragen, Aussagen oder Handlungsanweisung bzgl. Vorbereitung oder Ausführung eines (Gedanken-)Versuchs	a) ... gibt eine oder fragt nach einer Handlungsanweisung zu Aufbau / Durchführung eines (Gedanken-)Versuchs aus dem Lernmaterial.*	a) – „Du musst jetzt in beide Becher gleich viel Wasser schütten!“ – „3, 2, 1... los!“ – „Ihr müsst die Papierkegel gleichzeitig fallen lassen!“
		b) ... beschreibt oder erfragt den Aufbau / Ablauf eines geplanten (Gedanken-)Versuchs aus dem Lernmaterial.*	b) – „Der Trichter muss jetzt oben festgemacht werden und dann können wir dort Wasser einfüllen!“ – „Wir könnten jetzt statt der Papierkegel auch mal mit Radiergummis fallen lassen!“ – „Ich nehme jetzt mal Schlitten 1...“
		c) ... wiederholt wörtlich oder sinngemäß eine experimentbezogene Handlungsanweisung aus dem Lernmaterial.	c) – „Da stand doch eben, du musst beide Trinkhalme gleich tief eintauchen“
– *Kann auch Versuche betreffen, die nicht explizit im Lernmaterial angelegt sind, wenn Bezug zu Lernmaterial erkennbar ist.			
– Sollte dann parallel zu >Methodisch-??< kodiert werden, wenn beispielsweise Handlungsanweisung deutliche Bezüge zu Eigenschaften / Merkmalen / Regeln aus den Bereichen FH, PU, AI enthalten. Beispiel: „Achte darauf, dass du beide Kegel gleich hoch hältst!“ → „>Vorb. / Durchf. (verbal)< + >Methodisch-PU/vermutet<“			
>Sonstiges<	Code wird genutzt, um verbale Beiträge zu markieren, die nicht von anderen Codes aus Kategorie 1, 3 oder 4 erfasst werden.	a) ... äußert sich zu Rechtschreibung, Grammatik oder anderen, nicht inhaltlichen Auffälligkeiten des Materials usw.	a) – „Da fehlt ein Komma!“ – „Das hier ist doch ein Druckfehler!“ – „Schau mal, der hat ja einen komischen Namen“
		b) Lösung einer Aufgabe wird nur wiedergegeben / beschrieben, aber nicht vorgelesen (z. B. bei Ankreuzaufgaben).	b) – [Zeigt auf Lösungskarte]: „Das Kreuz ist schon mal richtig“ – [Schaut auf Lösungskarte]: „Ok, a), c) und f)“ – Nennt Begriff/Schlüsselwort von der Lösungskarte.
		c) ... äußert einen unvollständigen Satz, der nicht genügend Informationen für eine andere Code-Zuweisung enthält.	c) – „Jetzt müssen wir da!“ – „Weißt Du noch, was wir!“
d) Mehrere Codes aus Kategorie 3 können dem Beitrag mit gleicher Plausibilität zugewiesen werden. (Entscheidung nicht möglich.)		d) – „Was hatte ich vorhin nochmal gesagt?“ – ...	
– Das sind nicht die einzigen Indikatoren, sondern nur häufig auftretende Fälle aus dieser Kategorie!			
Kategorie 0: Karte Kategorie 1: Time on Task [>Off-Task<] Kategorie 2: Aktivität (nonverbal) [>Schreiben<, >Vorb. / Durchf. (nonverbal)<, >Sonstiges<] Kategorie 3: Beitrag (verbal) [>Vorlesen<, >Organisation<, >Fachinhaltlich<, >Methodisch-FH<, >Methodisch-PU<, >Methodisch-AI<, >Vorb. / Durchf. (verbal)<, >Sonstiges<] Kategorie 4: Erleben [>Negativ<, >Indifferent<, >Positiv<]			

Seite 9

Kategorie 4: Erleben

Codes werden schülerweise vergeben und können parallel zu Codes aus allen Kategorien vergeben werden.

- Erlebensäußerungen werden nur in Bezug auf die Bearbeitungsprozesse kodiert.
- Erlebensäußerungen können nur dann parallel zu **>Off-Task<** kodiert werden, wenn sie sich explizit auf das Lern- oder Experimentiermaterial beziehen (sehr selten!).
- **Keine** rein nonverbalen Hinweise auf Erleben (z. B. lachen, Arme in die Luft reißen, ...) kodieren.

Codes	Beschreibung	Indikatoren Die Schülerin / der Schüler ...	Beispiele
>Negativ<	Verbale Äußerungen, die auf Frustration, Unzufriedenheit, Überforderung oder Resignation schließen lassen	a) ... äußert Unzufriedenheit über zeitlichen Umfang von Aufgaben / Versuchen [auch in Fragen möglich]. b) ... äußert sich bzgl. fehlender eigener Fähigkeiten oder zu hoher Anforderung einer Aufgabe [auch in Fragen möglich]. c) ... äußert Langeweile bzgl. Aufgabe / Versuch [auch in Fragen möglich].	a) – „Das dauert ja ewig bis wir fertig sind“ – „Das sind noch so viele Aufgaben, das dauert mir zu lange“ b) – „Ich kann sowas nicht“ – „Das ist viel zu schwer“ – „Ich bin einfach zu doof“ c) – „Das macht mir keinen Spaß“ – „Findet ihr das auch so langweilig?“
>Indifferent<	Als Restekategorie genutzt, um Beiträge zu erfassen, die sich nicht eindeutig >Negativ< oder >Positiv< zuordnen lassen, gleichzeitig aber eher erlebensbezogen sind, als dass sie in Kategorie 3 abgebildet werden. Dazu gehören z. B. ironische Beiträge, wenn nicht aufgrund von Betonung o. Ä. eine sichere Zuweisung erfolgen kann.	a) ... äußert Freude über Bewältigung einer Aufgabe / korrekte Vorhersage / über korrekte Lösung einer Aufgabe [auch in Fragen möglich]. b) ... hebt die eigenen Fähigkeiten bzw. die Fähigkeiten der Gruppe hervor oder lobt diese [auch in Fragen möglich]. c) ... äußert Freude an Tätigkeit oder Interesse an Ergebnissen [auch in Fragen möglich].	a) – „Yeah, richtig“ – „Stehst du, da hatte ich recht!“ – „Hab ich's nicht gesagt?“ – „Boah, wir haben alles richtig!“ b) – „Wir sind einfach die besten!“ – „Ich hab's drauf!“ c) – „Fandet ihr den Versuch auch so cool?“ – „Ich will unbedingt wissen, was dann passiert.“

Kategorie 0: Karte | Kategorie 1: Time on Task [**>Off-Task<**] | Kategorie 2: Aktivität (nonverbal) [**>Schreibens<**, **>Vorb.** / **Durchf.** (nonverbal)] | Kategorie 3: Beitrag (verbal) [**>Vorlesens<**, **>Organisation<**, **>Fachinhaltlich<**, **>Methodisch-FH<**, **>Methodisch-FH<**, **>Methodisch-PU<**, **>Methodisch-AL<**, **>Vorb.** / **Durchf.** (verbal)] | Kategorie 4: Erleben [**>Negativ<**, **>Indifferent<**, **>Positiv<**]

B.2 Gruppierung der Lernenden

Die Teams wurden nach ihren mittleren Komeptenzzuwächsen gruppiert.

Team	Zuwachs- Gruppierung	Analysiert für		
		Einheit 1	Einheit 2	Einheit 3
A	Hoch	x	x	
B	Hoch	x	x	(x)
C	Niedrig		x	
D	Hoch	x	(x)	x
E	Mittel		x	
F	Mittel	x	x	(x)
G	Niedrig	x	(x)	x
H	Hoch	x		x
I	Mittel	x		
J	Niedrig	x		x
K	Niedrig	x		
L	Niedrig	x		x
M	Mittel	x		x
Q	Niedrig	x	x	(x)
R	Niedrig	x	x	(x)
S	Mittel	x	x	(x)
T	Hoch	x	x	x
U	Hoch	x	x	x
V	Mittel	x	x	x

Anmerkungen. Eingeklammerte Markierungen bedeuten, dass eine Person fehlt.

Abbildung B.1: Gruppierung der Teams nach mittleren Kompetenzzuwächsen der jeweils zugehörigen Personen sowie Einheiten, für die Analysen vorgenommen wurden.

B.3 Bearbeitungsdauern

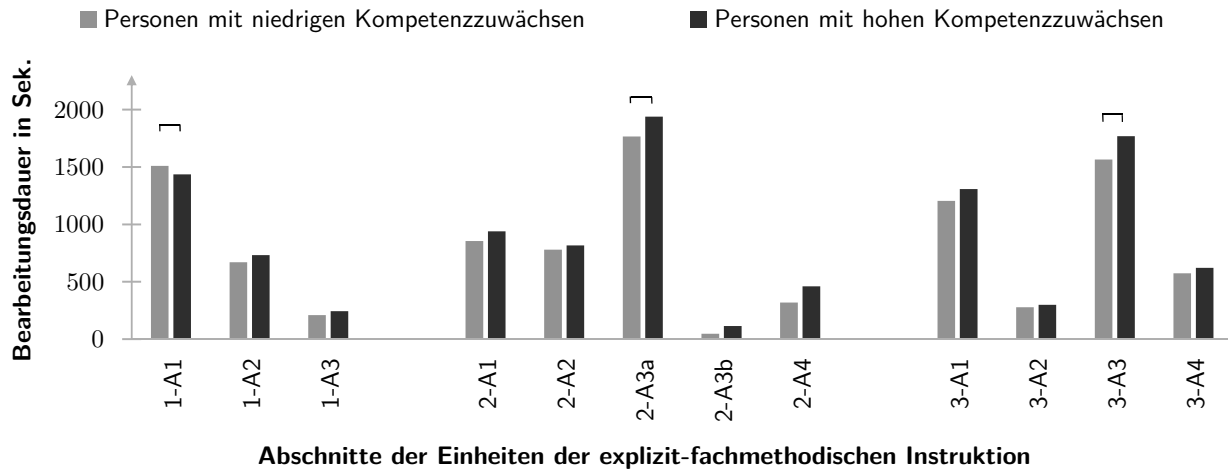


Abbildung B.2: Mittelwerte der Bearbeitungsdauern zu den einzelnen Karten der ersten Einheit der explizit-fachmethodischen Instruktion für Personen mit hohen vs. niedrigen Kompetenzzuwächsen. (Fehlerbalken entsprechen einer Standardabweichung des jeweiligen Mittelwerts; Abschnitte mit Unterschieden mit jeweils deutlichsten Unterschieden sind mit einer Klammer markiert.)

Tabelle B.1: Übersicht über die Vergleiche der Abschnittsbearbeitungsdauern für alle drei Einheiten (Personen mit hohen vs. niedrigen Kompetenzzuwächsen).

Vergleich		Zuwachs Niedrig			Zuwachs Hoch							Unterschied	
Einheit	Abschnitt	N_N	MW	SD	N_H	MW	SD	korr. df	$t(df)$	p	r	Absolut	Relativ
1	1	8	1509,4	220,6	12	1435,0	218,8	16,295	0,743	0,467	0,172	-74,4	-4,93%
	2		669,6	138,1		731,0	190,0	16,506	-0,784	0,443	0,182	61,4	8,40%
	3		208,4	59,3		243,5	28,5	9,567	-1,782	0,092	0,387	35,1	14,43%
2	1	6	854,3	136,7	11	939,1	141,4	8,408	136,71	0,251	0,295	84,8	9,03%
	2		778,8	149,4		816,5	77,5	7,331	149,38	0,498	0,177	37,7	4,62%
	3a		1764,8	216,6		1938,2	116,3	13,059	216,58	0,046	0,490	173,4	8,95%
	3b		46,3	50,0		113,1	72,9	14,007	50,04	0,065	0,457	66,7	59,02%
	4		317,8	215,5		459,1	189,3	7,720	215,46	0,181	0,341	141,3	30,78%
3	1	9	1204,7	242,3	10	1307,3	313,1	14,276	-0,791	0,44	0,189	102,6	7,85%
	2		277,6	115,1		298,9	105,7	16,996	-0,42	0,68	0,101	21,3	7,12%
	3		1565,1	205,3		1768,3	154,1	11,646	-2,457	0,025	0,512	203,3	11,49%
	4		573,0	195,9		620,4	234,3	14,599	-0,475	0,641	0,114	47,4	7,63%

Anmerkungen. Welch-Test (korr df = korrigierte Freiheitsgrade); $N_{H/N}$ =Stichprobengröße für Gruppe der Personen mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

B Ausführlichere Dokumentation zu Aktivitäten

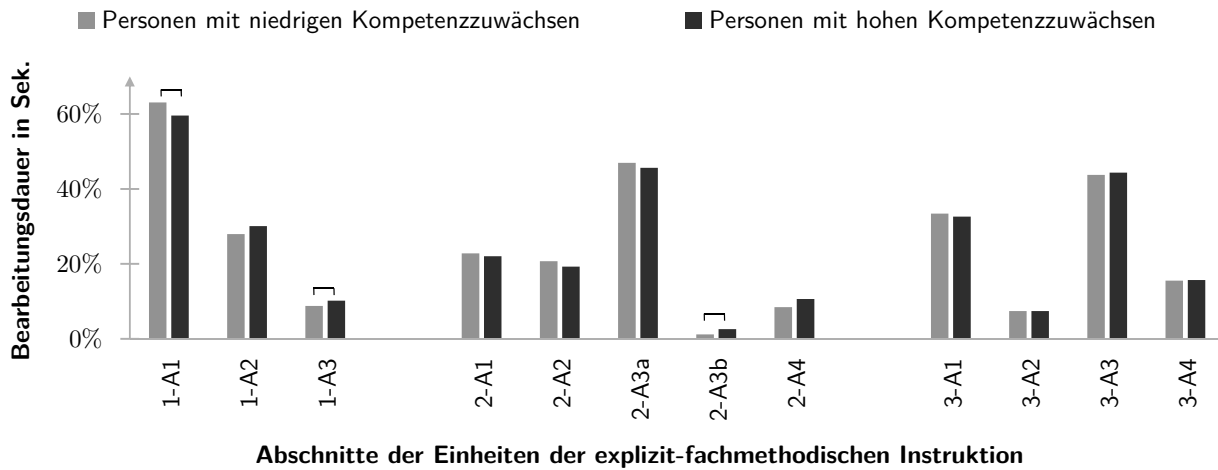


Abbildung B.3: Mittelwerte der *relativen* Anteile der Abschnittsbearbeitungsdauern an den jeweiligen Gesamtbearbeitungsdauern der Einheiten der explizit-fachmethodischen Instruktionsvariante für Personen mit hohen vs. niedrigen Kompetenzzuwächsen (Fehlerbalken entsprechen einer Standardabweichung des jeweiligen Mittelwerts; Abschnitte mit Unterschieden mit jeweils deutlichsten Unterschieden sind mit einer Klammer markiert).

Tabelle B.2: Übersicht über die Vergleiche der *relativen* Anteile der Abschnittsbearbeitungsdauern an der Gesamtbearbeitungsdauer der jeweiligen Einheit (Personen mit hohen vs. niedrigen Kompetenzzuwächsen).

Vergleich		Zuwachs Niedrig			Zuwachs Hoch									Unterschied	
Einheit	Abschnitt	N_N	MW	SD	N_H	MW	SD	korr. df	t (df)	p	r	Absolut	Relativ		
1	1	8	63,05%	5,23%	12	59,53%	5,83%	16,295	1,405	0,179	0,329	-3,51%	-5,57%		
	2		27,93%	5,06%		30,07%	5,77%	16,506	-0,87	0,394	-0,210	2,14%	7,11%		
	3		8,75%	2,39%		10,16%	1,25%	9,567	-1,54	0,156	-0,446	1,42%	13,94%		
2	1	6	22,79%	3,44%	11	22,00%	2,68%	8,408	0,027	0,639	0,165	-0,79%	-3,46%		
	2		20,68%	3,16%		19,23%	2,04%	7,331	0,02	0,344	0,350	-1,45%	-6,99%		
	3a		46,93%	1,98%		45,59%	2,60%	13,059	0,026	0,257	0,312	-1,34%	-2,85%		
	3b		1,16%	1,09%		2,57%	1,61%	14,007	0,016	0,051	-0,495	1,40%	54,67%		
3	4	9	8,43%	5,72%	10	10,60%	3,98%	7,720	0,04	0,434	-0,285	2,17%	20,43%		
	1		33,37%	3,78%		32,60%	6,85%	14,276	0,308	0,763	0,081	-0,77%	-2,31%		
	2		7,40%	1,92%		7,41%	2,11%	16,996	-0,01	0,992	-0,003	0,01%	0,13%		
	3		43,70%	3,51%		44,33%	1,80%	11,646	-0,484	0,637	-0,140	0,63%	1,42%		
	4		15,53%	3,51%		15,66%	6,11%	14,599	-0,059	0,954	-0,015	0,13%	0,84%		

Anmerkungen. Welch-Test (korr df = korrigierte Freiheitsgrade); $N_{H/N}$ =Stichprobengröße für Gruppe der Personen mit gemittelt hohen/niedrigen Kompetenzzuwächsen; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

B.4 Mittlere Anteile von Aktivitäten

Tabelle B.3: Students t-Tests zu den Gesamtbearbeitungsdauern im Vergleich von Personen mit unterschiedlichen Kompetenzzuwächsen, die die jeweilige Einheit der implizit-fachmethodischen Instruktionsvariante bearbeitet haben.

Einheit	Implizit			Explizit			df	$t(df)$	p	r	Unterschied	
	N_N	MW	SD	N_H	MW	SD					Absolut	Relativ
1	18	180,6	108,6	33	210,1	100,8	49	-0,97	0,335	0,138	29,5	14,05%
2	18	321,2	155,7	19	428,7	247,4	35	-1,572	0,125	0,257	107,5	25,08%
3	15	428,4	232,7	22	305,1	179,8	35	1,817	0,078	0,294	-123,3	-28,77%
Ges	15	914,2	438,9	9	922,5	481,9	22	-0,043	0,966	0,009	8,3	0,90%
1	6	541,7	113,3	11	630,2	188,8	15	-1,042	0,314	0,26	88,5	14,05%
2	6	963,6	160,9	7	1163,7	290,3	11	-1,497	0,163	0,411	200,1	17,20%
3	6	1071,1	400,5	8	839,2	187,2	12	1,453	0,172	0,387	-231,9	-21,65%
Ges	6	2576,3	601,1	4	2462,5	596,5	8	0,294	0,776	0,103	-113,9	-4,42%

Anmerkungen. Students t-Test; $N_{E/I}$ = Stichprobengröße für Gruppe der Teams, die die explizit-/implizit-fachmethodische Instruktionsvariante bearbeitet haben; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

Tabelle B.4: Students t-Tests zu den Gesamtbearbeitungsdauern im Vergleich von Personen mit unterschiedlichen Kompetenzzuwächsen, die die jeweilige Einheit der implizit-fachmethodischen Instruktionsvariante bearbeitet haben.

Einheit	Implizit			Explizit			df	$t(df)$	p	r	Unterschied	
	N_N	MW	SD	N_H	MW	SD					Absolut	Relativ
1	18	152,2	113,7	33	65,2	64,9	49	3,491	0,001	0,446	-87,0	-57,15%
2	18	450,0	323,2	21	147,0	114,5	37	4,019	0	0,551	-303,0	-67,34%
3	18	896,4	684,3	24	414,2	258,2	40	3,174	0,003	0,449	-482,1	-53,79%
Ges	15	370,0	231,8	9	278,2	103,9	22	1,116	0,277	0,231	-91,8	-24,82%
1	6	282,3	257,7	11	179,4	42,3	15	1,326	0,205	0,324	-102,8	-36,44%
2	6	918,8	252,5	7	249,2	152,7	11	5,893	0	0,871	-669,5	-72,87%
3	6	1651,1	497,3	8	821,5	143,6	12	4,528	0,001	0,794	-829,6	-50,25%
Ges	6	2852,1	445,0	4	1226,1	133,8	8	6,974	0,000	0,927	-1626,0	-57,01%

Anmerkungen. Students t-Test; $N_{E/I}$ = Stichprobengröße für Gruppe der Teams, die die explizit-/implizit-fachmethodische Instruktionsvariante bearbeitet haben; Unterschied Absolut = Differenz der Mittelwerte; Unterschied Relativ = Differenz der Mittelwerte geteilt durch größeren der beiden Mittelwerte.

B.5 Clusteranalysen zu Aktivitätsprofilen

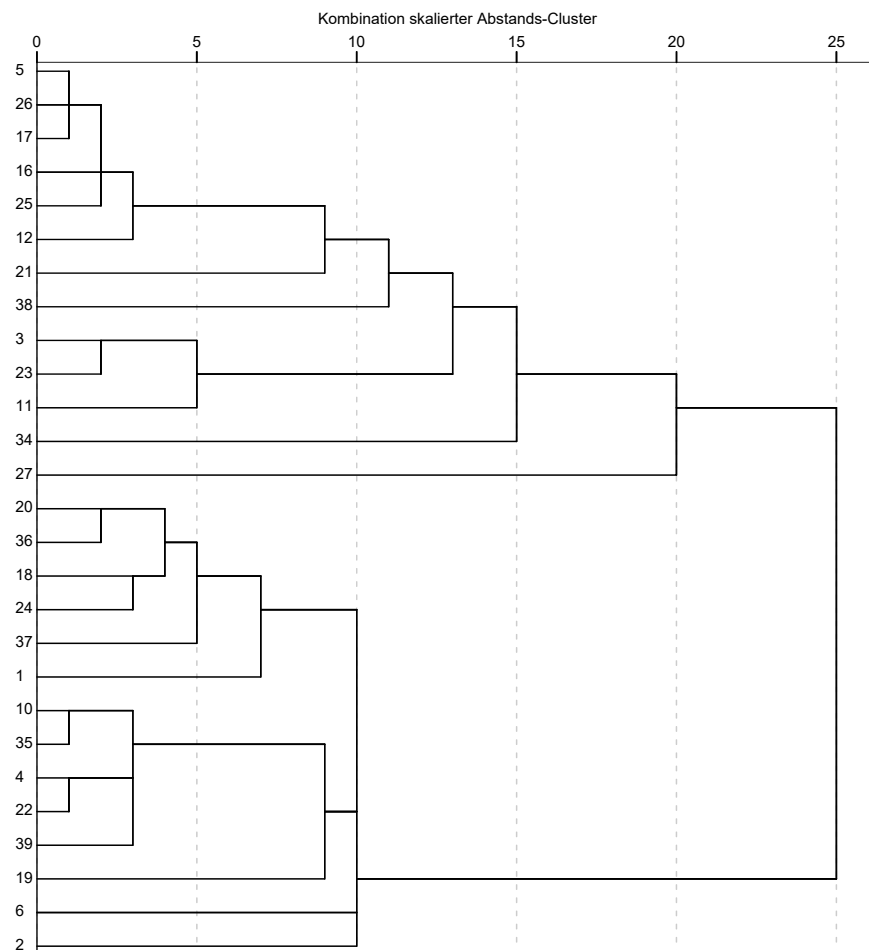


Abbildung B.4: Dendrogramm zum Prozess des Clusterings der Personen, die Einheit 1 bis zum Ende bearbeitet haben (Pearson-Korrelation für nächstgelegene Nachbarn).

B.5 Clusteranalysen zu Aktivitätsprofilen

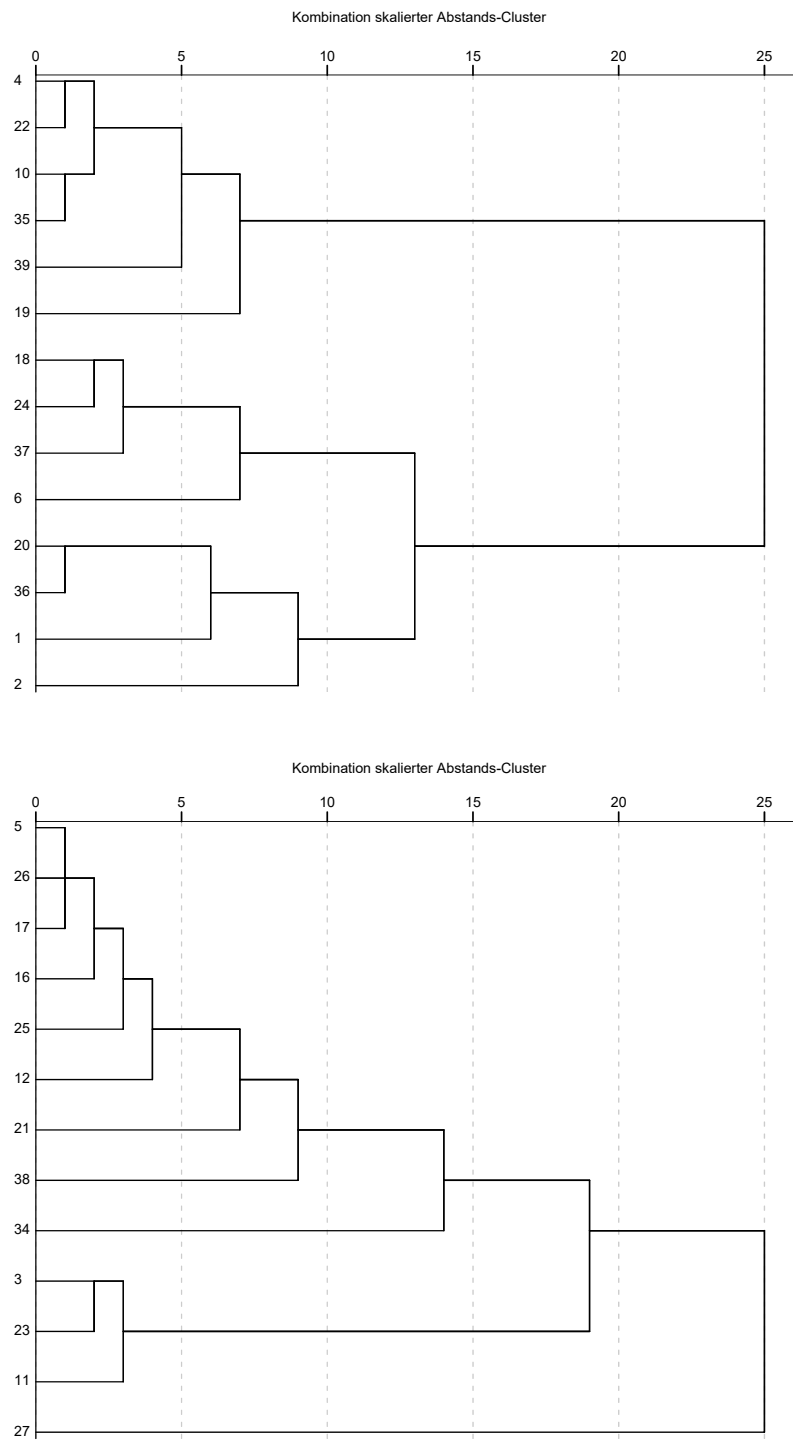


Abbildung B.5: Dendrogramm zum Prozess des Clusterings der Personen, die Einheit 1 bis zum Ende bearbeitet haben und niedrige (oben) bzw. hohen (unten) Anteile der Aktivität VorbereitenDurchführenVerbal aufweisen (Pearson-Korrelation für Verlinkung innerhalb der Cluster).

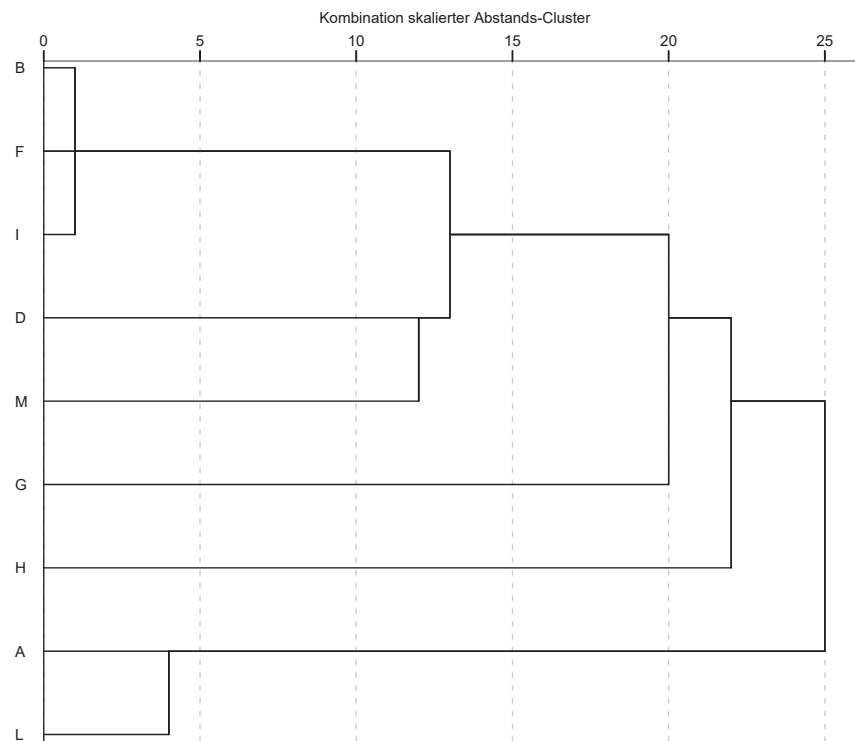


Abbildung B.6: Dendrogramm zum Prozess des Clusterings der Teams, die Einheit 1 bis zum Ende bearbeitet haben (Pearson-Korrelation für nächstgelegene Nachbarn).

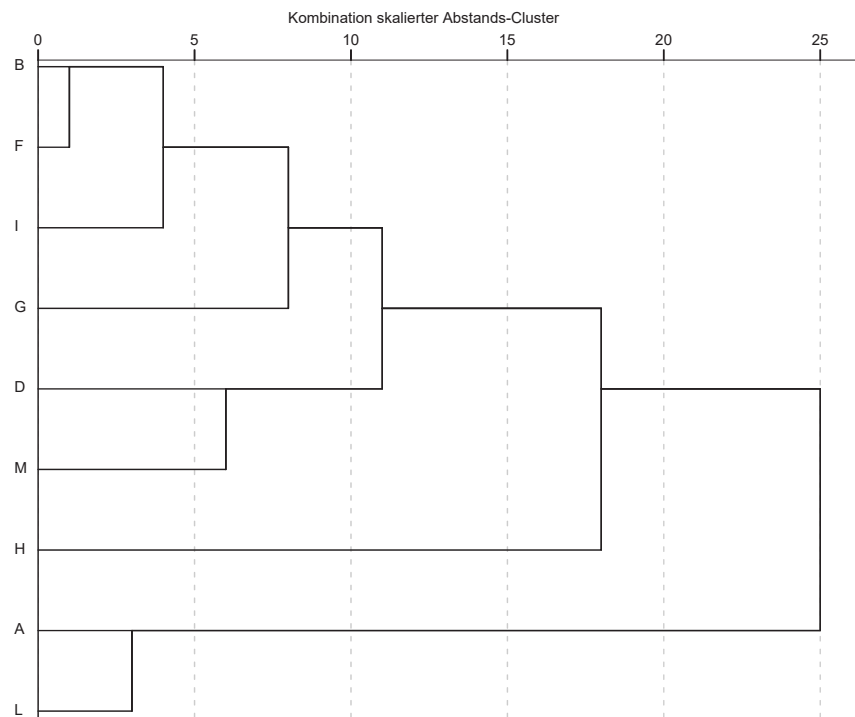


Abbildung B.7: Dendrogramm zum Prozess des Clusterings der Teams, die Einheit 1 bis zum Ende bearbeitet haben (Pearson-Korrelation für Verlinkung innerhalb der Cluster).

C Für Vorstellungen der Lernenden untersuchte Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen

21 Eine naturwissenschaftliche Fragestellung?

AUFGABE

Nicht jede Fragestellung, die etwas mit Physik, Chemie oder Biologie zu tun hat, ist automatisch eine *naturwissenschaftliche* Fragestellung. Anders herum können auch alltägliche Fragen naturwissenschaftliche Fragen sein.

Bei welchen der folgenden Fragestellungen handelt es sich Ihrer Einschätzung nach um naturwissenschaftliche Fragestellungen?

☐	Ist naturwissenschaftliches Arbeiten ein spannendes Thema für Schüler/innen?
☐	Wird Kaffee mit Milch schneller kalt als Kaffee ohne Milch?
☐	Hat der Luftdruck einen Einfluss darauf, bei welcher Temperatur Wasser kocht?
☐	Beeinflusst die Größe eines Hundes, wie oft er im Durchschnitt pro Stunde bellt?
☐	Schmeckt Mineralwasser mit Kohlensäure besser als Leitungswasser?
☐	Von welchen Faktoren hängt die Endgeschwindigkeit eines Fallschirmspringes ab?
☐	Sollten Schüler/innen etwas darüber lernen, welche Faktoren die Periodendauer eines Fadenpendels beeinflussen?

22 Eine naturwissenschaftliche Fragestellung?

KONTROLLE

Überprüfen Sie Ihre Einschätzungen!

☐	Ist naturwissenschaftliches Arbeiten ein spannendes Thema für Schüler/innen?
☒	Wird Kaffee mit Milch schneller kalt als Kaffee ohne Milch?
☒	Hat der Luftdruck einen Einfluss darauf, bei welcher Temperatur Wasser kocht?
☒	Beeinflusst die Größe eines Hundes, wie oft er im Durchschnitt pro Stunde bellt?
☐	Schmeckt Mineralwasser mit Kohlensäure besser als Leitungswasser?
☒	Von welchen Faktoren hängt die Endgeschwindigkeit eines Fallschirmspringes ab?
☐	Sollten Schüler/innen etwas darüber lernen, welche Faktoren die Periodendauer eines Fadenpendels beeinflussen?

Stellen Sie sich vor, Sie sollten jemandem erklären, was eine naturwissenschaftliche Fragestellung von einer nicht-naturwissenschaftlichen Fragestellung unterscheidet. Was würden Sie sagen? Die Beispiele oben können Ihnen bei Ihren Überlegungen helfen!

Abbildung C.1: Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen (Karten 1-21 und 1-22).

23 Eine naturwissenschaftliche Fragestellung?

AUFGABE

Es kann nicht immer eindeutig entschieden werden, ob eine Fragestellung „naturwissenschaftlich“ ist oder nicht. **Ein zentrales Merkmal einer naturwissenschaftlichen Fragestellung ist jedoch, dass sie mit naturwissenschaftlichen Methoden untersucht werden kann.**

Naturwissenschaftliche Methoden sind Messungen/Beobachtungen/Zählungen von objektivierbaren Ereignissen. Nicht objektivierbare Ereignisse sind z. B. Geschmack, Meinung, Einstellung usw.

Beispiel-Frage 1: *Hat der Luftdruck einen Einfluss darauf, bei welcher Temperatur Wasser kocht?*

Der Luftdruck kann mit einem Barometer *gemessen* werden. Die Temperatur von Wasser am Siedepunkt kann mit einem Thermometer *gemessen* werden. Das Ergebnis beider Messungen ist unabhängig davon, wer die Messungen durchführt und deshalb *objektiv*.

✓ **Es handelt sich um eine naturwissenschaftliche Fragestellung.**

Beispiel-Frage 2: *Schmeckt Mineralwasser mit Kohlensäure besser als Leitungswasser?*

Ob etwas gut schmeckt oder nicht, ist *nicht objektiv messbar*, sondern hängt von dem Geschmack der befragten Personen ab.

✗ **Es handelt sich *nicht* um eine naturwissenschaftliche Fragestellung.**

24 Eine naturwissenschaftliche Fragestellung?

AUFGABE

Warum ist „*Wird Kaffee mit Milch schneller kalt als Kaffee ohne Milch?*“ eine naturwissenschaftliche Fragestellung?

Warum ist „*Ist naturwissenschaftliches Arbeiten ein spannendes Thema für Schüler/innen?*“ keine naturwissenschaftliche Fragestellung?

Abbildung C.2: Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen (Karten 1-23 und 1-24).

Bei welchen der folgenden Fragestellungen handelt es sich Ihrer Einschätzung nach eher um naturwissenschaftliche Fragestellungen? Versuchen Sie, Ihre Einschätzung kurz zu begründen.

A: Gibt es einen Zusammenhang zwischen der Farbe einer Flüssigkeit und ihrer Dichte?

- ☐ Das ist eine naturwissenschaftliche Fragestellung, weil...
- ☐ Das ist keine naturwissenschaftliche Fragestellung, weil...

B: Wie viele Eier legt eine Henne im Durchschnitt pro Woche?

- ☐ Das ist eine naturwissenschaftliche Fragestellung, weil...
- ☐ Das ist keine naturwissenschaftliche Fragestellung, weil...

C: Muss man ein Ei bei Vollmond länger kochen, damit es hart wird?

- ☐ Das ist eine naturwissenschaftliche Fragestellung, weil...
- ☐ Das ist keine naturwissenschaftliche Fragestellung, weil...

D: Wie viel Energie benötigt man, um einen Liter Wasser von 20° C auf 80° C zu erwärmen?

- ☐ Das ist eine naturwissenschaftliche Fragestellung, weil...
- ☐ Das ist keine naturwissenschaftliche Fragestellung, weil...

Abbildung C.3: Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen (Karten 1-25 und 1-26).

27 Eine naturwissenschaftliche Fragestellung? (Teil 2)

KONTROLLE

Alle vier Fragen sind eher naturwissenschaftliche Fragestellungen, denn sie können alle mit naturwissenschaftlichen Methoden untersucht werden.

Manche Fragen werden Ihnen dennoch komisch vorgekommen sein, weil Sie die Antwort bereits sicher zu wissen glauben, z. B.: *Muss man ein Ei bei Vollmond länger kochen, damit es hart wird?*

Auch solche „unsinnigen“ Fragen können naturwissenschaftliche Fragestellungen sein, wenn sie mit naturwissenschaftlichen Methoden untersucht werden können.

Überlegen Sie für die Fragen, die Sie nicht als naturwissenschaftlich eingestuft haben: Mit welcher naturwissenschaftlichen Methode könnte man diesen Fragen nachgehen?

28 Exkurs: Die Welt der naturwissenschaftlichen Fragen

INFO

Vielleicht ist Ihnen bereits aufgefallen, dass es in der gesamten Lerneinheit bis jetzt nur um „Was...?“ und um „Wie...?“ Fragen gegangen ist. Es gibt aber auch noch eine Vielzahl andere Fragen.

„Was...?“ und „Wie...?“ Fragen

- Was passiert, wenn man zwei Lämpchen in Reihe schaltet?
- Wie hängt die Härte eines Eies mit der Kochzeit zusammen?
- ...

Andere naturwissenschaftliche Fragen

- Warum schwingt ein Pendel manchmal langsamer und manchmal schneller?
- Wieso ist der Himmel blau?
- ...

Warum sind für naturwissenschaftliche Untersuchungen „Was...?“ und „Wie...?“ Fragen in der Regel besser geeignet?

Tipp: Überlegen Sie sich für jede der Fragen in den Kästen oben, wie Sie ihr nachgehen würden. Fällt Ihnen für alle Fragen sofort eine Antwort ein?

Abbildung C.4: Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen (Karten 1-27 und 1-28).

D Für Vorstellungen der Lernenden untersuchte Sequenz zur Unterscheidung von Beobachtung und Deutung

19

Teil 2

Beobachtungen und Deutungen

20 Beobachtung oder Deutung?

AUFGABE

Stefanie, Erkan und David lassen eine Holzkugel (Radius 5 cm) und eine Feder (Länge 10 cm) gleichzeitig aus einem Meter Höhe auf den Boden fallen.

Sie diskutieren anschließend darüber, was man bei diesem Versuch beobachten konnte.

Stefanie: *Die Holzkugel und die Feder haben sich gleichzeitig in Bewegung gesetzt.*

Erkan: *Ja, und die Holzkugel ist zuerst auf dem Boden aufgekommen.*

Stefanie: *Stimmt, die Feder war langsamer, weil sie viel leichter ist als die Holzkugel.*

David: *Die Feder hat doch eine ganz andere Oberfläche als die Holzkugel.*

Erkan: *Genau, wegen ihrer Oberfläche kommt sie viel später unten an.*

Stefanie: *Aber sowohl die Holzkugel als auch die Feder werden doch mit der Erdbeschleunigung g beschleunigt.*

...

Diskutieren Sie: Bei welchen Aussagen handelt es sich um Beobachtungen?

Abbildung D.1: Sequenz zur Unterscheidung von Beobachtung und Deutung (Karten 3-19 und 3-20).

21 Beobachtung oder Deutung?

KONTROLLE

Haben Sie die folgenden Aussagen als Beobachtungen identifiziert?

Beobachtungen	<i>Die Holzkugel und die Feder haben sich gleichzeitig in Bewegung gesetzt.</i>
	<i>Ja, und die Holzkugel ist zuerst auf dem Boden aufgekommen.</i>
	<i>Die Feder hat doch eine ganz andere Oberfläche als die Holzkugel.</i>

Bei den folgenden Aussagen handelt es sich um Deutungen:

Deutungen	<i>Stimmt, die Feder war langsamer, weil sie viel leichter ist als die Holzkugel.</i>
	<i>Genau, wegen ihrer Oberfläche kommt sie viel später unten an.</i>
	<i>Aber sowohl die Holzkugel als auch die Feder werden doch mit der Erdbeschleunigung g beschleunigt.</i>

Haben Sie eine Idee, woran man erkennt, dass es sich um eine Deutung handelt?

22 Beobachtungen

INFO

Beobachtungen: Bei der Beobachtung geht es ausschließlich darum, zu beschreiben, was passiert und nicht darum, *warum* oder *wieso* es passiert. Beobachtungen müssen daher intersubjektiv überprüfbar sein. Das bedeutet, dass verschiedene Personen, die das gleiche Phänomen beobachten, zu gleichen Aussagen gelangen sollten.

Beispiel: Die Aussage „Ja, und die Holzkugel ist zuerst auf dem Boden aufgekommen“ ist eine Beobachtung, denn...

- ... sie beschreibt ausschließlich was passiert ist.
- ... alle Personen, die das Auftreffen der Holzkugel beobachtet haben, werden mit großer Wahrscheinlichkeit in dieser Aussage übereinstimmen.

Abbildung D.2: Sequenz zur Unterscheidung von Beobachtung und Deutung (Karten 3-21 und 3-22).

Deutungen: Eine Deutung geht immer über die Beschreibung von dem, was passiert ist, hinaus, z. B. durch eine zusätzliche Erklärung oder eine Verallgemeinerung.

- **Erklärung:** Macht eine Aussage darüber, **warum** etwas passiert sein könnte.
- **Verallgemeinerung:** Macht aus den Beobachtungen eine Regel, die auch für andere (meist ähnliche) Fälle gilt.

Deutungen sind **nicht intersubjektiv überprüfbar**, weil verschiedene Personen zu unterschiedlichen Erklärungen und Verallgemeinerungen kommen können.

Beispiel: Der zweite Teil der Aussage „*Stimmt, die Feder war langsamer, weil sie viel leichter ist als die Holzkugel.*“ ist eine Deutung, denn...

- ...sie beschreibt nicht nur, was passiert (Feder ist langsamer als Holzkugel), sondern auch, **warum es passiert** (Masse).
- ...**verschiedene Personen**, die diesen Versuch beobachten, können **unterschiedlicher Meinung** darüber sein, warum die Holzkugel schneller fällt als die Feder.

Hinweis: Deutungen werden manchmal auch als Interpretationen oder Schlussfolgerungen bezeichnet.

Stefanie, Erkan und David gehen über den Parkplatz ihrer Schule und sehen ein seltsam abgestelltes Auto (Abb. rechts). Da an ihrer Schule häufig Parkplatzmangel herrscht, beginnt eine hitzige Debatte.



- Stefanie: Spinnt der? Kein Wunder, dass man hier nie einen Parkplatz findet, wenn die Leute so schlecht einparken können, wie der da!
- Erkan: Ja, der steht wirklich blöd. Der hatte es bestimmt eilig, passiert mir auch manchmal.
- David: So ein Unfug! Wer sein Auto so abstellt, kann einfach nicht parken. Dem sollte man am besten den Führerschein gleich wieder abnehmen.
- Erkan: Vielleicht standen da ja vorher andere Autos nebendran und der Fahrer konnte gar nicht anders parken.
- Stefanie: Woher willst du das denn wissen? Der kann halt nicht parken, das ist alles!

...

Diskutieren Sie:

- a) Was ist in dieser Diskussion eine Beobachtung, was ist eine Deutung?
- b) Warum wäre es bei dieser Diskussion hilfreich, zwischen Beobachtungen und Deutungen zu unterscheiden?

Bildquelle: <http://listsoplenty.com/pix/more-outrageous-parking-fails> [abgerufen 20.09.13]

Abbildung D.3: Sequenz zur Unterscheidung von Beobachtung und Deutung (Karten 3-23 und 3-24).

25 Ungünstig geparkt...

KONTROLLE

Stefanie, Erkan und David sind sich in ihrer Beobachtung einig: Das Auto steht im Moment so, dass es zwei Parkplätze gleichzeitig belegt. Sie streiten sich um die Deutung, weil sie verschiedene Erklärungen haben, weshalb das Auto so dort abgestellt wurde (wenig Zeit, schlechter Autofahrer, andere geparkte Autos).

Auch Naturwissenschaftler/innen streiten sich häufig nicht um die Beobachtungen in einer Untersuchung, sondern um deren Deutung.
Welche der unterschiedlichen Deutungen richtig ist, kann dann z. B. mit weiteren Untersuchungen herausgefunden werden.

Abbildung D.4: Sequenz zur Unterscheidung von Beobachtung und Deutung (Karte 3-25).

E Transkripte der Bearbeitungen zu den beiden Instruktionsextrakten (ad Kapitel 6)

Dieser Anhang wurde ausgelagert und findet sich in dem Online-Forschungsdaten-Dokument „Transkripte und Diagramme zur Dissertation ‚Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten‘“ unter der folgenden Adresse: <http://dx.doi.org/10.22029/jlupub-542>. Nachfolgend finden sich nur die Transkriptionsregeln.

E.1 Transkriptionsregeln

Die Transkripte wurden nach den einfachen, erweiterten Transkriptionsregeln von Dresing und Pehl (2017, S. 25–32, insb. S. 29) erstellt. Abweichend wurden von den studentischen Hilfskräften institutsinterne vergleichbare Regeln verwendet. Eine Übersicht über die Regeln findet sich in Tabelle E.1 (auf S. 550).

E.2 Transkripte der Bearbeitungen zum Instruktionsextrakt für Kennzeichen von naturwissenschaftlichen Fragen

Dieser Anhang wurde ausgelagert und findet sich in dem Online-Forschungsdaten-Dokument „Transkripte und Diagramme zur Dissertation ‚Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten‘“ unter der folgenden Adresse: <http://dx.doi.org/10.22029/jlupub-542> auf den Seiten 3–63.

E.3 Transkripte der Bearbeitungen zum Instruktionsextrakt für die Unterscheidung von Beobachtung und Deutung

Dieser Anhang wurde ausgelagert und findet sich in dem Online-Forschungsdaten-Dokument „Transkripte und Diagramme zur Dissertation ‚Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten‘“ unter der folgenden Adresse: <http://dx.doi.org/10.22029/jlupub-542> auf den Seiten 64–86.

Tabelle E.1: Übersicht über die verwendeten Transkriptionsregeln als Orientierungshilfe.

Zeichen (Transk.büro)	Zeichen [Institut]	Bedeutung/Regel [Regel im Institut]	Beispiel
#hh:mm:ss-m#	#hh:mm:ss-m#	Zeitmarken („Enter“) - sind bei jedem Sprecherwechsel/-ende [-beginn] zu setzen - können regelmäßig (~ 10s) auch ohne Sprecherwechsel ergänzt werden	-
(.), (..), (...)		kurze Pausen (Anzahl der Punkte etwa Länge in Sekunden) [ca. 1-3s]	S1: Ich denke, (...) dass hier nichts passiert. #00:17:32-2#S1: So flott kann's weitergehen! #00:17:35-1# 5s #00:17:40-3#S2: Aufgabe 3.
(Xs)	Xs	längere Pausen (ab 4s) → Angabe der Länge durch Abzählen oder mithilfe von (leeren) Zeitmarken zu Beginn/am Ende der Pause	S1: Das ist eine Kraft.
S1:	S1:	Äußerung von S1	S1: Das ist doch ein (unv.) ..?...
(unv.)	..?..	Aussage/Wort beim Transkribieren nicht verstanden	S1: Das war die (Geschwindigkeit?).
(xxx?)	{xxx}	vermutete Äußerung (falls nicht eindeutig verständlich)	S1: Ich mach das. (stellt Kiste auf Tisch) Also.
(xxx)	(xxx)	Handlungsbeschreibungen, Geräusche und Beobachtungen (nicht verbal, allerdings Hervorhebung von lautem/lachendem Sprechen)	S2: (schreit) Ich hab's! (jubelt)
[xxx]	[xxx]	Kommentare zum Transkript, z. B. - Deutungen - Ansprechen einer bestimmten Person	S1: Das ist ja einfach. [möglicherweise Ironie] S1: Mhm. [zustimmend] oder S1: Ne. [ablehnend] S1: Kannst du [Tina] mir den Stift geben?
„Xxx“ o. (liest)	„Xxx“	vorgelesener Text	S1: Ich lese. „Baut nun den Versuch auf.“
,	,	Satzzeichen	S1: Ne. [ablehnend] Die a) ist falsch definitiv falsch!
.	.	- sind entsprechend der Intonation/Sprechmelodie bewusst zu setzen	S2: Das ist richtig, ne? [nach Zustimmung fragend]
!	!	- Kurzwörter wie z. B. „ne“ sind auf ihre Bedeutung hin zu prüfen und mit dem entsprechenden Satzzeichen oder [xxx] zu ergänzen	S1: Ich ha/ S2: Mach du das!
?	?	Unvollständiges Wort oder Unterbrechung eines Satzes (Satzzeichen bewusst weglassen)	S1: Also ich //würde ja sagen// S2: //Das ist glaub' ich// //wärmer.// S1: //Wenn du// das sagst.
Zu/ Zusage	-	Überlappung; paralleler Sprech-/Handungsverlauf.	S1: Ich denke, das ist <S2: Kälter!> genauso warm.
//XXX//	<Xxx>	Text in Vergleichszeichen überlappt bei beiden Sprechern. Bei mehreren Überlappung wird das Zeichen zur Unterscheidbarkeit gedoppelt	S1: Das wird da /immer so sein.
	<S2: Xxx>	Einwurf in eine Aussage (der sprechenden Person), die diese nicht dauerhaft unterbricht – max. ~3 Wörter, sonst Sprecherwechsel	S1: Das gilt da _ann immer?
IMMER	/Xxx	Betonung eines Wortes, deutliche Hebung der Stimme	S1: Wir sollen das Massstück mit einem Kilo anhängen. (hängt es an) Die Kraft ist dann zehn Newton.
nein_in	nein_in	Dehnung eines Wortes an der mit „_“ markierten Stelle	S2: (misst) Um 2,5 cm hat sich die Feder gedehnt. Wenn wir das einsetzen, bekommen wir (schreibt) D=10N/2,5cm=4 N/cm.

F Ausführlichere Dokumentationen für die Methoden und Ergebnisse zu den Vorstellungen von Lernenden im Bearbeitungsprozess (ad Kapitel 6)

F.1 Übersicht über Ideen-Oberkategorien und Ideen-Kategorien

Tabelle F.1: Übersicht über die Ideen-Oberkategorien und zugehörige Ideen-Kategorien für Kennzeichen von naturwissenschaftlichen Fragen. *Teil 1 von 2.*

Ideen-Oberkategorie	Ideen-Kategorie
Fachbezug	Hat zu tun mit Nawi
	Hat mit Biologie zu tun
	Hat mit Physik zu tun
	Hat mit Chemie zu tun
	Hat mit Astronomie zu tun
	Hat mit Pädagogik zu tun
	Hat mit Psychologie zu tun
	Hat zu tun mit Zahlen
	Alltägliches
	Wissenschaftliches
NawiEigenschaften	Naturwiss. Eigenschaften
	Arbeit mit Stoffen
	NaWiPhänomen
	NaturBezug
Persönlich	Persönlich
	Meinung
	Individuelle Ergebnisse
	Vorlieben
Empfindungen	Empfindungen / Gefühle
	Geschmack
MethodeSpezifisch	Umfrage
	Beobachtung möglich
	Messen
	Zählen
MethodeUnspezifisch	Naturwiss. Methode
	Versuch
	Paramter festlegbar
Prüfen	Prüfen
	Nachrechnen
	Untersuchen
Belegen	Faktisch belegbar
	Naturwissenschaftlich begründbar

Tabelle F.1: Übersicht über die Ideen-Oberkategorien und zugehörige Ideen-Kategorien für Kennzeichen von naturwissenschaftlichen Fragen. *Teil 2 von 2.*

Ideen-Oberkategorie	Ideen-Kategorie
BeweisenNachweisen	Beweisen Nachweisen
AntwortAusschlag	Es gibt Antwort Eindeutiges Ergebnis
Effekt	Effekt vorhanden Ergebnis entspricht Realität
Zusammenhang	Einfluss [wird untersucht] Erklärung gesucht
Objektiv	Objektiv Egal, wer experimentiert Allgemein
Subjektiv	Subjektiv Verschiedene Wahrnehmungen Diskutierbar
Klang	Hört sich so an [als ob naturw.] Klingt schlau Sinnlos Normale Frage
PF/AF	AF-Naturwissenschaftlich PF-Naturwissenschaftlich

Tabelle F.2: Übersicht über die Ideen-Oberkategorien und zugehörige Ideen-Kategorien für die Unterscheidung von Beobachtung und Deutung.

Ideen-Oberkategorie	Ideen-Kategorie
Meinung	Meinung Vorurteile
ObjektivSubjektiv	Objektiv Subjektiv
Sehen	Sehen
Beobachten	Beobachten
Beschreiben	Beschreiben
Messen	Messen
FaktenSicher	Feststellen TatsachenFakten Sicher Wissen
AuswertungAntwort	Auswertung Antwort Verallgemeinerung
BegründungErklärung	Erlären Begründung Weil Warum
Interpretation	Interpretation deuten
Vermutung	Behauptung Vermutung Theorie
Sonstiges	Prozess IdeeUnklar

F.2 Generierung der Ideen-Oberkategorien

Tabelle F.3: Ideen-Oberkategorien und zugehörige Kategorien der Formulierung (Kennzeichen von naturwissenschaftlichen Fragen).

Ideen-Oberkategorien	Formulierung				
	n+	n−	u+	u−	nn
Empfindungen	1	7	10	0	2
Persönlich	1	8	41	6	6
Subjektiv	3	4	11	1	6
Objektiv	28	1	2	5	5
MethodeSpezifisch	98	0	7	8	2
MethodeUnspezifisch	36	0	1	2	0
NawiEigenschaften	23	0	0	3	1
Zusammenhang	4	1	1	2	1
Effekt	3	0	0	7	3
Prüfen	28	0	0	0	6
Belegen	7	0	0	0	0
BeweisenNachweisen	29	0	0	2	1
AntwortAusschlag	15	2	0	1	1
PF/AF	6	0	1	4	4
Fachbezug	28	2	7	5	9
Klang	2	0	4	0	6

Anmerkungen. Eingetragen ist jeweils die Anzahl der Äußerungen, für die sowohl die jeweilige Ideen-Oberkategorie (Zeile) als auch die jeweilige Kategorie der Formulierung (Spalte) zugewiesen wurde. Grau hinterlegt sind die Spalten zu der Hauptdiagonalen von Tabelle 6.2, die überwiegt und anhand derer die Formulierung der Beschreibung der jeweiligen Ideen-Oberkategorie vorgenommen wurde.

Tabelle F.4: Ideen-Oberkategorien und zugehörige Kategorien der Formulierung (Unterscheidung von Beobachtung und Deutung).

Ideen-Oberkategorien	Formulierung				
	b+	b–	d+	d–	nn
Sehen	10	2	1	1	0
Beobachten	10	1	0	0	0
FaktenSicher	2	0	0	3	0
ObjektivSubjektiv	3	0	1	1	0
Beschreiben	4	0	0	1	0
Messen	0	0	0	0	1
BegründungErklärung	1	9	17	0	2
Vermutung	0	6	7	0	1
Interpretation	0	1	3	0	1
AuswertungAntwort	0	2	3	0	1
Meinung	0	0	1	0	2
Sonstiges	1	1	0	0	1

Anmerkungen. Eingetragen ist jeweils die Anzahl der Äußerungen, für die sowohl die jeweilige Ideen-Oberkategorie (Zeile) als auch die jeweilige Kategorie der Formulierung (Spalte) zugewiesen wurde. Grau hinterlegt sind die Spalten zu der Zelle von Tabelle 6.3, die überwiegt und anhand derer die Formulierung der Beschreibung der jeweiligen Ideen-Oberkategorie vorgenommen wurde.

Tabelle F.5: Zur Trennung der Ideen-Oberkategorien für Kennzeichen von naturwissenschaftlichen Fragen: Ähnlichkeitsmatrix (nach Rogers und Tanimoto) für die verschiedenen Personen, die zeigt, dass verschiedene Personen keine sehr ähnliche Muster aufweisen, die sich dann als Gesamtum als eine Vorstellung interpretieren lassen könnten. Ein Dendrogramm findet sich in Abbildung F.1.

Person	A1	A2	A3	G21	G19	G20	B4	B6	B5	N40	N41	N42	C9	C8	C7	O43	O44	O45	H23	H22	H24	M37	M38	M39	P46	P47	P48	F17	F18	F16	I23	I26	I27	
A1	1.00	0.68	0.39	0.39	0.33	0.45	0.39	0.45	0.33	0.39	0.60	0.33	0.39	0.33	0.33	0.52	0.39	0.68	0.52	0.88	0.33	0.23	0.52	0.33	0.23	0.39	0.28	0.28	0.28	0.33	0.45			
A2	0.68	1.00	0.45	0.60	0.52	0.52	0.33	0.28	0.28	0.33	0.52	0.52	0.33	0.33	0.52	0.45	0.45	0.33	0.60	0.60	0.60	0.19	0.28	0.45	0.39	0.28	0.45	0.33	0.45	0.23	0.28	0.28		
A3	0.39	0.45	1.00	0.45	0.52	0.39	0.45	0.39	0.39	0.33	0.52	0.68	0.45	0.45	0.52	0.45	0.60	0.45	0.33	0.45	0.33	0.45	0.52	0.60	0.39	0.52	0.60	0.45	0.33	0.33	0.52	0.68		
G21	0.39	0.60	0.45	1.00	0.88	0.68	0.33	0.28	0.28	0.33	0.52	0.68	0.45	0.45	0.23	0.39	0.33	0.33	0.45	0.45	0.33	0.28	0.39	0.45	0.28	0.19	0.33	0.33	0.45	0.33	0.39	0.39		
G19	0.33	0.52	0.88	1.00	0.60	0.39	0.33	0.33	0.33	0.28	0.45	0.60	0.52	0.28	0.45	0.28	0.39	0.39	0.39	0.39	0.39	0.28	0.45	0.52	0.33	0.23	0.39	0.28	0.39	0.39	0.45	0.45		
G20	0.45	0.52	0.39	0.68	0.60	1.00	0.39	0.33	0.33	0.19	0.45	0.33	0.39	0.28	0.33	0.39	0.39	0.39	0.52	0.39	0.39	0.33	0.33	0.45	0.33	0.39	0.52	0.39	0.52	0.52	0.60	0.33		
B4	0.39	0.33	0.45	0.33	0.39	0.39	1.00	0.68	0.68	0.23	0.52	0.39	0.78	0.78	0.68	0.45	0.60	0.45	0.45	0.33	0.39	0.68	0.45	0.52	0.33	0.39	0.52	0.39	0.52	0.23	0.23	0.60	0.52	
B6	0.45	0.28	0.39	0.28	0.33	0.33	0.68	1.00	0.78	0.39	0.45	0.23	0.68	0.52	0.45	0.52	0.68	0.52	0.39	0.39	0.52	0.45	0.45	0.52	0.45	0.45	0.19	0.39	0.19	0.10	0.39	0.45	0.60	
B5	0.33	0.28	0.39	0.28	0.33	0.33	0.68	0.78	1.00	0.39	0.45	0.23	0.68	0.52	0.45	0.52	0.68	0.52	0.39	0.39	0.39	0.45	0.60	0.39	0.45	0.45	0.19	0.39	0.28	0.19	0.39	0.45	0.45	
N40	0.39	0.33	0.33	0.23	0.28	0.19	0.23	0.39	0.39	1.00	0.52	0.39	0.23	0.33	0.28	0.33	0.45	0.45	0.23	0.33	0.45	0.28	0.19	0.33	0.28	0.23	0.33	0.33	0.23	0.23	0.28	0.39	0.39	
N41	0.60	0.52	0.52	0.39	0.45	0.45	0.52	0.45	0.45	0.52	1.00	0.60	0.52	0.68	0.45	0.39	0.68	0.39	0.52	0.39	0.52	0.33	0.33	0.52	0.45	0.33	0.39	0.28	0.28	0.39	0.45	0.45		
N42	0.33	0.52	0.68	0.52	0.60	0.33	0.39	0.23	0.23	0.39	0.60	1.00	0.39	0.52	0.60	0.28	0.39	0.28	0.28	0.28	0.28	0.45	0.52	0.33	0.33	0.52	0.33	0.33	0.52	0.39	0.28	0.33	0.45	
C9	0.39	0.33	0.45	0.45	0.52	0.39	0.78	0.68	0.68	0.23	0.52	0.39	1.00	0.60	0.52	0.45	0.60	0.45	0.45	0.33	0.33	0.52	0.68	0.60	0.52	0.39	0.33	0.33	0.33	0.14	0.14	0.45	0.39	0.52
C8	0.39	0.33	0.45	0.23	0.28	0.28	0.78	0.52	0.52	0.33	0.68	0.52	0.60	1.00	0.68	0.45	0.60	0.33	0.45	0.33	0.33	0.39	0.52	0.45	0.52	0.33	0.33	0.33	0.23	0.23	0.45	0.39	0.39	
C7	0.33	0.52	0.52	0.39	0.45	0.33	0.68	0.45	0.45	0.28	0.45	0.60	0.52	0.68	1.00	0.52	0.52	0.39	0.39	0.39	0.28	0.33	0.60	0.52	0.60	0.60	0.39	0.39	0.39	0.39	0.39	0.33	0.33	
O43	0.52	0.45	0.45	0.33	0.28	0.39	0.45	0.68	0.52	0.33	0.39	0.28	0.45	0.45	0.52	1.00	0.60	0.60	0.60	0.45	0.45	0.60	0.39	0.60	0.52	0.52	0.23	0.45	0.23	0.23	0.28	0.39	0.33	
O44	0.52	0.45	0.60	0.33	0.39	0.39	0.60	0.68	0.68	0.45	0.68	0.39	0.60	0.60	0.52	0.60	1.00	0.60	0.60	0.60	0.60	0.39	0.39	0.45	0.52	0.52	0.33	0.45	0.33	0.23	0.23	0.45	0.52	
O45	0.39	0.33	0.45	0.33	0.39	0.39	0.45	0.68	0.52	0.45	0.39	0.28	0.45	0.33	0.39	0.60	0.60	1.00	0.33	0.33	0.33	0.45	0.28	0.39	0.45	0.52	0.33	0.45	0.23	0.23	0.45	0.39	0.52	
H23	0.68	0.60	0.33	0.45	0.39	0.52	0.45	0.39	0.39	0.23	0.52	0.28	0.45	0.45	0.39	0.45	0.60	0.33	1.00	0.78	0.60	0.39	0.28	0.33	0.39	0.28	0.33	0.33	0.33	0.45	0.39	0.28		
H22	0.52	0.60	0.33	0.45	0.39	0.39	0.33	0.39	0.39	0.33	0.39	0.28	0.33	0.33	0.33	0.39	0.45	0.60	0.33	1.00	0.60	0.28	0.19	0.23	0.28	0.28	0.33	0.45	0.33	0.33	0.39	0.28		
H24	0.88	0.60	0.45	0.33	0.28	0.39	0.33	0.52	0.39	0.45	0.52	0.28	0.33	0.33	0.28	0.60	0.60	0.45	0.60	0.60	1.00	0.28	0.19	0.45	0.28	0.28	0.33	0.33	0.23	0.23	0.39	0.52		
M37	0.33	0.19	0.28	0.28	0.33	0.33	0.39	0.45	0.45	0.28	0.33	0.23	0.52	0.39	0.33	0.39	0.39	0.28	0.39	0.28	0.28	0.45	0.52	0.45	0.33	0.28	0.28	0.33	0.33	0.23	0.33	0.39	0.28	
M38	0.23	0.28	0.52	0.39	0.45	0.33	0.68	0.45	0.60	0.19	0.33	0.45	0.68	0.52	0.60	0.39	0.39	0.39	0.45	0.33	0.39	0.60	0.52	0.52	0.23	0.45	0.23	0.23	0.23	0.28	0.39	0.33	0.45	
M39	0.52	0.45	0.60	0.45	0.52	0.39	0.45	0.52	0.39	0.33	0.52	0.52	0.60	0.45	0.52	0.60	0.45	0.45	0.33	0.23	0.45	0.52	1.00	0.52	0.39	0.45	0.45	0.39	0.52	0.23	0.23	0.28	0.52	
P46	0.33	0.39	0.39	0.28	0.33	0.45	0.52	0.45	0.45	0.28	0.45	0.33	0.52	0.52	0.60	0.52	0.52	0.52	0.39	0.28	0.28	0.45	0.45	1.00	0.78	0.52	0.39	0.28	0.39	0.52	0.33	0.23		
P47	0.23	0.28	0.52	0.19	0.23	0.33	0.52	0.45	0.45	0.28	0.33	0.33	0.39	0.52	0.60	0.52	0.52	0.52	0.28	0.28	0.28	0.33	0.45	0.39	0.28	0.33	0.33	0.33	0.45	0.39	0.52	0.45	0.33	
P48	0.39	0.45	0.60	0.33	0.39	0.39	0.33	0.19	0.19	0.23	0.39	0.52	0.33	0.33	0.39	0.23	0.33	0.33	0.33	0.33	0.23	0.28	0.39	0.45	0.52	1.00	0.33	0.45	0.45	0.45	0.39	0.39		
F17	0.28	0.33	0.45	0.33	0.28	0.52	0.45	0.39	0.52	0.23	0.28	0.39	0.33	0.33	0.33	0.45	0.45	0.45	0.33	0.33	0.33	0.28	0.52	0.23	0.39	0.52	0.33	1.00	0.45	0.45	0.52	0.39	0.39	
F18	0.28	0.45	0.45	0.33	0.39	0.39	0.23	0.19	0.28	0.33	0.28	0.39	0.14	0.23	0.39	0.23	0.23	0.23	0.33	0.45	0.33	0.28	0.28	0.23	0.39	0.28	0.33	0.45	0.45	0.45	0.52	0.28		
F16	0.28	0.45	0.33	0.45	0.39	0.52	0.23	0.10	0.19	0.33	0.28	0.39	0.14	0.23	0.39	0.23	0.23	0.23	0.33	0.33	0.23	0.28	0.28	0.23	0.39	0.39	0.45	0.45	0.60	1.00	0.45	0.39	0.19	
I23	0.28	0.23	0.33	0.33	0.39	0.52	0.60	0.39	0.39	0.23	0.39	0.28	0.45	0.45	0.39	0.23	0.45	0.45	0.45	0.33	0.23	0.39	0.39	0.23	0.52	0.52	0.45	0.45	0.33	0.45	1.00	0.68	0.39	
I26	0.33	0.28	0.52	0.39	0.45	0.60	0.52	0.45	0.45	0.28	0.45	0.33	0.39	0.39	0.33	0.28	0.52	0.39	0.39	0.39	0.39	0.33	0.33	0.28	0.33	0.45	0.39	0.52	0.52	0.39	0.68	1.00	0.60	
I27	0.45	0.28	0.68	0.39	0.45	0.33	0.52	0.60	0.45	0.39	0.45	0.45	0.52	0.39	0.33	0.39	0.52	0.52	0.28	0.28	0.52	0.33	0.45	0.52	0.23	0.33	0.39	0.39	0.28	0.19	0.39	0.60	1.00	

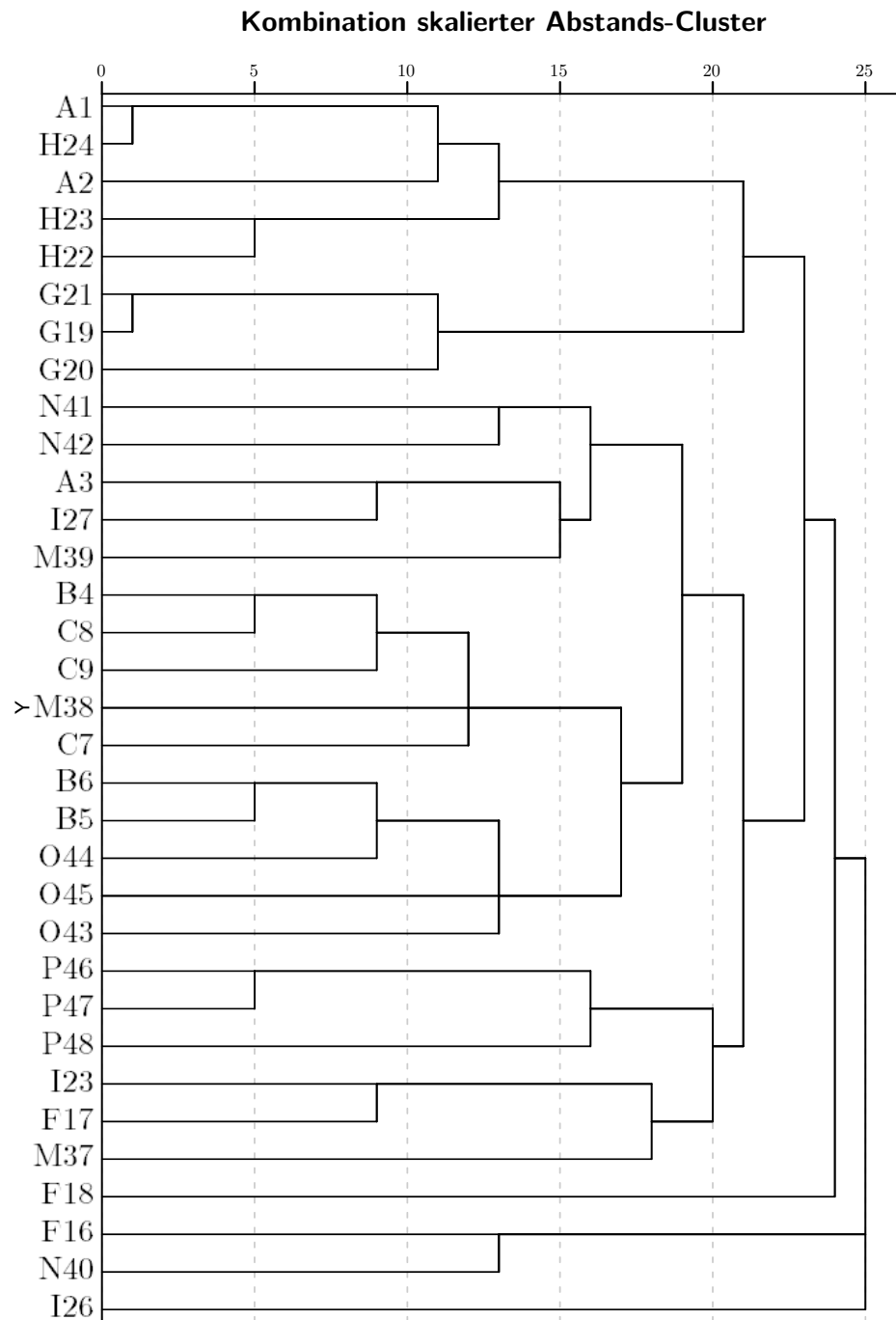


Abbildung F.1: Zur Trennung der Ideen-Oberkategorien für Kennzeichen von naturwissenschaftlichen Fragen: Dendrogram zur Clusteranalyse für die verschiedenen Personen, aus dem hervorgeht, dass verschiedene Personen keine sehr ähnliche Muster aufweisen, die sich dann als Gesamtum als eine Vorstellung interpretieren lassen könnten.

Tabelle F.6: Zur Trennung der Ideen-Oberkategorien für die Unterscheidung von Beobachtung und Deutung: Ähnlichkeitsmatrix (nach Rogers und Tanimoto) für die verschiedenen Personen, die zeigt, dass verschiedene Personen keine sehr ähnliche Muster aufweisen, die sich dann als Gesamtum als eine Vorstellung interpretieren lassen könnten. Ein Dendrogramm findet sich in Abbildung F.2.

Person	A1	A2	A3	E14	G21	K32	K31	K33	B4	B5	C7	C8	H23	H22	H24	L34	L36	L35	M37	M38	M39	P48	F16	F17	F18
A1	1,00	0,60	0,60	0,60	0,50	0,60	0,41	0,41	0,60	0,60	0,60	0,60	0,41	0,26	0,50	0,41	0,50	0,26	0,50	0,50	0,50	0,50	0,71	0,60	0,60
A2	0,60	1,00	0,71	0,50	0,85	0,71	0,71	0,50	0,71	0,50	0,71	0,71	0,33	0,33	0,26	0,71	0,85	0,50	0,60	0,41	0,60	0,60	0,60	1,00	0,71
A3	0,60	0,71	1,00	0,71	0,60	0,50	0,50	0,50	0,50	0,33	0,50	0,50	0,50	0,33	0,41	0,50	0,60	0,33	0,41	0,26	0,60	0,60	0,41	0,71	0,50
E14	0,60	0,50	0,71	1,00	0,41	0,71	0,50	0,71	0,71	0,50	0,50	0,50	0,50	0,50	0,60	0,33	0,41	0,33	0,41	0,26	0,60	0,60	0,41	0,50	0,33
G21	0,50	0,85	0,60	0,41	1,00	0,60	0,60	0,60	0,60	0,41	0,85	0,60	0,41	0,26	0,20	0,60	0,71	0,60	0,50	0,33	0,50	0,50	0,50	0,85	0,60
K32	0,60	0,71	0,50	0,71	0,60	1,00	0,71	0,71	1,00	0,71	0,71	0,71	0,33	0,50	0,41	0,50	0,60	0,50	0,60	0,41	0,60	0,60	0,60	0,71	0,50
K31	0,41	0,71	0,50	0,50	0,60	0,71	1,00	0,50	0,71	0,71	0,50	0,50	0,33	0,50	0,41	0,71	0,85	0,71	0,41	0,41	0,60	0,60	0,41	0,71	0,50
K33	0,41	0,50	0,50	0,71	0,60	0,71	0,50	1,00	0,71	0,50	0,71	0,50	0,50	0,50	0,41	0,33	0,41	0,50	0,41	0,26	0,60	0,60	0,41	0,50	0,33
B4	0,60	0,71	0,50	0,71	0,60	1,00	0,71	0,71	1,00	0,71	0,71	0,71	0,33	0,50	0,41	0,50	0,60	0,50	0,60	0,41	0,60	0,60	0,60	0,71	0,50
B5	0,60	0,50	0,33	0,50	0,41	0,71	0,71	0,50	0,71	1,00	0,50	0,50	0,33	0,50	0,60	0,50	0,60	0,50	0,41	0,60	0,60	0,60	0,60	0,50	0,50
C7	0,60	0,71	0,50	0,50	0,85	0,71	0,50	0,71	0,71	0,50	1,00	0,71	0,50	0,33	0,26	0,50	0,60	0,50	0,60	0,41	0,60	0,60	0,60	0,71	0,50
C8	0,60	0,71	0,50	0,50	0,60	0,71	0,50	0,50	0,71	0,50	0,71	1,00	0,33	0,50	0,26	0,50	0,60	0,33	0,26	0,41	0,60	0,60	0,60	0,71	0,50
H23	0,41	0,33	0,50	0,50	0,41	0,33	0,33	0,50	0,33	0,33	0,50	0,33	1,00	0,50	0,60	0,33	0,41	0,33	0,26	0,26	0,60	0,60	0,26	0,33	0,20
H22	0,26	0,33	0,33	0,50	0,26	0,50	0,50	0,50	0,50	0,50	0,33	0,50	0,50	1,00	0,60	0,33	0,41	0,33	0,26	0,26	0,60	0,60	0,26	0,33	0,20
H24	0,50	0,26	0,41	0,60	0,20	0,41	0,41	0,41	0,41	0,60	0,26	0,26	0,60	0,60	1,00	0,26	0,33	0,26	0,20	0,33	0,50	0,50	0,33	0,26	0,26
L34	0,41	0,71	0,50	0,33	0,60	0,50	0,71	0,33	0,50	0,50	0,50	0,50	0,33	0,33	0,26	1,00	0,85	0,50	0,41	0,41	0,60	0,60	0,41	0,71	0,50
L36	0,50	0,85	0,60	0,41	0,71	0,60	0,85	0,41	0,60	0,60	0,60	0,60	0,41	0,41	0,33	0,85	1,00	0,60	0,50	0,50	0,71	0,71	0,50	0,85	0,60
L35	0,26	0,50	0,33	0,33	0,60	0,50	0,71	0,50	0,50	0,50	0,50	0,33	0,33	0,33	0,26	0,50	0,60	1,00	0,41	0,41	0,41	0,41	0,41	0,50	0,33
M37	0,50	0,60	0,41	0,41	0,50	0,60	0,41	0,41	0,60	0,41	0,60	0,60	0,26	0,26	0,20	0,41	0,50	0,41	1,00	0,71	0,50	0,50	0,71	0,60	0,60
M38	0,50	0,41	0,26	0,26	0,33	0,41	0,41	0,26	0,41	0,60	0,41	0,41	0,26	0,26	0,33	0,41	0,50	0,41	0,71	1,00	0,50	0,50	0,71	0,41	0,60
M39	0,50	0,60	0,60	0,60	0,50	0,60	0,60	0,60	0,60	0,60	0,60	0,60	0,60	0,60	0,50	0,60	0,71	0,41	0,50	1,00	1,00	0,50	0,60	0,41	0,60
P48	0,50	0,60	0,60	0,60	0,50	0,60	0,60	0,60	0,60	0,60	0,60	0,60	0,60	0,60	0,50	0,60	0,71	0,41	0,50	1,00	1,00	0,50	0,60	0,41	0,60
F16	0,71	0,60	0,41	0,41	0,50	0,60	0,41	0,41	0,60	0,60	0,60	0,60	0,26	0,26	0,33	0,41	0,50	0,41	0,71	0,71	0,50	0,50	1,00	0,60	0,60
F17	0,60	1,00	0,71	0,50	0,85	0,71	0,71	0,50	0,71	0,50	0,71	0,71	0,33	0,33	0,26	0,71	0,85	0,50	0,60	0,41	0,60	0,60	0,60	1,00	0,71
F18	0,60	0,71	0,50	0,33	0,60	0,50	0,50	0,33	0,50	0,50	0,50	0,50	0,20	0,20	0,26	0,50	0,60	0,33	0,60	0,60	0,41	0,41	0,60	0,71	1,00

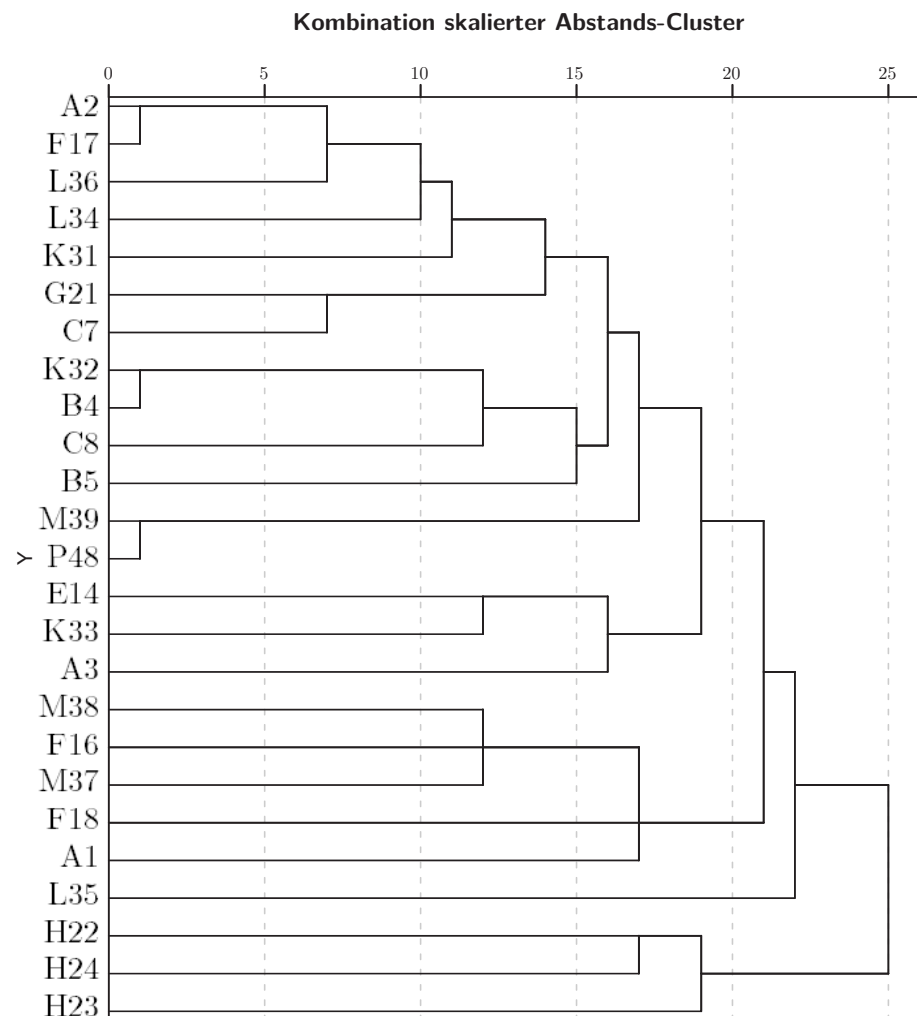


Abbildung F.2: Zur Trennung der Ideen-Oberkategorien für die Unterscheidung von Beobachtung und Deutung: Dendrogram zur Clusteranalyse für die verschiedenen Personen, aus dem hervorgeht, dass verschiedene Personen keine sehr ähnliche Muster aufweisen, die sich dann als Gesamtum als eine Vorstellung interpretieren lassen könnten.

F.3 Allgemeinheit der Äußerungen, aus denen die Vorstellungen rekonstruiert werden

Tabelle F.7: Generalisierungen pro Person für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen in Relation zu den insgesamt vorgenommenen Einschätzungen sowie der Anzahl der Transkriptabsätze pro Person.

Person	Generalisierungen	Anteil der Generalisierungen	Anzahl aller Einschätzungen	Anzahl der Transkriptabsätze	Generalisierungen außerhalb von 1-22-ii
A1	4	0,19	21	59	2
A2	1	0,06	17	50	0
A3	1	0,10	10	46	1
B4	0	0,00	5	29	0
B5	2	0,14	14	46	0
B6	1	0,07	14	52	1
C7	2	0,12	17	74	2
C8	6	0,35	17	69	3
C9	0	0,00	8	51	0
F16	3	0,16	19	83	2
F17	0	0,00	8	37	0
F18	1	0,05	19	74	0
G19	0	0,00	14	65	0
G20	0	0,00	9	42	0
G21	1	0,04	24	70	1
H22	0	0,00	15	67	0
H23	1	0,04	23	71	0
H24	1	0,09	11	51	0
I25	4	0,29	14	36	0
I26	1	0,07	14	39	0
I27	1	0,20	5	13	0
M37	0	0,00	16	62	0
M38	0	0,00	20	59	0
M39	0	0,00	6	24	0
N40	7	0,28	25	65	1
N41	4	0,13	30	65	1
N42	1	0,06	16	39	1
O43	2	0,12	17	57	0
O44	2	0,10	21	67	0
O45	1	0,04	24	64	0
P46	2	0,11	19	32	0
P47	3	0,23	13	39	0
P48	0	0,00	7	17	0
SUMME	52	–	512	–	15
MW	1,576	0,102	15,515	51,939	–
SD	1,741	0,093	6,116	17,278	–

Anmerkung. Bei Generalisierungen sind Werte fett gesetzt, die größer sind als der Mittelwert über alle Personen.

Tabelle F.8: Generalisierungen pro Person für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung in Relation zu den insgesamt vorgenommenen Einschätzungen sowie der Anzahl der Transkriptabsätze pro Person.

Person	Generalisierungen	Anteil der Generalisierungen	Anzahl aller Einschätzungen
A1	1	0,13	8
A2	0	0,00	1
A3	0	0,00	7
A0 = E14	1	0,25	4
G21	1	0,14	7
G19	0	0,00	3
G20	0		0
B4	0	0,00	6
B6	0	0,00	5
B5	2	0,40	5
K32	2	0,29	7
K31	0		0
K33	1	0,09	11
C9	0	0,00	3
C8	0	0,00	5
C7	1	0,13	8
L34	1	0,07	14
L36	1	0,08	13
L35	1	0,14	7
H23	0	0,00	4
H22	1	0,14	7
H24	1	0,25	4
M37	0	0,00	3
M38	3	0,30	10
M39	0		0
P46	1	0,10	10
P47	0		0
P48	0	0,00	6
F17	1	0,11	9
F18	0		0
F16	1	0,10	10
SUMME	21	–	178
MW	0,667	0,118	5,800
SD	0,745	0,110	3,851

Anmerkung. Bei Generalisierungen sind Werte weiß hinterlegt und gleichzeitig fett gesetzt, die Größer sind als der Mittelwert über alle Personen.

F.4 Entscheidungen der Lernenden und zugehörige Vorstellungen

F.4.1 Ausführliche Diskussion von Teams für Kennzeichen von naturwissenschaftlichen Fragen

Die beiden Teams A (oben rechts) und O (mittig links) stellen die Randpunkte des in Abbildung 6.16 dargestellten Trends dar, dass Teams mit im Mittel höherer Angemessenheit der Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen auch im Mittel korrektere Entscheidungen in diesem Inhaltsbereich treffen. Sie werden nachfolgend ausführlich besprochen.

Bei Team O haben alle Personen negative Werte beim Maß für die Angemessenheit der Ideen-Oberkategorien (siehe Tabelle F.11), was bei keinem anderen Team der Fall ist. Gleichzeitig fällt O45 durch einen deutlich niedrigeren Wert für die Angemessenheit auch innerhalb des Teams O auf. Die Korrektheit der Einschätzungen ist ebenfalls nicht gleichmäßig über die Personen verteilt. Vielmehr hat O45 einen deutlich schlechteren Wert für die Korrektheit der Einschätzungen als die anderen beiden Teammitglieder und nimmt acht der insgesamt zehn inkorrekten Einschätzungen des Teams vor. Die Einschätzungen der anderen beiden Personen sind aber in ihrer Anzahl nicht genug, um die Fehlentscheidungen von O45 im Teammaß auszugleichen.

Auf inhaltlicher Ebene könnte für das Team O die Deutung formuliert werden, dass die vorrangig unangemessenen Vorstellungen aller Personen aus dem Team die inkorrekten Entscheidungen von O45 begünstigen. Eine genauere Betrachtung der inkorrekten Einschätzungen von O45 zeigt allerdings, dass sieben von acht im Kontext der Aufgabe 1-26-C vorkommen. Gleichzeitig kann anhand des Transkriptes des zur Aufgabe gehörigen Diskurses illustriert werden, wie Vorstellungen und Entscheidungen mehrerer Personen im Prozess zueinander stehen.

O45: Ist //keine naturwissenschaftliche Frage.//
O44: [zu voriger Aufgabe]
O45: //Das ist eine na/ keine// naturwissenschaftliche Frage, weil/
O43: (..) Hä?
O44: Weil/ //(unv.)//
O45: //Muss man// ein Ei bei Vollmond länger kochen, damit es hart wird?
(7 Abs. fachinhaltlicher Diskurs)
O45: //Ei bei Vollmond// länger kochen, bis es hart ist? (..) Nein, das ist keine
(lachend) naturwissenschaftliche Frage.
O44: Ja, ich //(unv.), weil/ Was soll man/ //
O43: //Doch, du könntest es natürlich wieder/ //
O44: Was soll man denn weil dahinter schreiben? //Also/ //
O45: //Ja. // (.) Weil man es nicht nachweisen kann. (..) Oder?

F Ausführlichere Dokumentationen (zu Vorstellungen)

- O44: Ohne Scheiß jetzt/
O45: Weil es einfach nicht so //IST.//
O43: //Doch.// Du kannst es machen. Du kannst ja dann/ Kannst eins bei //Vollmond kochen oder eins bei äh/ //
O44: //Ja, theoretisch (unv.) Theoretisch kann// man es ja machen, aber //es ist so nicht/ //
O43: //Äh also wenn halt// abnehmender Mond oder zunehmender Mond ist und dann kannst du ja da rausfinden, dass das nicht/ Also ist es eine/ theoretisch gesehen eine naturwissenschaftliche.
O45: //Aber/ //
O44: //Also im// Endeffekt kann man es/ Ja. //Ja,//
O43: //Ja.//
O44: nachprüfen, //aber es ist/ //
O43: //Ja.// Da weiß ich es nicht. Es ist eine Fr/ naturwissenschaftliche Frage, aber äh es ist/ hat nichts damit zu tun. Also du/ Ja, weil du das wieder mit naturwissenschaftlichen Methoden herausarbeiten kannst.
O45: Ja, nein. Wir machen //einfach/ Ja und nein.//
O44: //Ja.//
O45: (..) Weil/
O44: Man kann beides nachtesten/ Mach hier so eine Klammer für beide und schreibe hin: Weil man es nachtesten kann.
O43: (...) Nachtesten?
O44: Nein, nein, nein, nein, nein. (zeigt auf Blatt, auf das O45 schreibt) Ich meine [eine Klammer] über die zwei Fragen, weil //wir/ //
O45: //Warum// das denn?
O44: Ja, weil da wahrscheinlich wieder dieselbe Antwort //kommen kann.//
O45: //(lachend) Nein.//
O44: Dann nicht.
O45: (.) Weil man es testen kann, aber unwahrscheinlich ist.
O44: (8) Dann machen wir die Klammer und das Kreuz auch. (greift nach ihrem Mäppchen)
O45: Nein, nein. N/ Das ist ja für beide. Ist doch richtig.
O43: Nein, das/ Es //ist/ //
O44: //Ja, aber// das macht ja keinen Sinn, das ist KEINE naturwissenschaftliche //Fragestellung, weil man es testen kann. Aber äh wahrscheinlich ist.//
O45: //Ja, theoretisch ist es eine, aber// (.) praktisch ist es sozusagen keine.

Die Äußerungen »Weil man es nicht nachweisen kann« und »Weil es einfach nicht so IST« lassen sich zunächst so deuten, dass O45 ausdrücken möchte, dass die einzuschätzende Beispielfrage nicht naturwissenschaftlich sei, weil es den dort benannten Zusammenhang in der Realität nicht gebe. Daher wurde auch die Ideen-Kategorie +ErgebnisEntsprichtRealität vergeben, die zur Ideen-Oberkategorie Effekt sortiert

wurde. In einem zweiten Schritt kann allerdings auch die Verbindung des Wortes »nachweisen« (Abs. 140) in der Äußerung von O45 mit dem Wort »nachprüfen« (in Abs. 149) von O44 als Ausgangspunkt genommen werden. Die Schülerin O44 nutzt in der analysierten Sequenz weitere Worte wie »nachtesten« oder »testen«, die hier scheinbar synonym gebraucht werden (mit der Ideen-Oberkategorie Prüfen kodiert). Die Schülerin O45 verbindet mit diesen Worten allerdings scheinbar eine andere Vorstellung als O44, was daran deutlich wird, wie die Worte in Abs. 140 und Abs. 149 gebraucht werden. Einmal ist – in Anlehnung an die Sprache der Schülerinnen – Nachprüfbarkeit eines Phänomens dadurch gegeben, dass das Phänomen in der Realität vorliegt und durch Prüfungen gezeigt werden kann; einmal ist Nachprüfbarkeit eines potentiell vorliegenden Phänomens dadurch gegeben, dass es Methoden gibt, durch die geprüft werden kann, ob das Phänomen in der Realität gegeben ist. In beiden Fällen ist keine angemessene Vorstellung zu Kennzeichen von naturwissenschaftlichen Fragen vorhanden, was im ersten Fall klar ist, im zweiten Fall daher kommt, dass es auch Phänomene gibt, die nicht mit naturwissenschaftlichen Methoden (sondern mit nicht-naturwissenschaftlichen Methoden) geprüft werden können. Die Verbindung in der Ähnlichkeit der Worte illustriert einen Weg, auf dem die Vorstellungen von gewissen Personen die Vorstellungen anderer Personen prägen könnten – nicht in einer direkten, linearen Weise, sondern auf Umwegen der Wortassoziation. Mit den verschiedenen Vorstellungen von O44 und O45 scheinen auch verschiedene Einschätzungen der Beispielfrage verbunden zu sein. Zwar können die unangemessenen Vorstellungen des gesamten Teams eventuell als metaphorischer Nährboden für inkorrekte Entscheidungen O45 angesehen werden. Der wesentlichere Faktor für die wiederholt inkorrekte Entscheidung von O45 dürfte aber in der spezifischen Vorstellung auszumachen sein. Dieses Beispiel zeigt also neben der Verwobenheit von Vorstellungen verschiedener Personen auch, dass Vorstellungen durchaus in manchen Fällen – die Erläuterungen auf S. 315 gewissermaßen relativierend – in erklärender Weise für die Korrektheit mancher Einschätzungen hinzugezogen werden können.

Daran, dass sieben von acht inkorrekten Einschätzungen von O45 im Kontext der Aufgabe 1-26-C vorkommen, wird auch eine methodische Einschränkung deutlich, die daher kommt, dass Einschätzungen in verschiedenen Transkriptabsätzen als verschiedene Einschätzungen gezählt werden. Aus dem aufgeführten Transkript und dem zugehörigen Video geht nun ergänzend hervor, dass die Wiederholungen der Einschätzungen von O45 in diesem Fall scheinbar nicht (nur) vorgenommen werden, um die Einschätzung zu unterstreichen, sondern auch (bzw. vorrangig) um die

Aufmerksamkeit der anderen beiden Personen zu erlangen. Diese Unklarheit sorgt dafür, dass es nur mit deutlicher Einschränkung möglich ist, die Anzahl inkorrekt eingeschätzter Fragen als Maß für die Anzahl verschiedener inkorrekt eingeschätzter Fragen oder als Maß für die wiederholt inkorrekte Einschätzung gegebenenfalls inkorrekt eingeschätzter Fragen zu nutzen.

Bei Team A, dem anderen Randpunkt des Trends, haben alle Personen hohe bis sehr hohe Werte der Angemessenheit der Ideen-Oberkategorien sowie der Korrektheit der Einschätzungen. Allerdings fällt die Schülerin A3 dadurch heraus, dass zwar der Werte für die Angemessenheit der Ideen-Oberkategorien deutlich höher, der Wert für die Korrektheit der Einschätzungen jedoch deutlich geringer ist als bei den anderen beiden Gruppenmitgliedern. Im Gegensatz zu den anderen beiden Gruppenmitgliedern äußert A3 allerdings auch nur eine einzige Ideen-Oberkategorie (achtmal *MethodeSpezifisch*) im Zusammenhang mit Einschätzungen (A1 äußert sieben, A2 sechs verschiedene), was als Hinweis darauf gedeutet werden könnte, dass die Anzahl verschiedener angemessener Vorstellungen nicht ohne Bedeutung ist. Ferner kommt für A3 die Ideen-Oberkategorie bei insgesamt nur 5 Teilaufgaben (nur nach der Explizierung) vor, wohingegen die verschiedenen Ideen-Oberkategorien der anderen Personen über 10 bzw. 13 Teilaufgaben (vor und nach der Explizierung) verteilt zugewiesen wurden. Bemerkenswert ist auch, dass A3 deutlich häufiger fachinhaltliche Vorstellungen aktiviert als die anderen Personen aus ihrem Team (u. a. bei 1-26-C, der Aufgabe zum Vollmond und dem Eierkochen).

Die beiden Teams I (oben mittig) und M (unten mittig) scheinen Ausreißer aus dem Trend darzustellen. Für Team M findet sich innerhalb des Teams, dass Personen mit höherem Wert bei der Korrektheit der Einschätzungen auch einen höheren Wert für die Angemessenheit der Ideen-Oberkategorien aufweisen. Während die beiden männlichen Personen jeweils sehr niedrige Werte aufweisen (Korrektheit: 0.18 und 0.17, Angemessenheit: -0.4 und -0.8), fällt tae04w heraus, weil sie jeweils einen sehr deutlich besseren Wert hat (Korrektheit: 0.8, Angemessenheit: 0.76). Im Team I werden nur 24 Einschätzungen getroffen, die allesamt korrekt sind (die anderen Teams treffen alle mehr als 33 Einschätzungen). Die niedrige Angemessenheit im Team I scheint darauf zurückzuführen zu sein, dass für I27 alle drei genutzten Ideen-Oberkategorien (eine doppelt) als unangemessen einzustufen sind, so dass sich für I27 ein Wert der Angemessenheit von -1.33 ergibt (die beiden anderen Personen haben Werte von 0.13 und 0.29). An den beiden Fällen wird deutlich, dass die Heterogenität bei der kleinen Größe der Teams deutlich ins Gewicht fallen kann, insbesondere

wenn nur eine Anzahl von Einschätzungen (oder verschiedenen Ideen-Oberkategorien) auftritt.¹⁴²

F.4.2 Ausführliche Diskussion zum zeitlichen Umfeld von ausgewählten Entscheidungen zur Unterscheidung von Beobachtung und Deutung

In Unterabschnitt 6.7.3 (ab S. 324) ist berichtet, dass sich der Zusammenhang der Korrektheit der Entscheidungen mit der Angemessenheit der Vorstellungen nicht gut untersuchen lässt. Statt Zeit auf statistische Analysen zu verwenden, wurde unter anderem der Kontext der wenigen aufgefundenen inkorrekten Entscheidungen zur Unterscheidung von Beobachtung und Deutung untersucht. Die Ergebnisse sind nicht auf die Forschungsfrage bezogen und daher an dieser Stelle im Anhang dokumentiert.

Eine Betrachtung des Kontextes der inkorrekten Einschätzungen liefert unter anderem das interessante Ergebnis, dass auf inkorrekte Einschätzungen häufig korrekte Einschätzungen anderer Personen folgen, gegen die kein Widerspruch erhoben wird. Daraus muss – insbesondere bei dem schnellen Gesprächsfluss der Gruppenarbeiten – nicht notwendigerweise folgen, dass die erste Person ihre inkorrekte Einschätzung revidiert und sich der Einschätzung der anderen Person anschließt. Auch das Fehlen von Begründungselementen für die Einschätzungen, die sich in kodierten Ideen-Oberkategorien widerspiegeln würden, spricht dafür, dass zumindest in einigen Fällen die inkorrekte Einschätzung einfach im weiteren Verlauf des Gruppengesprächs vernachlässigt wird, obwohl sie aus Forschungsperspektive wichtige Hinweise auf Probleme der Person liefert. Andererseits findet sich in zwei Fällen auch ein ausführlicherer Diskurs, in dem die inkorrekte Einschätzung durch die Person selbst revidiert wird. Im ersten Fall (P46 und P48, Abs. 17–22) stellt sich die Situation so dar:

P48: (..) Ja, ich denke, dass das Vierte dann jetzt auch, weil das ist ja/

P46: Ja/

P48: Ja, oder?

P46: Na obwohl, das kann man auch beobachten, dass sie/ (.., gestikuliert)
verschiedene Oberflächen haben. Also/

P48: Ja, stimmt eigentlich.

P46: Und das ist dann eher keine Begründung. Er sagt ja nicht äh: Die
Oberflächen haben einen Einfluss drauf in welche (?Zeit es unten ankommt?). (..)
[Weiter mit nächstem Beispielsatz.]

¹⁴²Die daran direkt anschließende Vermutung, dass Teams mit größerer Heterogenität auch stärker vom Trend abweichen, bestätigt sich. Eine grafische Betrachtung der Entfernung der Teams von der Trendlinie in Abbildung 6.16 legt nahe, dass Teams mit größerer Heterogenität (operationalisiert als Standardabweichung über die jeweiligen drei Personen) auch stärker vom Trend abweichen.

F Ausführlichere Dokumentationen (zu Vorstellungen)

Die Zustimmung von P46 in Abs. 10 wird von P46 selbst in Abs. 12 und ausführlicher in Abs. 14 revidiert. Auch P48 revidiert die eigene Einschätzung aus Abs. 9 später in Abs. 13. Der Auslöser scheint die Äußerung »Ja, oder?« von P48 zu sein, weshalb es schwer ist, den Verlauf des Gesprächs hin zu einer korrekten Entscheidung einer der Personen zuzuweisen, obwohl nur P46 in diesem Kontext Ideen-Oberkategorien zugewiesen sind. An dem Transkript wird zudem deutlich, dass inkorrekte Einschätzungen einzelner Personen teilweise instantan inkorrekte Einschätzungen anderer Personen nach sich ziehen können (Abs. 9 und dann Abs. 18).

In einem zweiten Fall der Revision einer inkorrekten Einschätzung ist eine ausführlichere Kommentierung des Diskurses notwendig. Das Team H spricht (Abs. 17–34) über den Beispielsatz »Die Feder hat doch eine ganz andere Oberfläche als die Holzkugel« (3-20-4, im Transkript »die Vier«):

H22: Hm (bejahend). (...) Die Vier ja. Die Fünf nicht und die //Sechs.//

H23: //Die// Vier?

H22: Nein, die Vier nicht.

H23: Ach so.

H24: Oder warte.

H22: Doch, die Vier //(unv.)//

H23: //Ich meine// eigentlich kann man ja //(unv.)//

H22: //Das ist doch eine Beobachtung,// das ist/

H23: Aber es ist keine Beobachtung eines Prozesses. //Weiß ich nicht, ob das

H22: //Ja, stimmt eigentlich.//

H24: //Ja.//

H23: dann/

H22: Ja, (?steht das?).

H23: Ja, aber/ Aber/

H24: (zeigt auf Blatt) Dann

H22: Das steht da nicht.

H23: (.) Aber dann ist das hier [zweite Hälfte von 3-20-3: »Stimmt, die Feder war langsamer, *weil sie viel leichter ist als die Holzkugel*«] auch eine Beobachtung, weil man ja feststellt/

H24: (unv.)

H22: (..) Aber das ist schon die Begründung, da ist eine Begründung mit einer Beobachtung zusammengemischt. Was wieder auf //eine Begründung schließt.//

H23: //Da soll eine// weitere Beobachtung als Begründung fungieren.

H22: (...) Ja und dadurch wird die Beobachtung zu einer Begründung und deshalb ist es keine Beobachtung. (..) Es ist so eine Halbbeobachtung. Ist gut. Würde sie nicht als Begründung da stehen, (..) wäre es einfacher. Okay. (...)

Zunächst schätzt H22 den Beispielsatz 3-20-4 als Beobachtung ein (Abs. 17), ändert auf Nachfrage von H23 (Abs. 18) die Einschätzung (Abs. 19), bevor sie – in Reaktion auf das Zögern von H24 – wieder zur ursprünglichen Einschätzung zurückkehrt. Ohne (eigene oder fremde) inhaltliche Argumente wechselt H22 hin und her. Die Wankelmütigkeit der Schülerinnen wird auch im weiteren Verlauf deutlich. Leider ist akustisch nicht verständlich, was H23 in Abs. 23 sagt, in Abs. 25 wird dann aber auch ihre Unsicherheit deutlich (»Weiß nicht«).

Das Argument, dass im Beispielsatz keine Beobachtung eines Prozesses beschrieben wird (sondern nur die Beschaffenheit der Oberflächen), scheint H22 nur kurzfristig zu überzeugen (Abs. 26), weil sie kurz danach zu prüfen scheint, welche Art von Beobachtungen die Aufgabenstellung fokussiert, und dann darauf verweist, dass es nicht nur Beobachtungen von Prozessen sein müssen. Bemerkenswerterweise reagiert H23 darauf nicht durch eine einfache Zustimmung (die impliziert zu sein scheint, so dass die Einschätzung von 3-20-4 beendet und korrekt vorgenommen ist), sondern durch einen Verweis auf die vorher vorgenommene Einschätzung zu 3-20-3: »Stimmt, die Feder war langsamer, weil sie viel leichter ist als die Holzkugel«. Die zweite Hälfte dieser Beispielaussage müsse dann auch eine Beobachtung sein (vermutlich deshalb, weil es auch hier um statische Merkmale von Gegenständen geht). Diese Äußerung von H23 ist als eine Einschätzung kodiert, die die Erwartung übertrifft, weil die Satzteile getrennt werden (obwohl grundsätzlich nur Einschätzungen zu den kompletten Sätzen erwartet wurden).

In den nächsten Absätzen folgen dann Äußerungen (von H22 und H23), in denen Einschätzungen der einzelnen Bestandteile des Beispielsatzes vorgenommen werden und aus diesen Einschätzungen eventuell eine Einschätzung des gesamten Satzes abgeleitet wird. Kommentiert sieht dies so aus:

H22: (..) Aber das

[die zweite Hälfte von 3-20-3]

ist schon die Begründung,

da ist eine Begründung mit einer Beobachtung zusammengemischt

[in der zweiten Hälfte von 3-20-3 ist eine Beobachtung gleichzeitig eine Begründung]

[alternativ, aber in Kombination mit dem nachstehenden Satz nicht so wahrscheinlich,

insbesondere vor dem Hintergrund von Abs. 34: im kompletten Satz 3-20-3 sind eine

Beobachtung (»die Feder war langsamer«) und eine Begründung (»weil sie viel leichter ist als die Holzkugel«) gemeinsam vorhanden].

Was wieder auf //eine Begründung schließt.//

[Woraus folgt, dass die zweite Hälfte von 3-20-3 als eine Deutung (Begründung als Antonym zu Beobachtung hier vermutlich metonym für Deutung genutzt) eingeschätzt werden sollte.]

[Alternativ, zur obigen Alternative gehörig, aber eventuell unabhängig davon möglich: Woraus

F Ausführlichere Dokumentationen (zu Vorstellungen)

folgt, dass 3-20-3 als kompletter Beispielsatz als eine Deutung (Begründung als Metonym genutzt) eingeschätzt werden sollte.]

H23: //Da soll eine// weitere Beobachtung als Begründung fungieren.

H22: (...) Ja und dadurch wird die Beobachtung zu einer Begründung

[die Beobachtung wird als Begründung eingesetzt]

und deshalb ist es keine Beobachtung.

[Die Beobachtung ist nach wie vor eine Beobachtung, aber sie darf nicht als Beobachtung eingeschätzt werden, weil sie nicht als Beobachtung, sondern als Begründung eingesetzt wird.]

(..) Es ist so eine Halbbeobachtung. Ist gut. Würde sie nicht als Begründung da stehen, (..) wäre es einfacher.

[Dann wäre es klar eine Beobachtung.]

Okay. (...)

Das aufgeführte Transkript gibt Einblick in einen Fall, in dem die Lernenden differenzierter Urteilen als vom Instruktionsmaterial intendiert (wie zumeist leicht an den Lösungskarten des Materials festgemacht werden kann). Weitere Fälle von Einschätzungen, deren Korrektheit feingliedriger ist und daher quasi über der Erwartung liegt, sind einfacher gestrickt und beziehen sich alle auf Trennungen von Sinneinheiten, die an sich nicht so differenziert betrachtet werden sollten (und auch von den meisten Personen nicht so differenziert betrachtet wurden). Ein Großteil bezieht sich auf den Beispielsatz 3-20-3 (der »Stimmt, die Feder war langsamer, weil sie viel langsamer ist als die Holzkugel« lautet), beispielsweise:

A1: //Und die Feder war// langsamer.

A3: Ja, das schon, aber //weil sie viel leicht ist, ist nicht mehr Beobachtung.//

A2: //Aber nicht, weil sie viel leichter (unv.)//

A0=E14: //Nein.//

(Team A, Abs. 26–29)

Oder:

B5: //Feder war langsamer ist ein/ // äh ist eine //Beobachtung, aber das [3-20-3 als Gesamtes] ist eine Vermutung.//

(Team B, Abs. 26)

Das Vorkommen von Einschätzungen, die über der erwarteten Korrektheit liegen, könnte auch darauf beruhen, dass die Lernenden die Aufgabe missverstanden haben. Insbesondere das ausführliche Transkript zu Team H legt nahe, dass das Aufgabenverständnis eine Rolle spielt. Gleichwohl scheinen diese Missverständnisse der Aufgaben nicht so geartet zu sein, dass davon die grundsätzliche Korrektheit der Einschätzungen betroffen ist oder nicht mehr über die Unterscheidung von Beobach-

tung und Deutung nachgedacht wird. Dies spiegelt sich auch in den sehr wenigen fachinhaltlichen Überlegungen im Kontext der Aufgaben zur Unterscheidung von Beobachtung und Deutung wider (insgesamt nur 7, keine Häufungen bei spezifischen Aufgaben über die Teams hinweg).

Aus den Ausführungen zu den Entscheidungen und den zugehörigen Transkripten geht ebenfalls hervor, was schon im Kontext der Darlegung des methodischen Vorgehens angemerkt wurde: Einschätzungen sind keine stabilen Konstrukte, sondern können innerhalb weniger Sekunden geändert werden. Die endgültige Einschätzung einer Person geht nicht zwingend aus der letzten formulierten Einschätzung hervor (z. B. weil sie durch implizite Zustimmung zur Äußerung einer anderen Person die Einschätzung noch ändert) und zudem ist nicht klar, ob die letzte Einschätzung überhaupt die im Denken der Person prominente Einschätzung ist. Die methodische Entscheidung, jede Einschätzung einzeln zu kodieren, hat also den Vorteil, diese Fluktuation abbilden zu können; allerdings wird dieses Potential in der vorliegenden Arbeit nur zum Teil ausgeschöpft, weil zeitliche Verläufe oder Reihenfolgen von Entscheidungen innerhalb einzelner (Teil-)Aufgaben nicht in den Blick genommen werden.

F.4.3 Ergänzende Abbildungen und Tabellen zum Textteil

Tabelle F.9: Korrektheit der Einschätzungen und Angemessenheit der zugehörigen Ideen-Oberkategorien für Kennzeichen von naturwissenschaftlichen Fragen (gehört zu Tabelle 6.18).

Ideen-Oberkategorie		Korrektheit der Entscheidungen			SUMME
Name	Angemessenheit	Korrekt	Inkorrekt	Unklar	
Fachbezug	-1	29	2	1	32
NawiEigenschaften	-1	12	0	1	13
MethodeSpezifisch	2	72	0	2	74
MethodeUnspezifisch	1	20	0	1	21
Prüfen	-1	22	2	0	24
Belegen	-1	4	0	0	4
BeweisenNachweisen	-1	18	0	2	20
Objektiv	1	20	1	0	21
Subjektiv	1	14	0	0	14
Empfindungen	2	7	0	0	7
Persönlich	1	29	0	0	29
AntwortAusschlaggebend	-2	11	0	0	11
Effekt	-2	3	6	3	12
Zusammenhang	1	5	0	0	5
Klang	-2	4	1	0	5
PF	-2	5	2	0	7
SUMME		275	14	10	299

F.4 Entscheidungen der Lernenden und zugehörige Vorstellungen

Tabelle F.10: Korrekte, inkorrekte und unklare Entscheidungen sowie fachinhaltliche Äußerungen pro Teilaufgabe des Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen.

Aufgabe	Einschätzung				Fachinhaltliche Überlegung
	Korrekt	Inkorrekt	Unklar	SUMME	
1-21	4	0	0	4	0
1-21-i	20	0	0	20	0
1-21-ii	24	2	1	27	5
1-21-iii	20	0	0	20	5
1-21-iv	6	13	2	21	3
1-21-v	14	0	0	14	3
1-21-vi	15	3	0	18	0
1-21-vii	16	1	1	18	0
1-22	2	0	1	3	0
1-22-i-i	1	0	0	1	0
1-22-i-ii	2	0	0	2	0
1-22-i-iii	0	0	0	0	0
1-22-i-iv	2	0	2	4	5
1-22-i-v	0	0	0	0	0
1-22-i-vi	0	0	0	0	0
1-22-i-vii	0	0	0	0	0
1-22-ii	3	0	0	3	0
1-23	1	0	0	1	0
1-24	2	0	0	2	0
1-24-i	42	0	0	42	3
1-24-ii	42	0	0	42	0
1-25	2	0	0	2	0
1-25-A	34	2	3	39	3
1-25-B	32	2	0	34	5
1-26	0	0	0	0	0
1-26-C	32	15	1	48	15
1-26-D	35	0	0	35	0
1-27	0	0	0	0	0
1-28	0	0	0	0	0
1-29	0	0	0	0	0
1-30	0	0	0	0	0
1-31	0	0	0	0	0
1-31-1	3	0	0	3	0
1-31-2	24	4	0	28	0
1-31-3	1	0	0	1	0
1-32	0	0	0	0	0
1-33	0	0	0	0	0
2-12	16	0	2	18	0
SUMME	395	42	13	450	47

Tabelle F.11: Angemessenheit der Ideen-Oberkategorien der Personen zu Kennzeichen von naturwissenschaftlichen Fragen (gehört zu Abbildung 6.15).

	Ideen-Oberkategorie	Fachbezug	NawiEigenschaften	MethodeSpezifisch	MethodeUnspezifisch	Prüfen	Belegen	BeweisenNachweisen	Objektiv	Subjektiv	Empfindungen	Persönlich	AntwortAusschlaggebend	Effekt	Zusammenhang	Klang	PF	Sonstiges	Angemessenheits- maß
	Angemessenheit	-1	-1	2	1	-1	-1	-1	1	1	2	1	-2	-2	1	-2	-2	0	
Person																			
A1		1	1	14	0	0	0	0	0	0	3	1	0	1	2	0	0	0	1,43
A2		1	0	12	0	0	0	0	0	0	1	3	0	0	1	2	0	0	1,25
A3		0	0	8	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1,67
B4		0	1	0	1	0	0	1	0	0	0	2	0	0	0	0	0	0	0,20
B5		1	1	0	2	2	2	2	0	0	0	3	0	0	0	0	0	0	-0,23
B6		1	2	0	3	1	0	2	0	0	0	2	0	1	0	0	0	0	-0,25
C7		0	0	2	4	0	0	4	0	0	0	3	0	0	0	2	0	0	0,20
C8		0	2	3	2	0	0	6	0	0	0	3	0	0	0	0	1	0	0,06
C9		1	1	0	4	0	0	4	0	0	0	0	0	0	0	0	0	0	-0,20
F16		0	0	4	0	0	1	0	4	3	2	2	1	0	0	2	0	0	0,74
F17		0	0	0	1	3	1	0	0	1	1	1	0	0	0	0	0	0	0,13
F18		0	0	3	0	1	1	0	7	5	0	2	0	0	1	1	0	0	0,81
G19		4	0	0	0	0	0	0	4	0	0	0	0	0	0	1	0	0	-0,22
G20		1	0	0	0	0	0	0	4	1	1	2	0	0	0	0	0	0	0,89
G21		9	0	0	0	0	0	0	4	0	3	0	0	0	0	1	0	0	-0,06
H22		4	2	8	2	1	0	0	1	0	1	3	0	0	1	1	0	0	0,67
H23		7	4	12	1	0	0	0	1	0	1	2	0	0	2	0	0	0	0,70
H24		2	1	4	0	1	0	0	0	0	2	1	0	1	1	0	0	0	0,62
I25		0	1	0	2	0	0	0	6	1	0	2	4	0	0	0	0	0	0,13
I26		0	2	0	0	3	0	0	3	2	0	4	0	0	0	0	0	0	0,29
I27		0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	-1,33
M37		2	3	9	2	0	1	1	1	2	0	0	0	4	0	0	0	0	0,32
M38		0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	-0,33
M39		1	0	1	0	0	0	1	0	0	0	0	0	2	0	0	0	0	-0,80
N40		1	4	8	0	3	1	0	0	0	0	5	1	2	0	1	7	0	-0,30
N41		1	2	11	0	0	0	0	0	0	0	4	0	0	0	0	5	0	0,57
N42		0	0	2	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0,00
O43		4	0	1	2	4	0	4	0	0	1	3	0	1	0	0	0	0	-0,25
O44		4	2	2	2	11	0	0	0	0	0	1	0	0	0	0	0	0	-0,45
O45		4	0	0	2	2	0	0	0	0	0	1	2	2	0	0	0	0	-0,85
P46		1	0	4	3	0	0	4	0	3	0	2	6	0	0	0	0	0	-0,04
P47		0	0	1	2	1	0	5	0	2	0	1	2	0	0	0	0	0	-0,21
P48		0	0	2	0	0	0	0	0	2	0	0	1	0	1	0	0	0	0,83
SUMME		50	30	111	36	35	8	35	35	22	16	53	17	15	9	12	14	0	

Tabelle F.12: Korrektheit der von den Personen vorgenommenen Einschätzungen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen (gehört zu Abbildung 6.15).

Person	Einschätzungen			SUMME	Korrekt- heitsmaß
	Korrekt	Inkorrekt	Unklar		
A1	16	0	1	17	0,94
A2	16	0	0	16	1,00
A3	7	1	1	9	0,67
G21	20	3	0	23	0,74
G19	10	4	0	14	0,43
G20	8	1	0	9	0,78
B4	4	0	1	5	0,80
B6	11	3	0	14	0,57
B5	11	1	0	12	0,83
N40	15	1	2	18	0,78
N41	23	2	1	26	0,81
N42	13	2	0	15	0,73
C9	6	1	1	8	0,63
C8	9	2	0	11	0,64
C7	15	0	0	15	1,00
O43	13	1	0	14	0,86
O44	16	1	0	17	0,88
O45	12	8	2	22	0,18
H23	20	0	2	22	0,91
H22	13	2	0	15	0,73
H24	8	1	1	10	0,70
M37	20	9	0	29	0,38
M38	4	0	0	4	1,00
M39	3	2	1	6	0,17
P46	17	0	0	17	1,00
P47	9	1	0	10	0,80
P48	6	1	0	7	0,71
F17	7	1	0	8	0,75
F18	17	1	0	18	0,89
F16	12	3	1	16	0,56
I25	10	0	0	10	1,00
I26	12	0	0	12	1,00
I27	4	0	0	4	1,00
SUMME	387	52	14	453	

Tabelle F.13: Anzahl der fachinhaltlichen Äußerungen der Personen sowie Korrektheit der von den Personen vorgenommenen Einschätzungen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen.

Person	#Fachinhaltlich	#Einschätzungen	Korrektheitsmaß	#Transkriptabsätze
A1	4	17	0,94	59
A2	6	16	1,00	50
A3	5	9	0,67	46
B4	0	5	0,80	29
B5	4	12	0,83	46
B6	3	14	0,57	52
C7	5	15	1,00	74
C8	5	11	0,64	69
C9	1	8	0,63	51
F16	1	16	0,56	83
F17	3	8	0,75	37
F18	1	18	0,89	74
G19	0	14	0,43	65
G20	0	9	0,78	42
G21	1	23	0,74	70
H22	2	15	0,73	67
H23	0	22	0,91	71
H24	3	10	0,70	51
I25	0	10	1,00	36
I26	0	12	1,00	39
I27	0	4	1,00	13
M37	2	16	0,18	62
M38	0	19	0,80	59
M39	1	6	0,17	24
N40	0	18	0,78	65
N41	0	26	0,81	65
N42	0	15	0,73	39
O43	3	14	0,86	57
O44	1	17	0,88	67
O45	1	22	0,18	64
P46	0	17	1,00	32
P47	0	10	0,80	39
P48	0	7	0,71	17
Σ/MW	Σ: 52	Σ: 455	~0,74 (SD:0,23)	~51,94 (SD:17,28)

F.5 Diagramme zur zeitlichen Auflösung der Vorstellungen der einzelnen Personen

Tabelle F.14: Anzahl der fachinhaltlichen Äußerungen je Team sowie Korrektheit der vorgenommenen Einschätzungen je Team.

Team	Angemessenheit der Ideen-OK			Korrektheit der Entscheidungen				#Fachinhaltlich
	Angemessen	Unangemessen	Maß	Korrekt	Inkorrekt	Unklar	Maß	
A	45	7	1,40	39	1	2	0,90	12
G	19	16	0,14	35	6	0	0,71	1
B	12	16	-0,18	24	4	1	0,69	7
N	30	30	0,05	51	5	2	0,79	0
C	19	19	0,03	29	3	1	0,79	11
O	14	39	-0,49	39	10	2	0,57	5
H	43	24	0,67	41	3	3	0,81	5
M	16	14	0,27	24	11	1	0,36	3
P	23	20	0,02	32	2	0	0,88	0
F	37	10	0,70	35	5	1	0,73	5
I	20	13	0,06	24	0	0	1,00	0

F.5 Diagramme zur zeitlichen Auflösung der Vorstellungen der einzelnen Personen

Dieser Anhang wurde ausgelagert und findet sich in dem Online-Forschungsdaten-Dokument „Transkripte und Diagramme zur Dissertation ‚Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten‘“ unter der folgenden Adresse: <http://dx.doi.org/10.22029/jlupub-542> auf den Seiten 87–110.

F.6 Stabilität der Vorstellungen

Tabelle F.15: Übersicht über harte und weiche Wechsel und zugehörige Maße für Kennzeichen von naturwissenschaftlichen Fragen.

Person	weiche Wechsel	harte Wechsel	Maß für weiche Wechsel	Maß für harte Wechsel
A1	6	4	0,5000	0,3333
A2	6	3	0,6667	0,3333
A3	1	0	0,2500	0,0000
B4	3	3	1,0000	1,0000
B5	6	5	1,0000	0,8333
B6	5	5	0,8333	0,8333
C7	3	2	0,6000	0,4000
C8	4	2	0,6667	0,3333
C9	4	3	0,8000	0,6000
F16	5	2	1,0000	0,4000
F17	4	4	0,8000	0,8000
F18	8	5	0,8889	0,5556
G19	3	2	0,5000	0,3333
G20	1	1	0,5000	0,5000
G21	4	4	0,5000	0,5000
H22	6	3	0,6000	0,3000
H23	6	3	0,6667	0,3333
H24	5	3	0,7143	0,4286
I25	4	3	0,8000	0,6000
I26	4	4	0,6667	0,6667
I27	1	1	1,0000	1,0000
M37	6	6	0,8571	0,8571
M38	5	3	1,0000	0,6000
M39	1	1	1,0000	1,0000
N40	8	6	0,8889	0,6667
N41	6	3	0,5000	0,2500
N42	1	1	0,5000	0,5000
O43	5	2	0,6250	0,2500
O44	3	1	0,4286	0,1429
O45	6	5	1,0000	0,8333
P46	7	4	1,0000	0,5714
P47	4	3	0,8000	0,6000
P48	3	3	0,6000	0,6000
Mittelwert	4,364	3,030	0,732	0,544
Standardabw.	1,936	1,487	0,206	0,250

Tabelle F.16: Übersicht über die Ketten und freien Ideen-Oberkategorien pro Person für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen.

Person	Ketten	Konsequente	Sprünge	Parallele Ketten	Anzahl freier Oks
A1	3	0	1	1	5
A2	1	0	0	0	8
A3	1	0	0	0	1
B4	0	0	0	0	5
B5	2	0	0	1	7
B6	1	0	0	0	6
C7	3	2	0	0	3
C8	2	1	0	0	7
C9	2	1	0	0	3
F16	3	2	0	0	7
F17	1	0	0	0	5
F18	4	3	0	0	7
G19	2	1	0	0	3
G20	2	0	0	1	1
G21	1	0	0	0	3
H22	3	1	0	0	8
H23	3	2	0	0	7
H24	2	1	0	0	7
I25	4	0	1	2	5
I26	1	0	0	0	6
I27	0	0	0	0	3
M37	1	0	0	0	7
M38	2	1	0	0	8
M39	0	0	0	0	2
N40	2	0	1	0	14
N41	2	1	0	0	5
N42	1	0	0	0	1
O43	4	3	0	0	4
O44	3	1	0	0	3
O45	1	0	0	0	10
P46	3	1	1	0	12
P47	2	1	0	0	6
P48	1	0	0	0	3
Mittelwert	1,909	0,667	0,121	0,152	5,515
Standardabw.	1,111	0,876	0,326	0,435	2,966

Anmerkungen. Konsequente sind Ketten, die direkt aneinander anschließen (also um genau eine Teilaufgabe überlappen oder genau eine Instanz Abstand haben); Sprünge bestehen aus Ketten, die mehrere Teilaufgaben Abstand zueinander haben, bei denen keine Ideen-Kategorie kodiert ist; Parallele Ketten sind Ketten, die genau dieselben Teilaufgaben gemeinsam überdecken (also vermutlich darauf schließen lassen, dass ein gemeinsames Vorstellungskonstrukt hinter beiden Ideen-Oberkategorien stehen). Die Anzahl der freien OKs gibt an, wie viele Ideen-Oberkategorien zu Kodierungen gehören, die sich keiner Kette zuordnen lassen; diese sind also eigenständig und frei von zeitlicher Verbindung.

F Ausführlichere Dokumentationen (zu Vorstellungen)

Tabelle F.17: Übersicht über alle auftretenden Ketten von Ideen-Oberkategorien für Kennzeichen von naturwissenschaftlichen Fragen.

Person	Ideen-	Aufgaben- Instanzen-		Person	Ideen-	Aufgaben- Instanzen-	
	Oberkategorie	Länge	Länge		Oberkategorie	Länge	Länge
A1	Empf	2	3	I25	Objektiv	3	2
A1	Zsh	2	3	I25	AntwAus	3	2
A1	Mspez	11	7	I25	Objektiv	2	2
A2	Mspez	10	7	I25	AntwAus	2	2
A3	Mspez	10	5	I26	Prüfen	3	3
B5	Belegen	1	2	M37	Empf	2	2
B5	Persönlich	1	2	M38	NawiEig	2	2
B6	Persönlich	3	2	M38	Mspez	2	2
C7	Klang	5	2	N40	PF/AF	10	2
C7	BewNachw	3	2	N40	Mspez	2	3
C7	Persönlich	3	2	N41	PF/AF	10	3
C8	Mspez	2	2	N41	Mspez	6	8
C8	BewNachw	6	4	N42	Mspez	4	2
C9	Munsp	2	2	O43	Fachbezug	15	3
C9	BewNachw	5	2	O43	Persönlich	2	2
F16	Klang	7	2	O43	BewNachw	3	3
F16	Objektiv	2	2	O43	Munsp	3	2
F16	Mspez	1	2	O44	Fachbezug	17	4
F17	Prüfen	3	2	O44	NawiEig	3	2
F18	Persönlich	4	2	O44	Prüfen	6	4
F18	Objektiv	8	2	O45	Fachbezug	3	2
F18	Mspez	2	2	P46	AntwAus	3	2
F18	Subjektiv	2	2	P46	Subjektiv	2	2
G19	Fachbezug	16	3	P46	AntwAus	4	2
G19	Objektiv	4	3	P47	BewNachw	3	2
G20	Objektiv	2	2	P47	Subjektiv	2	2
G20	Persönlich	2	2	P48	Mspez	8	3
G21	Fachbezug	16	4	Mittelwert		4,8056	2,8333
H22	Fachbezug	2	2	Standardabw.		3,8066	1,4434
H22	Munsp	2	2	Anzahl der Ketten		36	
H22	Mspez	7	6				
H23	Fachbezug	4	2				
H23	NawiEig	5	2				
H23	Mspez	8	5				
H24	Empf	4	2				
H24	Mspez	6	4				

Tabelle F.18: Übersicht über weiche und harte Wechsel bei der Unterscheidung von Beobachtung und Deutung.

Person	weiche Wechsel	harte Wechsel	Maß für weiche Wechsel	Maß für harte Wechsel
A0 = E14	3	3	1	1
A1	1	1	1	1
A2	0	0	0	0
A3	0	0	0	0
B4	0	0	0	0
B5	3	3	0,75	0,75
B6	0	0	0	0
C7	1	1	1	1
C8	0	0	0	0
C9	0	0	0	0
F16	1	1	1	1
F17	x	x		
F18	1	1	1	1
G19	0	0	0	0
G20	0	0	0	0
G21	0	0		
H22	5	5	0,714285714	0,714285714
H23	0	0		
H24	4	4	1	1
K31	1	1	1	1
K32	1	1	0,25	0,25
K33	2	2	1	1
L34	1	1	1	1
L35	1	1	1	1
L36	0	0	0	0
M37	0	0		
M38	0	0	0	0
M39	0	0	0	0
P46	2	0	0,5	0
P47	x	x		
P48	2	2	1	1
Mittelwert	1,000	0,931	0,508	0,489
Standardabw.	1,313	1,311	0,467	0,477

F.7 Vorstellungen im zeitlichen Verlauf

F.7.1 Instruktionale Information zu Kennzeichen von naturwissenschaftlichen Fragen

Die Tabellen auf den folgenden Seiten weisen in der gedruckten Veröffentlichung zum Teil sehr kleine Schriftgrößen auf. In der digitalen Veröffentlichung kann diesem Umstand durch Vergrößern abgeholfen werden.

Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden (sortiert nach Angemessenheit der Ideen-Oberkategorien sowie nach Team-Zugehörigkeit der Personen).

[illegible]

Tabelle F.20: Punkte der einzelnen Personen jeweils vor bzw. nach der Explizierung zu Kennzeichen von naturwissenschaftlichen Fragen.

Person	Angemessenheit			Expliziert		
	Vor	Nach	Differenz	Vor	Nach	Differenz
A1	-1	2	3	0	1	1
A2	1	1	0	1	2	1
A3	0	1	1	0	1	1
B4	1	0	-1	1	2	1
B5	-1	-1	0	0	2	2
B6	-1	-2	-1	-1	2	3
C7	-1	2	3	1	2	1
C8	-3	2	5	0	2	2
C9	-2	0	2	-1	1	2
F16	1	5	4	3	3	0
F17	0	3	3	1	2	1
F18	3	2	-1	2	2	0
G19	-1	-2	-1	-1	0	1
G20	0	4	4	0	2	2
G21	-1	0	1	-1	1	2
H22	-1	0	1	-1	3	4
H23	1	3	2	0	3	3
H24	-1	4	5	0	3	3
I25	1	0	-1	3	2	-1
I26	3	-1	-4	2	1	-1
I27	-3	-1	2	0	0	0
M37	-1	-1	0	-1	2	3
M38	0	0	0	0	3	3
M39	-1	2	3	-1	1	2
N40	-4	-4	0	0	2	2
N41	-4	1	5	-1	2	3
N42	-2	2	4	0	1	1
O43	-1	1	2	-1	3	4
O44	-2	1	3	-1	2	3
O45	-1	-4	-3	-1	1	2
P46	-2	2	4	0	3	3
P47	-2	2	4	1	2	1
P48	-1	3	4	0	1	1
Mittelwert	-0,787879	0,818182	1,606061	0,121212	1,818182	1,696970
Standardabw.	1,609773	2,081004	2,334710	1,121622	0,833196	1,266942

F.7 Vorstellungen im zeitlichen Verlauf

Tabelle F.21: Punkte der einzelnen Personen jeweils vor, direkt nach, später nach und deutlich später nach der Explizierung zu Kennzeichen von naturwissenschaftlichen Fragen.

Person	Angemessenheit				Expliziert			
	Deutlich				Deutlich			
	Vor	Nach	Später	Später	Vor	Nach	Später	Später
A1	-1	2	2	2	0	1	1	1
A2	1	1	2	3	1	2	1	2
A3	0	2	2	1	0	1	1	1
G21	-1	0	0	0	-1	1	0	0
G19	-1	-2	0	0	-1	0	0	0
G20	0	4	0	0	0	2	0	0
B4	1	0	0	0	1	1	0	1
B6	-1	-2	0	0	-1	1	0	1
B5	-1	0	0	0	0	1	1	1
N40	-4	-4	0	0	0	2	0	0
N41	-4	1	0	-2	-1	2	0	0
N42	-2	2	0	0	0	1	0	0
C9	-2	0	0	0	-1	1	0	1
C8	-3	2	0	0	0	2	0	0
C7	-1	2	0	0	1	2	0	0
O43	-1	1	-1	0	-1	3	0	0
O44	-2	0	0	0	-1	1	1	0
O45	-1	-4	-1	0	-1	1	0	0
H23	1	4	2	-1	0	3	1	0
H22	-1	1	1	0	-1	3	1	0
H24	-1	4	2	0	0	3	1	0
M37	-1	-1	0	0	-1	1	0	1
M38	0	2	-1	0	0	3	0	1
M39	-1	2	0	0	-1	1	0	0
P46	-2	2	-3	0	0	3	0	0
P47	-2	2	-1	0	1	2	0	0
P48	-1	3	0	2	0	1	0	1
F17	0	2	-1	1	1	1	0	1
F18	3	3	0	0	2	2	1	0
F16	0	0	0	0	0	0	0	0
I25	1	0	0	0	3	2	0	0
I26	3	-1	1	0	2	1	1	0
I27	-3	-1	0	0	0	0	0	0
Mittelwert	-0,81818	0,8182	0,121212	0,181818	0,030303	1,545455	0,30303	0,363636
Standardabw.	1,585054	1,991718	1,037409	0,833196	0,999541	0,890724	0,459568	0,54038

Tabelle F.22: Prüfung, ob die absoluten Punktzahlen im zeitlichen Umfeld der Explizierung signifikant von Null verschieden sind (für Kennzeichen von naturwissenschaftlichen Fragen).

Phase	Zuwachs Niedrig			df	t (df)	p	r
	N	MW	SD				
<i>Angemessenheit der Ideen-Oberkategorien</i>							
Vor	33	-0,79	1,635	32	-2,769	0,009	0,440
Nach	33	1,00	2,208	32	2,602	0,014	-0,418
Später	33	0,09	1,071	32	0,487	0,629	-0,086
DeutlichSpäter	33	0,18	0,846	32	1,234	0,226	-0,213
<i>Übereinstimmung der Ideen-Oberkategorien mit instrukt. Information</i>							
Vor	33	0,12	1,139	32	0,611	0,545	-0,107
Nach	33	1,64	0,895	32	10,502	0,000	-0,880
Später	33	0,30	0,467	32	3,730	0,001	-0,550
DeutlichSpäter	33	0,36	0,549	32	3,807	0,001	-0,558
<i>Anmerkungen.</i> Students t-Test gegen Null; N=Stichprobengröße.							

Anmerkungen. Students t-Test gegen Null; N=Stichprobengröße.

F.7 Vorstellungen im zeitlichen Verlauf

Tabelle F.23: Punkte der einzelnen Personen jeweils vor, direkt nach, später nach und deutlich später nach der Explizierung zu Kennzeichen von naturwissenschaftlichen Fragen.

Person	Angemessenheit				Expliziert			
	Vor	Nach	Deutlich		Vor	Nach	Deutlich	
			Später	Später			Später	Später
A1	-0,0139	0,018	0,095	0,095	0,000	0,009	0,048	0,048
A2	0,014	0,009	0,095	0,143	0,014	0,018	0,048	0,095
A3	0,000	0,018	0,095	0,048	0,000	0,009	0,048	0,048
G21	-0,014	0,000	0,000	0,000	-0,014	0,009	0,000	0,000
G19	-0,014	-0,018	0,000	0,000	-0,014	0,000	0,000	0,000
G20	0,000	0,035	0,000	0,000	0,000	0,018	0,000	0,000
B4	0,014	0,000	0,000	0,000	0,014	0,009	0,000	0,048
B6	-0,014	-0,018	0,000	0,000	-0,014	0,009	0,000	0,048
B5	-0,014	0,000	0,000	0,000	0,000	0,009	0,048	0,048
N40	-0,056	-0,035	0,000	0,000	0,000	0,018	0,000	0,000
N41	-0,056	0,009	0,000	-0,095	-0,014	0,018	0,000	0,000
N42	-0,028	0,018	0,000	0,000	0,000	0,009	0,000	0,000
C9	-0,028	0,000	0,000	0,000	-0,014	0,009	0,000	0,048
C8	-0,042	0,018	0,000	0,000	0,000	0,018	0,000	0,000
C7	-0,014	0,018	0,000	0,000	0,014	0,018	0,000	0,000
O43	-0,014	0,009	-0,048	0,000	-0,014	0,026	0,000	0,000
O44	-0,028	0,000	0,000	0,000	-0,014	0,009	0,048	0,000
O45	-0,014	-0,035	-0,048	0,000	-0,014	0,009	0,000	0,000
H23	0,014	0,035	0,095	-0,048	0,000	0,026	0,048	0,000
H22	-0,014	0,009	0,048	0,000	-0,014	0,026	0,048	0,000
H24	-0,014	0,035	0,095	0,000	0,000	0,026	0,048	0,000
M37	-0,014	-0,009	0,000	0,000	-0,014	0,009	0,000	0,048
M38	0,000	0,018	-0,048	0,000	0,000	0,026	0,000	0,048
M39	-0,014	0,018	0,000	0,000	-0,014	0,009	0,000	0,000
P46	-0,028	0,018	-0,143	0,000	0,000	0,026	0,000	0,000
P47	-0,028	0,018	-0,048	0,000	0,014	0,018	0,000	0,000
P48	-0,014	0,026	0,000	0,095	0,000	0,009	0,000	0,048
F17	0,000	0,018	-0,048	0,048	0,014	0,009	0,000	0,048
F18	0,042	0,026	0,000	0,000	0,028	0,018	0,048	0,000
F16	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
I25	0,014	0,000	0,000	0,000	0,042	0,018	0,000	0,000
I26	0,042	-0,009	0,048	0,000	0,028	0,009	0,048	0,000
I27	-0,042	-0,009	0,000	0,000	0,000	0,000	0,000	0,000
Mittelwert	-0,01136	0,0072	0,005772	0,008658	0,000421	0,013557	0,01443	0,017316
Standardabw.	0,022015	0,017471	0,0494	0,039676	0,013883	0,007813	0,021884	0,025732

Tabelle F.24: Prüfung, ob die relativen Punktzahlen im zeitlichen Umfeld der Explizierung signifikant von Null verschieden sind (für Kennzeichen von naturwissenschaftlichen Fragen).

Phase	Zuwachs Niedrig			df	t (df)	p	r
	N	MW	SD				
<i>Angemessenheit der Ideen-Oberkategorien</i>							
Vor	33	-0,01	0,021	32	-2,766	0,009	0,439
Nach	33	0,01	0,019	32	2,599	0,014	-0,418
Später	33	0,00	0,049	32	0,486	0,630	-0,086
DeutlichSpäter	33	0,01	0,040	32	1,235	0,226	-0,213
<i>Übereinstimmung der Ideen-Oberkategorien mit instrukt. Information</i>							
Vor	33	0,00	0,015	32	0,613	0,544	-0,108
Nach	33	0,01	0,008	32	10,528	0,000	-0,881
Später	33	0,01	0,021	32	3,730	0,001	-0,550
DeutlichSpäter	33	0,02	0,026	32	3,807	0,001	-0,558
<i>Anmerkungen.</i> Students t-Test gegen Null; N=Stichprobengröße.							

Anmerkungen. Students t-Test gegen Null; N=Stichprobengröße.

F.7.2 Instruktionale Information zur Unterscheidung von Beobachtung und Deutung

Tabelle F.25: Punkte der einzelnen Personen jeweils vor bzw. nach der Explizierung zur Unterscheidung von Beobachtung und Deutung (absolut und relativ).

Person	Angemessenheit (absolut)		Angemessenheit (relativ)		Expliziert (absolut)		Expliziert (relativ)	
	Vor	Nach	Vor	Nach	Vor	Nach	Vor	Nach
A0 = E14	1	2	0,0233	0,1176	0	2	0,0000	0,1176
A1	0	0	0,0000	0,0000	0	0	0,0000	0,0000
A2	4	0	0,0930	0,0000	3	0	0,0698	0,0000
A3	0	0	0,0000	0,0000	0	0	0,0000	0,0000
B4	0	0	0,0000	0,0000	0	0	0,0000	0,0000
B5	4	4	0,0930	0,2353	0	1	0,0000	0,0588
B6	0	0	0,0000	0,0000	0	0	0,0000	0,0000
C7	0	0	0,0000	0,0000	0	2	0,0000	0,1176
C8	0	0	0,0000	0,0000	0	0	0,0000	0,0000
C9	0	0	0,0000	0,0000	0	0	0,0000	0,0000
F16	6	0	0,1395	0,0000	4	0	0,0930	0,0000
F17	0	0	0,0000	0,0000	0	0	0,0000	0,0000
F18	2	-1	0,0465	-0,0588	1	0	0,0233	0,0000
G19	0	0	0,0000	0,0000	0	0	0,0000	0,0000
G20	0	0	0,0000	0,0000	0	0	0,0000	0,0000
G21	-2	0	-0,0465	0,0000	0	0	0,0000	0,0000
H22	5	4	0,1163	0,2353	2	2	0,0465	0,1176
H23	4	0	0,0930	0,0000	4	0	0,0930	0,0000
H24	3	6	0,0698	0,3529	2	3	0,0465	0,1765
K31	4	0	0,0930	0,0000	0	0	0,0000	0,0000
K32	4	2	0,0930	0,1176	2	2	0,0465	0,1176
K33	3	-2	0,0698	-0,1176	2	0	0,0465	0,0000
L34	4	0	0,0930	0,0000	0	0	0,0000	0,0000
L35	2	2	0,0465	0,1176	0	1	0,0000	0,0588
L36	2	0	0,0465	0,0000	0	0	0,0000	0,0000
M37	2	0	0,0465	0,0000	2	0	0,0465	0,0000
M38	4	3	0,0930	0,1765	2	2	0,0465	0,1176
M39	0	0	0,0000	0,0000	0	0	0,0000	0,0000
P46	3	0	0,0698	0,0000	2	0	0,0465	0,0000
P47	0	0	0,0000	0,0000	0	0	0,0000	0,0000
P48	3	0	0,0698	0,0000	2	0	0,0465	0,0000
Mittelwert	1,871	0,645	0,04351	0,03795	0,903	0,484	0,02101	0,02846
Standardabw.	1,980	1,617	0,04604	0,09514	1,254	0,875	0,02915	0,05148

F.7.3 Kodierungen für die Karten der instruktionalen Information zur Unterscheidung von Beobachtung und Deutung

In Abbildung F.3 (auf S. 590) sind exemplarisch die dem Informationstext zur Unterscheidung von Beobachtung und Deutung zugewiesenen Kodierungen dargestellt.

	27	AI22
Beschreiben	28	Beobachtungen: Bei der Beobachtung geht es ausschließlich darum, zu beschreiben, was passiert und nicht darum, warum oder wieso es passiert.
BegründungErklärung		Beobachtungen müssen daher intersubjektiv überprüfbar sein. Das bedeutet, dass verschiedene Personen, die das gleiche Phänomen beobachten, zu gleichen Aussagen gelangen sollten.
ObjektivSubjektiv		29 Beispiel: Die Aussage „Ja, und die Holzkugel ist zuerst auf dem Boden aufgekomen“ ist eine Beobachtung, denn...
Beschreiben	30	▪ ... sie beschreibt ausschließlich was passiert ist .
ObjektivSubjektiv	31	▪ ... alle Personen , die das Auftreffen der Holzkugel beobachtet haben, werden mit großer Wahrscheinlichkeit in dieser Aussage übereinstimmen .
	32	AI23
Beschreiben	33	Deutungen: Eine Deutung geht immer über die Beschreibung von dem, was passiert ist, hinaus, z. B. durch eine zusätzliche Erklärung oder eine Verallgemeinerung.
BegründungErklärung	34	▪ Erklärung: Macht eine Aussage darüber, warum etwas passiert sein könnte.
AuswertungAntwort	35	▪ Verallgemeinerung: Macht aus den Beobachtungen eine Regel, die auch für andere (meist ähnliche) Fälle gilt.
ObjektivSubjektiv	36	Deutungen sind nicht intersubjektiv überprüfbar , weil verschiedene Personen zu unterschiedlichen Erklärungen und Verallgemeinerungen kommen können.
	37	Beispiel: Der zweite Teil der Aussage „Stimmt, die Feder war langsamer, weil sie viel leichter ist als die Holzkugel.“ ist eine Deutung, denn...
BegründungErklärung	38	▪ ...sie beschreibt nicht nur, was passiert (Feder ist langsamer als Holzkugel), sondern auch, warum es passiert (Masse).
Meinung	39	▪ ... verschiedene Personen , die diesen Versuch beobachten, können unterschiedlicher Meinung darüber sein, warum die Holzkugel schneller fällt als die Feder.
Interpretation	40	Hinweis: Deutungen werden manchmal auch als Interpretationen oder Schlussfolgerungen bezeichnet.
AuswertungAntwort		

Abbildung F.3: Exemplarische Darstellung der Kodierung des instruktionalen Informationstextes zur Unterscheidung von Beobachtung und Deutung.

G Weitere Auszüge aus dem Material der Instruktion

4 Viele Fragestellungen AUFGABE

Mit einem Versuch kann einer Vielzahl von Fragestellungen nachgegangen werden, hier z. B.:

- a) Was passiert, wenn mit zwei Trinkhalmen gleichzeitig und immer stärker in zwei unterschiedlich hoch gefüllte Gläser hineingeblasen wird?
- b) In welchem Glas steigen zuerst Blasen auf, wenn mit zwei Trinkhalmen gleichzeitig und immer stärker in zwei unterschiedlich hoch gefüllte Gläser hineingeblasen wird?

In der Tabelle sehen Sie Merkmale, die man bei Nataljas Versuch beobachten könnte.

a)	b)	Merkmal
		Wie groß die aufsteigenden Luftblasen sind.
		Wann Luftblasen im Glas aufsteigen.
		Wie sich das Gesicht des/der Pustenden während des Versuchs verändert.
		Wie lange der/die Pustende pusten kann.
		Wie warm sich das Wasser anfühlt.
		Welche Farbe das Wasser hat.

Welche Merkmale müsste Natalja beobachten, wenn sie einen Versuch zu Frage a) durchführt, welche bei einem Versuch zu Frage b)? Kreuzen Sie an!

5 Viele Fragestellungen KONTROLLE

Überprüfen Sie Ihre Einschätzungen!

a)	b)	Merkmal
X		Wie groß die aufsteigenden Luftblasen sind.
X	X	Wann Luftblasen im Glas aufsteigen.
X		Wie sich das Gesicht des/der Pustenden während des Versuchs verändert.
X		Wie lange der/die Pustende pusten kann.
X		Wie warm sich das Wasser anfühlt.
X		Welche Farbe das Wasser hat.

Falls Sie an einer Stelle anders gekreuzt haben, versuchen Sie zu klären, warum Sie das Kreuz (nicht) hätten setzen sollen.

Abbildung G.1: Auszug aus der ersten Einheit der explizit-fachmethodischen Instruktionsvariante: Präzise und allgemein formulierte Fragen. Wie auf Seite 441 in Absatz 7.2.1 erwähnt könnten Lernende in diesem Kontext über Variablenkontrolle sprechen (und tun es zum Teil auch). Teil 1 von 3 (Karten 1-04 und 1-05).

6

Präzise und allgemeine Fragen

INFO

Frage a) „Was passiert, wenn ...“ beschreibt *allgemein*, was mit dem Versuch untersucht werden soll. Die Frage ist wenig präzise und daher können viele verschiedene Merkmale beobachtet werden, z. B.:

- *Gesicht des Experimentators*
- *Farbe des Wassers*
- *Größe der Luftblase*
- ...

Frage b) „In welchem Glas steigen zuerst Blasen auf, wenn...“ beschreibt präzise, was mit dem Versuch untersucht werden soll; *das zu beobachtende Merkmal wird deutlich*.

7

Präzise oder allgemein?

AUFGABE

Entscheiden Sie, ob es sich bei den folgenden Fragen eher um präzise Fragen (PF), oder eher um allgemeine Fragen (AF) handelt.

PF	AF	Frage
		Kommen aus Trinkhalmen mit größerem Durchmesser auch größere Luftblasen?
		Was passiert, wenn man zu einem Lämpchen ein zweites Lämpchen in Reihe schaltet?
		Gibt es einen Zusammenhang zwischen der Spannung und der Stromstärke in einem einfachen Stromkreis (Batterie und Lämpchen)?
		Fallen Gegenstände mit großer Masse schneller zu Boden als Gegenstände mit kleiner Masse?
		Warum fühlt sich Holz in der Regel wärmer an als Metall?
		Wie hoch muss ein Spiegel mindestens sein, damit man sich ganz in ihm sehen kann?
		Wie erreicht man, dass sich die Schwingungsdauer eines Fadenpendels vergrößert?

Abbildung G.1: Auszug aus der ersten Einheit der explizit-fachmethodischen Instruktionsvariante: Präzise und allgemein formulierte Fragen. Wie auf Seite 441 in Absatz 7.2.1 erwähnt könnten Lernende in diesem Kontext über Variablenkontrolle sprechen (und tun es zum Teil auch). Teil 1 von 3 (Karten 1-06 und 1-07).

Überprüfen Sie Ihre Einschätzungen!

PF	AF	Frage
X		Kommen aus Trinkhalmen mit größerem Durchmesser auch größere Luftblasen?
	X	Was passiert, wenn man zu einem Lämpchen ein zweites Lämpchen in Reihe schaltet?
X		Gibt es einen Zusammenhang zwischen der Spannung und der Stromstärke in einem einfachen Stromkreis mit Batterie und Lämpchen?
X		Fallen Gegenstände mit großer Masse schneller zu Boden als Gegenstände mit kleiner Masse?
	X	Warum fühlt sich Holz in der Regel wärmer an als Metall?
X		Wie groß muss ein Spiegel mindestens sein, damit man sich ganz in ihm sehen kann?
	X	Wie erreicht man, dass sich die Schwingungsdauer eines Fadenpendels vergrößert?

Diskutieren Sie noch einmal die Fragen, bei denen Ihre Einschätzung von dieser Karte abweicht. *Tipp: Überlegen Sie, was von der Frage genau/nicht genau festgelegt wird.*

Vielleicht haben Sie sich bereits darüber gewundert, dass die Frage:

„Wie erreicht man, dass sich die Schwingungsdauer eines Fadenpendels vergrößert?“

als allgemeine Frage angekreuzt wurde.

Die Frage legt lässt offen, welche Größe in dem Versuch verändert werden soll, während das Merkmal Schwingungsdauer beobachtet wird.

Eine präzise Fragestellung muss möglichst genau festlegen, welches Merkmal verändert und welches Merkmal beobachtet werden soll.



Diskutieren Sie kurz:

- 1) Wieso ist es nicht immer möglich, eine Fragestellung präzise zu formulieren?
- 2) Welche Vorteile hat es, eine Fragestellung so präzise wie möglich zu formulieren?

Abbildung G.1: Auszug aus der ersten Einheit der explizit-fachmethodischen Instruktionsvariante: Präzise und allgemein formulierte Fragen. Wie auf Seite 441 in Absatz 7.2.1 erwähnt könnten Lernende in diesem Kontext über Variablenkontrolle sprechen (und tun es zum Teil auch). Teil 3 von 3 (Karten 1-08 und 1-09).

52 Zum Weiterdenken...

AUFGABE

- 2) Versuchen Sie mit den Begriffen „unabhängige Variable“ und „abhängige Variable“ zu erklären, wie man...
- a) ...eine präzise von einer allgemeinen Frage unterscheiden kann.
 - b) ...eine Frage, die zu einem bestimmten Versuch passt, von einer Frage unterscheiden kann, die nicht zu diesem Versuch passt.



53 Zum Weiterdenken...

KONTROLLE

- 1a) Vermutlich haben Sie Variablen wie Farbe der Küche oder Größe des Kochs bei dem Eier-Versuch nicht als unabhängige Variablen genannt, weil Sie sich sicher sind, dass diese Variablen keinen Einfluss auf die Härte eines gekochten Eies haben.
- 1b) Wenn man mit großer Sicherheit davon ausgehen kann, dass eine Variable keinen Einfluss auf die abhängige Variable hat, muss diese Variable bei der Durchführung nicht kontrolliert werden.
Gleichzeitig ist es aber kein Fehler, eine solche Variable dennoch zu kontrollieren!
- 2a) Eine präzise Frage legt im Gegensatz zu einer eher allgemeinen Frage genau fest, was die abhängige und die unabhängige Variable der Untersuchung sind.
- 2b) Eine passende Fragestellung fragt nach dem Einfluss der unabhängigen Variable des Versuchs auf die abhängige Variable des Versuchs.
Eine unpassende Fragestellung hingegen fragt nach einer unabhängigen und/oder abhängigen Variable, die in dem Versuch nicht geprüft wird.

Abbildung G.2: Auszug aus der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante (Karten 2-52 und 2-53): Verknüpfung von Fragen und Variablenkontrolle (siehe 440 in Absatz 7.2.1 für eine Diskussion).

Abbildungsverzeichnis

2.1	Befunde zu zwei Kategorien für Begründungen, die Schülerinnen und Schüler verschiedener Altersstufen für die Einschätzung von Fragen gegeben haben (Werte entnommen aus Millar et al., 1996).	41
2.2	Vorstellungen zum Unterscheiden von Beobachtungen und Deutungen (Zusammenstellung von Aussagen aus Petermann, 2017, Anhang B, ab S. 201).	43
4.1	Überblick über die zeitliche Struktur der Erhebungen für die zugrundeliegende Studie.	51
4.2	Sequenz von Explorationsaufgaben, Informationstext und Übungsaufgaben zu Kennzeichen von naturwissenschaftlichen Fragen aus der explizit-fachmethodischen Instruktionsvariante (Konzept ausformuliert auf Karte 1-23; graue Hinterlegungen kennzeichnen Anmerkungen).	56
4.3	Exemplarische Struktur von teilweise überlappenden Aufgabenserien.	57
4.4	Typische Kameraperspektiven, aus denen die Schülerinnen und Schüler bei der Bearbeitung der Instruktion gefilmt wurden	65
5.1	Übersicht über die Ebenen und Kategorien des Kodiermanuals.	68
5.2	Skizze zu einem Versuchsaufbau aus der ersten Instruktionseinheit beider Varianten.	69
5.3	Exemplarischer Auszug aus dem Kodiermanual für die Kategorie Fachinhaltlich mit Beschreibung, Indikatoren und Beispielen. Komplettes Manual in Anhang B.1 (ab S. 522).	72
5.4	Histogramm zur Länge der mit Fachmethodisch-Indiziert kodierten Segmente in Einheit 1 der explizit-fachmethodischen Instruktionsvariante.	75
5.5	Visualisierung von zwei zu vergleichenden kodierten Segmenten sowie deren Schnittmenge und deren Vereinigung.	77
5.6	Auszüge aus dem initialen Kodiertraining: (a) Zur Unterscheidung von fachmethodisch-indizierten und fachmethodisch-vermuteten Beiträgen (oben), (b) zur Anwendung der 3-Sekunden-Regel bei hoch-frequentiertem Sprecherwechsel (unten); die genutzten Bezeichnungen weichen leicht von den in der Arbeit genutzten Bezeichnungen ab.	86
5.7	Schematisches Mengen-Diagramm zur Darstellung der Ausgangslage bezüglich der für Personen vorliegenden Daten (Drei Personen eines exemplarischen fiktiven Teams in schwarz eingetragen).	89
5.8	Gruppierung der Video-Teilstichproben und grafischer Vergleich mit den jeweiligen Instruktionsvarianten-Gesamtstichproben.	97

5.8	Gruppierung der Video-Teilstichproben und grafischer Vergleich mit den jeweiligen Instruktionsvarianten-Teilstichproben.	98
5.9	Histogramme für die absoluten Anteile der Kategorie OffTask in Einheit 1 der explizit-fachmethodischen Instruktionsvariante.	107
5.10	Informationen auf der nächsten Seite.	109
5.10	Gesamtbearbeitungsdauer aufgetragen über dem teamweise gemittelten Kompetenzzuwachs.	110
5.11	Informationen auf der nächsten Seite.	111
5.11	Gesamtbearbeitungsdauer des Teams, in dem sich die Person befindet, aufgetragen über dem Kompetenzzuwachs der Personen.	112
5.12	Mittelwerte der <i>absoluten</i> Abschnittsbearbeitungsdauern der explizit-fachmethodischen Instruktionsvariante für Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen.	115
5.13	Mittelwerte der <i>relativen</i> Anteile der Abschnittsbearbeitungsdauern an den jeweiligen Gesamtbearbeitungsdauern der Einheiten der explizit-fachmethodischen Instruktionsvariante für Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen	117
5.14	Mittelwerte der Bearbeitungsdauern zu den einzelnen Karten der drei Einheiten der explizit-fachmethodischen Instruktionsvariante für Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen.	119
5.15	Illustrative Darstellung vom Aktivitätsprofil der Schülerin B6 als Balkendiagramm. (Der Anteil von OffTask ist untypisch im Vergleich zur Gesamtheit der Personen.)	123
5.16	Exemplarisches Dendrogramm, welches den Prozess des Clusterings darstellt.	130
5.17	Mittlere Anteile der betrachteten Aktivitäten von Personen bei der ersten Einheit der explizit-fachmethodischen Instruktionsvariante.	132
5.18	Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 1 der explizit-fachmethodischen Instruktionsvariante (Teil 1 der Abbildung: E1-P-P1-..., d. h. $VD_n \gg 0$).	134
5.18	Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 1 der explizit-fachmethodischen Instruktionsvariante (Teil 2 der Abbildung: E1-P-P2-..., d. h. $VD_n \geq 0$).	135
5.19	Dendrogramm zum Prozess des Clusterings der Personen, die Einheit 1 bearbeitet haben (Schritt 1: Pearson-Korrelation für nächstgelegene Nachbarn).	137
5.20	Dendrogramme zu den Prozessen des Clusterings in Schritt 2 zur ersten Einheit der Instruktion für Personen (Pearson-Korrelation für Verlinkung innerhalb der Cluster).	139
5.21	Mittlere Anteile der betrachteten Aktivitäten für Teams bei der ersten Einheit der explizit-fachmethodischen Instruktionsvariante.	141
5.22	Aktivitätsprofile aller Teams sortiert nach Clustern (Einheit 1).	142

5.23	Dendrogramm zum Prozess des Clusterings der Teams, die Einheit 1 bearbeitet haben (Schritt 1: Pearson-Korrelation für nächstgelegene Nachbarn).	144
5.24	Dendrogramm zum Prozess des Clusterings der Teams, die Einheit 1 bearbeitet haben, außer Ausreißer (Schritt 2: Pearson-Korrelation für Verlinkung innerhalb der Cluster).	145
5.25	Mittlere Anteile der betrachteten Aktivitäten von Personen bei der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante.	146
5.26	Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 2 der explizit-fachmethodischen Instruktionsvariante (Teil 1 der Abbildung).	148
5.26	Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 2 der explizit-fachmethodischen Instruktionsvariante (Teil 2 der Abbildung).	149
5.27	Mittlere Anteile der betrachteten Aktivitäten für Teams bei der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante.	151
5.28	Repräsentanten für die Aktivitätsprofiltypen für Teams bei Einheit 2 der explizit-fachmethodischen Instruktionsvariante.	152
5.29	Mittlere Anteile der betrachteten Aktivitäten von Personen bei der dritten Einheit der explizit-fachmethodischen Instruktionsvariante.	153
5.30	Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 3 der explizit-fachmethodischen Instruktionsvariante (Teil 1 der Abbildung: $VD_n \gg$ alle anderen).	154
5.30	Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Einheit 3 der explizit-fachmethodischen Instruktionsvariante (Teil 2 der Abbildung: $VD_n \geq FMi \gg$ alle anderen).	155
5.31	Mittlere Anteile der betrachteten Aktivitäten für Teams bei der dritten Einheit der explizit-fachmethodischen Instruktionsvariante.	157
5.32	Repräsentanten für die Aktivitätsprofiltypen für Teams bei Einheit 2 der explizit-fachmethodischen Instruktionsvariante.	158
5.33	Mittlere Anteile der betrachteten Aktivitäten von Personen bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante.	164
5.34	Relative zeitliche Anteile fachmethodischer Beiträge (Summe; indizierte; vermutete) aufgelöst nach Zugehörigkeit zu den Teilprozessen des experimentbezogenen Denkens und Arbeitens und nach den Einheiten der explizit-fachmethodischen Instruktionsvariante.	174
5.35	Streudiagramm für den zeitlichen Anteil fachmethodisch-indizierter Beiträge bei Einheit 2 und für Kompetenzzuwächse der Personen.	178
5.36	Streudiagramm für den zeitlichen Anteil fachmethodisch-indizierter Beiträge bei Einheit 1 und für Kompetenzzuwächse der Personen.	179
5.37	Streudiagramm für den zeitlichen Anteil fachmethodisch-vermuteter Beiträge bei Einheit 3 und für Kompetenzzuwächse der Personen.	180

5.38	Mittelwerte der relativen Anteile der fachmethodisch-indizierten (FMi) und -vermuteten (FMv) Beiträge an den jeweiligen Abschnittsbearbeitungsdauern für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante für Personen mit hohen (H) vs. niedrigen (N) Kompetenzzuwächsen.	185
5.39	Fachmethodische Beiträge (vermutet; indiziert) bei Experimentiersituationen aus den drei Einheiten (Einheit 1; Einheit 2; Einheit 3) aufgelöst für Personen mit hohen vs. niedrigen Kompetenzzuwächsen.	193
5.40	Karten 2-33 und 2-34 der explizit-fachmethodischen Instruktionsvariante . .	196
5.41	Mittelwerte der relativen Anteile der fachmethodisch-indizierten (FMi) und -vermuteten (FMv) Beiträge an der Gesamtbearbeitungsdauer von Einheit 2 der explizit-fachmethodischen Instruktionsvariante für Personen mit niedrigen (a) und hohen (b) Kompetenzzuwächsen.	197
5.42	Mittelwerte der relativen Anteile der fachmethodischen Beiträge an der Gesamtbearbeitungsdauer von Einheit 2 der explizit-fachmethodischen Instruktionsvariante für Personen mit hohen (ZuwachsN) vs. niedrigen (ZuwachsN) Kompetenzzuwächsen.	198
5.43	Mittelwerte der relativen Anteile der Bearbeitungsdauer der untersuchten Karten an der Gesamtbearbeitungsdauer von Einheit 2 der explizit-fachmethodischen Instruktionsvariante für Personen mit hohen (ZuwachsN) vs. niedrigen (ZuwachsN) Kompetenzzuwächsen.	200
5.44	Mittelwerte der relativen Anteile der Aktivitäten des verbalen (VDv) und des nonverbalen (VDn) Vorbereitens und Durchführens von Versuchen an der Gesamtbearbeitungsdauer von Einheit 2 der explizit-fachmethodischen Instruktionsvariante für Personen mit niedrigen (a) und hohen (b) Kompetenzzuwächsen.	200
5.45	Mittelwerte der Gesamtbearbeitungsdauern zu beiden Instruktionsvarianten	209
5.46	Mittelwerte der nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen und aller betrachteten verbalen Aktivitäten teamweise und personenweise im Vergleich der Instruktionsvarianten.	216
5.47	Mittelwerte der nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen und aller betrachteten verbalen Aktivitäten teamweise und personenweise im Vergleich der Instruktionsvarianten aufgeteilt nach den jeweils drei (zueinander parallelen) Instruktionseinheiten.	216
5.48	Auf den Anteil der betrachteten verbalen Aktivitäten normierte mittlere Anteile der für die Aktivitätsprofile betrachteten Aktivitäten von Personen bei der Bearbeitung der implizit-fachmethodischen bzw. der explizit-fachmethodischen Instruktionsvariante.	218
5.49	Mittlere Anteile der betrachteten Aktivitäten von Personen bei der Bearbeitung der implizit-fachmethodischen bzw. der explizit-fachmethodischen Instruktionsvariante.	219

5.50	Mittlere Anteile der betrachteten Aktivitäten von Personen bei der Bearbeitung der explizit-fachmethodischen Instruktionsvariante (hellgraue Balken gehören zu niedrigen, dunkelgraue zu hohen Kompetenzzuwächsen). Fehlerbalken geben die 95 %-Konfidenzintervalle der Mittelwerte an.	222
5.51	Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Zusammenfassung der Einheiten der implizit-fachmethodischen Instruktionsvariante (Teil 1 der Abbildung: VDn $\approx$ 20 %).	224
5.51	Repräsentanten für die verschiedenen Aktivitätsprofiltypen für Personen bei Zusammenfassung der Einheiten der implizit-fachmethodischen Instruktionsvariante (Teil 2 der Abbildung: c & d mit VDn $\approx$ 10 %, e mit VDn $\approx$ 5 %). . .	225
5.52	Zeitliche Anteile fachmethodischer Beiträge (Summe; indiziert; vermutete) aufgelöst nach Zugehörigkeit zu den Teilprozessen des experimentbezogenen Denkens und Arbeitens und nach den Einheiten der implizit-fachmethodischen Instruktionsvariante.	230
5.53	Mittelwerte der zeitlichen Anteile fachmethodisch-vermuteter Beiträge an der Gesamtbearbeitungsdauer aller drei Einheiten verglichen für Lernende aus beiden Instruktionsvarianten (27 Personen für explizit-, 18 Personen für implizit-fachmethodische Variante) gruppiert nach den erreichten Kompetenzzuwächsen.	234
5.54	Histogramm zum zeitlichen Anteil der fachmethodischen Beiträge an der Gesamtbearbeitungsdauer aller drei Einheiten aufgelöst nach Instruktionsvarianten (27 Personen für explizit-, 18 Personen für implizit-fachmethodische Variante).	235
5.55	Histogramme zum zeitlichen Anteil der fachmethodisch-indizierten und -vermuteten Beiträge an der Gesamtbearbeitungsdauer aller drei Einheiten aufgelöst nach Instruktionsvarianten.	236
6.1	Übersicht über das Vorgehen bei der qualitativen Inhaltsanalyse.	242
6.2	Transkript mit zwei Markierungen für Ideen-Kategorien.	248
6.3	Vorgehen auf Ebene 2 der qualitativen Inhaltsanalyse zum Modus der Äußerungen. (Sel = Selektionskriterium)	254
6.4	Nummerierung der (Teil-)Aufgaben auf Karte 1-21.	258
6.5	Einschätzungen mithilfe der Lösungen auf Karte 1-22 auf Korrektheit prüfen.	259
6.6	Übersicht über die Segmente der Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen.	270
6.7	Übersicht über die Karten und Teilaufgaben von Segment 1 aus dem Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen.	270
6.8	Übersicht über das eine Segmente des Instruktionsextrakts zur Unterscheidung von Beobachtung und Deutung (analog zu Abbildung 6.6).	272
6.9	Übersicht über die Karten und Teilaufgaben des Instruktionsextrakts zur Unterscheidung von Beobachtung und Deutung.	273

6.10	Karte 2-10 zur Illustration der Vernetzung von Konzepten zum Planen von Untersuchungen.	276
6.11	Relativer Anteil von allgemeinen Formulierungen an allen Instanzen aufgelöst für einzelne Ideen-Oberkategorien zu Kennzeichen von naturwissenschaftlichen Fragen.	296
6.12	Anzahl der allgemeinen Formulierungen bei Instanzen, die 1-22-ii zugehörig sind, aufgelöst für einzelne Ideen-Oberkategorien.	298
6.13	Relativer Anteil von allgemeinen Formulierungen an Instanzen abgesehen von solchen, die 1-22-ii zugehörig sind, aufgelöst für einzelne Ideen-Oberkategorien.	299
6.14	Relativer Anteil von allgemeiner Formulierungen an allen Instanzen aufgelöst für einzelne Ideen-Oberkategorien zur Unterscheidung von Beobachtung und Deutung.	302
6.15	Streudiagramm zur Korrektheit der Einschätzungen und Angemessenheit der Ideen-Oberkategorien für jede einzelne Person ($N = 33$).	321
6.16	Streudiagramm zur Korrektheit der Einschätzungen und Angemessenheit der Ideen-Oberkategorien für jedes Team ($N = 11$).	323
6.17	Anzahl der Personen, denen eine Ideen-Oberkategorie bei einer Teilaufgabe zugewiesen wurde (Kennzeichen von naturwissenschaftlichen Fragen). . . .	327
6.18	Anzahl der Personen, denen eine Ideen-Oberkategorie bei einer Teilaufgabe zugewiesen wurde (Unterscheidung von Beobachtung und Deutung).	327
6.19	Verteilung der Anzahl der Teilaufgaben, denen pro Person mindestens eine Ideen-Oberkategorie zugewiesen ist (a+b) sowie Verteilung der Anzahl der pro Person existierenden Transkriptabsätze für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen (c).	331
6.20	Ideen-Oberkategorien, die dem Schüler N41 bei den (Teil-)Aufgaben zugewiesen wurden (Ketten = hellgraue Linien).	345
6.21	Häufigkeiten der beiden Kettenlängen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen.	347
6.22	Ideen-Oberkategorien, die dem Schüler O43 bei den einzelnen (Teil-)Aufgaben zugewiesen wurden.	348
6.23	Ideen-Oberkategorien, die der Schülerin H22 bei den einzelnen (Teil-)Aufgaben zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden. . .	349
6.24	Ideen-Oberkategorien, die dem Schüler P46 bei den einzelnen (Teil-)Aufgaben zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden. . .	350
6.25	Ideen-Oberkategorien, die dem Schüler K32 bei den einzelnen (Teil-)Aufgaben zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden. . .	351
6.26	Text der Explizierung zu Kennzeichen von naturwissenschaftlichen Fragen. .	361
6.27	Mittlere absolute Punktzahlen für die Angemessenheit der Ideen-Oberkategorien (a) und die Übereinstimmungen der Ideen-Oberkategorien mit der Explizierung (b) vor der, direkt nach der, später nach der und deutlich später nach der Explizierung (für Kennzeichen von naturwissenschaftlichen Fragen).	363

6.28	Mittlere relative Punktzahlen für die Angemessenheit der Ideen-Oberkategorien (a) und die Übereinstimmungen der Ideen-Oberkategorien mit der instruktionalen Information (b) vor der, direkt nach der, später nach der und deutlich später nach der instruktionalen Information (für Kennzeichen von naturwissenschaftlichen Fragen).	366
6.29	Mittlere absolute (a, b) bzw. relative (c, d) Punktzahlen für die Angemessenheit der Ideen-Oberkategorien (a, c) und die Übereinstimmungen der Ideen-Oberkategorien mit der Explizierung (b, d) vor und nach der Explizierung (für die Unterscheidung von Beobachtung und Deutung).	373
6.30	Mittlere absolute (a, b) bzw. relative (c, d) Punktzahlen für die Angemessenheit der Ideen-Oberkategorien (a, c) und die Übereinstimmungen der Ideen-Oberkategorien mit der Explizierung (b, d) vor und nach der Explizierung bei korrigierter Stichprobe (für die Unterscheidung von Beobachtung und Deutung). <i>Teil 1 von 2 (Teil 2 auf S. 376)</i>	375
6.30	Mittlere absolute (a, b) bzw. relative (c, d) Punktzahlen für die Angemessenheit der Ideen-Oberkategorien (a, c) und die Übereinstimmungen der Ideen-Oberkategorien mit der Explizierung (b, d) vor und nach der Explizierung bei korrigierter Stichprobe (für die Unterscheidung von Beobachtung und Deutung). <i>Teil 2 von 2</i>	376
6.31	Text der Explizierung zur Unterscheidung von Beobachtung und Deutung. .	377
6.32	Schematische Übersicht über die Profile gehäuft auftretender Vorstellungen im zeitlichen Verlauf.	380
6.33	Ideen-Oberkategorien, die der Schülerin H24 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden. <i>Profil A</i> .	381
6.34	Ideen-Oberkategorien, die der Schülerin F18 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden. <i>Profil B</i> .	382
6.35	Ideen-Oberkategorien, die dem Schüler P46 bei den einzelnen (Teil-)Aufgaben zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden. <i>Profil C</i> .	383
6.36	Ideen-Oberkategorien, die der Schülerin A2 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden. Werden mindestens vierfach vorkommende Ideen-Oberkategorien als gehäufte Ideen-Oberkategorien gewertet, ist die Person dem Profil A zugehörig (nur <i>MethodeSpezifisch</i> nach der instruktionalen Information); werden dreifach vorkommende Ideen-Oberkategorien hinzugenommen, ist die Person dem Profil E zugehörig (<i>Persönlich</i> vor und nach, <i>MethodeSpezifisch</i> nach).	385
6.37	Ideen-Oberkategorien, die der Schülerin O45 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden. . .	390
6.38	Ideen-Oberkategorien, die der Schülerin H23 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden. . .	391
6.39	Ideen-Oberkategorien, die der Schülerin I25 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden. . .	394

6.40	Ideen-Oberkategorien, die dem Schüler P46 bei den einzelnen (Teil-)Aufgaben zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden. . .	395
6.41	Verteilung der Personen auf der Skala des Kompetenzzuwachses von Prä- zu Posttest für die Teilstichprobe zu Kennzeichen von naturwissenschaftlichen Fragen (oben, 33 Personen) und die Teilstichprobe zur Unterscheidung von Beobachtung und Deutung (unten, 28 Personen). Von l. n. r. in unterschiedlichen Grautönen: Negativer, niedriger, mittlerer und zuletzt hoher Zuwachs. . . .	401
B.1	Gruppierung der Teams nach mittleren Kompetenzzuwächsen der jeweils zugehörigen Personen sowie Einheiten, für die Analysen vorgenommen wurden.	532
B.2	Mittelwerte der Bearbeitungsdauern zu den einzelnen Karten der ersten Einheit der explizit-fachmethodischen Instruktion für Personen mit hohen vs. niedrigen Kompetenzzuwächsen. (Fehlerbalken entsprechen einer Standardabweichung des jeweiligen Mittelwerts; Abschnitte mit Unterschieden mit jeweils deutlichsten Unterschieden sind mit einer Klammer markiert.)	533
B.3	Mittelwerte der <i>relativen</i> Anteile der Abschnittsbearbeitungsdauern an den jeweiligen Gesamtbearbeitungsdauern der Einheiten der explizit-fachmethodischen Instruktionsvariante für Personen mit hohen vs. niedrigen Kompetenzzuwächsen (Fehlerbalken entsprechen einer Standardabweichung des jeweiligen Mittelwerts; Abschnitte mit Unterschieden mit jeweils deutlichsten Unterschieden sind mit einer Klammer markiert).	534
B.4	Dendrogramm zum Prozess des Clusterings der Personen, die Einheit 1 bis zum Ende bearbeitet haben (Pearson-Korrelation für nächstgelegene Nachbarn).	536
B.5	Dendrogramm zum Prozess des Clusterings der Personen, die Einheit 1 bis zum Ende bearbeitet haben und niedrige (oben) bzw. hohen (unten) Anteile der Aktivität VorbereitenDurchführenVerbal aufweisen (Pearson-Korrelation für Verlinkung innerhalb der Cluster).	537
B.6	Dendrogramm zum Prozess des Clusterings der Teams, die Einheit 1 bis zum Ende bearbeitet haben (Pearson-Korrelation für nächstgelegene Nachbarn).	538
B.7	Dendrogramm zum Prozess des Clusterings der Teams, die Einheit 1 bis zum Ende bearbeitet haben (Pearson-Korrelation für Verlinkung innerhalb der Cluster).	539
C.1	Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen (Karten 1-21 und 1-22).	541
C.2	Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen (Karten 1-23 und 1-24).	542
C.3	Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen (Karten 1-25 und 1-26).	543
C.4	Sequenz zu Kennzeichen von naturwissenschaftlichen Fragen (Karten 1-27 und 1-28).	544

D.1	Sequenz zur Unterscheidung von Beobachtung und Deutung (Karten 3-19 und 3-20).	545
D.2	Sequenz zur Unterscheidung von Beobachtung und Deutung (Karten 3-21 und 3-22).	546
D.3	Sequenz zur Unterscheidung von Beobachtung und Deutung (Karten 3-23 und 3-24).	547
D.4	Sequenz zur Unterscheidung von Beobachtung und Deutung (Karte 3-25). .	548
F.1	Zur Trennung der Ideen-Oberkategorien für Kennzeichen von naturwissenschaftlichen Fragen: Dendrogram zur Clusteranalyse für die verschiedenen Personen, aus dem hervorgeht, dass verschiedene Personen keine sehr ähnliche Muster aufweisen.	558
F.2	Zur Trennung der Ideen-Oberkategorien für die Unterscheidung von Beobachtung und Deutung: Dendrogram zur Clusteranalyse für die verschiedenen Personen, aus dem hervorgeht, dass verschiedene Personen keine sehr ähnliche Muster aufweisen.	560
F.3	Exemplarische Darstellung der Kodierung des instruktionalen Informationstextes zur Unterscheidung von Beobachtung und Deutung.	590
G.1	Auszug aus der ersten Einheit der explizit-fachmethodischen Instruktionsvariante: Präzise und allgemein formulierte Fragen. Wie auf Seite 441 in Absatz 7.2.1 erwähnt könnten Lernende in diesem Kontext über Variablenkontrolle sprechen (und tun es zum Teil auch). Teil 1 von 3 (Karten 1-04 und 1-05).	591
G.1	Auszug aus der ersten Einheit der explizit-fachmethodischen Instruktionsvariante: Präzise und allgemein formulierte Fragen. Wie auf Seite 441 in Absatz 7.2.1 erwähnt könnten Lernende in diesem Kontext über Variablenkontrolle sprechen (und tun es zum Teil auch). Teil 1 von 3 (Karten 1-06 und 1-07).	592
G.1	Auszug aus der ersten Einheit der explizit-fachmethodischen Instruktionsvariante: Präzise und allgemein formulierte Fragen. Wie auf Seite 441 in Absatz 7.2.1 erwähnt könnten Lernende in diesem Kontext über Variablenkontrolle sprechen (und tun es zum Teil auch). Teil 3 von 3 (Karten 1-08 und 1-09).	593
G.2	Auszug aus der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante (Karten 2-52 und 2-53): Verknüpfung von Fragen und Variablenkontrolle (siehe 440 in Absatz 7.2.1 für eine Diskussion).	594

Tabellenverzeichnis

2.1	Vorstellungen zum Beobachten und Deuten – Zusammenschau der Literatur entnommen aus Petermann (2017, S. 16–17).	42
4.1	Fähigkeiten zu den drei Teilprozessen des experimentbezogenen Denkens und Arbeitens (Inhalt entnommen aus Vorholzer, 2016, Tabelle 2.2).	53
4.2	Konzepte zu den drei Teilprozessen des experimentbezogenen Denkens und Arbeitens basierend auf Vorholzer (2016, Tab. 2.2) mit Ergänzungen aus Gabi (2018, S. 10–11).	54
4.3	Inhalte und Fokusse der Einheiten der expliziten und der impliziten Variante der Instruktion (basierend auf Vorholzer, 2016, S. 95).	58
4.4	Abschnitte der explizit-fachmethodischen Instruktionsvariante.	60
4.5	Abschnitte der implizit-fachmethodischen Instruktionsvariante.	61
4.6	Überblick über Personenanzahlen nach Instruktionsvariante sowie Geschlecht.	64
5.1	Auf einer ergänzenden Ebene des Kategoriensystems für die implizit-fachmethodische Instruktionsvariante eingesetzte Kategorien zum Inhalt fachmethodischer Beiträge.	73
5.2	Übersicht über die Beurteilerübereinstimmungsberechnung für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante.	85
5.3	Übersicht zu den Video-Teilstichproben: Anwesenheiten und Analysen.	94
5.4	Beispiel für Übersicht über die Ergebnisse eines statistischen Tests zum Vergleich zweier Gruppen hinsichtlich zweier Variablen.	103
5.5	Übersicht über die Vergleiche der Gesamtbearbeitungsdauern der drei Einheiten (Teams mit gemittelt hohen vs. niedrigen Kompetenzzuwächsen).	108
5.6	Übersicht über die Vergleiche der Gesamtbearbeitungsdauern der drei Einheiten (Personen mit hohen vs. niedrigen Kompetenzzuwächsen).	111
5.7	Übersicht über die Vergleiche der <i>absoluten</i> Abschnittsbearbeitungsdauern für alle drei Einheiten (Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen).	115
5.8	Übersicht über die Vergleiche der <i>relativen</i> Anteile der Abschnittsbearbeitungsdauern an der Gesamtbearbeitungsdauer der jeweiligen Einheit (Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen).	116
5.9	Übersicht über Karten der explizit-fachmethodischen Instruktionsvariante, für die Unterschiede mit $p < .2$ und $r > .5$ im Vergleich der Teams mit gemittelt hohen vs. gemittelt niedrigen Kompetenzzuwächsen auftreten.	121

5.10 Mittlerer Anteil <i>aller</i> verbalen Aktivitäten an der Gesamtbearbeitungsdauer bei den einzelnen Einheiten aufgelöst auf Ebene der Personen und auf Ebene der Teams.	124
5.11 Mittlerer Anteil <i>der in den Aktivitätsprofilen betrachteten</i> verbalen Aktivitäten an der Gesamtbearbeitungsdauer bei den einzelnen Einheiten aufgelöst auf Ebene der Personen und auf Ebene der Teams.	125
5.12 Aktivitätsprofile der Personen, die für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante analysiert wurden.	162
5.13 Übersicht über alle Aktivitätsprofiltypen und die Anzahl der Personen (bzw. Teams) mit (gemittelt) hohen und niedrigen Kompetenzzuwächsen.	166
5.14 Aktivitätsprofile über alle Einheiten hinweg gewichtet gemittelt für die 9 für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante untersuchten Personen.	171
5.15 Zeitliche Anteile fachmethodischer Beiträge (absolut) aufgelöst nach Zugehörigkeit zu den Teilprozessen des experimentbezogenen Denkens und Arbeitens und nach den Einheiten der explizit-fachmethodischen Instruktionsvariante.	173
5.16 Students t-Tests für die Gesamtdauer fachmethodischer Beiträge von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante.	176
5.17 Welch-Tests für die Gesamtdauer fachmethodischer Beiträge von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für jede der drei Einheiten der explizit-fachmethodischen Instruktionsvariante.	177
5.18 Übersicht über Ergebnisse zu ausgewählten Einheiten zum Zusammenhang von dem zeitlichen Anteil fachmethodischer Beiträge mit den Kompetenzzuwächsen für die einzelnen Personen, die die explizit-fachmethodische Instruktionsvariante bearbeitet haben.	179
5.19 Students t-Tests für die Gesamtdauer fachmethodischer Beiträge im Team von Personen mit hohen vs. niedrigen Kompetenzzuwächsen in der Gesamtschau über alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante hinweg.	181
5.20 Welch-Tests für die Gesamtdauer fachmethodischer Beiträge im Team von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für alle drei Einheiten der explizit-fachmethodischen Instruktionsvariante.	182
5.21 Welch-Tests für die Dauer fachmethodischer Beiträge (insgesamt als Summe der beiden Subkategorien) von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für jeden Abschnitt der drei Einheiten der explizit-fachmethodischen Instruktionsvariante.	186
5.22 Welch-Tests für die Dauer fachmethodisch- <i>indizierter</i> Beiträge von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für jeden Abschnitt der drei Einheiten der explizit-fachmethodischen Instruktionsvariante.	186

5.23	Welch-Tests für die Dauer fachmethodisch- <i>vermuteter</i> Beiträge von Personen mit hohen vs. niedrigen Kompetenzzuwächsen für jeden Abschnitt der drei Einheiten der explizit-fachmethodischen Instruktionsvariante.	187
5.24	Übersicht über die Karten der explizit-fachmethodischen Instruktionsvariante, für die sich signifikante Unterschiede mit großer Effektstärke beim Vergleich von Personen mit unterschiedlichen Kompetenzzuwächsen hinsichtlich einer der beiden Kategorien zu fachmethodischen Beiträgen ergeben.	189
5.25	Personenweise gepaarte t-Tests für die zeitlichen Anteile (der verschiedenen Kategorien) fachmethodischer Beiträge bei den Karten 33 und 34 der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante einzeln für die Gruppen von Personen mit niedrigen Kompetenzzuwächsen.	198
5.26	Students t-Tests zum Vergleich der Personen mit niedrigen und hohen Kompetenzzuwächsen bezüglich der Differenzen der zeitlichen Anteile der jeweiligen fachmethodischen Beiträge für die Karten 33 und 34 der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante.	198
5.27	Students t-Tests zum Vergleich von Lernenden mit niedrigen und hohen Kompetenzzuwächsen bezüglich der verbalen und nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen jeweils bei Karte 33 und Karte 34 der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante sowie bezüglich der Bearbeitungsdauern der Karten.	201
5.28	Personenweise gepaarte t-Tests für die zeitlichen Anteile verbaler und nonverbaler Aktivitäten des Vorbereitens und Durchführens von Versuchen (VDv/VDn) bei den Karten 33 und 34 der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante sowie die Bearbeitungsdauern der Karten (BearbD) einzeln betrachtet für die Gruppen von Personen mit niedrigen Kompetenzzuwächsen (ZuwachsN/ZuwachsH).	201
5.29	Students t-Tests zum Vergleich der Personen mit niedrigen und hohen Kompetenzzuwächsen bezüglich der Differenzen der zeitlichen Anteile der verbalen und nonverbalen Aktivitäten des Vorbereitens und Durchführens von Versuchen (VDv/VDn) für die Karten 33 und 34 der zweiten Einheit der explizit-fachmethodischen Instruktionsvariante sowie für deren Bearbeitungsdauern (BearbD).	201
5.30	Ergebnisse des Ratings dazu, ob die Lernenden bei Karte 2-34 den auf Karte 2-33 durchzuführenden Versuch (erneut) durchgeführt haben sowie auf der Videodurchsicht basierende Anmerkungen zu den Bearbeitungen der beiden Karten.	202
5.31	Students t-Tests zu den Gesamtbearbeitungsdauern der Einheiten für beide Instruktionsvarianten im Vergleich von Teams, die alle Einheiten bearbeitet haben.	209

5.32	Students t-Tests zu den Gesamtbearbeitungsdauern der Einheiten für beide Instruktionsvarianten im Vergleich von Teams, die die jeweilige Einheit bearbeitet haben.	209
5.33	Übersicht über die Bearbeitungsdauern zur implizit-fachmethodischen Instruktionsvariante.	211
5.34	Students t-Tests zu den Gesamtbearbeitungsdauern im Vergleich von Teams mit unterschiedlichen Kompetenzzuwächsen, die die jeweilige Einheit der implizit-fachmethodischen Instruktionsvariante bearbeitet haben.	211
5.35	Students t-Tests zu den Gesamtbearbeitungsdauern im Vergleich von Personen mit unterschiedlichen Kompetenzzuwächsen, die die jeweilige Einheit der implizit-fachmethodischen Instruktionsvariante bearbeitet haben.	212
5.36	Mittlerer Anteil der in den Aktivitätsprofilen betrachteten verbalen Aktivitäten an der Gesamtbearbeitungsdauer bei den einzelnen Einheiten aufgelöst auf Ebene der Personen und auf Ebene der Teams.	214
5.37	Übersicht über alle Aktivitätsprofiltypen und die Anzahl der Personen mit hohen und niedrigen Kompetenzzuwächsen. Die Anzahl der Personen mit mittleren Kompetenzzuwächsen ist nicht in der Tabelle aufgeführt.	227
5.38	Zeitliche Anteile fachmethodischer Beiträge (absolut) aufgelöst nach Zugehörigkeit zu den Teilprozessen des experimentbezogenen Denkens und Arbeitens und nach den Einheiten der explizit-fachmethodischen Instruktionsvariante.	229
5.39	Zusammenfassung der Ergebnisse dreier Kovarianzanalysen zum Zusammenhang des zeitlichen Anteils fachmethodischer Beiträge mit den Instruktionsvarianten unter Kontrolle der von den Lernenden jeweils erreichten Kompetenzzuwächse.	233
5.40	Übersicht über die Inhalte fachmethodischer Beiträge bei der Bearbeitung der implizit-fachmethodischen Instruktionsvariante.	239
6.1	Allgemeines Schema für Kategorien der Formulierung von Äußerungen. . .	251
6.2	Kategorien der Formulierung von Äußerungen für Kennzeichen von naturwissenschaftlichen Fragen.	253
6.3	Kategorien der Formulierung von Äußerungen für die Unterscheidung von Beobachtung und Deutung.	253
6.4	Kategorien zum Modus mit Beschreibungen und Ankerbeispielen – zu Teil 2a aus Abbildung 6.3.	255
6.5	Kategorien zum Modus mit Beschreibungen und Ankerbeispielen – zu Teil 2b aus Abbildung 6.3.	256
6.6	Kategorien zum Modus mit Beschreibungen und Ankerbeispielen – zu Teil 2c aus Abbildung 6.3.	256
6.7	Übersicht über die generierten Ideen-Kategorien für die verschiedenen Personen bei der Überprüfung der Beurteilerübereinstimmung.	263

6.8	Übersicht über die nach Teilen des Kategoriensystems getrennt berechneten κ -Werte der Beurteilerübereinstimmung im Vergleich der beteiligten Personen (κ nach Brennan & Prediger, 1981; Details in Unterabschnitt 5.1.2 ab S. 76).	263
6.9	Zeitungsumfang der jeweils analysierten Sequenzen für jedes Team (in Minuten und Sekunden).	281
6.10	Beschreibungen der Ideen-Oberkategorien zu Kennzeichen von naturwissenschaftlichen Fragen (und Anzahl der Personen (# P) bzw. Teams (# T) mit jeweiliger OK).	284
6.11	Häufigkeiten, mit denen verschiedene Ideen-Oberkategorien zu Kennzeichen von naturwissenschaftlichen Fragen gleichzeitig zugewiesen wurden.	287
6.12	Beschreibungen der Ideen-Oberkategorien zur Unterscheidung von Beobachtung und Deutung	290
6.13	Häufigkeiten, mit denen verschiedene Ideen-Oberkategorien zu Unterscheidung von Beobachtung und Deutung innerhalb desselben Transkriptabsatzes zugewiesen wurden.	291
6.14	Abgleich der Ideen-Oberkategorien mit dokumentierten Vorstellungen	293
6.15	Anzahl der allgemeinen Formulierungen pro Ideen-Oberkategorie für die Unterscheidung von Beobachtung und Deutung.	301
6.16	Einschätzung der Angemessenheit der Ideen-Oberkategorien zu Kennzeichen von naturwissenschaftlichen Fragen.	307
6.17	Korrektheit der Einschätzungen und Angabe, ob Ideen-Oberkategorie bei der Einschätzung kodiert wurde, für Kennzeichen von naturwissenschaftlichen Fragen.	314
6.18	Korrektheit der Einschätzung sowie Angemessenheit der je Instanz zugehörigen Ideen-Oberkategorie für Kennzeichen von naturwissenschaftlichen Fragen.	315
6.19	Korrektheit der Einschätzungen und zugehörige Ideen-Oberkategorien im Kontext der Aufgabe 1-26-C.	318
6.20	Korrektheit der Einschätzungen, die mit den einzelnen Ideen-Oberkategorien einhergehen, für die Unterscheidung von Beobachtung und Deutung.	324
6.21	Anzahl der verschiedenen Ideen-Oberkategorien und gehäuft auftretenden Ideen-Oberkategorien, die einer Person zugewiesen wurden (Kennzeichen von naturwissenschaftlichen Fragen).	334
6.22	Anzahl der verschiedenen Ideen-Oberkategorien und gehäuft auftretenden Ideen-Oberkategorien, die einem Team zugewiesen wurden, sowie Anzahl der Teilaufgaben, bei denen einem Team irgendwelche bzw. gehäuft auftretende Ideen-Oberkategorien zugewiesen wurden (Kennzeichen von naturwissenschaftlichen Fragen).	336
6.23	Anzahl der Ideen-Oberkategorien und gehäuft auftretenden Ideen-Oberkategorien, die einer Person zugewiesen wurden (Unterscheidung von Beobachtung und Deutung).	337

6.24	Anzahl der Personen bzw. Teams, für die eine Ideen-Oberkategorie stabil vorliegt (Kennzeichen von naturwissenschaftlichen Fragen).	339
6.25	Kodierte Formulierungen für Fachbezug bei der Aufgabenkarte 1-21.	340
6.26	Gehäuft auftretende und dreifach auftretende Ideen-Oberkategorien für die Unterscheidung von Beobachtung und Deutung.	343
6.27	Auflösung der Ideen-Oberkategorien, die den einzelnen Personen jeweils vor bzw. nach der instruktionalen Information zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden (sortiert nach Angemessenheit der Ideen-Oberkategorien sowie nach Team-Zugehörigkeit der Personen).	355
6.28	T-Tests zum Vergleich der absoluten Punktzahlen im zeitlichen Umfeld nach der instruktionalen Information mit der Punktzahl vor der instruktionalen Information (oben für Angemessenheit, unten für Übereinstimmung; für Kennzeichen von naturwissenschaftlichen Fragen).	363
6.29	T-Tests zum Vergleich der relativen Punktzahlen im zeitlichen Umfeld nach der instruktionalen Information mit der Punktzahl vor der instruktionalen Information (oben für Angemessenheit, unten für Übereinstimmung; für Kennzeichen von naturwissenschaftlichen Fragen).	366
6.30	Auflösung der Ideen-Oberkategorien, die den einzelnen Personen jeweils vor bzw. nach der Explizierung zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden (sortiert nach Angemessenheit der Ideen-Oberkategorien sowie nach Team-Zugehörigkeit der Personen).	369
6.31	Auflösung der Ideen-Oberkategorien, die den einzelnen Personen jeweils vor bzw. nach der Explizierung zur Unterscheidung von Beobachtung und Deutung zugewiesen wurden (sortiert nach Angemessenheit der Ideen-Oberkategorien sowie nach Team-Zugehörigkeit der Personen), nur für Personen, denen nach der Explizierung mindestens eine Ideen-Oberkategorie zugewiesen wurde.	371
6.32	Zuordnung der Personen zu den Profilen gehäuft auftretender Vorstellungen im zeitlichen Verlauf.	385
6.33	Häufigkeit der Ideen-Oberkategorien als profil-definierende Ideen-Oberkategorien aufgelöst nach Profilen gehäuft auftretender Ideen-Oberkategorien im zeitlichen Verlauf.	387
6.34	Anzahl der Ketten zum Instruktionsextrakt für Kennzeichen von naturwissenschaftlichen Fragen pro Person, aufgelöst danach, ob alle Instanzen der Kette vor der instruktionalen Information vorliegen oder danach oder mindestens eine davor und mindestens eine danach.	393
6.35	Anzahl der Ketten zum Instruktionsextrakt für die Unterscheidung von Beobachtung und Deutung pro Person, aufgelöst danach, ob alle Instanzen der Kette vor der instruktionalen Information vorliegen oder danach oder mindestens eine davor und mindestens eine danach.	396
6.36	Übersicht über die Gruppierung der Personen nach ihren Testwerten für die verschiedenen Testvariablen.	400

6.37	Übersicht über die Prozessvariablen, die für die Analysen zu F-Vor7 betrachtet werden.	402
6.38	Zusammenhänge von Test- und Prozessvariablen jeweils für Vergleiche der Gruppen mit niedrigen vs. hohen Testwerten.	404
6.38	Zusammenhänge von Test- und Prozessvariablen (Fortsetzung: Legende). . .	405
6.39	Auszug zu instruktionalen Informationen (=Explizierungen) aus Tabelle 6.38: <i>Zusammenhänge von Test- und Prozessvariablen jeweils für Vergleiche der Gruppen mit niedrigen vs. hohen Testwerten</i> (Legende siehe dort).	406
6.40	Schematische Darstellung des Zusammenhangs der einzeln betrachteten Prozessmerkmale mit Kompetenzzuwachs und fachmethodischer Eingangsvoraussetzung. Realisiert als fiktiver Auszug aus Tabelle 6.38: <i>Zusammenhänge von Test- und Prozessvariablen jeweils für Vergleiche der Gruppen mit niedrigen vs. hohen Testwerten</i> (Legende siehe dort).	409
7.1	Konzeptualisierungsniveaus basierend auf Rogge (2010, S. 108–109) adaptiert mit Beispielen zu fachmethodischen Beiträgen (Teil 1).	458
7.1	Konzeptualisierungsniveaus basierend auf Rogge (2010, S. 108–109) adaptiert mit Beispielen zu fachmethodischen Beiträgen (Teil 2).	459
7.2	Übersicht über die ausblicksartig mithilfe der Konzeptualisierungsniveaus analysierten Lernenden, insbesondere ihre Kompetenzzuwächse.	460
7.3	Übersicht über die Konzeptualisierungsniveaus, die die ausblicksartig untersuchten Lernenden erreicht haben.	461
7.4	Konzeptualisierungsniveaus der sechs ausblicksartig untersuchten Personen zeitlich aufgelöst ((ca. 10 Minuten Bearbeitungsdauer pro vertikalem grauem durchgezogenem Strich).	484
A.1	Übersicht über bei Vorholzer eingesetzte Instrumente (basierend auf Vorholzer, 2016, S. 90, Tabelle 9.2: »Zentrale Eigenschaften und statistische Kennwerte der eingesetzten Testinstrumente«).	517
A.2	Übersicht über die zu den Aktivitäten untersuchten Personen (Teil 1: explizit-fachmethodische Instruktionsvariante).	518
A.2	Übersicht über die zu den Aktivitäten untersuchten Personen (Teil 2: implizit-fachmethodische Instruktionsvariante).	519
A.3	Übersicht über die Testwerte der zu Vorstellungen zu Kennzeichen von naturwissenschaftlichen Fragen untersuchten Personen.	520
A.4	Übersicht über die Testwerte der zu Vorstellungen zur Unterscheidung von Beobachtung und Deutung untersuchten Personen.	521
B.1	Übersicht über die Vergleiche der Abschnittsbearbeitungsdauern für alle drei Einheiten (Personen mit hohen vs. niedrigen Kompetenzzuwächsen).	533

B.2	Übersicht über die Vergleiche der <i>relativen</i> Anteile der Abschnittsbearbeitungsdauern an der Gesamtbearbeitungsdauer der jeweiligen Einheit (Personen mit hohen vs. niedrigen Kompetenzzuwächsen).	534
B.3	Students t-Tests zu den Gesamtbearbeitungsdauern im Vergleich von Personen mit unterschiedlichen Kompetenzzuwächsen, die die jeweilige Einheit der implizit-fachmethodischen Instruktionsvariante bearbeitet haben.	535
B.4	Students t-Tests zu den Gesamtbearbeitungsdauern im Vergleich von Personen mit unterschiedlichen Kompetenzzuwächsen, die die jeweilige Einheit der implizit-fachmethodischen Instruktionsvariante bearbeitet haben.	535
E.1	Übersicht über die verwendeten Transkriptionsregeln als Orientierungshilfe.	550
F.1	Übersicht über die Ideen-Oberkategorien und zugehörige Ideen-Kategorien für Kennzeichen von naturwissenschaftlichen Fragen.	552
F.1	Übersicht über die Ideen-Oberkategorien und zugehörige Ideen-Kategorien für Kennzeichen von naturwissenschaftlichen Fragen. <i>Teil 2 von 2.</i>	553
F.2	Übersicht über die Ideen-Oberkategorien und zugehörige Ideen-Kategorien für die Unterscheidung von Beobachtung und Deutung.	554
F.3	Ideen-Oberkategorien und zugehörige Kategorien der Formulierung (Kennzeichen von naturwissenschaftlichen Fragen).	555
F.4	Ideen-Oberkategorien und zugehörige Kategorien der Formulierung (Unterscheidung von Beobachtung und Deutung).	556
F.5	Zur Trennung der Ideen-Oberkategorien für Kennzeichen von naturwissenschaftlichen Fragen: Ähnlichkeitsmatrix für die verschiedenen Personen, die zeigt, dass verschiedene Personen keine sehr ähnliche Muster aufweisen.	557
F.6	Zur Trennung der Ideen-Oberkategorien für die Unterscheidung von Beobachtung und Deutung: Ähnlichkeitsmatrix für die verschiedenen Personen, die zeigt, dass verschiedene Personen keine sehr ähnliche Muster aufweisen.	559
F.7	Generalisierungen pro Person für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen.	561
F.8	Generalisierungen pro Person für das Instruktionsextrakt zur Unterscheidung von Beobachtung und Deutung.	562
F.9	Korrektheit der Einschätzungen und Angemessenheit der zugehörigen Ideen-Oberkategorien für Kennzeichen von naturwissenschaftlichen Fragen (gehört zu Tabelle 6.18).	572
F.10	Korrekte, inkorrekte und unklare Entscheidungen sowie fachinhaltliche Äußerungen pro Teilaufgabe des Instruktionsextrakts zu Kennzeichen von naturwissenschaftlichen Fragen.	573
F.11	Angemessenheit der Ideen-Oberkategorien der Personen zu Kennzeichen von naturwissenschaftlichen Fragen (gehört zu Abbildung 6.15).	574

F.12	Korrektheit der von den Personen vorgenommenen Einschätzungen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen (gehört zu Abbildung 6.15).	575
F.13	Anzahl der fachinhaltlichen Äußerungen der Personen sowie Korrektheit der von den Personen vorgenommenen Einschätzungen für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen.	576
F.14	Anzahl der fachinhaltlichen Äußerungen je Team sowie Korrektheit der vorgenommenen Einschätzungen je Team.	577
F.15	Übersicht über harte und weiche Wechsel und zugehörige Maße für Kennzeichen von naturwissenschaftlichen Fragen.	578
F.16	Übersicht über die Ketten und freien Ideen-Oberkategorien pro Person für das Instruktionsextrakt zu Kennzeichen von naturwissenschaftlichen Fragen. . .	579
F.17	Übersicht über alle auftretenden Ketten von Ideen-Oberkategorien für Kennzeichen von naturwissenschaftlichen Fragen.	580
F.18	Übersicht über weiche und harte Wechsel bei der Unterscheidung von Beobachtung und Deutung.	581
F.19	Auflösung der Ideen-Oberkategorien, die den einzelnen Personen jeweils vor bzw. nach der Explizierung zu Kennzeichen von naturwissenschaftlichen Fragen zugewiesen wurden (sortiert nach Angemessenheit der Ideen-Oberkategorien sowie nach Team-Zugehörigkeit der Personen).	583
F.20	Punkte der einzelnen Personen jeweils vor bzw. nach der Explizierung zu Kennzeichen von naturwissenschaftlichen Fragen.	584
F.21	Punkte der einzelnen Personen jeweils vor, direkt nach, später nach und deutlich später nach der Explizierung zu Kennzeichen von naturwissenschaftlichen Fragen.	585
F.22	Prüfung, ob die absoluten Punktzahlen im zeitlichen Umfeld der Explizierung signifikant von Null verschieden sind (für Kennzeichen von naturwissenschaftlichen Fragen).	586
F.23	Punkte der einzelnen Personen jeweils vor, direkt nach, später nach und deutlich später nach der Explizierung zu Kennzeichen von naturwissenschaftlichen Fragen.	587
F.24	Prüfung, ob die relativen Punktzahlen im zeitlichen Umfeld der Explizierung signifikant von Null verschieden sind (für Kennzeichen von naturwissenschaftlichen Fragen).	588
F.25	Punkte der einzelnen Personen jeweils vor bzw. nach der Explizierung zur Unterscheidung von Beobachtung und Deutung (absolut und relativ). . . .	589

Danksagungen

Mein erster Dank gebührt Frau Prof. Dr. Claudia von Aufschnaiter. Du hast Dich nicht nur um alle Organisation bezüglich meiner Arbeitsstelle gekümmert, sondern vor allem das Projekt und die Promotion mit Freude, Zeiteinsatz und Geduld betreut. Du hast mich ermutigt, kritisiert, beraten und ertragen. Für alles, was ich am Institut von und mit Dir lernen durfte – angefangen bei Forschung und Lehre bis hin zum angemessenen Umgang mit Viren aller Art – bin ich Dir von Herzen dankbar.

Ich danke Herrn Prof. Dr. Andreas Vorholzer von Herzen für die intensive Betreuung und die gute Zusammenarbeit. Du hast mich zu Beginn in alles eingeführt und hattest auch später immer ein offenes Ohr für mich. Deine zurückhaltend formulierte Kritik hat mich angespornt. Deine inhaltlichen und persönlichen Anmerkungen haben oft längere Denk- und Lernprozesse angeregt, von denen ich noch lange profitieren werde. Danke.

Ein großes Dankeschön geht auch an Frau Prof. Dr. Claudia Höhne und Herrn Prof. Dr. Alexander Eitel für den umfangreichen Zeiteinsatz in der Prüfungskommission.

Danke sage ich auch den aktuellen und ehemaligen Mitgliedern der Arbeitsgruppe. Ihr habt mich unterstützt, ausgehalten, entlastet und erheitert. Außerdem konnte ich von Eurer Arbeit in Forschung und Lehre sowie Eurer Positionierung zu Mensch und Welt viel lernen. Verena, danke, dass ich Dich jederzeit im Büro ansprechen durfte. Du hast sehr viel geholfen.

Den Hilfskräften und Examenskandidaten danke ich für alle eingesetzten Arbeitsstunden. Ihr hattet zwischenzeitlich das Kodieren von Videos ähnlich satt wie ich, und ich hoffe, dass nicht nur ich im Verlauf die Freude daran zurückgewonnen habe.

Ich danke meinen Eltern für die liebevolle Unterstützung in allen Bereichen des Lebens. Ihr wart immer da, wenn wir Gießener Euch brauchten. Auch Gespräche über das Projekt und das Promovieren hatten ihren Platz, wenn sie nötig waren.

Von Herzen und voller Liebe danke ich meiner Ehefrau! Kiki, Du liebst mich bedingungslos, ermahnst mich schonungslos, ermutigst mich vorbehaltslos und beschenkst mich endlos. Danke, dass Du mich die ganze Zeit ertragen hast und mir geholfen hast, immer wieder Kraft zu tanken und zu rekalisieren.

Mein letzter und umfassender Dank gilt meinem Gott und Herrn, Jesus Christus. Du hast mir all die obengenannten Menschen und weitere geschenkt. Du hast die Umstände bereitet, so dass ich meine Freude am Erforschen Deiner Schöpfung ausleben durfte. Du hast mir Dich selbst und durch den Heiligen Geist auch Gemeinschaft mit dem Vater geschenkt und ich bin dankbar und voller Freude, die Ewigkeit mit Dir zu verbringen und Dein Wesen und Deine Welt zu erforschen.

Bisher erschienene Bände der Reihe „*Studien zum Physik- und Chemielernen*“

ISSN 1614-8967 (vormals *Studien zum Physiklernen* ISSN 1435-5280)

- 1 Helmut Fischler, Jochen Peuckert (Hrsg.): Concept Mapping in fachdidaktischen Forschungsprojekten der Physik und Chemie
ISBN 978-3-89722-256-4 40.50 EUR
- 2 Anja Schoster: Bedeutungsentwicklungsprozesse beim Lösen algorithmischer Physikaufgaben. *Eine Fallstudie zu Lernprozessen von Schülern im Physiknachhilfeunterricht während der Bearbeitung algorithmischer Physikaufgaben*
ISBN 978-3-89722-045-4 40.50 EUR
- 3 Claudia von Aufschnaiter: Bedeutungsentwicklungen, Interaktionen und situatives Erleben beim Bearbeiten physikalischer Aufgaben
ISBN 978-3-89722-143-7 40.50 EUR
- 4 Susanne Haerberlen: Lernprozesse im Unterricht mit Wasserstromkreisen. *Eine Fallstudie in der Sekundarstufe I*
ISBN 978-3-89722-172-7 40.50 EUR
- 5 Kerstin Haller: Über den Zusammenhang von Handlungen und Zielen. *Eine empirische Untersuchung zu Lernprozessen im physikalischen Praktikum*
ISBN 978-3-89722-242-7 40.50 EUR
- 6 Michaela Horstendahl: Motivationale Orientierungen im Physikunterricht
ISBN 978-3-89722-227-4 50.00 EUR
- 7 Stefan Deylitz: Lernergebnisse in der Quanten-Atomphysik. *Evaluation des Bremer Unterrichtskonzepts*
ISBN 978-3-89722-291-5 40.50 EUR
- 8 Lorenz Huckle: Handlungsregulation und Wissenserwerb in traditionellen und computergestützten Experimenten des physikalischen Praktikums
ISBN 978-3-89722-316-5 50.00 EUR
- 9 Heike Theyßen: Ein Physikpraktikum für Studierende der Medizin. *Darstellung der Entwicklung und Evaluation eines adressatenspezifischen Praktikums nach dem Modell der Didaktischen Rekonstruktion*
ISBN 978-3-89722-334-9 40.50 EUR
- 10 Annette Schick: Der Einfluß von Interesse und anderen selbstbezogenen Kognitionen auf Handlungen im Physikunterricht. *Fallstudien zu Interessenhandlungen im Physikunterricht*
ISBN 978-3-89722-380-6 40.50 EUR
- 11 Roland Berger: Moderne bildgebende Verfahren der medizinischen Diagnostik. *Ein Weg zu interessanterem Physikunterricht*
ISBN 978-3-89722-445-2 40.50 EUR

- 12 Johannes Werner: Vom Licht zum Atom. *Ein Unterrichtskonzept zur Quantenphysik unter Nutzung des Zeigermodells*
ISBN 978-3-89722-471-1 40.50 EUR
- 13 Florian Sander: Verbindung von Theorie und Experiment im physikalischen Praktikum. *Eine empirische Untersuchung zum handlungsbezogenen Vorverständnis und dem Einsatz grafikorientierter Modellbildung im Praktikum*
ISBN 978-3-89722-482-7 40.50 EUR
- 14 Jörn Gerdes: Der Begriff der physikalischen Kompetenz. *Zur Validierung eines Konstruktes*
ISBN 978-3-89722-510-7 40.50 EUR
- 15 Malte Meyer-Arndt: Interaktionen im Physikpraktikum zwischen Studierenden und Betreuern. *Feldstudie zu Bedeutungsentwicklungsprozessen im physikalischen Praktikum*
ISBN 978-3-89722-541-1 40.50 EUR
- 16 Dietmar Höttecke: Die Natur der Naturwissenschaften historisch verstehen. *Fachdidaktische und wissenschaftshistorische Untersuchungen*
ISBN 978-3-89722-607-4 40.50 EUR
- 17 Gil Gabriel Mavanga: Entwicklung und Evaluation eines experimentell- und phänomenorientierten Optikcurriculums. *Untersuchung zu Schülervorstellungen in der Sekundarstufe I in Mosambik und Deutschland*
ISBN 978-3-89722-721-7 40.50 EUR
- 18 Meike Ute Zastrow: Interaktive Experimentieranleitungen. *Entwicklung und Evaluation eines Konzeptes zur Vorbereitung auf das Experimentieren mit Messgeräten im Physikalischen Praktikum*
ISBN 978-3-89722-802-3 40.50 EUR
- 19 Gunnar Friege: Wissen und Problemlösen. *Eine empirische Untersuchung des wissenszentrierten Problemlösens im Gebiet der Elektrizitätslehre auf der Grundlage des Experten-Novizen-Vergleichs*
ISBN 978-3-89722-809-2 40.50 EUR
- 20 Erich Starauschek: Physikunterricht nach dem Karlsruher Physikkurs. *Ergebnisse einer Evaluationsstudie*
ISBN 978-3-89722-823-8 40.50 EUR
- 21 Roland Paatz: Charakteristika analogiebasierten Denkens. *Vergleich von Lernprozessen in Basis- und Zielbereich*
ISBN 978-3-89722-944-0 40.50 EUR
- 22 Silke Mikelskis-Seifert: Die Entwicklung von Metakzepten zur Teilchenvorstellung bei Schülern. *Untersuchung eines Unterrichts über Modelle mithilfe eines Systems multipler Repräsentationsebenen*
ISBN 978-3-8325-0013-9 40.50 EUR
- 23 Brunhild Landwehr: Distanzen von Lehrkräften und Studierenden des Sachunterrichts zur Physik. *Eine qualitativ-empirische Studie zu den Ursachen*
ISBN 978-3-8325-0044-3 40.50 EUR

- 24 Lydia Murmann: Physiklernen zu Licht, Schatten und Sehen. *Eine phänomenografische Untersuchung in der Primarstufe*
ISBN 978-3-8325-0060-3 40.50 EUR
- 25 Thorsten Bell: Strukturprinzipien der Selbstregulation. *Komplexe Systeme, Elementarisierungen und Lernprozessstudien für den Unterricht der Sekundarstufe II*
ISBN 978-3-8325-0134-1 40.50 EUR
- 26 Rainer Müller: Quantenphysik in der Schule
ISBN 978-3-8325-0186-0 40.50 EUR
- 27 Jutta Roth: Bedeutungsentwicklungsprozesse von Physikerinnen und Physikern in den Dimensionen Komplexität, Zeit und Inhalt
ISBN 978-3-8325-0183-9 40.50 EUR
- 28 Andreas Saniter: Spezifika der Verhaltensmuster fortgeschrittener Studierender der Physik
ISBN 978-3-8325-0292-8 40.50 EUR
- 29 Thomas Weber: Kumulatives Lernen im Physikunterricht. *Eine vergleichende Untersuchung in Unterrichtsgängen zur geometrischen Optik*
ISBN 978-3-8325-0316-1 40.50 EUR
- 30 Markus Rehm: Über die Chancen und Grenzen moralischer Erziehung im naturwissenschaftlichen Unterricht
ISBN 978-3-8325-0368-0 40.50 EUR
- 31 Marion Budde: Lernwirkungen in der Quanten-Atom-Physik. *Fallstudien über Resonanzen zwischen Lernangeboten und SchülerInnen-Vorstellungen*
ISBN 978-3-8325-0483-0 40.50 EUR
- 32 Thomas Reyer: Oberflächenmerkmale und Tiefenstrukturen im Unterricht. *Exemplarische Analysen im Physikunterricht der gymnasialen Sekundarstufe*
ISBN 978-3-8325-0488-5 40.50 EUR
- 33 Christoph Thomas Müller: Subjektive Theorien und handlungsleitende Kognitionen von Lehrern als Determinanten schulischer Lehr-Lern-Prozesse im Physikunterricht
ISBN 978-3-8325-0543-1 40.50 EUR
- 34 Gabriela Jonas-Ahrend: Physiklehrervorstellungen zum Experiment im Physikunterricht
ISBN 978-3-8325-0576-9 40.50 EUR
- 35 Dimitrios Stavrou: Das Zusammenspiel von Zufall und Gesetzmäßigkeiten in der nicht-linearen Dynamik. *Didaktische Analyse und Lernprozesse*
ISBN 978-3-8325-0609-4 40.50 EUR
- 36 Katrin Engeln: Schülerlabors: authentische, aktivierende Lernumgebungen als Möglichkeit, Interesse an Naturwissenschaften und Technik zu wecken
ISBN 978-3-8325-0689-6 40.50 EUR
- 37 Susann Hartmann: Erklärungsvielfalt
ISBN 978-3-8325-0730-5 40.50 EUR

- 38 Knut Neumann: Didaktische Rekonstruktion eines physikalischen Praktikums für Physiker
ISBN 978-3-8325-0762-6 40.50 EUR
- 39 Michael Späth: Kontextbedingungen für Physikunterricht an der Hauptschule. *Möglichkeiten und Ansatzpunkte für einen fachübergreifenden, handlungsorientierten und berufsorientierten Unterricht*
ISBN 978-3-8325-0827-2 40.50 EUR
- 40 Jörg Hirsch: Interesse, Handlungen und situatives Erleben von Schülerinnen und Schülern beim Bearbeiten physikalischer Aufgaben
ISBN 978-3-8325-0875-3 40.50 EUR
- 41 Monika Hüther: Evaluation einer hypermedialen Lernumgebung zum Thema Gasgesetze. *Eine Studie im Rahmen des Physikpraktikums für Studierende der Medizin*
ISBN 978-3-8325-0911-8 40.50 EUR
- 42 Maike Tesch: Das Experiment im Physikunterricht. *Didaktische Konzepte und Ergebnisse einer Videostudie*
ISBN 978-3-8325-0975-0 40.50 EUR
- 43 Nina Nicolai: Skriptgeleitete Eltern-Kind-Interaktion bei Chemiehausaufgaben. *Eine Evaluationsstudie im Themenbereich Säure-Base*
ISBN 978-3-8325-1013-8 40.50 EUR
- 44 Antje Leisner: Entwicklung von Modellkompetenz im Physikunterricht
ISBN 978-3-8325-1020-6 40.50 EUR
- 45 Stefan Rumann: Evaluation einer Interventionsstudie zur Säure-Base-Thematik
ISBN 978-3-8325-1027-5 40.50 EUR
- 46 Thomas Wilhelm: Konzeption und Evaluation eines Kinematik/Dynamik-Lehrgangs zur Veränderung von Schülervorstellungen mit Hilfe dynamisch ikonischer Repräsentationen und graphischer Modellbildung – mit CD-ROM
ISBN 978-3-8325-1046-6 45.50 EUR
- 47 Andrea Maier-Richter: Computerunterstütztes Lernen mit Lösungsbeispielen in der Chemie. *Eine Evaluationsstudie im Themenbereich Löslichkeit*
ISBN 978-3-8325-1046-6 40.50 EUR
- 48 Jochen Peuckert: Stabilität und Ausprägung kognitiver Strukturen zum Atombegriff
ISBN 978-3-8325-1104-3 40.50 EUR
- 49 Maik Walpuski: Optimierung von experimenteller Kleingruppenarbeit durch Strukturierungshilfen und Feedback
ISBN 978-3-8325-1184-5 40.50 EUR
- 50 Helmut Fischler, Christiane S. Reiners (Hrsg.): Die Teilchenstruktur der Materie im Physik- und Chemieunterricht
ISBN 978-3-8325-1225-5 34.90 EUR
- 51 Claudia Eysel: Interdisziplinäres Lehren und Lernen in der Lehrerbildung. *Eine empirische Studie zum Kompetenzerwerb in einer komplexen Lernumgebung*
ISBN 978-3-8325-1238-5 40.50 EUR

- 52 Johannes Günther: Lehrerfortbildung über die Natur der Naturwissenschaften. *Studien über das Wissenschaftsverständnis von Grundschullehrkräften*
ISBN 978-3-8325-1287-3 40.50 EUR
- 53 Christoph Neugebauer: Lernen mit Simulationen und der Einfluss auf das Problemlösen in der Physik
ISBN 978-3-8325-1300-9 40.50 EUR
- 54 Andreas Schnirch: Gendergerechte Interessen- und Motivationsförderung im Kontext naturwissenschaftlicher Grundbildung. *Konzeption, Entwicklung und Evaluation einer multimedial unterstützten Lernumgebung*
ISBN 978-3-8325-1334-4 40.50 EUR
- 55 Hilde Köster: Freies Explorieren und Experimentieren. *Eine Untersuchung zur selbstbestimmten Gewinnung von Erfahrungen mit physikalischen Phänomenen im Sachunterricht*
ISBN 978-3-8325-1348-1 40.50 EUR
- 56 Eva Heran-Dörr: Entwicklung und Evaluation einer Lehrerfortbildung zur Förderung der physikdidaktischen Kompetenz von Sachunterrichtslehrkräften
ISBN 978-3-8325-1377-1 40.50 EUR
- 57 Agnes Szabone Varnai: Unterstützung des Problemlösens in Physik durch den Einsatz von Simulationen und die Vorgabe eines strukturierten Kooperationsformats
ISBN 978-3-8325-1403-7 40.50 EUR
- 58 Johannes Rethfeld: Aufgabenbasierte Lernprozesse in selbstorganisationsoffenem Unterricht der Sekundarstufe I zum Themengebiet ELEKTROSTATIK. *Eine Feldstudie in vier 10. Klassen zu einer kartenbasierten Lernumgebung mit Aufgaben aus der Elektrostatik*
ISBN 978-3-8325-1416-7 40.50 EUR
- 59 Christian Henke: Experimentell-naturwissenschaftliche Arbeitsweisen in der Oberstufe. *Untersuchung am Beispiel des HIGHSEA-Projekts in Bremerhaven*
ISBN 978-3-8325-1515-7 40.50 EUR
- 60 Lutz Kasper: Diskursiv-narrative Elemente für den Physikunterricht. *Entwicklung und Evaluation einer multimedialen Lernumgebung zum Erdmagnetismus*
ISBN 978-3-8325-1537-9 40.50 EUR
- 61 Thorid Rabe: Textgestaltung und Aufforderung zu Selbsterklärungen beim Physiklernen mit Multimedia
ISBN 978-3-8325-1539-3 40.50 EUR
- 62 Ina Glemnitz: Vertikale Vernetzung im Chemieunterricht. *Ein Vergleich von traditionellem Unterricht mit Unterricht nach Chemie im Kontext*
ISBN 978-3-8325-1628-4 40.50 EUR
- 63 Erik Einhaus: Schülerkompetenzen im Bereich Wärmelehre. *Entwicklung eines Testinstruments zur Überprüfung und Weiterentwicklung eines normativen Modells fachbezogener Kompetenzen*
ISBN 978-3-8325-1630-7 40.50 EUR

- 64 Jasmin Neuroth: Concept Mapping als Lernstrategie. *Eine Interventionsstudie zum Chemielernen aus Texten*
ISBN 978-3-8325-1659-8 40.50 EUR
- 65 Hans Gerd Hegeler-Burkhart: Zur Kommunikation von Hauptschülerinnen und Hauptschülern in einem handlungsorientierten und fächerübergreifenden Unterricht mit physikalischen und technischen Inhalten
ISBN 978-3-8325-1667-3 40.50 EUR
- 66 Karsten Rincke: Sprachentwicklung und Fachlernen im Mechanikunterricht. *Sprache und Kommunikation bei der Einführung in den Kraftbegriff*
ISBN 978-3-8325-1699-4 40.50 EUR
- 67 Nina Strehle: Das Ion im Chemieunterricht. *Alternative Schülervorstellungen und curriculare Konsequenzen*
ISBN 978-3-8325-1710-6 40.50 EUR
- 68 Martin Hopf: Problemorientierte Schülerexperimente
ISBN 978-3-8325-1711-3 40.50 EUR
- 69 Anne Beerenwinkel: Fostering conceptual change in chemistry classes using expository texts
ISBN 978-3-8325-1721-2 40.50 EUR
- 70 Roland Berger: Das Gruppenpuzzle im Physikunterricht der Sekundarstufe II. *Eine empirische Untersuchung auf der Grundlage der Selbstbestimmungstheorie der Motivation*
ISBN 978-3-8325-1732-8 40.50 EUR
- 71 Giuseppe Colicchia: Physikunterricht im Kontext von Medizin und Biologie. *Entwicklung und Erprobung von Unterrichtseinheiten*
ISBN 978-3-8325-1746-5 40.50 EUR
- 72 Sandra Winheller: Geschlechtsspezifische Auswirkungen der Lehrer-Schüler-Interaktion im Chemieanfangsunterricht
ISBN 978-3-8325-1757-1 40.50 EUR
- 73 Isabel Wahser: Training von naturwissenschaftlichen Arbeitsweisen zur Unterstützung experimenteller Kleingruppenarbeit im Fach Chemie
ISBN 978-3-8325-1815-8 40.50 EUR
- 74 Claus Brell: Lernmedien und Lernerfolg - reale und virtuelle Materialien im Physikunterricht. *Empirische Untersuchungen in achten Klassen an Gymnasien (Laborstudie) zum Computereinsatz mit Simulation und IBE*
ISBN 978-3-8325-1829-5 40.50 EUR
- 75 Rainer Wackermann: Überprüfung der Wirksamkeit eines Basismodell-Trainings für Physiklehrer
ISBN 978-3-8325-1882-0 40.50 EUR
- 76 Oliver Tepner: Effektivität von Aufgaben im Chemieunterricht der Sekundarstufe I
ISBN 978-3-8325-1919-3 40.50 EUR

- 77 Claudia Geyer: Museums- und Science-Center-Besuche im naturwissenschaftlichen Unterricht aus einer motivationalen Perspektive. *Die Sicht von Lehrkräften und Schülerinnen und Schülern*
ISBN 978-3-8325-1922-3 40.50 EUR
- 78 Tobias Leonhard: Professionalisierung in der Lehrerbildung. *Eine explorative Studie zur Entwicklung professioneller Kompetenzen in der Lehrererstausbildung*
ISBN 978-3-8325-1924-7 40.50 EUR
- 79 Alexander Kauertz: Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben
ISBN 978-3-8325-1925-4 40.50 EUR
- 80 Regina Hübinger: Schüler auf Weltreise. *Entwicklung und Evaluation von Lehr-/Lernmaterialien zur Förderung experimentell-naturwissenschaftlicher Kompetenzen für die Jahrgangsstufen 5 und 6*
ISBN 978-3-8325-1932-2 40.50 EUR
- 81 Christine Waltner: Physik lernen im Deutschen Museum
ISBN 978-3-8325-1933-9 40.50 EUR
- 82 Torsten Fischer: Handlungsmuster von Physiklehrkräften beim Einsatz neuer Medien. *Fallstudien zur Unterrichtspraxis*
ISBN 978-3-8325-1948-3 42.00 EUR
- 83 Corinna Kieren: Chemiehausaufgaben in der Sekundarstufe I des Gymnasiums. *Fragebogenerhebung zur gegenwärtigen Praxis und Entwicklung eines optimierten Hausaufgabendesigns im Themenbereich Säure-Base*
978-3-8325-1975-9 37.00 EUR
- 84 Marco Thiele: Modelle der Thermohalinen Zirkulation im Unterricht. *Eine empirische Studie zur Förderung des Modellverständnisses*
ISBN 978-3-8325-1982-7 40.50 EUR
- 85 Bernd Zinn: Physik lernen, um Physik zu lehren. *Eine Möglichkeit für interessanteren Physikunterricht*
ISBN 978-3-8325-1995-7 39.50 EUR
- 86 Esther Klaes: Außerschulische Lernorte im naturwissenschaftlichen Unterricht. *Die Perspektive der Lehrkraft*
ISBN 978-3-8325-2006-9 43.00 EUR
- 87 Marita Schmidt: Kompetenzmodellierung und -diagnostik im Themengebiet Energie der Sekundarstufe I. *Entwicklung und Erprobung eines Testinventars*
ISBN 978-3-8325-2024-3 37.00 EUR
- 88 Gudrun Franke-Braun: Aufgaben mit gestuften Lernhilfen. *Ein Aufgabenformat zur Förderung der sachbezogenen Kommunikation und Lernleistung für den naturwissenschaftlichen Unterricht*
ISBN 978-3-8325-2026-7 38.00 EUR
- 89 Silke Klos: Kompetenzförderung im naturwissenschaftlichen Anfangsunterricht. *Der Einfluss eines integrierten Unterrichtskonzepts*
ISBN 978-3-8325-2133-2 37.00 EUR

- 90 Ulrike Elisabeth Burkard: Quantenphysik in der Schule. *Bestandsaufnahme, Perspektiven und Weiterentwicklungsmöglichkeiten durch die Implementation eines Medienservers*
ISBN 978-3-8325-2215-5 43.00 EUR
- 91 Ulrike Gromadecki: Argumente in physikalischen Kontexten. *Welche Geltungsgründe halten Physikanfänger für überzeugend?*
ISBN 978-3-8325-2250-6 41.50 EUR
- 92 Jürgen Bruns: Auf dem Weg zur Förderung naturwissenschaftsspezifischer Vorstellungen von zukünftigen Chemie-Lehrenden
ISBN 978-3-8325-2257-5 43.50 EUR
- 93 Cornelius Marsch: Räumliche Atomvorstellung. *Entwicklung und Erprobung eines Unterrichtskonzeptes mit Hilfe des Computers*
ISBN 978-3-8325-2293-3 82.50 EUR
- 94 Maja Brückmann: Sachstrukturen im Physikunterricht. *Ergebnisse einer Videostudie*
ISBN 978-3-8325-2272-8 39.50 EUR
- 95 Sabine Fechner: Effects of Context-oriented Learning on Student Interest and Achievement in Chemistry Education
ISBN 978-3-8325-2343-5 36.50 EUR
- 96 Clemens Nagel: eLearning im Physikalischen Anfängerpraktikum
ISBN 978-3-8325-2355-8 39.50 EUR
- 97 Josef Riese: Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften
ISBN 978-3-8325-2376-3 39.00 EUR
- 98 Sascha Bernholt: Kompetenzmodellierung in der Chemie. *Theoretische und empirische Reflexion am Beispiel des Modells hierarchischer Komplexität*
ISBN 978-3-8325-2447-0 40.00 EUR
- 99 Holger Christoph Stawitz: Auswirkung unterschiedlicher Aufgabenprofile auf die Schülerleistung. *Vergleich von Naturwissenschafts- und Problemlöseaufgaben der PISA 2003-Studie*
ISBN 978-3-8325-2451-7 37.50 EUR
- 100 Hans Ernst Fischer, Elke Sumfleth (Hrsg.): nwu-essen – 10 Jahre Essener Forschung zum naturwissenschaftlichen Unterricht
ISBN 978-3-8325-3331-1 40.00 EUR
- 101 Hendrik Härtig: Sachstrukturen von Physikschulbüchern als Grundlage zur Bestimmung der Inhaltsvalidität eines Tests
ISBN 978-3-8325-2512-5 34.00 EUR
- 102 Thomas Grüß-Niehaus: Zum Verständnis des Löslichkeitskonzeptes im Chemieunterricht. *Der Effekt von Methoden progressiver und kollaborativer Reflexion*
ISBN 978-3-8325-2537-8 40.50 EUR

- 103 Patrick Bronner: Quantenoptische Experimente als Grundlage eines Curriculums zur Quantenphysik des Photons
ISBN 978-3-8325-2540-8 36.00 EUR
- 104 Adrian Voßkühler: Blickbewegungsmessung an Versuchsaufbauten. *Studien zur Wahrnehmung, Verarbeitung und Usability von physikbezogenen Experimenten am Bildschirm und in der Realität*
ISBN 978-3-8325-2548-4 47.50 EUR
- 105 Verena Tobias: Newton'sche Mechanik im Anfangsunterricht. *Die Wirksamkeit einer Einführung über die zweidimensionale Dynamik auf das Lehren und Lernen*
ISBN 978-3-8325-2558-3 54.00 EUR
- 106 Christian Rogge: Entwicklung physikalischer Konzepte in aufgabenbasierten Lernumgebungen
ISBN 978-3-8325-2574-3 45.00 EUR
- 107 Mathias Ropohl: Modellierung von Schülerkompetenzen im Basiskonzept Chemische Reaktion. *Entwicklung und Analyse von Testaufgaben*
ISBN 978-3-8325-2609-2 36.50 EUR
- 108 Christoph Kulgemeyer: Physikalische Kommunikationskompetenz. *Modellierung und Diagnostik*
ISBN 978-3-8325-2674-0 44.50 EUR
- 109 Jennifer Olszewski: The Impact of Physics Teachers' Pedagogical Content Knowledge on Teacher Actions and Student Outcomes
ISBN 978-3-8325-2680-1 33.50 EUR
- 110 Annika Ohle: Primary School Teachers' Content Knowledge in Physics and its Impact on Teaching and Students' Achievement
ISBN 978-3-8325-2684-9 36.50 EUR
- 111 Susanne Mannel: Assessing scientific inquiry. *Development and evaluation of a test for the low-performing stage*
ISBN 978-3-8325-2761-7 40.00 EUR
- 112 Michael Plomer: Physik physiologisch passend praktiziert. *Eine Studie zur Lernwirksamkeit von traditionellen und adressatenspezifischen Physikpraktika für die Physiologie*
ISBN 978-3-8325-2804-1 34.50 EUR
- 113 Alexandra Schulz: Experimentierspezifische Qualitätsmerkmale im Chemieunterricht. *Eine Videostudie*
ISBN 978-3-8325-2817-1 40.00 EUR
- 114 Franz Boczianowski: Eine empirische Untersuchung zu Vektoren im Physikunterricht der Mittelstufe
ISBN 978-3-8325-2843-0 39.50 EUR
- 115 Maria Ploog: Internetbasiertes Lernen durch Textproduktion im Fach Physik
ISBN 978-3-8325-2853-9 39.50 EUR

- 116 Anja Dhein: Lernen in Explorier- und Experimentiersituationen. *Eine explorative Studie zu Bedeutungsentwicklungsprozessen bei Kindern im Alter zwischen 4 und 6 Jahren*
ISBN 978-3-8325-2859-1 45.50 EUR
- 117 Irene Neumann: Beyond Physics Content Knowledge. *Modeling Competence Regarding Nature of Scientific Inquiry and Nature of Scientific Knowledge*
ISBN 978-3-8325-2880-5 37.00 EUR
- 118 Markus Emden: Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens. *Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I*
ISBN 978-3-8325-2867-6 38.00 EUR
- 119 Birgit Hofmann: Analyse von Blickbewegungen von Schülern beim Lesen von physikbezogenen Texten mit Bildern. *Eye Tracking als Methodenwerkzeug in der physikdidaktischen Forschung*
ISBN 978-3-8325-2925-3 59.00 EUR
- 120 Rebecca Knobloch: Analyse der fachinhaltlichen Qualität von Schüleräußerungen und deren Einfluss auf den Lernerfolg. *Eine Videostudie zu kooperativer Kleingruppenarbeit*
ISBN 978-3-8325-3006-8 36.50 EUR
- 121 Julia Hostenbach: Entwicklung und Prüfung eines Modells zur Beschreibung der Bewertungskompetenz im Chemieunterricht
ISBN 978-3-8325-3013-6 38.00 EUR
- 122 Anna Windt: Naturwissenschaftliches Experimentieren im Elementarbereich. *Evaluation verschiedener Lernsituationen*
ISBN 978-3-8325-3020-4 43.50 EUR
- 123 Eva Kölbach: Kontexteinflüsse beim Lernen mit Lösungsbeispielen
ISBN 978-3-8325-3025-9 38.50 EUR
- 124 Anna Lau: Passung und vertikale Vernetzung im Chemie- und Physikunterricht
ISBN 978-3-8325-3021-1 36.00 EUR
- 125 Jan Lamprecht: Ausbildungswege und Komponenten professioneller Handlungskompetenz. *Vergleich von Quereinsteigern mit Lehramtsabsolventen für Gymnasien im Fach Physik*
ISBN 978-3-8325-3035-8 38.50 EUR
- 126 Ulrike Böhm: Förderung von Verstehensprozessen unter Einsatz von Modellen
ISBN 978-3-8325-3042-6 41.00 EUR
- 127 Sabrina Dollny: Entwicklung und Evaluation eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Chemielehrkräften
ISBN 978-3-8325-3046-4 37.00 EUR
- 128 Monika Zimmermann: Naturwissenschaftliche Bildung im Kindergarten. *Eine integrative Längsschnittstudie zur Kompetenzentwicklung von Erzieherinnen*
ISBN 978-3-8325-3053-2 54.00 EUR

- 129 Ulf Saballus: Über das Schlussfolgern von Schülerinnen und Schülern zu öffentlichen Kontroversen mit naturwissenschaftlichem Hintergrund. *Eine Fallstudie*
ISBN 978-3-8325-3086-0 39.50 EUR
- 130 Olaf Krey: Zur Rolle der Mathematik in der Physik. *Wissenschaftstheoretische Aspekte und Vorstellungen Physiklernender*
ISBN 978-3-8325-3101-0 46.00 EUR
- 131 Angelika Wolf: Zusammenhänge zwischen der Eigenständigkeit im Physikunterricht, der Motivation, den Grundbedürfnissen und dem Lernerfolg von Schülern
ISBN 978-3-8325-3161-4 45.00 EUR
- 132 Johannes Börlin: Das Experiment als Lerngelegenheit. *Vom interkulturellen Vergleich des Physikunterrichts zu Merkmalen seiner Qualität*
ISBN 978-3-8325-3170-6 45.00 EUR
- 133 Olaf Uhden: Mathematisches Denken im Physikunterricht. *Theorieentwicklung und Problemanalyse*
ISBN 978-3-8325-3170-6 45.00 EUR
- 134 Christoph Gut: Modellierung und Messung experimenteller Kompetenz. *Analyse eines large-scale Experimentiertests*
ISBN 978-3-8325-3213-0 40.00 EUR
- 135 Antonio Rueda: Lernen mit ExploMultimedial in kolumbianischen Schulen. *Analyse von kurzzeitigen Lernprozessen und der Motivation beim länderübergreifenden Einsatz einer deutschen computergestützten multimedialen Lernumgebung für den naturwissenschaftlichen Unterricht*
ISBN 978-3-8325-3218-5 45.50 EUR
- 136 Krisztina Berger: Bilder, Animationen und Notizen. *Empirische Untersuchung zur Wirkung einfacher visueller Repräsentationen und Notizen auf den Wissenserwerb in der Optik*
ISBN 978-3-8325-3238-3 41.50 EUR
- 137 Antony Crossley: Untersuchung des Einflusses unterschiedlicher physikalischer Konzepte auf den Wissenserwerb in der Thermodynamik der Sekundarstufe I
ISBN 978-3-8325-3275-8 40.00 EUR
- 138 Tobias Viering: Entwicklung physikalischer Kompetenz in der Sekundarstufe I. *Validierung eines Kompetenzentwicklungsmodells für das Energiekonzept im Bereich Fachwissen*
ISBN 978-3-8325-3277-2 37.00 EUR
- 139 Nico Schreiber: Diagnostik experimenteller Kompetenz. *Validierung technologiegestützter Testverfahren im Rahmen eines Kompetenzstrukturmodells*
ISBN 978-3-8325-3284-0 39.00 EUR
- 140 Sarah Hundertmark: Einblicke in kollaborative Lernprozesse. *Eine Fallstudie zur reflektierenden Zusammenarbeit unterstützt durch die Methoden Concept Mapping und Lernbegleitbogen*
ISBN 978-3-8325-3251-2 43.00 EUR

- 141 Ronny Scherer: Analyse der Struktur, Messinvarianz und Ausprägung komplexer Problemlösekompetenz im Fach Chemie. *Eine Querschnittstudie in der Sekundarstufe I und am Übergang zur Sekundarstufe II*
ISBN 978-3-8325-3312-0 43.00 EUR
- 142 Patricia Heitmann: Bewertungskompetenz im Rahmen naturwissenschaftlicher Problemlöseprozesse. *Modellierung und Diagnose der Kompetenzen Bewertung und analytisches Problemlösen für das Fach Chemie*
ISBN 978-3-8325-3314-4 37.00 EUR
- 143 Jan Fleischhauer: Wissenschaftliches Argumentieren und Entwicklung von Konzepten beim Lernen von Physik
ISBN 978-3-8325-3325-0 35.00 EUR
- 144 Nermin Özcan: Zum Einfluss der Fachsprache auf die Leistung im Fach Chemie. *Eine Förderstudie zur Fachsprache im Chemieunterricht*
ISBN 978-3-8325-3328-1 36.50 EUR
- 145 Helena van Vorst: Kontextmerkmale und ihr Einfluss auf das Schülerinteresse im Fach Chemie
ISBN 978-3-8325-3321-2 38.50 EUR
- 146 Janine Cappell: Fachspezifische Diagnosekompetenz angehender Physiklehrkräfte in der ersten Ausbildungsphase
ISBN 978-3-8325-3356-4 38.50 EUR
- 147 Susanne Bley: Förderung von Transferprozessen im Chemieunterricht
ISBN 978-3-8325-3407-3 40.50 EUR
- 148 Cathrin Blaes: Die übungsgestützte Lehrerpräsentation im Chemieunterricht der Sekundarstufe I. *Evaluation der Effektivität*
ISBN 978-3-8325-3409-7 43.50 EUR
- 149 Julia Suckut: Die Wirksamkeit von piko-OWL als Lehrerfortbildung. Eine Evaluation zum Projekt *Physik im Kontext* in Fallstudien
ISBN 978-3-8325-3440-0 45.00 EUR
- 150 Alexandra Dorschu: Die Wirkung von Kontexten in Physikkompetenztestaufgaben
ISBN 978-3-8325-3446-2 37.00 EUR
- 151 Jochen Scheid: Multiple Repräsentationen, Verständnis physikalischer Experimente und kognitive Aktivierung: *Ein Beitrag zur Entwicklung der Aufgabenkultur*
ISBN 978-3-8325-3449-3 49.00 EUR
- 152 Tim Plasa: Die Wahrnehmung von Schülerlaboren und Schülerforschungszentren
ISBN 978-3-8325-3483-7 35.50 EUR
- 153 Felix Schoppmeier: Physikkompetenz in der gymnasialen Oberstufe. *Entwicklung und Validierung eines Kompetenzstrukturmodells für den Kompetenzbereich Umgang mit Fachwissen*
ISBN 978-3-8325-3502-5 36.00 EUR

- 154 Katharina Groß: Experimente alternativ dokumentieren. *Eine qualitative Studie zur Förderung der Diagnose- und Differenzierungskompetenz in der Chemielehrerbildung*
ISBN 978-3-8325-3508-7 43.50 EUR
- 155 Barbara Hank: Konzeptwandelprozesse im Anfangsunterricht Chemie. *Eine quasixperimentelle Längsschnittstudie*
ISBN 978-3-8325-3519-3 38.50 EUR
- 156 Katja Freyer: Zum Einfluss von Studieneingangsvoraussetzungen auf den Studienerfolg Erstsemesterstudierender im Fach Chemie
ISBN 978-3-8325-3544-5 38.00 EUR
- 157 Alexander Rachel: Auswirkungen instruktionaler Hilfen bei der Einführung des (Ferro-)Magnetismus. *Eine Vergleichsstudie in der Primar- und Sekundarstufe*
ISBN 978-3-8325-3548-3 43.50 EUR
- 158 Sebastian Ritter: Einfluss des Lerninhalts Nanogrößeneffekte auf Teilchen- und Teilchenmodellvorstellungen von Schülerinnen und Schülern
ISBN 978-3-8325-3558-2 36.00 EUR
- 159 Andrea Harbach: Problemorientierung und Vernetzung in kontextbasierten Lernaufgaben
ISBN 978-3-8325-3564-3 39.00 EUR
- 160 David Obst: Interaktive Tafeln im Physikunterricht. *Entwicklung und Evaluation einer Lehrerfortbildung*
ISBN 978-3-8325-3582-7 40.50 EUR
- 161 Sophie Kirschner: Modellierung und Analyse des Professionswissens von Physiklehrkräften
ISBN 978-3-8325-3601-5 35.00 EUR
- 162 Katja Stief: Selbstregulationsprozesse und Hausaufgabenmotivation im Chemieunterricht
ISBN 978-3-8325-3631-2 34.00 EUR
- 163 Nicola Meschede: Professionelle Wahrnehmung der inhaltlichen Strukturierung im naturwissenschaftlichen Grundschulunterricht. *Theoretische Beschreibung und empirische Erfassung*
ISBN 978-3-8325-3668-8 37.00 EUR
- 164 Johannes Maximilian Barth: Experimentieren im Physikunterricht der gymnasialen Oberstufe. *Eine Rekonstruktion übergeordneter Einbettungsstrategien*
ISBN 978-3-8325-3681-7 39.00 EUR
- 165 Sandra Lein: Das Betriebspraktikum in der Lehrerbildung. *Eine Untersuchung zur Förderung der Wissenschafts- und Technikbildung im allgemeinbildenden Unterricht*
ISBN 978-3-8325-3698-5 40.00 EUR
- 166 Veranika Maiseyenka: Modellbasiertes Experimentieren im Unterricht. *Praxistauglichkeit und Lernwirkungen*
ISBN 978-3-8325-3708-1 38.00 EUR

- 167 Christoph Stolzenberger: Der Einfluss der didaktischen Lernumgebung auf das Erreichen geforderter Bildungsziele am Beispiel der W- und P-Seminare im Fach Physik
ISBN 978-3-8325-3708-1 38.00 EUR
- 168 Pia Altenburger: Mehrebenenregressionsanalysen zum Physiklernen im Sachunterricht der Primarstufe. *Ergebnisse einer Evaluationsstudie.*
ISBN 978-3-8325-3717-3 37.50 EUR
- 169 Nora Ferber: Entwicklung und Validierung eines Testinstruments zur Erfassung von Kompetenzentwicklung im Fach Chemie in der Sekundarstufe I
ISBN 978-3-8325-3727-2 39.50 EUR
- 170 Anita Stender: Unterrichtsplanung: Vom Wissen zum Handeln.
Theoretische Entwicklung und empirische Überprüfung des Transformationsmodells der Unterrichtsplanung
ISBN 978-3-8325-3750-0 41.50 EUR
- 171 Jenna Koenen: Entwicklung und Evaluation von experimentunterstützten Lösungsbeispielen zur Förderung naturwissenschaftlich-experimenteller Arbeitsweisen
ISBN 978-3-8325-3785-2 43.00 EUR
- 172 Teresa Henning: Empirische Untersuchung kontextorientierter Lernumgebungen in der Hochschuldidaktik. *Entwicklung und Evaluation kontextorientierter Aufgaben in der Studieneingangsphase für Fach- und Nebenfachstudierende der Physik*
ISBN 978-3-8325-3801-9 43.00 EUR
- 173 Alexander Pusch: Fachspezifische Instrumente zur Diagnose und individuellen Förderung von Lehramtsstudierenden der Physik
ISBN 978-3-8325-3829-3 38.00 EUR
- 174 Christoph Vogelsang: Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenz von (angehenden) Physiklehrkräften. *Zusammenhangsanalysen zwischen Lehrerkompetenz und Lehrerperformanz*
ISBN 978-3-8325-3846-0 50.50 EUR
- 175 Ingo Brebeck: Selbstreguliertes Lernen in der Studieneingangsphase im Fach Chemie
ISBN 978-3-8325-3859-0 37.00 EUR
- 176 Axel Eghtessad: Merkmale und Strukturen von Professionalisierungsprozessen in der ersten und zweiten Phase der Chemielehrerbildung. *Eine empirisch-qualitative Studie mit niedersächsischen Fachleiter_innen der Sekundarstufenlehrkräfte*
ISBN 978-3-8325-3861-3 45.00 EUR
- 177 Andreas Nehring: Wissenschaftliche Denk- und Arbeitsweisen im Fach Chemie. Eine kompetenzorientierte Modell- und Testentwicklung für den Bereich der Erkenntnisgewinnung
ISBN 978-3-8325-3872-9 39.50 EUR
- 178 Maike Schmidt: Professionswissen von Sachunterrichtslehrkräften. Zusammenhangsanalyse zur Wirkung von Ausbildungshintergrund und Unterrichtserfahrung auf das fachspezifische Professionswissen im Unterrichtsinhalt „Verbrennung“
ISBN 978-3-8325-3907-8 38.50 EUR

- 179 Jan Winkelmann: Auswirkungen auf den Fachwissenszuwachs und auf affektive Schülermerkmale durch Schüler- und Demonstrationsexperimente im Physikunterricht
ISBN 978-3-8325-3915-3 41.00 EUR
- 180 Iwen Kobow: Entwicklung und Validierung eines Testinstrumentes zur Erfassung der Kommunikationskompetenz im Fach Chemie
ISBN 978-3-8325-3927-6 34.50 EUR
- 181 Yvonne Gramzow: Fachdidaktisches Wissen von Lehramtsstudierenden im Fach Physik. Modellierung und Testkonstruktion
ISBN 978-3-8325-3931-3 42.50 EUR
- 182 Evelin Schröter: Entwicklung der Kompetenzerwartung durch Lösen physikalischer Aufgaben einer multimedialen Lernumgebung
ISBN 978-3-8325-3975-7 54.50 EUR
- 183 Inga Kallweit: Effektivität des Einsatzes von Selbsteinschätzungsbögen im Chemieunterricht der Sekundarstufe I. *Individuelle Förderung durch selbstreguliertes Lernen*
ISBN 978-3-8325-3965-8 44.00 EUR
- 184 Andrea Schumacher: Paving the way towards authentic chemistry teaching. *A contribution to teachers' professional development*
ISBN 978-3-8325-3976-4 48.50 EUR
- 185 David Woitkowski: Fachliches Wissen Physik in der Hochschulausbildung. *Konzeptualisierung, Messung, Niveaubildung*
ISBN 978-3-8325-3988-7 53.00 EUR
- 186 Marianne Korner: Cross-Age Peer Tutoring in Physik. *Evaluation einer Unterrichtsmethode*
ISBN 978-3-8325-3979-5 38.50 EUR
- 187 Simone Nakoinz: Untersuchung zur Verknüpfung submikroskopischer und makroskopischer Konzepte im Fach Chemie
ISBN 978-3-8325-4057-9 38.50 EUR
- 188 Sandra Anus: Evaluation individueller Förderung im Chemieunterricht. *Adaptivität von Lerninhalten an das Vorwissen von Lernenden am Beispiel des Basiskonzeptes Chemische Reaktion*
ISBN 978-3-8325-4059-3 43.50 EUR
- 189 Thomas Roßbegalle: Fachdidaktische Entwicklungsforschung zum besseren Verständnis atmosphärischer Phänomene. *Treibhauseffekt, saurer Regen und stratosphärischer Ozonabbau als Kontexte zur Vermittlung von Basiskonzepten der Chemie*
ISBN 978-3-8325-4059-3 45.50 EUR
- 190 Kathrin Steckenmesser-Sander: Gemeinsamkeiten und Unterschiede physikbezogener Handlungs-, Denk- und Lernprozesse von Mädchen und Jungen
ISBN 978-3-8325-4066-1 38.50 EUR
- 191 Cornelia Geller: Lernprozessorientierte Sequenzierung des Physikunterrichts im Zusammenhang mit Fachwissenserwerb. *Eine Videostudie in Finnland, Deutschland und der Schweiz*
ISBN 978-3-8325-4082-1 35.50 EUR

- 192 Jan Hofmann: Untersuchung des Kompetenzaufbaus von Physiklehrkräften während einer Fortbildungsmaßnahme
ISBN 978-3-8325-4104-0 38.50 EUR
- 193 Andreas Dickhäuser: Chemiespezifischer Humor. *Theoriebildung, Materialentwicklung, Evaluation*
ISBN 978-3-8325-4108-8 37.00 EUR
- 194 Stefan Korte: Die Grenzen der Naturwissenschaft als Thema des Physikunterrichts
ISBN 978-3-8325-4112-5 57.50 EUR
- 195 Carolin Hülsmann: Kurswahlmotive im Fach Chemie. Eine Studie zum Wahlverhalten und Erfolg von Schülerinnen und Schülern in der gymnasialen Oberstufe
ISBN 978-3-8325-4144-6 49.00 EUR
- 196 Caroline Körbs: Mindeststandards im Fach Chemie am Ende der Pflichtschulzeit
ISBN 978-3-8325-4148-4 34.00 EUR
- 197 Andreas Vorholzer: Wie lassen sich Kompetenzen des experimentellen Denkens und Arbeitens fördern? *Eine empirische Untersuchung der Wirkung eines expliziten und eines impliziten Instruktionsansatzes*
ISBN 978-3-8325-4194-1 37.50 EUR
- 198 Anna Katharina Schmitt: Entwicklung und Evaluation einer Chemielehrerfortbildung zum Kompetenzbereich Erkenntnisgewinnung
ISBN 978-3-8325-4228-3 39.50 EUR
- 199 Christian Maurer: Strukturierung von Lehr-Lern-Sequenzen
ISBN 978-3-8325-4247-4 36.50 EUR
- 200 Helmut Fischler, Elke Sumfleth (Hrsg.): Professionelle Kompetenz von Lehrkräften der Chemie und Physik
ISBN 978-3-8325-4523-9 34.00 EUR
- 201 Simon Zander: Lehrerfortbildung zu Basismodellen und Zusammenhänge zum Fachwissen
ISBN 978-3-8325-4248-1 35.00 EUR
- 202 Kerstin Arndt: Experimentierkompetenz erfassen.
Analyse von Prozessen und Mustern am Beispiel von Lehramtsstudierenden der Chemie
ISBN 978-3-8325-4266-5 45.00 EUR
- 203 Christian Lang: Kompetenzorientierung im Rahmen experimentalchemischer Praktika
ISBN 978-3-8325-4268-9 42.50 EUR
- 204 Eva Cauet: Testen wir relevantes Wissen? *Zusammenhang zwischen dem Professionswissen von Physiklehrkräften und gutem und erfolgreichem Unterrichten*
ISBN 978-3-8325-4276-4 39.50 EUR
- 205 Patrick Löffler: Modellanwendung in Problemlöseaufgaben. *Wie wirkt Kontext?*
ISBN 978-3-8325-4303-7 35.00 EUR

- 206 Carina Gehlen: Kompetenzstruktur naturwissenschaftlicher Erkenntnisgewinnung im Fach Chemie
ISBN 978-3-8325-4318-1 43.00 EUR
- 207 Lars Oettinghaus: Lehrerüberzeugungen und physikbezogenes Professionswissen. *Vergleich von Absolventinnen und Absolventen verschiedener Ausbildungswege im Physikreferendariat*
ISBN 978-3-8325-4319-8 38.50 EUR
- 208 Jennifer Petersen: Zum Einfluss des Merkmals Humor auf die Gesundheitsförderung im Chemieunterricht der Sekundarstufe I. *Eine Interventionsstudie zum Thema Sonnenschutz*
ISBN 978-3-8325-4348-8 40.00 EUR
- 209 Philipp Straube: Modellierung und Erfassung von Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung bei (Lehramts-) Studierenden im Fach Physik
ISBN 978-3-8325-4351-8 35.50 EUR
- 210 Martin Dickmann: Messung von Experimentierfähigkeiten. *Validierungsstudien zur Qualität eines computerbasierten Testverfahrens*
ISBN 978-3-8325-4356-3 41.00 EUR
- 211 Markus Bohlmann: Science Education. Empirie, Kulturen und Mechanismen der Didaktik der Naturwissenschaften
ISBN 978-3-8325-4377-8 44.00 EUR
- 212 Martin Draude: Die Kompetenz von Physiklehrkräften, Schwierigkeiten von Schülerinnen und Schülern beim eigenständigen Experimentieren zu diagnostizieren
ISBN 978-3-8325-4382-2 37.50 EUR
- 213 Henning Rode: Prototypen evidenzbasierten Physikunterrichts. *Zwei empirische Studien zum Einsatz von Feedback und Blackboxes in der Sekundarstufe*
ISBN 978-3-8325-4389-1 42.00 EUR
- 214 Jan-Henrik Kechel: Schülerschwierigkeiten beim eigenständigen Experimentieren. *Eine qualitative Studie am Beispiel einer Experimentieraufgabe zum Hooke'schen Gesetz*
ISBN 978-3-8325-4392-1 55.00 EUR
- 215 Katharina Fricke: Classroom Management and its Impact on Lesson Outcomes in Physics. *A multi-perspective comparison of teaching practices in primary and secondary schools*
ISBN 978-3-8325-4394-5 40.00 EUR
- 216 Hannes Sander: Orientierungen von Jugendlichen beim Urteilen und Entscheiden in Kontexten nachhaltiger Entwicklung. *Eine rekonstruktive Perspektive auf Bewertungskompetenz in der Didaktik der Naturwissenschaft*
ISBN 978-3-8325-4434-8 46.00 EUR
- 217 Inka Haak: Maßnahmen zur Unterstützung kognitiver und metakognitiver Prozesse in der Studieneingangsphase. *Eine Design-Based-Research-Studie zum universitären Lernzentrum Physiktreff*
ISBN 978-3-8325-4437-9 46.50 EUR

- 218 Martina Brandenburger: Was beeinflusst den Erfolg beim Problemlösen in der Physik?
Eine Untersuchung mit Studierenden
ISBN 978-3-8325-4409-6 42.50 EUR
- 219 Corinna Helms: Entwicklung und Evaluation eines Trainings zur Verbesserung der Erklärqualität von Schülerinnen und Schülern im Gruppenpuzzle
ISBN 978-3-8325-4454-6 42.50 EUR
- 220 Viktoria Rath: Diagnostische Kompetenz von angehenden Physiklehrkräften. *Modellierung, Testinstrumentenentwicklung und Erhebung der Performanz bei der Diagnose von Schülervorstellungen in der Mechanik*
ISBN 978-3-8325-4456-0 42.50 EUR
- 221 Janne Krüger: Schülerperspektiven auf die zeitliche Entwicklung der Naturwissenschaften
ISBN 978-3-8325-4457-7 45.50 EUR
- 222 Stefan Mutke: Das Professionswissen von Chemiereferendarinnen und -referendaren in Nordrhein-Westfalen. *Eine Längsschnittstudie*
ISBN 978-3-8325-4458-4 37.50 EUR
- 223 Sebastian Habig: Systematisch variierte Kontextaufgaben und ihr Einfluss auf kognitive und affektive Schülerfaktoren
ISBN 978-3-8325-4467-6 40.50 EUR
- 224 Sven Liepertz: Zusammenhang zwischen dem Professionswissen von Physiklehrkräften, dem sachstrukturellen Angebot des Unterrichts und der Schülerleistung
ISBN 978-3-8325-4480-5 34.00 EUR
- 225 Elina Platova: Optimierung eines Laborpraktikums durch kognitive Aktivierung
ISBN 978-3-8325-4481-2 39.00 EUR
- 226 Tim Reschke: Lesegeschichten im Chemieunterricht der Sekundarstufe I zur Unterstützung von situationalem Interesse und Lernerfolg
ISBN 978-3-8325-4487-4 41.00 EUR
- 227 Lena Mareike Walper: Entwicklung der physikbezogenen Interessen und selbstbezogenen Kognitionen von Schülerinnen und Schülern in der Übergangsphase von der Primar- in die Sekundarstufe. *Eine Längsschnittanalyse vom vierten bis zum siebten Schuljahr*
ISBN 978-3-8325-4495-9 43.00 EUR
- 228 Stefan Anthofer: Förderung des fachspezifischen Professionswissens von Chemielehramtsstudierenden
ISBN 978-3-8325-4498-0 39.50 EUR
- 229 Marcel Bullinger: Handlungsorientiertes Physiklernen mit instruierten Selbsterklärungen in der Primarstufe. *Eine experimentelle Laborstudie*
ISBN 978-3-8325-4504-8 44.00 EUR
- 230 Thomas Amenda: Bedeutung fachlicher Elementarisierungen für das Verständnis der Kinematik
ISBN 978-3-8325-4531-4 43.50 EUR

- 231 Sabrina Milke: Beeinflusst *Priming* das Physiklernen?
Eine empirische Studie zum Dritten Newtonschen Axiom
ISBN 978-3-8325-4549-4 42.00 EUR
- 232 Corinna Erfmann: Ein anschaulicher Weg zum Verständnis der elektromagnetischen Induktion. *Evaluation eines Unterrichtsvorschlags und Validierung eines Leistungsdiagnoseinstruments*
ISBN 978-3-8325-4550-5 49.50 EUR
- 233 Hanne Rautenstrauch: Erhebung des (Fach-)Sprachstandes bei Lehramtsstudierenden im Kontext des Faches Chemie
ISBN 978-3-8325-4556-7 40.50 EUR
- 234 Tobias Klug: Wirkung kontextorientierter physikalischer Praktikumsversuche auf Lernprozesse von Studierenden der Medizin
ISBN 978-3-8325-4558-1 37.00 EUR
- 235 Mareike Bohrmann: Zur Förderung des Verständnisses der Variablenkontrolle im naturwissenschaftlichen Sachunterricht
ISBN 978-3-8325-4559-8 52.00 EUR
- 236 Anja Schödl: FALKO-Physik – Fachspezifische Lehrerkompetenzen im Fach Physik. *Entwicklung und Validierung eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Physiklehrkräften*
ISBN 978-3-8325-4553-6 40.50 EUR
- 237 Hilda Scheuermann: Entwicklung und Evaluation von Unterstützungsmaßnahmen zur Förderung der Variablenkontrollstrategie beim Planen von Experimenten
ISBN 978-3-8325-4568-0 39.00 EUR
- 238 Christian G. Strippel: Naturwissenschaftliche Erkenntnisgewinnung an chemischen Inhalten vermitteln. *Konzeption und empirische Untersuchung einer Ausstellung mit Experimentierstation*
ISBN 978-3-8325-4577-2 41.50 EUR
- 239 Sarah Rau: Durchführung von Sachunterricht im Vorbereitungsdienst. *Eine längsschnittliche, videobasierte Unterrichtsanalyse*
ISBN 978-3-8325-4579-6 46.00 EUR
- 240 Thomas Plotz: Lernprozesse zu nicht-sichtbarer Strahlung. *Empirische Untersuchungen in der Sekundarstufe 2*
ISBN 978-3-8325-4624-3 39.50 EUR
- 241 Wolfgang Aschauer: Elektrische und magnetische Felder. *Eine empirische Studie zu Lernprozessen in der Sekundarstufe II*
ISBN 978-3-8325-4625-0 50.00 EUR
- 242 Anna Donhauser: Didaktisch rekonstruierte Materialwissenschaft. *Aufbau und Konzeption eines Schülerlabors für den Exzellenzcluster Engineering of Advanced Materials*
ISBN 978-3-8325-4636-6 39.00 EUR

- 243 Katrin Schüller: Lernen mit Lösungsbeispielen im Chemieunterricht. *Einflüsse auf Lernerfolg, kognitive Belastung und Motivation*
ISBN 978-3-8325-4640-3 42.50 EUR
- 244 Timo Fleischer: Untersuchung der chemischen Fachsprache unter besonderer Berücksichtigung chemischer Repräsentationen
ISBN 978-3-8325-4642-7 46.50 EUR
- 245 Rosina Steininger: Concept Cartoons als Stimuli für Kleingruppendiskussionen im Chemieunterricht. *Beschreibung und Analyse einer komplexen Lerngelegenheit*
ISBN 978-3-8325-4647-2 39.00 EUR
- 246 Daniel Rehfeldt: Erfassung der Lehrqualität naturwissenschaftlicher Experimentalpraktika
ISBN 978-3-8325-4590-1 40.00 EUR
- 247 Sandra Puddu: Implementing Inquiry-based Learning in a Diverse Classroom: Investigating Strategies of Scaffolding and Students' Views of Scientific Inquiry
ISBN 978-3-8325-4591-8 35.50 EUR
- 248 Markus Bliersbach: Kreativität in der Chemie. *Erhebung und Förderung der Vorstellungen von Chemielehramtsstudierenden*
ISBN 978-3-8325-4593-2 44.00 EUR
- 249 Lennart Kimpel: Aufgaben in der Allgemeinen Chemie. *Zum Zusammenspiel von chemischem Verständnis und Rechenfähigkeit*
ISBN 978-3-8325-4618-2 36.00 EUR
- 250 Louise Bindel: Effects of integrated learning: explicating a mathematical concept in inquiry-based science camps
ISBN 978-3-8325-4655-7 37.50 EUR
- 251 Michael Wenzel: Computereinsatz in Schule und Schülerlabor. *Einstellung von Physiklehrkräften zu Neuen Medien*
ISBN 978-3-8325-4659-5 38.50 EUR
- 252 Laura Muth: Einfluss der Auswertephase von Experimenten im Physikunterricht. *Ergebnisse einer Interventionsstudie zum Zuwachs von Fachwissen und experimenteller Kompetenz von Schülerinnen und Schülern*
ISBN 978-3-8325-4675-5 36.50 EUR
- 253 Annika Fricke: Interaktive Skripte im Physikalischen Praktikum. *Entwicklung und Evaluation von Hypermedien für die Nebenfachausbildung*
ISBN 978-3-8325-4676-2 41.00 EUR
- 254 Julia Haase: Selbstbestimmtes Lernen im naturwissenschaftlichen Sachunterricht. *Eine empirische Interventionsstudie mit Fokus auf Feedback und Kompetenzerleben*
ISBN 978-3-8325-4685-4 38.50 EUR
- 255 Antje J. Heine: Was ist Theoretische Physik? *Eine wissenschaftstheoretische Betrachtung und Rekonstruktion von Vorstellungen von Studierenden und Dozenten über das Wesen der Theoretischen Physik*
ISBN 978-3-8325-4691-5 46.50 EUR

- 256 Claudia Meinhardt: Entwicklung und Validierung eines Testinstruments zu Selbstwirksamkeitserwartungen von (angehenden) Physiklehrkräften in physikdidaktischen Handlungsfeldern
ISBN 978-3-8325-4712-7 47.00 EUR
- 257 Ann-Kathrin Schlüter: Professionalisierung angehender Chemielehrkräfte für einen Gemeinsamen Unterricht
ISBN 978-3-8325-4713-4 53.50 EUR
- 258 Stefan Richtberg: Elektronenbahnen in Feldern. Konzeption und Evaluation einer webbasierten Lernumgebung
ISBN 978-3-8325-4723-3 49.00 EUR
- 259 Jan-Philipp Burde: Konzeption und Evaluation eines Unterrichtskonzepts zu einfachen Stromkreisen auf Basis des Elektronengasmodells
ISBN 978-3-8325-4726-4 57.50 EUR
- 260 Frank Finkenberg: Flipped Classroom im Physikunterricht
ISBN 978-3-8325-4737-4 42.50 EUR
- 261 Florian Treisch: Die Entwicklung der Professionellen Unterrichtswahrnehmung im Lehr-Lern-Labor Seminar
ISBN 978-3-8325-4741-4 41.50 EUR
- 262 Desiree Mayr: Strukturiertheit des experimentellen naturwissenschaftlichen Problemlöseprozesses
ISBN 978-3-8325-4757-8 37.00 EUR
- 263 Katrin Weber: Entwicklung und Validierung einer Learning Progression für das Konzept der chemischen Reaktion in der Sekundarstufe I
ISBN 978-3-8325-4762-2 48.50 EUR
- 264 Hauke Bartels: Entwicklung und Bewertung eines performanznahen Videovignetten-tests zur Messung der Erklärfähigkeit von Physiklehrkräften
ISBN 978-3-8325-4804-9 37.00 EUR
- 265 Karl Marniok: Zum Wesen von Theorien und Gesetzen in der Chemie. *Begriffsanalyse und Förderung der Vorstellungen von Lehramtsstudierenden*
ISBN 978-3-8325-4805-6 42.00 EUR
- 266 Marisa Holzapfel: Fachspezifischer Humor als Methode in der Gesundheitsbildung im Übergang von der Primarstufe zur Sekundarstufe I
ISBN 978-3-8325-4808-7 50.00 EUR
- 267 Anna Stolz: Die Auswirkungen von Experimentiersituationen mit unterschiedlichem Öffnungsgrad auf Leistung und Motivation der Schülerinnen und Schüler
ISBN 978-3-8325-4781-3 38.00 EUR
- 268 Nina Ulrich: Interaktive Lernaufgaben in dem digitalen Schulbuch eChemBook. *Einfluss des Interaktivitätsgrads der Lernaufgaben und des Vorwissens der Lernenden auf den Lernerfolg*
ISBN 978-3-8325-4814-8 43.50 EUR

- 269 Kim-Alessandro Weber: Quantenoptik in der Lehrerfortbildung. *Ein bedarfsgeprägtes Fortbildungskonzept zum Quantenobjekt „Photon“ mit Realexperimenten*
ISBN 978-3-8325-4792-9 55.00 EUR
- 270 Nina Skorsetz: Empathisierer und Systematisierer im Vorschulalter. *Eine Fragebogen- und Videostudie zur Motivation, sich mit Naturphänomenen zu beschäftigen*
ISBN 978-3-8325-4825-4 43.50 EUR
- 271 Franziska Kehne: Analyse des Transfers von kontextualisiert erworbenem Wissen im Fach Chemie
ISBN 978-3-8325-4846-9 45.00 EUR
- 272 Markus Elsholz: Das akademische Selbstkonzept angehender Physiklehrkräfte als Teil ihrer professionellen Identität. *Dimensionalität und Veränderung während einer zentralen Praxisphase*
ISBN 978-3-8325-4857-5 37.50 EUR
- 273 Joachim Müller: Studienerfolg in der Physik. *Zusammenhang zwischen Modellierungskompetenz und Studienerfolg*
ISBN 978-3-8325-4859-9 35.00 EUR
- 274 Jennifer Dörschelln: Organische Leuchtdioden. *Implementation eines innovativen Themas in den Chemieunterricht*
ISBN 978-3-8325-4865-0 59.00 EUR
- 275 Stephanie Strelow: Beliefs von Studienanfängern des Kombi-Bachelors Physik über die Natur der Naturwissenschaften
ISBN 978-3-8325-4881-0 40.50 EUR
- 276 Dennis Jaeger: Kognitive Belastung und aufgabenspezifische sowie personenspezifische Einflussfaktoren beim Lösen von Physikaufgaben
ISBN 978-3-8325-4928-2 50.50 EUR
- 277 Vanessa Fischer: Der Einfluss von Interesse und Motivation auf die Messung von Fach- und Bewertungskompetenz im Fach Chemie
ISBN 978-3-8325-4933-6 39.00 EUR
- 278 René Dohrmann: Professionsbezogene Wirkungen einer Lehr-Lern-Labor-Veranstaltung. *Eine multimethodische Studie zu den professionsbezogenen Wirkungen einer Lehr-Lern-Labor-Blockveranstaltung auf Studierende der Bachelorstudiengänge Lehramt Physik und Grundschulpädagogik (Sachunterricht)*
ISBN 978-3-8325-4958-9 40.00 EUR
- 279 Meike Bergs: Can We Make Them Use These Strategies? *Fostering Inquiry-Based Science Learning Skills with Physical and Virtual Experimentation Environments*
ISBN 978-3-8325-4962-6 39.50 EUR
- 280 Marie-Therese Hauerstein: Untersuchung zur Effektivität von Strukturierung und Binnendifferenzierung im Chemieunterricht der Sekundarstufe I. *Evaluation der Strukturierungshilfe Lernleiter*
ISBN 978-3-8325-4982-4 42.50 EUR

- 281 Verena Zucker: Erkennen und Beschreiben von formativem Assessment im naturwissenschaftlichen Grundschulunterricht. *Entwicklung eines Instruments zur Erfassung von Teilfähigkeiten der professionellen Wahrnehmung von Lehramtsstudierenden*
ISBN 978-3-8325-4991-6 38.00 EUR
- 282 Victoria Telser: Erfassung und Förderung experimenteller Kompetenz von Lehrkräften im Fach Chemie
ISBN 978-3-8325-4996-1 50.50 EUR
- 283 Kristine Tschirschky: Entwicklung und Evaluation eines gedächtnisorientierten Aufgabendesigns für Physikaufgaben
ISBN 978-3-8325-5002-8 42.50 EUR
- 284 Thomas Elert: Course Success in the Undergraduate General Chemistry Lab
ISBN 978-3-8325-5004-2 41.50 EUR
- 285 Britta Kalthoff: Explizit oder implizit? *Untersuchung der Lernwirksamkeit verschiedener fachmethodischer Instruktionen im Hinblick auf fachmethodische und fachinhaltliche Fähigkeiten von Sachunterrichtsstudierenden*
ISBN 978-3-8325-5013-4 37.50 EUR
- 286 Thomas Dickmann: Visuelles Modellverständnis und Studienerfolg in der Chemie. *Zwei Seiten einer Medaille*
ISBN 978-3-8325-5016-5 44.00 EUR
- 287 Markus Sebastian Feser: Physiklehrkräfte korrigieren Schülertexte. *Eine Explorationsstudie zur fachlich-konzeptuellen und sprachlichen Leistungsfeststellung und -beurteilung im Physikunterricht*
ISBN 978-3-8325-5020-2 49.00 EUR
- 288 Matylda Dudzinska: Lernen mit Beispielaufgaben und Feedback im Physikunterricht der Sekundarstufe 1. *Energieerhaltung zur Lösung von Aufgaben nutzen*
ISBN 978-3-8325-5025-7 47.00 EUR
- 289 Ines Sonnenschein: Naturwissenschaftliche Denk- und Arbeitsprozesse Studierender im Labor
ISBN 978-3-8325-5033-2 52.00 EUR
- 290 Florian Simon: Der Einfluss von Betreuung und Betreuenden auf die Wirksamkeit von Schülerlaborbesuchen. *Eine Zusammenhangsanalyse von Betreuungsqualität, Betreuermerkmalen und Schülerlaborzielen sowie Replikationsstudie zur Wirksamkeit von Schülerlaborbesuchen*
ISBN 978-3-8325-5036-3 49.50 EUR
- 291 Marie-Annette Geyer: Physikalisch-mathematische Darstellungswechsel funktionaler Zusammenhänge. *Das Vorgehen von SchülerInnen der Sekundarstufe 1 und ihre Schwierigkeiten*
ISBN 978-3-8325-5047-9 46.50 EUR
- 292 Susanne Digel: Messung von Modellierungskompetenz in Physik. *Theoretische Herleitung und empirische Prüfung eines Kompetenzmodells physikspezifischer Modellierungskompetenz*
ISBN 978-3-8325-5055-4 41.00 EUR

- 293 Sönke Janssen: Angebots-Nutzungs-Prozesse eines Schülerlabors analysieren und gestalten. *Ein design-based research Projekt*
ISBN 978-3-8325-5065-3 57.50 EUR
- 294 Knut Wille: Der Productive Failure Ansatz als Beitrag zur Weiterentwicklung der Aufgabenkultur
ISBN 978-3-8325-5074-5 49.00 EUR
- 295 Lisanne Kraeva: Problemlösestrategien von Schülerinnen und Schülern diagnostizieren
ISBN 978-3-8325-5110-0 59.50 EUR
- 296 Jenny Lorentzen: Entwicklung und Evaluation eines Lernangebots im Lehramtsstudium Chemie zur Förderung von Vernetzungen innerhalb des fachbezogenen Professionswissens
ISBN 978-3-8325-5120-9 39.50 EUR
- 297 Micha Winkelmann: Lernprozesse in einem Schülerlabor unter Berücksichtigung individueller naturwissenschaftlicher Interessenstrukturen
ISBN 978-3-8325-5147-6 48.50 EUR
- 298 Carina Wöhlke: Entwicklung und Validierung eines Instruments zur Erfassung der professionellen Unterrichtswahrnehmung angehender Physiklehrkräfte
ISBN 978-3-8325-5149-0 43.00 EUR
- 299 Thomas Schubatzky: Das Amalgam Anfangs-Elektrizitätslehreunterricht. *Eine multiperspektivische Betrachtung in Deutschland und Österreich*
ISBN 978-3-8325-5159-9 50.50 EUR
- 300 Amany Annaggar: A Design Framework for Video Game-Based Gamification Elements to Assess Problem-solving Competence in Chemistry Education
ISBN 978-3-8325-5150-6 52.00 EUR
- 301 Alexander Engl: CHEMIE PUR – Unterrichten in der Natur: *Entwicklung und Evaluation eines kontextorientierten Unterrichtskonzepts im Bereich Outdoor Education zur Änderung der Einstellung zu „Chemie und Natur“*
ISBN 978-3-8325-5174-2 59.00 EUR
- 302 Christin Marie Sajons: Kognitive und motivationale Dynamik in Schülerlaboren. *Kontextualisierung, Problemorientierung und Autonomieunterstützung der didaktischen Struktur analysieren und weiterentwickeln*
ISBN 978-3-8325-5155-1 56.00 EUR
- 303 Philipp Bitzenbauer: Quantenoptik an Schulen. *Studie im Mixed-Methods Design zur Evaluation des Erlanger Unterrichtskonzepts zur Quantenoptik*
ISBN 978-3-8325-5123-0 59.00 EUR
- 304 Malte S. Ubben: Typisierung des Verständnisses mentaler Modelle mittels empirischer Datenerhebung am Beispiel der Quantenphysik
ISBN 978-3-8325-5181-0 43.50 EUR
- 305 Wiebke Kuske-Janßen: Sprachlicher Umgang mit Formeln von LehrerInnen im Physikunterricht am Beispiel des elektrischen Widerstandes in Klassenstufe 8
ISBN 978-3-8325-5183-4 47.50 EUR

- 306 Kai Bliesmer: Physik der Küste für außerschulische Lernorte. *Eine Didaktische Rekonstruktion*
ISBN 978-3-8325-5190-2 58.00 EUR
- 307 Nikola Schild: Eignung von domänenspezifischen Studieneingangsvariablen als Prädiktoren für Studienerfolg im Fach und Lehramt Physik
ISBN 978-3-8325-5226-8 42.00 EUR
- 308 Daniel Averbeck: Zum Studienerfolg in der Studieneingangsphase des Chemiestudiums. *Der Einfluss kognitiver und affektiv-motivationaler Variablen*
ISBN 978-3-8325-5227-5 51.00 EUR
- 309 Martina Strübe: Modelle und Experimente im Chemieunterricht. *Eine Videostudie zum fachspezifischen Lehrerwissen und -handeln*
ISBN 978-3-8325-5245-9 45.50 EUR
- 310 Wolfgang Becker: Auswirkungen unterschiedlicher experimenteller Repräsentationen auf den Kenntnisstand bei Grundschulkindern
ISBN 978-3-8325-5255-8 50.00 EUR
- 311 Marvin Rost: Modelle als Mittel der Erkenntnisgewinnung im Chemieunterricht der Sekundarstufe I. *Entwicklung und quantitative Dimensionalitätsanalyse eines Testinstruments aus epistemologischer Perspektive*
ISBN 978-3-8325-5256-5 44.00 EUR
- 312 Christina Kobl: Förderung und Erfassung der Reflexionskompetenz im Fach Chemie
ISBN 978-3-8325-5259-6 41.00 EUR
- 313 Ann-Kathrin Beretz: Diagnostische Prozesse von Studierenden des Lehramts – *eine Videostudie in den Fächern Physik und Mathematik*
ISBN 978-3-8325-5288-6 45.00 EUR
- 314 Judith Breuer: Implementierung fachdidaktischer Innovationen durch das Angebot materialgestützter Unterrichtskonzeptionen. *Fallanalysen zum Nutzungsverhalten von Lehrkräften am Beispiel des Münchener Lehrgangs zur Quantenmechanik*
ISBN 978-3-8325-5293-0 50.50 EUR
- 315 Michaela Oettle: Modellierung des Fachwissens von Lehrkräften in der Teilchenphysik. *Eine Delphi-Studie*
ISBN 978-3-8325-5305-0 57.50 EUR
- 316 Volker Brüggemann: Entwicklung und Pilotierung eines adaptiven Multistage-Tests zur Kompetenzerfassung im Bereich naturwissenschaftlichen Denkens
ISBN 978-3-8325-5331-9 40.00 EUR
- 317 Stefan Müller: Die Vorläufigkeit und soziokulturelle Eingebundenheit naturwissenschaftlicher Erkenntnisse. *Kritische Reflexion, empirische Befunde und fachdidaktische Konsequenzen für die Chemielehrer*innenbildung*
ISBN 978-3-8325-5343-2 63.00 EUR
- 318 Laurence Müller: Alltagsentscheidungen für den Chemieunterricht erkennen und Entscheidungsprozesse explorativ begleiten
ISBN 978-3-8325-5379-1 59.00 EUR

- 319 Lars Ehlert: Entwicklung und Evaluation einer Lehrkräftefortbildung zur Planung von selbstgesteuerten Experimenten
ISBN 978-3-8325-5393-71 41.50 EUR
- 320 Florian Seiler: Entwicklung und Evaluation eines Seminarkonzepts zur Förderung der experimentellen Planungskompetenz von Lehramtsstudierenden im Fach Chemie
ISBN 978-3-8325-5397-5 47.50 EUR
- 321 Nadine Boele: Entwicklung eines Messinstruments zur Erfassung der professionellen Unterrichtswahrnehmung von (angehenden) Chemielehrkräften hinsichtlich der Lernunterstützung
ISBN 978-3-8325-5402-6 46.50 EUR
- 322 Franziska Zimmermann: Entwicklung und Evaluation digitalisierungsbezogener Kompetenzen von angehenden Chemielehrkräften
ISBN 978-3-8325-5410-1 49.50 EUR
- 323 Lars-Frederik Weiß: Der Flipped Classroom in der Physik-Lehre. *Empirische Untersuchungen in Schule und Hochschule*
ISBN 978-3-8325-5418-7 51.00 EUR
- 324 Tilmann Steinmetz: Kumulatives Lehren und Lernen im Lehramtsstudium Physik. *Theorie und Evaluation eines Lehrkonzepts*
ISBN 978-3-8325-5421-7 51.00 EUR
- 325 Kübra Nur Celik: Entwicklung von chemischem Fachwissen in der Sekundarstufe I. *Validierung einer Learning Progression für die Basiskonzepte „Struktur der Materie“, „Chemische Reaktion“ und „Energie“ im Kompetenzbereich Umgang mit Fachwissen*
ISBN 978-3-8325-5431-6 55.00 EUR
- 326 Matthias Ungermann: Förderung des Verständnisses von Nature of Science und der experimentellen Kompetenz im Schüler*innen-Labor Physik in Abgrenzung zum Regelunterricht
ISBN 978-3-8325-5442-2 55.50 EUR
- 327 Christoph Hoyer: Multimedial unterstütztes Experimentieren im webbasierten Labor zur Messung, Visualisierung und Analyse des Feldes eines Permanentmagneten
ISBN 978-3-8325-5453-8 45.00 EUR
- 328 Tobias Schüttler: Schülerlabore als interesselördernde authentische Lernorte für den naturwissenschaftlichen Unterricht nutzen
ISBN 978-3-8325-5454-5 50.50 EUR
- 329 Christopher Kurth: Die Kompetenz von Studierenden, Schülerschwierigkeiten beim eigenständigen Experimentieren zu diagnostizieren
ISBN 978-3-8325-5457-6 58.50 EUR
- 330 Dagmar Michna: Inklusiver Anfangsunterricht Chemie *Entwicklung und Evaluation einer Unterrichtseinheit zur Einführung der chemischen Reaktion*
ISBN 978-3-8325-5463-7 49.50 EUR
- 331 Marco Seiter: Die Bedeutung der Elementarisierung für den Erfolg von Mechanikunterricht in der Sekundarstufe I
ISBN 978-3-8325-5471-2 66.00 EUR

- 332 Jörn Hägele: Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten. *Videobasierte Analysen zu Aktivitäten und Vorstellungen von Schülerinnen und Schülern der gymnasialen Oberstufe bei der Bearbeitung von fachmethodischer Instruktion*
ISBN 978-3-8325-5476-7 56.50 EUR
- 333 Erik Heine: Wissenschaftliche Kontroversen im Physikunterricht. *Explorationsstudie zum Umgang von Physiklehrkräften und Physiklehramtsstudierenden mit einer wissenschaftlichen Kontroverse am Beispiel der Masse in der Speziellen Relativitätstheorie*
ISBN 978-3-8325-5478-1 48.50 EUR
- 334 Simon Goertz: Module und Lernzirkel der Plattform FLeXKom zur Förderung experimenteller Kompetenzen in der Schulpraxis *Verlauf und Ergebnisse einer Design-Based Research Studie*
ISBN 978-3-8325-5494-1 66.50 EUR
- 335 Christina Toschka: Lernen mit Modellexperimenten *Empirische Untersuchung der Wahrnehmung und des Denkens in Analogien beim Umgang mit Modellexperimenten*
ISBN 978-3-8325-5495-8 50.00 EUR
- 336 Alina Behrendt: Chemiebezogene Kompetenzen in der Übergangsphase zwischen dem Sachunterricht der Primarstufe und dem Chemieunterricht der Sekundarstufe I
ISBN 978-3-8325-5498-9 40.50 EUR

Alle erschienenen Bücher können unter der angegebenen ISBN direkt online (<http://www.logos-verlag.de>) oder per Fax (030 - 42 85 10 92) beim Logos Verlag Berlin bestellt werden.

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf, Hans Niedderer, Mathias Ropohl und Elke Sumfleth

Die Reihe umfasst inzwischen eine große Zahl von wissenschaftlichen Arbeiten aus vielen Arbeitsgruppen der Physik- und Chemiedidaktik und zeichnet damit ein gültiges Bild der empirischen physik- und chemiedidaktischen Forschung im deutschsprachigen Raum.

Die Herausgeber laden daher Interessenten zu neuen Beiträgen ein und bitten sie, sich im Bedarfsfall an den Logos-Verlag oder an ein Mitglied des Herausgeberteams zu wenden.

Kontaktadressen:

Univ.-Prof. Dr. Martin Hopf
Universität Wien,
Österreichisches Kompetenzzentrum
für Didaktik der Physik,
Porzellangasse 4, Stiege 2,
1090 Wien, Österreich,
Tel. +43-1-4277-60330,
e-mail: martin.hopf@univie.ac.at

Prof. Dr. Hans Niedderer
Inst. f. Didaktik der Naturwissenschaften,
Abt. Physikdidaktik,
FB Physik/ Elektrotechnik,
Universität Bremen,
Postfach 33 04 40, 28334 Bremen
Tel. 0421-218 4695 (Sekretariat),
e-mail: niedderer@physik.uni-bremen.de

Prof. Dr. Mathias Ropohl
Didaktik der Chemie,
Fakultät für Chemie,
Universität Duisburg-Essen,
Schützenbahn 70, 45127 Essen,
Tel. 0201-183 2704,
e-mail: mathias.ropohl@uni-due.de

Prof. Dr. Elke Sumfleth
Didaktik der Chemie,
Fakultät für Chemie,
Universität Duisburg-Essen,
Schützenbahn 70, 45127 Essen
Tel. 0201-183 3757/3761,
e-mail: elke.sumfleth@uni-due.de

Experimentieren kann nicht nur dazu beitragen, fachinhaltliche Erkenntnisse zu gewinnen, sondern auch selbst Lerngegenstand sein. Wenig geklärt ist bisher, wie beim Experimentieren gelernt wird, worauf zu achten ist. Unklar ist auch, welche fachmethodischen Vorstellungen Lernende beim Planen, Durchführen und Auswerten aktivieren.

Diesen Fragen wurde mithilfe einer schriftlichen Instruktion nachgegangen, die Lernende (Jg. 11) anregt, nicht nur Experimente durchzuführen, sondern auch über fachmethodische Grundlagen nachzudenken. Die Arbeitsphasen wurden für 42 Schülergruppen videografiert und kategorienbasiert ausgewertet, um Aussagen über Aktivitätsprofile, Vorstellungen und Kompetenzveränderungen zu generieren.

Die Ergebnisse belegen die Komplexität von Lernprozessen und zeigen die Schwierigkeit, individuelle Kompetenzzuwächse auf bestimmte Prozesse zurückzuführen. Zum Beispiel konnten hohe Kompetenzzuwächse sowohl für Lernende mit häufigen fachmethodischen Beiträgen als auch für solche mit nur geringen fachmethodischen Aktivitäten festgestellt werden. Des Weiteren gehen experimentell angemessene Aktivitäten in erkennbaren Anteilen mit fachmethodisch unpassenden Vorstellungen einher, wie z. B. bei einer korrekt vorgenommenen, aber falsch begründeten Trennung zwischen einer Beobachtung und einer Deutung.

Die Arbeit liefert damit Hinweise, dass Lernende neben fachinhaltlichen auch fachmethodische Vorstellungen in den Unterricht mitbringen. Diese gilt es z. B. durch explizite und wiederholte Thematisierung zu adressieren.

Logos Verlag Berlin

ISBN 978-3-8325-5476-7