

DISSERTATION

**FÖRDERUNG SELBSTREGULIERTEN SCHREIBENS
IN DER GRUNDSCHULE**

Sandra Budde

2010

**FÖRDERUNG SELBSTREGULIERTEN SCHREIBENS
IN DER GRUNDSCHULE:**

ENTWICKLUNG UND ÜBERPRÜFUNG EINES
UNTERRICHTSINTEGRIERTEN INTERVENTIONSPROGRAMMS ZUR
FÖRDERUNG DER SCHREIBKOMPETENZ
VON VIERTKLÄSSLERN

Dissertation zur Erlangung des Grades Dr. phil.
des Fachbereichs Psychologie und Sportwissenschaften
der Justus-Liebig-Universität Gießen

vorgelegt von
Sandra Budde

Gießen

2010

Gutachter:

Prof. Dr. Joachim C. Brunstein

Prof. Dr. Cornelia Glaser

Datum der Disputation:

07. Juli 2010

„Die Wissenschaft fängt eigentlich erst da an interessant zu werden,
wo sie aufhört.“

Justus von Liebig (1803-1873)

Danksagung

Die Arbeit wurde im Rahmen des Forschungsnetzwerks „Empirische Unterrichts- und Bildungsforschung“ der Justus-Liebig-Universität Gießen angefertigt und gefördert. Ich möchte mich herzlich bei allen bedanken, die mich bei meinem Promotionsvorhaben unterstützt und begleitet haben. Mein größter Dank gilt Prof. Dr. Joachim C. Brunstein und Prof. Dr. Cornelia Glaser für die intensive und hervorragende Betreuung der Arbeit und für die vielen konstruktiven Anregungen während meiner gesamten Promotionszeit. Des Weiteren danke ich den Trainern, Tutoren und studentischen Hilfskräften, ohne die die Umsetzung der Projekte in dieser Form nicht möglich gewesen wäre. Ferner danke ich meiner Familie und meinen Freunden für ihr Interesse an meiner Arbeit, aber auch für die Ablenkung und Aufmunterung gerade in der Endphase. Insbesondere für die vielen wertvollen und hilfreichen Anmerkungen zu meiner Arbeit möchte ich mich bedanken.

Zuletzt gilt mein ganz besonderer Dank meinen Eltern, die mich während meines bisherigen Berufswegs fortwährend unterstützt und gefördert haben.

INHALTSVERZEICHNIS

ABBILDUNGSVERZEICHNIS	V
TABELLENVERZEICHNIS	IX
1 EINLEITUNG	1
2 THEORETISCHER HINTERGRUND	4
2.1 PERFORMANZPROZESSE DES SCHREIBENS	5
2.1.1 <i>Kognitive Modelle der Textproduktion</i>	5
2.1.2 <i>Ontogenese der Schreibkompetenz</i>	11
2.1.3 <i>Zusammenfassung</i>	22
2.2 SELBSTREGULATION BEIM PRODUZIEREN EINES TEXTES.....	24
2.2.1 <i>Selbstreguliertes Lernen</i>	24
2.2.2 <i>Förderung der Selbstregulation</i>	29
2.2.3 <i>Selbstreguliertes Schreiben</i>	32
2.2.4 <i>Zusammenfassung</i>	43
2.3 FÖRDERUNG DER SCHREIBKOMPETENZ	44
2.3.1 <i>Schreibprozessansatz zur Förderung der Schreibkompetenz</i>	44
2.3.2 <i>Kognitive Strategieinstruktion beim Schreiben</i>	49
2.3.3 <i>Komponentenanalyse</i>	64
2.3.4 <i>Zusammenfassung</i>	69
2.4 IMPLEMENTIERUNG VON SCHREIBPROGRAMMEN IN DEN REGELUNTERRICHT	70
2.4.1 <i>Implementationsforschung</i>	72
2.4.2 <i>Standards zur Sicherung der Qualität der Implementierung</i>	81
2.4.3 <i>Implementationsforschung im Bereich Schreiben</i>	91
2.4.4 <i>Zusammenfassung</i>	95
3 FRAGESTELLUNG UND HYPOTHESEN.....	97
3.1 ZIELSETZUNG	97
3.2 GRUNDPRINZIPIEN DES EMPIRISCHEN VORGEHENS.....	98
3.3 ÜBERSICHT ÜBER DIE STUDIEN.....	101
3.4 UNTERRICHTSINTEGRIERTES SELBSTREGULATORISCHES AUFSATZTRAINING (USAT).....	104
3.4.1 <i>Schreibstrategien</i>	104

3.4.2	<i>Selbstregulatorische Methoden</i>	105
3.4.3	<i>Instruktionsprinzipien</i>	106
4	LEHRERBEFRAGUNG	108
4.1	ZIELE DER LEHRERBEFRAGUNG	108
4.2	METHODE.....	108
4.3	ERGEBNISSE	109
4.4	DISKUSSION	111
5	PILOTSTUDIE	114
5.1	VORARBEITEN	114
5.2	FRAGESTELLUNG UND HYPOTHESEN	117
5.3	METHODE.....	118
5.3.1	<i>Untersuchungsdesign, Ablauf und Stichprobe</i>	118
5.3.2	<i>Unterrichtintegriertes Selbstregulatorisches Aufsatztraining (USAT)</i>	119
5.3.3	<i>Genauigkeit der Implementation</i>	139
5.3.4	<i>Testungsmaterial</i>	140
5.4	ERGEBNISSE	142
5.5	DISKUSSION	143
6	STUDIE I: IMPLEMENTIERUNG DURCH TRAINER	146
6.1	FRAGESTELLUNG UND HYPOTHESEN	146
6.2	METHODE.....	148
6.2.1	<i>Untersuchungsdesign und Ablauf</i>	148
6.2.2	<i>Stichprobe</i>	150
6.2.3	<i>Trainingsbedingungen</i>	151
6.2.4	<i>Genauigkeit der Implementation</i>	173
6.2.5	<i>Testungsmaterial</i>	180
6.3	ERGEBNISSE	186
6.3.1	<i>Statistische Analysen</i>	186
6.3.2	<i>Zusätzliche Analysen für gruppenrandomisierte Designs</i>	191
6.3.3	<i>Schreibleistung</i>	206
6.3.4	<i>Strategienahe Schreibmaße</i>	210
6.3.5	<i>Transferleistungen</i>	216
6.3.6	<i>Maße der praktischen Signifikanz</i>	218

6.3.7	<i>Differentielle Betrachtung der Bedingungseffekte in Abhängigkeit von den Kontrollvariablen</i>	219
6.4	DISKUSSION	229
6.4.1	<i>Zentrale Befunde der Studie I</i>	229
6.4.2	<i>Einschränkungen und Schlussfolgerungen für Studie II</i>	232
6.4.3	<i>Zusammenfassung</i>	235
7	STUDIE II: IMPLEMENTIERUNG DURCH LEHRER	237
7.1	FRAGESTELLUNG UND HYPOTHESEN	237
7.2	METHODE.....	239
7.2.1	<i>Untersuchungsdesign und Ablauf</i>	239
7.2.2	<i>Stichprobe</i>	240
7.2.3	<i>Trainingsbedingungen</i>	242
7.2.4	<i>Genauigkeit der Implementation</i>	266
7.2.5	<i>Testungsmaterial</i>	275
7.3	ERGEBNISSE	280
7.3.1	<i>Statistische Analysen</i>	280
7.3.2	<i>Zusätzliche Analysen für gruppenrandomisierte Designs</i>	282
7.3.3	<i>Schreibleistung</i>	295
7.3.4	<i>Strategienahe Schreibmaße</i>	300
7.3.5	<i>Maße der praktischen Signifikanz</i>	305
7.3.6	<i>Differentielle Betrachtung der Bedingungseffekte in Abhängigkeit von den Kontrollvariablen</i>	306
7.4	DISKUSSION	316
7.4.1	<i>Zentrale Befunde der Studie II</i>	316
7.4.2	<i>Einschränkungen</i>	321
7.4.3	<i>Zusammenfassung</i>	326
8	GESAMTDISKUSSION	328
8.1	DISKUSSION DER HYPOTHESEN	328
8.1.1	<i>Implementierung in den Regelunterricht</i>	329
8.1.2	<i>Effektivitätsnachweis selbstregulatorischer Strategien</i>	330
8.2	ANMERKUNGEN ZUR QUALITÄT DER IMPLEMENTIERUNGEN	335
8.3	GRENZEN UND PERSPEKTIVEN.....	338
8.4	PRAKTISCHE IMPLIKATIONEN.....	343

9	ZUSAMMENFASSUNG	345
	LITERATUR.....	348
	ANHANG.....	366
A1	LEHRERBEFRAGUNG.....	366
A2	TESTUNGEN.....	366
A3	SCHREIBANLÄSSE IM TRAINING.....	366
A4	PILOTSTUDIE	366
	<i>A4.1 Gruppenarbeiten.....</i>	<i>366</i>
	<i>A4.2 Trainingsmanuale.....</i>	<i>366</i>
	<i>A4.3 Hausaufgaben.....</i>	<i>366</i>
	<i>A4.4 Übungen</i>	<i>366</i>
A5	STUDIE I.....	366
	<i>A5.1 USAT.....</i>	<i>366</i>
	<i>A5.2 WW.....</i>	<i>366</i>
	<i>A5.3 Implementationsgenauigkeit</i>	<i>366</i>
A6	STUDIE II.....	367
	<i>A6.1 USAT.....</i>	<i>367</i>
	<i>A6.2 LUG.....</i>	<i>367</i>
	<i>A6.3 Implementationsgenauigkeit</i>	<i>367</i>
A7	SCHREIBSTRATEGIEN.....	367
	<i>A7.1 Inhalt.....</i>	<i>367</i>
	<i>A7.2 Ausdruck.....</i>	<i>367</i>
	<i>A7.3 Zusammenhang.....</i>	<i>367</i>
	<i>A7.4 Planen.....</i>	<i>367</i>
	<i>A7.5 Überarbeiten.....</i>	<i>367</i>

Abbildungsverzeichnis

Abbildung 1:	Allgemeines kognitives Schreibmodell, adaptiert von Hayes und Flower (1980a, S. 393).....	6
Abbildung 2:	Überarbeitetes kognitives Schreibmodell, adaptiert von Hayes (1996, S. 4).....	8
Abbildung 3:	Beziehung zwischen den Komponenten des Arbeitsgedächtnisses und den Schreibphasen, adaptiert von Kellogg (1996, S. 58).....	9
Abbildung 4:	Knowledge Telling-Strategie, adaptiert von Bereiter und Scardamalia (1987, S. 8).....	17
Abbildung 5:	Knowledge Transforming-Strategie, adaptiert von Bereiter und Scardamalia (1987, S. 12).....	18
Abbildung 6:	Entwicklungsstufen der Schreibkompetenz, adaptiert von Kellogg (2008, S. 4).....	19
Abbildung 7:	Not so simple view of writing, adaptiert von Berninger und Winn (2006, S. 97).....	21
Abbildung 8:	Drei-Schichten-Modell der Selbstregulation, adaptiert von Boekaerts (1999, S. 449).....	25
Abbildung 9:	Phasen und Subprozesse der Selbstregulation, adaptiert von Zimmerman (2002, S. 67).....	27
Abbildung 10:	Zyklisches Trainingsmodell der Selbstregulation, adaptiert von Zimmerman (1998, S. 83).....	30
Abbildung 11:	Story Grammar Strategy, adaptiert von Harris und Graham (1996, S. 49).....	56
Abbildung 12:	AHA-7W-Strategie des SAT-Programms, adaptiert von Glaser (2005, S. 88).....	63
Abbildung 13:	Rahmenmodell der Verbesserung des Unterrichts, adaptiert von Helmke (2009, S. 309).....	71
Abbildung 14:	Zeitliche Reihenfolge der Studien.....	104
Abbildung 15:	3-Schritte-Technik-Poster der USAT-Bedingung.....	120
Abbildung 16:	7W-Fragen-Poster.....	121
Abbildung 17:	Ausdruckskarte.....	122
Abbildung 18:	Zusammenhangssymbol.....	122

Abbildung 19:	7W-Protokoll I.....	124
Abbildung 20:	7W-Protokoll II.	125
Abbildung 21:	Leistungsprotokoll.....	126
Abbildung 22:	Rückseite des Leistungsprotokolls.....	127
Abbildung 23:	Geschichtenplan.	128
Abbildung 24:	Checklisten zur Überprüfung des Inhalts, Ausdrucks und Zusammenhangs einer Geschichte.	129
Abbildung 25:	Poster zur Würdigung der besten Teams.....	130
Abbildung 26:	Untersuchungsdesign der Studie I.....	149
Abbildung 27:	Gemeinsamkeiten und Unterschiede der USAT- und der WW-Bedingung.	153
Abbildung 28:	Trickposter der Studie I und II.	157
Abbildung 29:	3-Schritte-Technik-Poster der WW- und der LUG-Bedingung.	158
Abbildung 30:	Button Bester Bergsteiger (USAT-Bedingung) und Größter Schreiberfan (WW-Bedingung).....	161
Abbildung 31:	Optimierte Checklisten I und II der Studie I und II.	162
Abbildung 32:	Optimiertes 7W-Protokoll II der Studie I und II.	163
Abbildung 33:	Optimiertes Leistungsprotokoll der Studie I und II.	164
Abbildung 34:	Textlupe (WW-Bedingung).....	170
Abbildung 35:	Geschichtenelemente der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).....	193
Abbildung 36:	Geschichtenelemente der WW-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).....	194
Abbildung 37:	Geschichtenqualität der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).....	195
Abbildung 38:	Geschichtenqualität der WW-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).....	195
Abbildung 39:	Wortanzahl der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).	196
Abbildung 40:	Wortanzahl der WW-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).	197
Abbildung 41:	Planen Geschichtenelemente als Funktion von Messzeitpunkt und Klasse (Studie I).	198
Abbildung 42:	Planen Struktur als Funktion von Messzeitpunkt und Klasse (Studie I)....	199

Abbildung 43: Revisionen der USAT-Bedingung in Abhängigkeit von Messzeitpunkt und Klasse (Studie I).....	200
Abbildung 44: Revisionen der WW-Bedingung in Abhängigkeit von Messzeitpunkt und Klasse (Studie I).....	200
Abbildung 45: Wissen der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).....	201
Abbildung 46: Wissen der WW-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).....	202
Abbildung 47: Erlebniserzählungen Geschichtenelemente als Funktion von Messzeitpunkt und Klasse (Studie I).....	203
Abbildung 48: Erlebniserzählungen Geschichtenqualität als Funktion von Messzeitpunkt und Klasse (Studie I).....	203
Abbildung 49: Geschichtenelemente in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).....	207
Abbildung 50: Geschichtenqualität in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).....	209
Abbildung 51: Wortanzahl in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).....	210
Abbildung 52: Planen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).....	212
Abbildung 53: Revisionen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).....	214
Abbildung 54: Wissen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I)....	215
Abbildung 55: Erlebniserzählungen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).....	217
Abbildung 56: Untersuchungsdesign der Studie II.	239
Abbildung 57: Gemeinsamkeiten und Unterschiede der USAT- und der LUG-Bedingung.....	244
Abbildung 58: AHA-7W-Poster und Merkkärtchen der LUG-Bedingung.....	262
Abbildung 59: Ausdrucksgewürze der LUG-Bedingung.....	263
Abbildung 60: Arbeitsblatt zum Planen einer Geschichte der LUG-Bedingung.....	264
Abbildung 61: Geschichtenelemente der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).....	284
Abbildung 62: Geschichtenelemente der LUG-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).....	285

Abbildung 63: Geschichtenqualität der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II)	286
Abbildung 64: Geschichtenelemente der LUG-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II)	287
Abbildung 65: Wortanzahl der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).....	288
Abbildung 66: Wortanzahl der LUG-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).....	288
Abbildung 67: Planen der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).....	290
Abbildung 68: Planen der LUG-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).....	290
Abbildung 69: Revisionen der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).....	292
Abbildung 70: Revisionen der LUG-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).....	292
Abbildung 71: Wissen der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).....	293
Abbildung 72: Wissen der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).....	294
Abbildung 73: Geschichtenelemente in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II).....	297
Abbildung 74: Geschichtenqualität in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II).....	298
Abbildung 75: Wortanzahl in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II)	300
Abbildung 76: Planen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II)....	301
Abbildung 77: Revisionen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II)	303
Abbildung 78: Wissen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II)	304

Tabellenverzeichnis

Tabelle 1:	Beim Schreiben beteiligte Komponenten des Arbeitsgedächtnisses, adaptiert von Kellogg (1999, S. 46).	10
Tabelle 2:	Zusammenhang zwischen den drei Stufen der Rechtschreibentwicklung und der Verknüpfung phonologischer, orthografischer und morphologischer Wortformen, adaptiert von Berninger et al. (2009, S. 36).	14
Tabelle 3:	Effektstärken zur Beurteilung des Einflusses von Strategieinstruktion auf Leistung, Lernstrategien und Affekt.	31
Tabelle 4:	Selbstregulatorische Prozesse erfolgreicher Schriftsteller, adaptiert von Zimmerman und Risemberg (1997, S. 79).	33
Tabelle 5:	Wissenschaftlich fundierte Empfehlungen für den Aufsatzunterricht.	36
Tabelle 6:	Effektstärken und PND's zur Beurteilung des Einflusses von SRSD auf die Aufsatzqualität.	61
Tabelle 7:	Vergleich psychologischer Experimente mit dem Design-Based Research-Ansatz, adaptiert von Barab und Squire (2004, S. 4).	74
Tabelle 8:	Übersicht über die Studien.	102
Tabelle 9:	Inhalte der Trainingseinheiten der Vorarbeiten.	115
Tabelle 10:	Selbstregulatorische Materialien und deren Funktionen.	123
Tabelle 11:	Inhaltlicher Überblick der Pilotstudie.	131
Tabelle 12:	Mittelwerte, Standardabweichungen und T-Werte für alle abhängigen Variablen getrennt nach Messzeitpunkt (Pilotstudie).	142
Tabelle 13:	Beschreibung der Stichprobe getrennt nach Bedingung (Studie I).	151
Tabelle 14:	Prototypische Trainingsinhalte der USAT- und der WW-Bedingung.	152
Tabelle 15:	Vergleich der USAT- und der WW-Bedingung.	154
Tabelle 16:	Angabe der Anteile sehr guter Umsetzungen getrennt nach Faktoren, Trainer und Bedingung (Studie I).	177
Tabelle 17:	Mittelwerte, Standardabweichung und T-Werte für Inhalt, Materialien, Methoden und Neuheit getrennt nach Bedingung (Studie I).	179
Tabelle 18:	Bewertungskriterien der Geschichtenelemente.	182
Tabelle 19:	Bewertungskriterien der Geschichtenqualität (Studie I).	184
Tabelle 20:	Interrater-Reliabilitäten (Studie I).	185

Tabelle 21: Signifikanzniveaus für den Bedingungseffekt nach Berücksichtigung der Kontrollvariablen Rechtschreibung, Klassengröße, Geschlecht, Nationalität und Sprache (Studie I).....	188
Tabelle 22: Geschätzte Intra-Klassen-Korrelationen (ICC's) getrennt nach Messzeitpunkt und Bedingung (Studie I).....	191
Tabelle 23: Mittelwerte und Standardabweichungen für Geschichtenelemente getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).....	192
Tabelle 24: Mittelwerte und Standardabweichungen für Geschichtenqualität getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).....	194
Tabelle 25: Mittelwerte und Standardabweichungen für Wortanzahl getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).....	196
Tabelle 26: Mittelwerte und Standardabweichungen für Planen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).....	197
Tabelle 27: Mittelwerte und Standardabweichungen für Revisionen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).....	199
Tabelle 28: Mittelwerte und Standardabweichungen für Wissen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).....	201
Tabelle 29: Mittelwerte und Standardabweichungen für Erlebniserzählungen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).....	202
Tabelle 30: Statistische Kennwerte (Wald Z) für den Einfluss Klasse innerhalb Bedingung bzw. Schüler innerhalb Klasse sowie F-Werte für den Haupteffekt Bedingung (Studie I).....	205
Tabelle 31: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Geschichtenelemente getrennt nach Bedingung und Messzeitpunkt (Studie I).....	206
Tabelle 32: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Geschichtenqualität getrennt nach Bedingung und Messzeitpunkt (Studie I).....	208
Tabelle 33: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Wortanzahl getrennt nach Bedingung und Messzeitpunkt (Studie I).....	209
Tabelle 34: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Planen getrennt nach Bedingung und Messzeitpunkt (Studie I).....	211

Tabelle 35: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Revisionen getrennt nach Bedingung und Messzeitpunkt (Studie I).....	213
Tabelle 36: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Wissen getrennt nach Bedingung und Messzeitpunkt (Studie I).....	214
Tabelle 37: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Erlebniserzählungen getrennt nach Bedingung und Messzeitpunkt (Studie I).....	216
Tabelle 38: Effektstärken nach Cohen's d für Posttest und Follow-Up mit jeweils Klasse und Schüler als Analyseeinheit (Studie I).....	218
Tabelle 39: Mittelwerte und Standardabweichungen für Mädchen und Jungen getrennt nach Messzeitpunkt und Bedingung (Studie I).....	219
Tabelle 40: Ergebnisse der Analysen zur Prüfung des Faktors Geschlecht als Moderator (Studie I).....	221
Tabelle 41: Mittelwerte und Standardabweichungen für rechtschreibschwache und rechtschreibstarke Schüler getrennt nach Messzeitpunkt und Bedingung (Studie I).....	223
Tabelle 42: Ergebnisse der Analysen zur Prüfung des Faktors Rechtschreibung als Moderator (Studie I).....	224
Tabelle 43: Mittelwerte und Standardabweichungen für deutschsprechende und nicht-deutsch-sprechende Schüler getrennt nach Messzeitpunkt und Bedingung (Studie I).....	226
Tabelle 44: Ergebnisse der Analysen zur Prüfung des Faktors Sprache als Moderator (Studie I).....	227
Tabelle 45: Beschreibung der Stichprobe getrennt nach Bedingung (Studie II).....	241
Tabelle 46: Prototypische Trainingsinhalte der USAT- und der LUG-Bedingung.....	242
Tabelle 47: Vergleich der USAT- und der LUG-Bedingung.....	244
Tabelle 48: Inhalte des USAT-Workshops.....	249
Tabelle 49: Inhalte des LUG-Workshops.....	258
Tabelle 50: Ergebnisse der Workshop-Evaluationen getrennt nach Bedingung (Studie II).....	268
Tabelle 51: Mittelwerte, Standardabweichungen und T-Werte für die Unterrichtsqualität getrennt nach Bedingung (Studie II).....	270

Tabelle 52: Prozentwerte für die inhaltliche Vollständigkeit in den ersten drei Trainingseinheiten der USAT-Bedingung (Studie II).....	272
Tabelle 53: Prozentwerte für selbstregulatorische Aussagen in den ersten drei Trainingseinheiten der USAT-Bedingung (Studie II).....	272
Tabelle 54: Ergebnisse der Trainings-Evaluationen getrennt nach Bedingung (Studie II).	274
Tabelle 55: Bewertungskriterien der Geschichtenqualität Sprache (Studie II).....	277
Tabelle 56: Bewertungskriterien der Geschichtenqualität Struktur (Studie II).....	278
Tabelle 57: Interrater-Reliabilitäten (Studie II).....	278
Tabelle 58: Geschätzte Intra-Klassen-Korrelationen (ICC's) getrennt nach Messzeitpunkt und Bedingung (Studie II).	283
Tabelle 59: Mittelwerte und Standardabweichungen für Geschichtenelemente getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).	284
Tabelle 60: Mittelwerte und Standardabweichungen für Geschichtenqualität getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).	285
Tabelle 61: Mittelwerte und Standardabweichungen für Wortanzahl getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).....	287
Tabelle 62: Mittelwerte und Standardabweichungen für Planen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).....	289
Tabelle 63: Mittelwerte und Standardabweichungen für Revisionen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).....	291
Tabelle 64: Mittelwerte und Standardabweichungen für Wissen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).....	293
Tabelle 65: Statistische Kennwerte (Wald Z) für den Einfluss Klasse innerhalb Bedingung bzw. Schüler innerhalb Klasse sowie F-Werte für den Haupteffekt Bedingung (Studie II).....	295
Tabelle 66: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Geschichtenelemente getrennt nach Bedingung und Messzeitpunkt (Studie II).	296
Tabelle 67: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Geschichtenqualität getrennt nach Bedingung und Messzeitpunkt (Studie II).	297

Tabelle 68: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Wortzahl getrennt nach Bedingung und Messzeitpunkt (Studie II).....	299
Tabelle 69: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Planen getrennt nach Bedingung und Messzeitpunkt (Studie II).....	300
Tabelle 70: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Revisionen getrennt nach Bedingung und Messzeitpunkt (Studie II).....	302
Tabelle 71: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Wissen getrennt nach Bedingung und Messzeitpunkt (Studie II).....	303
Tabelle 72: Effektstärken nach Cohen's d für Posttest und Follow-Up mit jeweils Klasse und Schüler als Analyseeinheit (Studie II).....	305
Tabelle 73: Mittelwerte und Standardabweichungen für Mädchen und Jungen getrennt nach Messzeitpunkt und Bedingung (Studie II).....	306
Tabelle 74: Ergebnisse der Analysen zur Prüfung des Faktors Geschlecht als Moderator (Studie II).....	307
Tabelle 75: Mittelwerte und Standardabweichungen für rechtschreibschwache und rechtschreibstarke Schüler getrennt nach Messzeitpunkt und Bedingung (Studie II).....	308
Tabelle 76: Ergebnisse der Analysen zur Prüfung des Faktors Rechtschreibung als Moderator (Studie II).....	309
Tabelle 77: Mittelwerte und Standardabweichungen für sprachverständnisschwache und sprachverständnisstarke Schüler getrennt nach Messzeitpunkt und Bedingung (Studie II).....	311
Tabelle 78: Ergebnisse der Analysen zur Prüfung des Faktors Sprachverständnis als Moderator (Studie II).....	312
Tabelle 79: Mittelwerte und Standardabweichungen für deutschsprechende und nicht-deutschsprechende Schüler getrennt nach Messzeitpunkt und Bedingung (Studie II).....	314
Tabelle 80: Ergebnisse der Analysen zur Prüfung des Faktors Sprache als Moderator (Studie II).....	315

1 Einleitung

Schreiben, Lesen und Rechnen sind traditionelle Kulturtechniken, die als elementar für das schulische Lernen sowie für die Teilnahme am gesellschaftlichen Leben insgesamt anzusehen sind. Insbesondere die Sprache kann als „Königsweg zur Bildung“ (Schwanitz, 2002, S.525) bezeichnet werden, die den „Menschen erst zum Menschen macht“ (Schwanitz, 2002, S.523).

Die Sprache -und damit auch die Schrift- fungiert als Kommunikationsmittel innerhalb einer Gesellschaft und erlaubt es, die eigenen Gedanken und Wahrnehmungen anderen zugänglich zu machen. Auf diese Weise ist Sprache zudem Ausdruck der eigenen Identität und charakterisiert durch Stil, Wortschatz und Intonation den Besitzer und Anwender der Sprache.

Des Weiteren erschließt sich das Verständnis über die Welt allein durch Sprache. Während das Lesen die Wissensaneignung ermöglicht, ist die Schrift als Grundlage allen Wissens anzusehen, da sie erst die Speicherung von Informationen über Jahrhunderte hinweg gewährleistet. Mit Wissen sind dabei nicht nur die Inhalte von Lehrbüchern gemeint, sondern obendrein alltägliche Dinge wie Liedtexte, Filme, Werbesprüche, Artikelbeschreibungen etc., die Kennzeichen einer Kultur oder einer bestimmten Generation sind. Auch die Grundpfeiler des gesellschaftlichen Zusammenlebens basieren auf schriftlichen Dokumenten wie Gesetzestexten. Ferner ist in der heutigen Zeit der Zugang zu einem Arbeitsplatz nur durch schriftliche Bewerbungen möglich. Eine gut bezahlte Beschäftigung ist zudem nur dann erreichbar, wenn eine Ausbildung oder ein Studium absolviert wurde, was mit dem Durchlaufen mehrerer schriftlicher Tests und dem Anfertigen von schriftlichen Arbeiten einhergeht. Auch im Beruf stellt die schriftliche Ausdrucksfähigkeit eine Voraussetzung für die Karriere dar. Folglich ist das Schreiben ein integraler Bestandteil unserer heutigen Gesellschaft und für den Erfolg in und außerhalb der Schule wesentlich (Troia, 2009b).

Deutsche Lehrpläne der Primar- und Sekundarstufe verfolgen deswegen das explizite Ziel, die Schreibkompetenz der Schüler¹ zu fördern. In den USA ergaben jedoch national repräsentative Erhebungen, dass nur ein geringer Anteil der Schüler tatsächlich kompetent schreibt (Persky, Daane, & Jin, 2003). In Deutschland löste der „Pisa-Schock“ ein erhöhtes öffentliches Interesse an dem Schulunterricht aus und führte zu Forderungen nach Maßnahmen zur Optimierung der Qualität in der Bildung, so dass auch im Bereich Schreiben Forschungsarbeiten zur Förderung der Schreibkompetenz durchgeführt wurden. In kontrollierten experimentellen Studien mit Kleingruppen erwiesen sich zwar Schreibtrainings

¹ Die männliche Form „Schüler“ schließt Schülerinnen mit ein.

als sehr erfolgreich, es fehlten bislang jedoch Ansätze, diese Programme im Rahmen des regulären Deutschunterrichts mit ganzen Klassen zu integrieren.

Dies stellt jedoch einen Widerspruch zu dem eigentlichen Zweck der Forschung im Bereich Pädagogischer Psychologie dar. Die wissenschaftliche Diskussion von Erfolgen in kontrollierten experimentellen Studien garantiert zwar die eigene Forschungskarriere, sie verfolgt jedoch nicht das explizite gesellschaftliche Ziel, welchem sich insbesondere die Pädagogische Psychologie verpflichtet fühlt, nämlich die Schaffung von Brücken zwischen Theorie und Praxis (Paris & Paris, 2001). Demnach sollte das Ergebnis der Pädagogischen Psychologie sein, praktischen Nutzen denjenigen zu bringen, die Tag für Tag mit den Schwierigkeiten des Unterrichts und der Schule zu kämpfen haben. Folglich sollte sich jeder pädagogisch-psychologisch Forschende regelmäßig folgende Frage stellen: Welchen praktischen Nutzen hat meine Forschung bislang leisten können?

Die vorliegende Arbeit intendiert die Entwicklung eines Trainingsprogramms zur Förderung der Schreibkompetenz von Viertklässlern, welches theoretisch fundiert ist und sich bereits in streng kontrollierten experimentellen Studien als effektiv erwiesen hat. Im Gegensatz zu vorangegangenen Untersuchungen soll überprüft werden, ob das Trainingsprogramm auch unter den realen Bedingungen des schulischen Alltags und von Lehrern selbst zu bewältigen ist. Dieses Vorgehen impliziert methodische Herausforderungen wie die Konstanthaltung von Störfaktoren, da mit natürlich vorkommenden Stichproben gearbeitet wird. Die Ziele der Arbeit wurden schrittweise verfolgt.

Im nachfolgenden zweiten Kapitel werden zunächst die theoretischen Grundlagen, auf denen das Schreibtraining basiert, skizziert. Dazu zählen Modelle der Schreibkompetenz und der Schreibentwicklung sowie Überlegungen zu Zusammenhängen zwischen Schreiben und selbstreguliertem Lernen. Überdies werden in diesem Kapitel verschiedene Programme zur Förderung der Schreibkompetenz vorgestellt und abschließend Betrachtungen über die Implementierungsmöglichkeiten von wissenschaftlich erarbeiteten Interventionen angestellt.

Die theoretischen Überlegungen münden in forschungsleitenden Fragen der vorliegenden Arbeit, welche im dritten Kapitel herausgearbeitet werden. Zudem wird eine Übersicht über die Studien der Arbeit gegeben und es werden die wesentlichen Inhalte des Schreibtrainings erörtert.

Die erste Untersuchung ist eine Lehrerbefragung, die sich mit der Frage nach den aktuellen Inhalten und Methoden des Aufsatzunterrichts in der vierten Klasse beschäftigt. Die Fragestellung, Methode, Ergebnisse und Diskussion werden im vierten Kapitel skizziert.

Gegenstand des nächsten Kapitels ist die Beschreibung einer Pilotstudie, die der erstmaligen Umsetzung des Schreibtrainings in einer vierten Klasse diente. Das methodische Vorgehen, die Ergebnisse sowie Schlussfolgerungen für weitere Untersuchungen werden in dem fünften Kapitel beschrieben.

In den darauf folgenden Kapiteln sechs und sieben werden die beiden Hauptstudien der vorliegenden Arbeit präsentiert. In der ersten Studie wurde das Schreibtraining durch geschulte Trainer² und in der nachfolgenden zweiten Studie von Lehrern³ in den Unterricht vierter Klassen implementiert. Die beiden Kapitel ermöglichen, die Ziele, die methodische Umsetzung sowie die statistische Auswertung der beiden Studien nachzuvollziehen. Zusätzlich werden die Ergebnisse dargestellt und abschließend diskutiert.

Die Diskussion sämtlicher Ergebnisse erfolgt im achten Kapitel. Darüber hinaus werden Einschränkungen der Arbeit erörtert und darauf aufbauend zukünftige Forschungsideen skizziert. Das Kapitel schließt mit praktischen Anmerkungen für die Integration des Schreibtrainings in die Lehrerausbildung.

Das letzte bzw. neunte Kapitel fasst die zentralen Befunde der Arbeit zusammen.

Die oben formulierte Frage nach dem praktischen Nutzen der Forschung diene als Leitgedanke der gesamten Arbeit. Nach Analyse und Interpretation der Ergebnisse des vorliegenden Dissertationsprojekts ist ein klarer praktischer Beitrag für Lehrer und Schulen erkennbar.

² Die männliche Form „Trainer“ schließt Trainerinnen mit ein.

³ Die männliche Form „Lehrer“ schließt Lehrerinnen mit ein.

2 Theoretischer Hintergrund

Die Entwicklung der kognitiven Schreibforschung nahm ihren Ursprung in den USA und ist historisch in die für die Lernpsychologie relevante kognitive Wende der 60er Jahre einzuordnen. Gemäß Nystrand (2006) stellten jedoch erst mehrere Artikel in der *Newsweek*, *Times* und im *New Yorker* Mitte bis Ende der 70er Jahre einen Wendepunkt in der Schreibforschung dar. Diese kritisierten den damals in den öffentlichen Schulen praktizierten offenen Unterricht zum Erlernen von Schreiben und führten zu einer öffentlichen Debatte um geeignete Instruktionsprinzipien zur Verbesserung der Schreibleistungen. Folglich zielten bereits die ersten empirischen Forschungsarbeiten im Bereich Schreiben auf die Entwicklung adäquater Förderungsmaßnahmen ab.

Dennoch war es ein weiter Weg von den ersten Ansätzen bis zur Entwicklung von erfolgreichen Förderprogrammen. Während ursprünglich vor allem untersucht wurde, wie das Textprodukt am Ende aussehen sollte, begann man Anfang der 80er Jahre mit der Untersuchung, wie ein durchschnittlicher Schreiber vorgeht und welche kognitiven Prozesse dabei ablaufen. Die Pionierarbeit hierzu wurde von Hayes und Flower (1980a, 1980b) geleistet, deren einflussreiches Schreibmodell aus den 80er Jahren in dem ersten Abschnitt des Theorieteils, Performanzmodelle des Schreibens, vorgestellt werden. In den 90er Jahren kritisierte Kellogg (1996) das Modell von Hayes und Flower und entwickelte basierend auf den Theorien der Informationsverarbeitung ein eigenes Schreibmodell. Sein Modell, welches ebenfalls im ersten Abschnitt dargestellt wird, berücksichtigt die beim Schreiben zur Verfügung stehenden kognitiven Ressourcen. In den 80er Jahren wurden ebenfalls immer mehr Modelle zur Beschreibung der Ontogenese der Schreibkompetenz hervorgebracht, worauf im darauffolgenden Abschnitt eingegangen wird. Der Ansatz von Bereiter und Scardamalia (1987) ist dabei der bekannteste.

Durch die Entwicklung neuer Forschungsmethoden in den 90er Jahren setzte sich die Überzeugung durch, dass das Arbeitsgedächtnis beim Schreiben eine zentrale Rolle spielt und dass die Entlastung des Arbeitsgedächtnisses das vornehmliche Ziel erfolgreicher Interventionsmaßnahmen sein sollte. In diesem Zusammenhang kristallisierte sich die Rolle der Selbstregulation als zentral heraus, so dass den selbstregulatorischen Prozessen bei der Textproduktion ein separater Abschnitt gewidmet wird.

In dem anschließenden Abschnitt werden Schreibprogramme skizziert. Eine besonders effiziente Möglichkeit wurde bereits Anfang der 90er von Graham und Harris (1996) mit dem „Self-Regulated Strategy Development“-Programm entwickelt, welcher die kombinierte

Vermittlung selbstregulatorischer Techniken mit kognitiven Schreibstrategien beinhaltet. Aufbauend auf diesem Ansatz konzipierte Glaser (2005) ein selbstregulatorisches Aufsatztraining.

In dem letzten Abschnitt des Theorieteils werden schließlich wichtige Überlegungen bei der Implementierung von Interventionsmaßnahmen angeführt, um valide Aussagen über die Wirksamkeit treffen zu können.

2.1 Performanzprozesse des Schreibens

In empirischen Studien mit dualer oder dreifacher Aufgabenstellung konnte belegt werden, dass Schreiben ein hoch komplexer und für den Schreibenden anstrengender Prozess ist. In einer zusammenfassenden Schilderung der Befunde zeigten Piolat, Olive und Kellogg (2004), dass Schachspielen am ehesten den kognitiven Anforderungen beim Schreiben gleichkommt. Schreiben und Schachspielen ist gemeinsam, dass eine große Menge an Informationen aus dem Langzeitgedächtnis abgerufen wird, konzeptuell geplant wird und Lösungen für ein Problem entwickelt werden. Dagegen benötigten Probanden, die einen Text lasen, nur halb so viel Zeit wie schreibende Probanden, um auf eine parallel ablaufende Aufgabe, das Drücken einer Taste beim Ertönen eines Beep-Tones, zu reagieren (Olive & Kellogg, 2002).

Um Trainingsprogramme zur Förderung der Schreibkompetenz zu entwickeln, ist es notwendig, Modelle über die beim Schreiben beteiligten mentalen Aktivitäten sowie deren Koordination und Regulation zu spezifizieren. Das Verfassen eines Textes kann dabei als eine Aufgabe des Problemlösens verstanden werden.

Alamargot und Chaquoy (2001) definierten Textproduktion als eine in sich abgeschlossene und komplexe Aktivität, bei der durch die Implementierung verschiedener mentaler Prozesse eine große Menge an Wissen zur Erreichung eines spezifischen Ziels verarbeitet wird. Sie unterschieden dabei vor allem vier Wissensbereiche; (a) Bereichswissen, (b) linguistisches Wissen, (c) pragmatisches Wissen und (d) prozedurales Wissen. Die Autoren wiesen darauf hin, dass bislang kein Schreibmodell existiert, welches sowohl den gesamten Schreibprozess als auch sämtliche Wissensbereiche adäquat abbildet.

2.1.1 Kognitive Modelle der Textproduktion

Die kognitive Wende der 60er Jahre setzte den Ablauf kognitiver Prozesse beim Schreiben in den Mittelpunkt der Schreibforschung. Das 1980 von Hayes and Flower entwickelte Schreibmodell leistete einen großen Beitrag dazu, dass Schreiben heute als ein Problemlöseprozess und nicht als ein rein linguistischer Prozess betrachtet wird. Die meiste Kritik an diesem Modell erfolgte durch Kellogg (1996), was schließlich eine Revision des

ursprünglichen Modells durch Hayes und Flower zur Folge hatte. Im folgenden Abschnitt werden die kognitiven Modelle von Hayes und Flower sowie das Modell von Kellogg vorgestellt.

2.1.1.1 Modelle von Hayes und Flower (1980) und Hayes (1996)

Hayes und Flower verwendeten Protokolle des „Lauten Denkens“, um das Lösen einer Schreibaufgabe und die dabei ablaufenden Prozesse und Schwierigkeiten zu analysieren. Die Protokolle wurden nach der „top down“-Methode ausgewertet: Hayes und Flower untersuchten zunächst den kompletten Schreibvorgang und identifizierten anschließend die einzelnen Subprozesse. Auf diese Weise unterschieden Hayes und Flower zunächst drei Komponenten: das Langzeitgedächtnis, den Schreibprozess und die Schreibumwelt. Das Langzeitgedächtnis umfasst themenspezifisches Wissen, kommunikatives Wissen und bereits abgespeicherte Schreibpläne. Unter Schreibumwelt verstanden Hayes und Flower (1980a, 1980b) neben der Schreibaufgabe und dem bereits formulierten Text alle sich außerhalb des Schreibers befindlichen Einflüsse, wie z.B. der ernste Gesichtsausdruck des Lehrers, welche erhebliche Auswirkungen auf die Schreibmotivation ausüben. Eine vereinfachte Darstellung des Modells ist in Abbildung 1 skizziert.

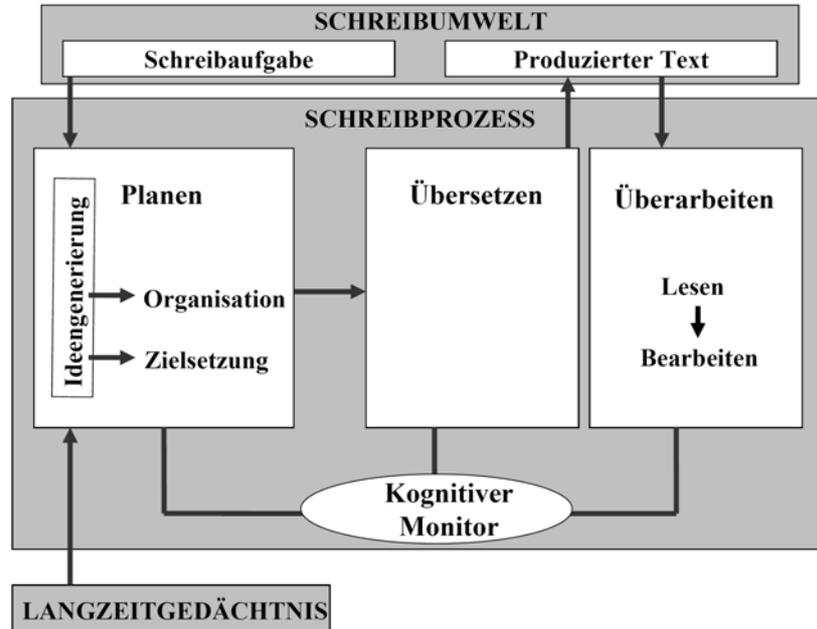


Abbildung 1: Allgemeines kognitives Schreibmodell, adaptiert von Hayes und Flower (1980a, S. 393).

In den weiteren Analysen fokussierten Hayes und Flower (1980a, 1980b) den Schreibprozess und unterteilten diesen in die drei Hauptprozesse des Planens, Schreibens und Überarbeitens, denen wiederum viele weitere Prozesse untergeordnet sind.

Ziel des Planungsprozesses ist die Entwicklung eines Schreibplans. Wie in Abbildung 1 erkenntlich, kann entweder ein bereits vollständig vorliegender Plan aus dem Langzeitgedächtnis abgerufen oder ein neuer Plan erarbeitet werden. Dazu werden zunächst unter Berücksichtigung der Schreibaufgabe Ideen aus dem Langzeitgedächtnis abgerufen („*Ideengenerierung*“) und anschließend sortiert („*Organisation*“). Schließlich werden Kriterien zur Beurteilung der Angemessenheit des anzufertigenden Textes bezüglich der Intention formuliert („*Zielsetzung*“). Der fertige Schreibplan ermöglicht es dem Schreiber, mit der Textproduktion („*Übersetzen*“) zu beginnen. Hierbei wird grammatikalisch-syntaktisches Wissen sowie die abgespeicherten Schreibpläne und Ideen aus dem Langzeitgedächtnis in Sprache übersetzt. Hayes und Flower (1980b) gingen davon aus, dass dieses Wissen als Propositionen im Langzeitgedächtnis gespeichert ist und durch lexikalische und grammatikalische Prozesse in korrekte Sätze umgeformt wird. Der letzte Schreibprozess, das Überarbeiten, dient der Verbesserung des geschriebenen Textes und umfasst die Subprozesse des Lesens und Bearbeitens. Das Überarbeiten stellt eine systematische Überprüfung des geschriebenen Textes dar. Der Schreiber kontrolliert, ob der verfasste Text hinsichtlich semantischer, linguistischer und pragmatischer Eigenschaften dem ursprünglichen Schreibziel angemessen ist. Während das „Über“-arbeiten“ ein abschließender Prozess darstellt, kann der Subprozess des „Be“-arbeitens“ während der gesamten Schreibphase auftreten und somit andere Prozesse unterbrechen. Die Kontrolle und Regulation des Auftretens der einzelnen Subprozesse wird durch den kognitiven Monitor gesteuert. Dieser repräsentiert eine metakognitive Instanz, da er für die zielgerichtete Ausführung der am Schreiben beteiligten Prozesse verantwortlich ist. Die einzelnen Schreibprozesse werden in Abhängigkeit von den gesetzten Zielen initiiert oder unterbrochen, wodurch individuelle Unterschiede im Schreiben erklärt werden können.

Auch wenn das Schreibmodell von Hayes und Flower (1980a, 1980b) nur anhand des Denkprotokolls eines einzelnen Schreibexperten evaluiert wurde, ist es die erste systematische und präzise Darstellung von Schreibprozessen und dient noch heute als Grundlage für viele Untersuchungen. Bereits damals waren sich Hayes und Flower (1980a) jedoch bewusst, dass ihr Modell nur eingeschränkt gültig sein kann. Wesentliche Kritik erfuhr dieses Modell in zweierlei Hinsicht: einerseits wurde die unspezifische und deskriptive Natur des Modells bemängelt und andererseits die Vernachlässigung individueller Unterschiede, des kollaborativen Schreibens und des Schreibmediums (Stift vs. Computer) (Alamargot & Chanquoy, 2001). Daraufhin entwickelte Hayes (1996) ein revidiertes Modell, welches in Abbildung 2 dargestellt ist.

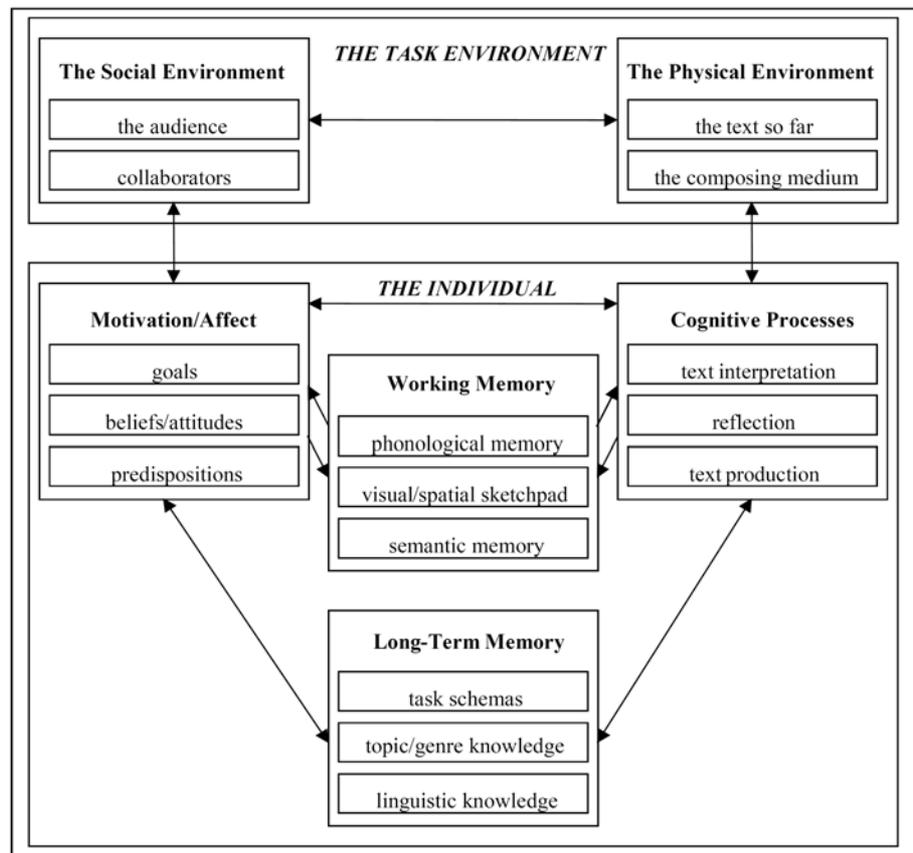


Abbildung 2: Überarbeitetes kognitives Schreibmodell, adaptiert von Hayes (1996, S. 4).

Das neue Modell umfasst die beiden Hauptkomponenten Aufgabenumfeld und Textproduzent. Das Aufgabenumfeld bezieht alle außerhalb des Schreibers liegenden Einflüsse mit ein und ist untergliedert in a) soziale Umwelt, d.h. der Leser und weitere am Schreibprozess Beteiligte sowie b) situative Faktoren, d.h. der bisher verfasste Text und das Schreibmedium.

Der Aufbau der zweiten Hauptkomponente, des Textproduzenten, ist dagegen vielschichtiger und berücksichtigt die Bereiche Motivation/Affekt (Ziele, Prädispositionen, Einstellungen, Kosten-Nutzen-Berechnungen), kognitive Prozesse (Textinterpretation, Reflexion, Textproduktion), Langzeitgedächtnis (Aufgabenschemata, inhaltliches Wissen, leserbezogenes Wissen, linguistisches Wissen, genrespezifisches Wissen) und Arbeitsgedächtnis (phonologisches Gedächtnis, visuell-räumliches Gedächtnis, semantisches Gedächtnis) (Hayes, 1996).

Ein wesentlicher Unterschied zu dem 1980er Modell bestand in der zentralen Rolle des Arbeitsgedächtnisses, welches größtenteils dem Arbeitsgedächtnis-Modell von Baddeley (1986) entspricht. Dieses stellt in dem Modell von Hayes eine entscheidende Ressource dar, auf die der Schreiber während des gesamten Schreibprozesses zugreifen kann. Die weiteren drei Komponenten, Motivation/Affekt, kognitive Prozesse, Langzeitgedächtnis, sind in

Abbildung 2 um das Arbeitsgedächtnis angeordnet und ebenfalls miteinander verknüpft. So beeinflussen z.B. die Ziele, welche Informationen aus dem Langzeitgedächtnis abgerufen und während der Textproduktionsphase unter Verwendung des semantischen Arbeitsgedächtnisses in Sätze transformiert werden.

Im Vergleich zum alten Modell taucht 1996 kein kognitiver Monitor auf. Die ursprünglichen Planungs-, Übersetzungs- und Überarbeitungsphasen wurden durch allgemeine Prozesse der Reflexion, der Textproduktion und der Textinterpretation ersetzt, die jedoch nicht mehr nach einem klaren zeitlichen Ablauf auftreten. Dennoch findet gemäß Hayes (1996) eine Koordination der einzelnen kognitiven Prozesse statt, die durch unterschiedliche Komponenten, vor allem von den Aufgabenschemata des Langzeitgedächtnisses übernommen wird. Abgesehen von dem Arbeitsgedächtnis, welches in dem überarbeiteten Modell hinzugefügt wurde, stellt das neue Schreibmodell von Hayes größtenteils eine Rekonzeptualisierung und eine Neudefinierung der Komponenten des Modells von 1980 dar. Insgesamt ist festzustellen, dass die meisten Bestandteile des frühen Modells auch in dem überarbeiteten Modell wiederzufinden sind.

2.1.1.2 Modell von Kellogg (1996)

Unabhängig von Hayes entwickelte Kellogg (1996) zeitgleich ein kognitives Schreibmodell, welches ebenfalls dem Arbeitsgedächtnis eine zentrale Rolle zuweist. Kellogg war einer der Hauptkritiker des Modells von Hayes und Flower (1980a, 1980b). Er forderte, dass die kognitiven Ressourcen, die während der einzelnen Schreibphasen benötigt werden, stärker berücksichtigt werden. Daraufhin erarbeitete er ein eigenes Schreibmodell, welches vereinfacht in Abbildung 3 dargestellt ist.

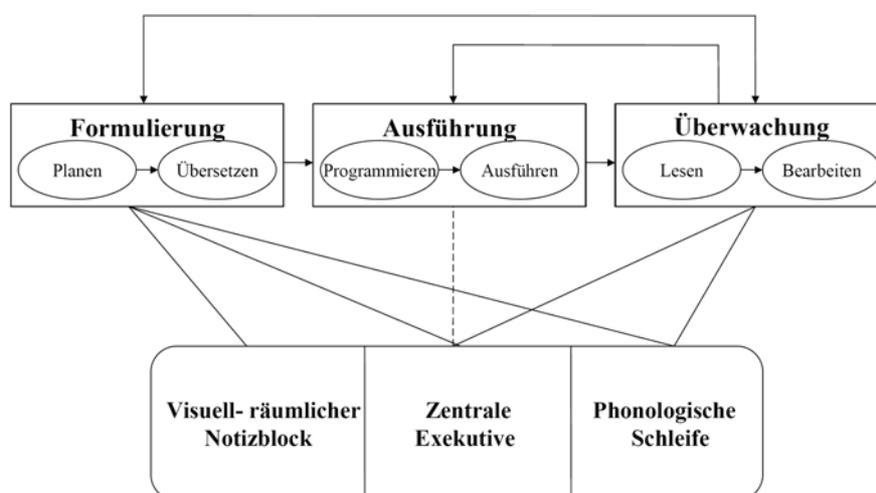


Abbildung 3: Beziehung zwischen den Komponenten des Arbeitsgedächtnisses und den Schreibphasen, adaptiert von Kellogg (1996, S. 58).

Wie auch Baddeley (1986) ging Kellogg (1996) von drei Komponenten des Arbeitsgedächtnisses aus. Die phonologische Schleife ist für die Speicherung und Änderung sprachlicher Informationen zuständig. Der visuell-räumliche Notizblock verarbeitet visuell-räumliche Informationen. Die zentrale Exekutive ist für die Steuerung der beiden Speichermodule sowie für komplexe kognitive Prozesse wie logisches Denken verantwortlich.

Diesen drei Komponenten des Arbeitsgedächtnisses ordnete Kellogg einzelne Schreibphasen zu. Er unterschied zunächst die drei Hauptprozesse des Schreibens (Formulieren, Ausführen, Überarbeiten), denen jeweils zwei Subprozesse untergeordnet waren. Insgesamt sind die drei Phasen mit denen von Hayes und Flower (1980a, 1980b) postulierten Schreibphasen des Planens, Schreibens und Überarbeitens vergleichbar. Im Gegensatz zum kognitiven Schreibmodell von Hayes (1996), in dem der Schreiber in allen Schreibphasen in gleicher Weise auf das Arbeitsgedächtnis zugreift, macht Kelloggs Modell Aussagen darüber, welche spezifischen Komponenten des Arbeitsgedächtnisse bei welcher Schreibphase eine wesentliche Rolle spielen. Diese Überlegungen bieten wichtige Anhaltspunkte für die Konzeptualisierung von Fördermaßnahmen. Eine spezifische Zuordnung der Arbeitsgedächtniskomponenten zu den einzelnen Schreibprozessphasen wurde von Kellogg 1999 noch einmal genauer spezifiziert und ist in Tabelle 1 aufgelistet.

Tabelle 1: Beim Schreiben beteiligte Komponenten des Arbeitsgedächtnisses, adaptiert von Kellogg (1999, S. 46).

		Komponenten des Arbeitsgedächtnisses		
		Visuell-räumlicher Notizblock	Zentrale Exekutive	Phonologische Schleife
Basisprozesse				
Formulierung	Planen	X	X	
	Übersetzen		X	X
Ausführung	Programmieren		⌘	
	Ausführen			
Überwachung	Lesen		X	X
	Bearbeiten	X		

Kellogg (1999) ging davon aus, dass beim Planen insbesondere der visuell-räumliche Notizblock benötigt wird, um Konzepte aus dem Langzeitgedächtnis abzurufen. Die zentrale Exekutive ist während des Planens dafür verantwortlich, die durchgeführten Prozesse in geeignete Schreibpläne umzuwandeln. Der Übersetzungsprozess greift dagegen auf die

phonologische Schleife zurück, wo Wörter und syntaktische Strukturen verarbeitet werden. Die zentrale Exekutive hilft dem Schreiber wiederum dabei, angemessene Wörter auszuwählen und korrekte Sätze zu bilden.

Der zweite Hauptprozess in Kelloggs Modell unterscheidet sich von den anderen Prozessen dahingehend, dass hierbei individuelle Unterschiede auftreten. Die gestrichelte Linie in Abbildung 3 und das hervorgehobene X in Tabelle 1 weisen darauf hin, dass die zentrale Exekutive beim Ausführen nur dann notwendig ist, wenn die motorische Ausführung des Schreibens noch nicht automatisiert ist.

Die Überwachung des Schreibprozesses beansprucht während des Lesens die zentrale Exekutive und die phonologische Schleife. Während des Bearbeitens ist gemäß Kellogg nur noch der visuell-räumliche Notizblock zuständig.

Kelloggs Modell ist ein einflussreiches Schreibmodell, welches viele Forschungsarbeiten anregte. Dennoch wurden bislang nicht sämtliche von ihm postulierten Hypothesen eindeutig belegt. In einer Serie von mehreren Experimenten überprüften Levy und Marek (1999) Kelloggs Modell. Stützende Befunde ermittelten sie für folgende Annahmen: (a) die Beteiligung der phonologischen Schleife beim Verfassen von Sätzen, (b) keine Beteiligung der phonologischen Schleife beim Abschreiben von Texten und (c) keine Beteiligung der phonologischen Schleife beim Identifizieren von Rechtschreib- und Grammatikfehlern. In einer zusammenfassenden Darstellung zu den Befunden zweifelt Hayes (2006) jedoch die letzte Feststellung an. Er berichtet, dass eine experimentelle Manipulation des Zugriffs zur phonologischen Schleife zwar keinen Einfluss auf die absolute Fehlerrate, jedoch auf die Fertigkeit der Probanden, die Fehler zu korrigieren, hat. Folglich sind weitere Untersuchungen notwendig, die einerseits die widersprüchlichen Ergebnisse in Bezug zur phonologischen Schleife klären und andererseits die spezifische Rolle des visuell-räumlichen Notizblocks und der exekutiven Kontrolle beim Schreiben analysieren.

2.1.2 Ontogenese der Schreibkompetenz

In dem kognitiven Schreibmodell von Kellogg (1996, 1999) wurde bereits darauf hingewiesen, dass bei noch nicht voll entwickelten graphomotorischen Fertigkeiten Arbeitsgedächtniskapazitäten notwendig sind, um das Schreiben auszuführen, welche dadurch bei anderen Schreibprozessen fehlen. Aus diesem Grund werden in dem folgenden Abschnitt Theorien zur Erklärung der Entwicklung basaler Schreibfertigkeiten wie graphomotorische und orthografische Strategien vorgestellt. Anschließend wird das Schreibentwicklungsmodell von Bereiter und Scardamalia (1987) erläutert. Dieses ist nicht nur aus historischer Sicht als erstes Modell zur Erklärung der Entwicklung der Textproduktionskompetenz relevant,

sondern beeinflusst noch heute die Schreibforschung. Zuletzt wird die Rolle des Arbeitsgedächtnisses bei der Veränderung vom Schreibnovizen zum Schreibexperten diskutiert.

2.1.2.1 Entwicklung basaler Schreibfertigkeiten

Unter dem Begriff „Schreiben“ sind vielfältige Fertigkeiten gefasst. Auch wenn der Fokus der vorliegenden Arbeit auf höheren Prozessen der Textproduktion liegt, wird im Folgenden ein kurzer Abriss der Entwicklung basaler Schreibfertigkeiten aufgezeigt.

Obwohl Abbott und Berninger (1993) anhand von Pfadanalysen nachweisen konnten, dass graphomotorische, orthografische und Textproduktionsfertigkeiten separate Prozesse darstellen, gibt es viele Belege dafür, dass Rechtschreibprobleme und eine nicht-flüssige Handschrift mit Schwierigkeiten beim Verfassen von Texten einhergehen (Berninger & Amtmann, 2003; Graham & Harris, 2003). Graphomotorische und orthografische Prozesse beanspruchen erhebliche kognitive Ressourcen, wenn sie nicht automatisiert ablaufen, so dass der Textgenerierung nicht mehr genügend Ressourcen zur Verfügung stehen. Dies hat zur Folge, dass Schreibnovizen weniger planen und revidieren sowie einen qualitativ schlechteren Text produzieren (McCutchen, 2006).

Demnach ist ein automatisierter Ablauf der Graphomotorik und der Orthographie für die Effektivität von Trainingsmaßnahmen zur Förderung der Textproduktion hilfreich (Berninger & Amtmann, 2003; Graham & Harris, 2000). Graham, Harris und Fink (2000) führten eine Studie mit Erstklässlern durch, in der graphomotorische Schwierigkeiten mit der Fertigkeit, einen Text zu produzieren, interferierten, so dass ein kombiniertes Training von Graphomotorik und Textproduktionsfertigkeiten zu deutlich besseren Schreibprodukten führte als ein Training der phonologischen Bewusstheit. Ähnliche Ergebnisse ergaben sich für die Rechtschreibung. Zweiklässler, die ein Rechtschreibtraining absolvierten, schrieben im Vergleich zu Zweitklässlern, die eine Matheförderung erhielten, flüssiger. Dieser Effekt war jedoch sechs Wochen nach dem Training nicht mehr nachzuweisen. Gleichzeitig unterschieden sich die beiden Trainingsbedingungen nicht hinsichtlich der Textqualität und der Textlänge (Graham, Harris, & Fink Chorzempa, 2002). In einer Studie mit Drittklässlern konnten ebenfalls positive Effekte eines kombinierten Rechtschreib- und Schreibstrategietrainings auf die Rechtschreibfertigkeiten und auf die Textqualität nachgewiesen werden (Berninger, Vaughan, Abbott, Begay, Coleman, Curtin, Hawkins et al., 2002).

Graham und Harris (2003) fassten die Ergebnisse zur Orthographie und Graphomotorik in drei Punkten zusammen. Erstens berichteten sie, dass Unterschiede

bezüglich des in einer vorgegebenen Zeit produzierten Textumfangs in der Primar- und Mittelstufe jeweils zu 66% durch eine flüssige Handschrift und 41% durch die Beherrschung der Rechtschreibung aufgeklärt werden konnten (Graham, Berninger, Abbott, Abbott, & Whitaker, 1997). Für die Textqualität wurde jedoch nur noch 25% bzw. 42% der Varianz durch Handschrift und Rechtschreibung aufgeklärt (Graham et al., 1997). Zweitens erhöhte sich die Textqualität, wenn Schüler ihre Pläne und Texte nicht selbst aufschrieben, sondern diktierten (De La Paz & Graham, 1997). Drittens führte ein Training der graphomotorischen Fertigkeiten nicht nur zu einer verbesserten Handschrift, sondern zusätzlich zu qualitativ besseren Texten (Berninger, Vaughan, Abbott, Abbott, Woodruff Rogan, Brooks et al., 1997). Es wurde geschlussfolgert, dass Kinder mit schlechter Handschrift das Schreiben generell vermeiden und somit weder ihre Handschrift verbessern, noch das Schreiben eines Textes erlernen.

In einer zusammenfassenden Metaanalyse ermittelten Graham, Olinghouse und Harris (2009) einen moderaten Effekt ($ES = .59$, berechnet nach Cohen's d) für die Vermittlung von Transkriptionsfertigkeiten auf die Schreibqualität. Daraus folgt, dass eine genauere Betrachtung der Entwicklung basaler Schreibfertigkeiten bei der Beschäftigung mit dem Aufsatzschreiben zwingend erforderlich ist.

Die ersten Schreiberfahrungen machen Kinder bereits ab dem dritten Lebensjahr in Form von undefinierbaren grafischen Produkten, die von den üblichen Bildern zu unterscheiden sind (Tolchinsky, 2006). Auch wenn noch keine einheitlichen Grundzüge vorliegen, kann aufgrund der Positionierung auf dem Blatt sowie von der Intention des Kindes zu schreiben, von einem Schreibprodukt ausgegangen werden. Die Mehrzahl der Vorschulkinder verfügt des Weiteren über ausreichend Buchstaben-Wissen, um den eigenen Namen sowie ein bis drei Wörter und mehrere Zahlen zu schreiben (Molfese, Beswick, Molnar, & Jacobi-Vessels, 2006). Folglich ist der Erwerb von Schreibfertigkeiten nicht nur auf schulisches Lernen zurückzuführen, sondern stellt einen natürlichen Entwicklungsprozess dar.

Bezüglich der Entwicklung von Rechtschreibfertigkeiten besteht bislang in der Literatur Uneinigkeit darüber, ob die Entwicklung von der Phonologie über die Orthographie zur Morphologie phasenweise oder gleichzeitig abläuft. In einer Langzeitstudie zeigten sich signifikante phonologische, orthografische und morphologische Unterschiede zwischen guten, durchschnittlichen und schlechten Rechtschreibern von der ersten bis zur siebten Klasse, was gegen eine phasenweise Entwicklung spricht (Berninger, Garcia, & Abbott, 2009). Gleichzeitig ist der Zuwachs der phonologischen und orthografischen Fertigkeiten von der

ersten bis zur dritten Klasse am größten, während die morphologischen Fertigkeiten noch weit nach der vierten Klasse ansteigen (Berninger et al., 2009). Ein Modell der Entwicklung von Rechtschreibfertigkeiten ist in Tabelle 2 abgebildet.

Tabelle 2: Zusammenhang zwischen den drei Stufen der Rechtschreibentwicklung und der Verknüpfung phonologischer, orthografischer und morphologischer Wortformen, adaptiert von Berninger et al. (2009, S. 36).

Phasen	Verknüpfung phonologischer, orthografischer und morphologischer Wortformen
phonologisch	<p>Schrittweise Enkodierung gesprochener Sprache in geschriebene Sprache</p> <ol style="list-style-type: none"> 1. Abbilden von Phonemen in Buchstaben (einzelne Buchstaben oder ein-/zweisilbige Wörter); 2. Abbilden von gesprochenen Anfangsreimen in geschriebene Anfangsreime; 3. Abbilden von ausgesprochenen ganzen Wörtern in komplett buchstabierte Wörter; 4. Abbilden von Phonemen zu Buchstabeneinheiten, die morphologischen Einheiten zugeordnet sind (z.B. die englische Endung –ing steht für Präsens, während die Endung –ed für die Vergangenheit steht; Identifikation von Pluralformen).
orthografisch	<p>Erstellen eines autonomen orthografischen Lexikons im Langzeitgedächtnis</p> <ol style="list-style-type: none"> 1. Das Gedächtnis enthält alle für die Buchstabierung notwendigen Buchstaben eines spezifischen Wortes und verknüpft die Schreibweise mit der Aussprache und der morphologischen Struktur des Wortes. 2. Der Zugriff auf das Gedächtnis erfolgt automatisch und ohne den Enkodierungsprozess der phonologischen Stufe.
morphologisch	<p>Lernen und Anwenden der morphophonemischen Rechtschreibregeln und der Transformationsprozesse in der Muttersprache</p> <ol style="list-style-type: none"> 1. Buchstabieren morphologisch komplexer Wörter, speziell mit lateinischem oder griechischem Wortstamm; 2. Beherrschung des Transformationsprozesses, bei dem morphologisch komplexe Wörter von morphologisch einfachen Wörtern abgeleitet oder generiert werden oder morphologisch komplexe Wörter in einzelne Bestandteile zerlegt werden.

Bezüglich des Zeitpunkts spezifischer Fördermaßnahmen schlagen Berninger et al. (2009) zur Behebung phonologischer und orthografischer Schwierigkeiten die zweite bis vierte Klasse vor. Trainings zur Förderung von morphologischen Fertigkeiten zeigen jedoch die größten Effekte bei Schülern der Klassen sieben bis neun (Berninger et al., 2009).

In mehreren Metaanalysen wurde der Einfluss des Grammatikunterrichts auf die Qualität der Schreibprodukte untersucht. Es ergaben sich dabei unterschiedliche Ergebnisse. Während Graham und Perin (2007a, 2007b) keinen positiven Einfluss des Grammatikunterrichts feststellen konnten, zeigte sich in einer Metaanalyse mit Einzelfalluntersuchungen (Rogers & Graham, 2008) eine Verbesserung der Grammatik in Texten, nachdem ein entsprechendes Training durchgeführt wurde. Die Unterschiede können darauf zurückgeführt werden, dass die Einzelfalluntersuchungen ausschließlich Schüler mit Schreibschwierigkeiten betrachteten, während die Metaanalysen von Graham und Perin (2007a, 2007b) vornehmlich Studien mit durchschnittlich begabten Schreibern analysierten.

Insgesamt lässt sich aus den Befunden zu dem Zusammenhang zwischen Aufsatzschreiben und Graphomotorik bzw. Orthographie Folgendes schließen:

- Die Befunde zur Orthographie sind uneindeutig. Es fehlen zwingende Beweise dafür, dass ein Schreibtraining, welches gleichzeitig Rechtschreibung- und Textproduktionsfertigkeiten schult, zu besseren Ergebnissen führt, als ein reines Rechtschreib- bzw. Textproduktionstraining. Dennoch beeinflussen orthografische Kenntnisse die Textqualität und müssen folglich berücksichtigt werden.
- Im Gegensatz dazu ist ausreichend belegt, dass graphomotorische Schwierigkeiten die Fähigkeit, Aufsätze zu produzieren, erheblich einschränken. Aus diesem Grund sollte ein Trainingsprogramm zur Förderung der Schreibkompetenz erst nach Erreichen einer flüssigen Handschrift beginnen. Es ist davon auszugehen, dass Viertklässler ausreichend flüssig transkribieren können (Berninger et al., 2009; Graham & Harris, 2000; McCutchen, 2006). Auch in dem Entwicklungsmodell von Fitzgerald und Shanhan (2000), welches den Erwerb von Lesen und Schreiben integriert, ist in der vierten Klasse bereits der Aufbau basaler Schreibfertigkeiten für einfache und komplexe Wörter sowie die Aneignung orthografischen und morphologischen Wissens abgeschlossen.

2.1.2.2 Modell von Bereiter und Scardamalia (1987)

In seinen früheren Arbeiten ging Bereiter (1980) von einem Stufenmodell der Schreibentwicklung aus. Er betrachtete die Entwicklung der Schreibkompetenz als die Aneignung einer Vielzahl von Wissenssystemen und Fertigkeiten, die von der flüssigen Produktion von Sprache bis hin zur Beherrschung von Schreibkonventionen und reflektiertem Denken führten. Obwohl er nicht von einer strengen Abfolge der einzelnen Stufen ausging, nahm Bereiter (1980) aufgrund von begrenzten Informationsverarbeitungskapazitäten eine nacheinander ablaufende Reihenfolge der Stufen an. Er unterschied folgende fünf

Entwicklungsstufen: assoziatives Schreiben, regelkonformes Schreiben, leserbezogenes Schreiben, selbstreflektiertes Schreiben und epistemisches Schreiben, welches den Höhepunkt der Schreibentwicklung darstellt. Unter epistemischen Schreiben versteht Bereiter (1980) die Fähigkeit, nicht nur den Prozess des Schreibens auszuführen (assoziatives Schreiben) und ein nach Schreibkonventionen korrektes Schreibprodukt zu erstellen (regelkonformes Schreiben), sondern auch den Leser zu berücksichtigen (leserbezogenes Schreiben), einen eigenen Stil und kritisches Urteilsvermögen zu entwickeln (selbstreflektiertes Schreiben), welches schließlich reflektiertes Denken integriert, so dass das Schreiben selbst zu neuen Erkenntnisgewinnen führt (epistemisches Schreiben).

Basierend auf Novizen-Experten-Vergleichen und Protokollen des „Lauten Denkens“ entwickelten Bereiter und Scardamalia (1987) schließlich ein Entwicklungsmodell der Schreibkompetenz, welches von der Existenz zweier Schreibstrategien zur Kontrolle des Prozesses der Inhaltsgenerierung ausgeht. Während Schreibnovizen nur über die *Knowledge Telling-Strategie* verfügen, welche mit der Stufe des assoziativen Schreibens vergleichbar ist, können Schreibexperten flexibel zwischen der *Knowledge Telling-* und *Knowledge Transforming-Strategie* wählen, so dass dieses Vorgehen dem epistemischen Schreiben gleichkommt (Bereiter, Burtis, & Scardamalia, 1988). Die beiden Strategien werden als Extreme eines Kontinuums angesehen, so dass die Schreibkompetenzentwicklung progressiv via verschiedener Zwischen-Strategien erfolgt.

Knowledge Telling-Strategie: Abbildung 4 illustriert die *Knowledge Telling-Strategie*. Im Zentrum steht der *Knowledge Telling*-Prozess, welcher von drei Komponenten beeinflusst wird: die mentale Repräsentation der Aufgabe sowie das inhalts- und sprachbezogene Wissen aus dem Langzeitgedächtnis.

Die Generierung des Textinhalts erfolgt in sieben Stufen. Als erstes werden unter Heranziehung des Langzeitgedächtnisses und in Abhängigkeit von der Schreibaufgabe das Thema und die Textart des zu schreibenden Textes bestimmt. Anschließend werden mögliche Textinhalte abgerufen und auf deren Angemessenheit überprüft. Erst wenn die Prüfung bestanden wird, werden die Textinhalte aufgeschrieben. Der verfasste Text wird schließlich hinsichtlich der Vollständigkeit kontrolliert. Die *Knowledge Telling-Strategie* wird so lange durchgeführt bis die Schreibaufgabe erledigt ist.

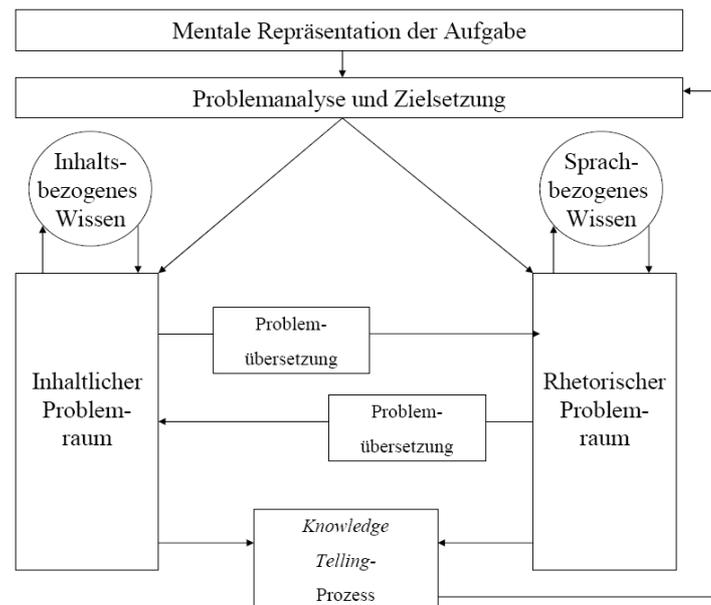


Abbildung 4: Knowledge Telling-Strategie, adaptiert von Bereiter und Scardamalia (1987, S. 8).

Kennzeichnend für die *Knowledge Telling-Strategie* ist, dass Wissen aus dem Langzeitgedächtnis ohne jegliche Reorganisation oder Bearbeitung in einen Text transformiert wird. Sie stellt somit eine einfache und ökonomische Strategie dar, die vor allem Schreibnovizen benutzen. Um die Strategie jedoch anwenden zu können, muss der Schreiber über inhalts- und sprachbezogenes Wissen verfügen und fähig sein, die Angemessenheit der abgerufenen Inhalte zu überprüfen, so dass meistens erst Kinder im Alter von neun bis zehn Jahren die *Knowledge Telling-Strategie* zum Einsatz bringen (Alamargot & Chanquoy, 2001). Aber auch Schreibexperten können bei einfachen Schreibaufgaben, wie das Schreiben von Weihnachtskarten, die *Knowledge Telling-Strategie* anwenden.

Knowledge Transforming: Die *Knowledge Transforming-Strategie* stellt eine Erweiterung der *Knowledge Telling-Strategie* dar, indem Wissen nicht nur abgerufen und hinsichtlich seiner Angemessenheit überprüft, sondern zusätzlich auch organisiert und aufgearbeitet wird. Schreibexperten verfügen über diese Strategie und wenden sie bei komplexen und schwierigen Schreibaufgaben an.

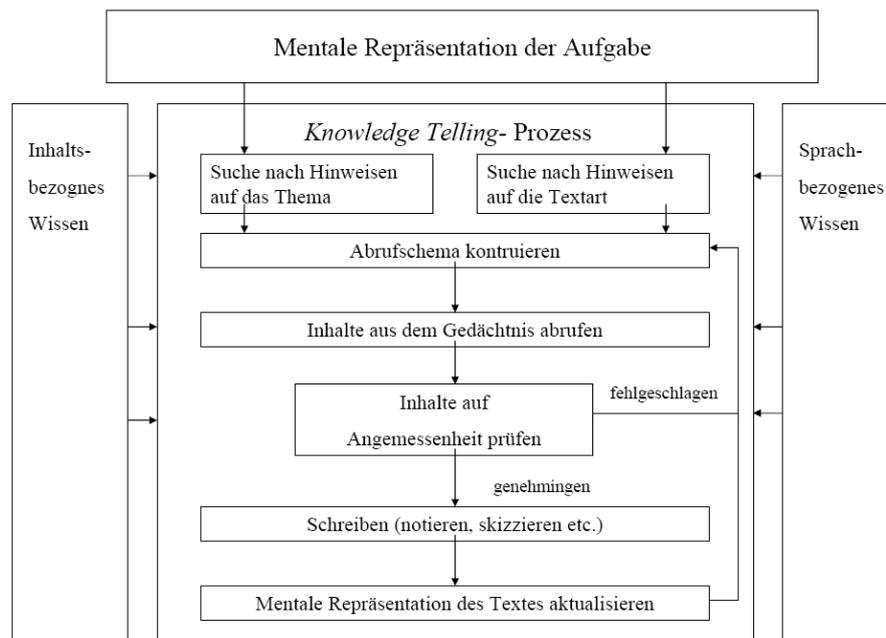


Abbildung 5: Knowledge Transforming-Strategie, adaptiert von Bereiter und Scardamalia (1987, S. 12).

Abbildung 5 illustriert diese Strategie. Ein Vergleich mit Abbildung 4 ergibt, dass die *Knowledge Transforming-Strategie* die Komponenten der *Knowledge Telling-Strategie* mit einschließt. Am Anfang steht wieder die mentale Repräsentation der Aufgabe. Auch die beiden Komponenten des Langzeitgedächtnisses, das inhalts- und sprachbezogene Wissen, sind in dem Modell integriert. Weiter steht am Ende der Strategie der *Knowledge Telling-Prozess*. Zwischen der mentalen Repräsentation der Aufgabe und des *Knowledge Telling-Prozesses* ist jedoch der Prozess der Problemanalyse und der Zielsetzung geschaltet. Dies hat zur Folge, dass das aus dem Langzeitgedächtnis abgerufene Wissen zur Lösung des Problems und zur Erreichung der Ziele reorganisiert wird. Im Einzelnen impliziert dies ein Planungsprozess, welcher die beiden Problembereiche Inhalt, d.h. was soll geschrieben werden, und Sprache, d.h. an wen und wie soll geschrieben werden, kombiniert. Der Austausch zwischen den beiden Problembereichen in Form von Problemübersetzungen ist gemäß Scardamalia und Bereiter (1984) der Kernpunkt erfolgreichen und reflexiven Schreibens und stellt somit einen wesentlichen Ansatzpunkt für mögliche Fördermaßnahmen dar.

Eine weitere wichtige Komponente der *Knowledge Transforming-Strategie* ist die Feedback-Schleife von dem *Knowledge Telling-Prozess* zurück zur Problemanalyse. Diese Rückmeldung ermöglicht eine Reflexion über den geschriebenen Text hinsichtlich der gesetzten Ziele, was eine Überarbeitung des Textes zur Folge hat. Dementsprechend ist die Feedback-Schleife mit dem kognitiven Monitor des Modells von Hayes und Flower (1980a) vergleichbar.

Der progressive Übergang von *Knowledge Telling* zu *Knowledge Transforming* sollte in der Regel erst im Jugendalter vollzogen werden (Alamargot & Chanquoy, 2001). Für Bereiter und Scardamalia (1987) ist das fortgeschrittene Alter dadurch zu erklären, dass die *Knowledge Transforming-Strategie* eine Reifung der Planungsfertigkeiten und eine hohe Gedächtnisspanne voraussetzt.

Erweiterung durch Kellogg (2008): Kelloggs (2008) Erweiterungen des Modells von Bereiter und Scardamalia (1987) betreffen vornehmlich eine dritte Entwicklungsstufe, das *Knowledge Crafting*, und die Betonung des Arbeitsgedächtnisses bei der Schreibentwicklung.

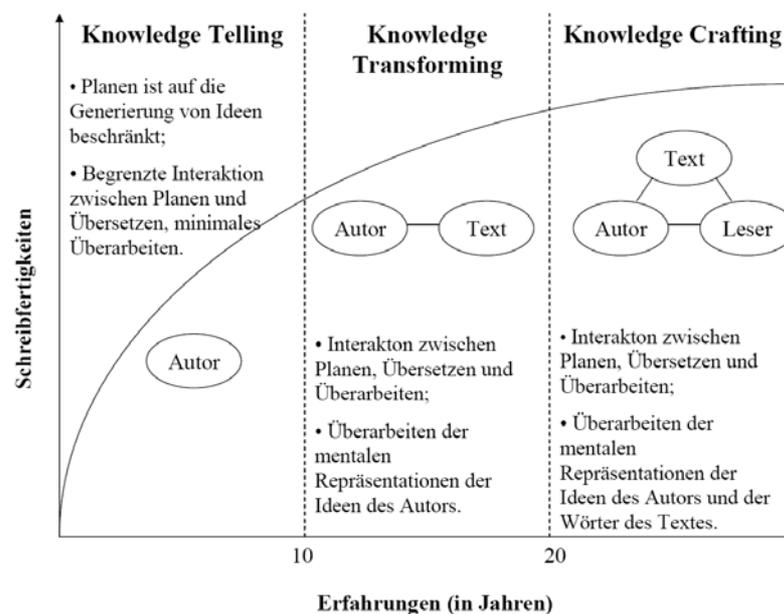


Abbildung 6: Entwicklungsstufen der Schreibkompetenz, adaptiert von Kellogg (2008, S. 4).

Abbildung 6 zeigt eine grafische Darstellung des Entwicklungsmodells von Kellogg (2008). Demnach können in der ersten Stufe, *Knowledge Telling*, nur die Ideen des Autors im Arbeitsgedächtnis gehalten und folglich in das Schreibprodukt integriert werden. In der zweiten Stufe, *Knowledge Transforming*, sind Reifungs- und Lernprozesse bereits soweit fortgeschritten, dass auch der geschriebene Text im Arbeitsgedächtnis gespeichert und mit den Ideen des Autors verknüpft werden kann. Aus diesem Grund ist auf der zweiten Stufe, im Gegensatz zur ersten, das Überarbeiten des Textes möglich. Die Erwartungen des Lesers sowie dessen mögliche Interpretationen des Textes können schließlich erst in der dritten Stufe, dem *Knowledge Crafting*, beim Planen, Schreiben und Überarbeiten berücksichtigt werden. Auch wenn die dritte Entwicklungsstufe selten diskutiert wird, ist sie dennoch notwendig, um ein vollständiges Bild der Entwicklung vom Schreibnovizen zu einem geschickten professionellen Schreiber zu zeichnen. Das *Knowledge Crafting* erfordert ein hohes Maß an Bewusstsein für den potenziellen Leser, aber auch ein ausgeprägtes

Arbeitsgedächtnis und ein flexibel einsetzbares Aufmerksamkeitskontrollsystem, so dass diese Entwicklungsstufe noch nicht im frühen Erwachsenenalter abgeschlossen ist und nicht von jedem erreicht wird (Kellogg, 2008).

Zusammenfassend wird im Verlauf der Schreibentwicklung durch Reifung, Instruktion und Training immer mehr exekutive Kontrolle über die am Schreiben beteiligten kognitiven Prozesse gewonnen, so dass auf die vielfältigen Anforderungen, die die Bewältigung einer Schreibaufgabe umfasst, immer besser und vor allem adaptiver reagiert werden kann.

2.1.2.3 Not so simple view of writing: Das Modell von Berninger und Winn (2006)

Dass die Entlastung des Arbeitsgedächtnisses bei der Entwicklung vom Schreibnovizen zu einem Schreibexperten eine zentrale Rolle spielt, wird nicht nur von Kellogg (2008), sondern von vielen weiteren Schreibforschern vertreten. Eine zentrale Rolle spielt das Arbeitsgedächtnis ebenfalls bei dem *Simple view of writing-Modell* (Berninger & Amtmann, 2003), welches später von Berninger und Winn (2006) zu einem *Not so simple view of writing-Modell* erweitert wurde (siehe Abbildung 7). Obwohl bislang eine genaue Theorie der Schreibentwicklung fehlt (Alamargot & Chanquoy, 2001), stellt das *Simple view of writing* bzw. *Not so simple view of writing-Modell* einen Versuch dar, die in dem vorliegenden Abschnitt geschilderten Befunde zusammenzufassen.

In dem *Simple view of writing-* wie auch in der Erweiterung, dem *Not so simple view of writing-Modell*, wird der Schreibprozess als ein Dreieck angesehen, in dem basale (z.B. graphomotorische Strategien), prozessbezogene (z.B. Planungs- und Überarbeitungsstrategien) und produktbezogene (z.B. Auswahl geeigneter Wörter und Sätze) Komponenten zusammenspielen und von dem Kurzzeitgedächtnis koordiniert werden.

Berninger und Amtmann (2003) erklären in ihrer Darstellung des *Simple view of writing-Modells*, dass am Anfang der Schreibentwicklung die Automatisierung der Transkriptionsfertigkeiten eine entscheidende Rolle spielt. Im weiteren Verlauf sind jedoch exekutive Funktionen für den Erfolg der Textgenierung wichtiger. Da Schreibanfänger die Planungs-, Schreib- und Überarbeitungsprozesse noch nicht automatisiert ausführen können, sind diese auf externe Regulatoren wie Lehrer, Eltern und Peers verstärkt angewiesen. Mit zunehmendem Alter und durch geeignete Instruktionen findet jedoch ein Übergang von der Fremdregulation zur Selbstregulation statt (Berninger & Amtmann, 2003). Folglich muss eine erfolgreiche Schreibinstruktion die Anforderungen, die während des Schreibens an das Arbeitsgedächtnis gestellt werden, reduzieren. Dies kann entweder dadurch geschehen, dass

die einzelnen Komponenten des Schreibprozesses durch ein Gerüst, wie z.B. geeignete Materialien oder Mnemotechniken, gestützt werden oder die Komponenten über die Zeit hinweg oder auf verschiedene Personen aufgeteilt werden (McCutchen, 2006).

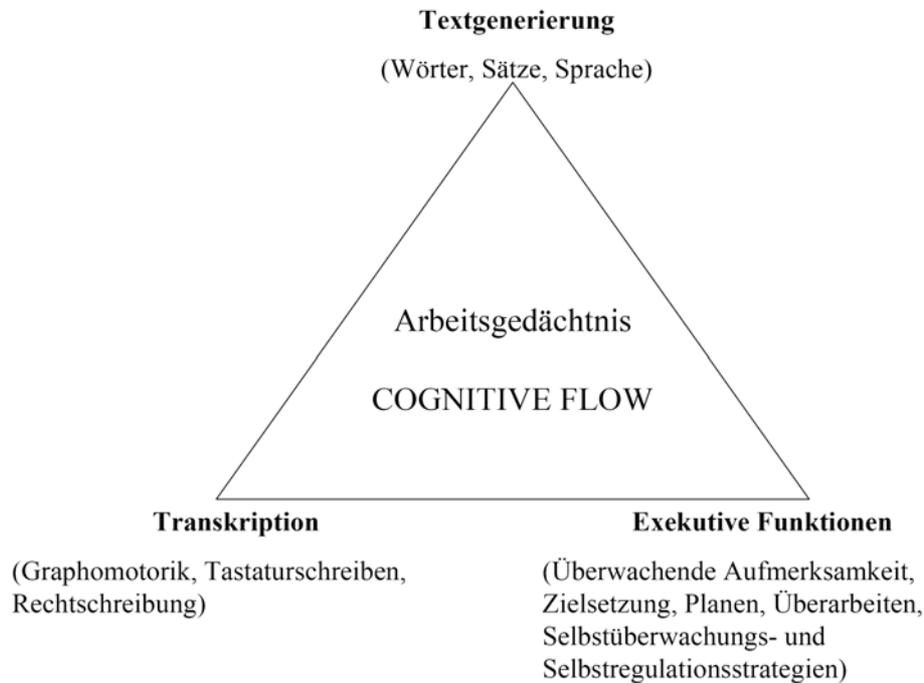


Abbildung 7: Not so simple view of writing, adaptiert von Berninger und Winn (2006, S. 97).

Das *Not so simple view of writing-Modell* wurde entwickelt, um den komplexen Ansprüchen im Schreibunterricht und den multiplen Prozessen, die während des Schreibenlernens ablaufen, gerecht zu werden (Berninger et al., 2009). Im Vergleich zum *Simple view of writing* werden die Rollen des Arbeitsgedächtnisses und der exekutiven Funktionen genauer beschrieben. Auch das Konzept des *Cognitive Flow's* wurde neu hinzugefügt. Hinsichtlich des Arbeitsgedächtnisses unterscheiden Berninger und ihre Kollegen (2009) folgende Komponenten in dem erweiterten Modell: (1) Speicher für orthografische, phonologische und morphologische Informationen, (2) phonologische Schleife für das Erlernen von gesprochener Sprache, (3) orthografische Schleife für das Erlernen von Rechtschreibung, (4) Exekutive, die das verbale Arbeitsgedächtnis mit dem nonverbalen Arbeitsgedächtnis und mit den exekutiven Funktionssystemen verbindet (Berninger et al., 2009). Bezüglich der exekutiven Funktionssysteme werden nicht nur *high level*-Prozesse wie Planen, Schreiben und Überarbeiten, sondern ebenfalls *low level*-Prozesse wie die Kontrolle der Aufmerksamkeit integriert.

2.1.3 Zusammenfassung

In dem vorangegangenen Abschnitt wurden verschiedene Schreibmodelle vorgestellt, wobei zwischen kognitiven Schreibmodellen und Schreibentwicklungsmodellen differenziert wurde.

Kognitive Schreibmodelle: Allen vorgestellten kognitiven Schreibmodellen ist gemeinsam, dass sie versuchen, das Vorgehen eines Schreibexperten zu erklären. Im Hinblick auf das Ziel dieses Dissertationsprojekts, ein erfolgreiches Trainingsprogramm zur Förderung der Schreibkompetenz zu entwickeln, sind die dargestellten Modelle in dem Sinne relevant, dass sie aufzeigen, welche Prozesse in einem Training gefördert werden müssen, um Schreibexpertise zu entwickeln. Da der Fokus der Arbeit auf der Identifikation einer gemeinsamen Basis der beschriebenen Schreibmodelle liegt, lässt sich diesbezüglich folgendes feststellen (Alamargot & Chanquoy, 2001):

- Es können drei Makroprozesse der Textproduktion unterschieden werden: (1) Planen des Inhalts, (2) Übersetzen des Inhalts in eine linguistische Form, (3) Überarbeiten oder Korrigieren des Inhalts. Diesen drei Makroprozessen sind wiederum Subprozesse untergeordnet.
- Die Aktivierung der Schreibprozesse erfolgt in Abhängigkeit von dem Arbeitsgedächtnis.
- Der gesamte Schreibprozess unterliegt Einschränkungen aufgrund begrenzter kognitiver Ressourcen des Arbeitsgedächtnisses.
- Der Ablauf der Prozesse wird zumeist als sequentiell angenommen, wobei die einzelnen Prozesse rekursiv sind und so lange durchgeführt werden bis das Schreibziel erreicht wird.
- Es wird zudem von einer Überwachungsinstanz im Sinne des kognitiven Monitors (Hayes & Flower, 1980a; Hayes & Flower, 1980b) ausgegangen, welcher die Ausführung der Schreibprozesse kontrolliert und reguliert.

Für die vorliegende Arbeit sind die oben genannten Gemeinsamkeiten als Basis für Förderungsansätze hilfreich, da diese mit den Schwierigkeiten von Schreibnovizen übereinstimmen (siehe Kapitel 2.2.3.2).

Ontogenese der Schreibkompetenz: Der Begriff des Schreibens umfasst nicht nur die Produktion eines Textes, wie es in der vorliegenden Arbeit vornehmlich verstanden wird, sondern schließt zudem die Beherrschung basaler Schreibfertigkeiten wie die Handschrift und die Rechtschreibung mit ein. Die Entwicklung dieser basalen Schreibfertigkeiten steht am Anfang des Lernprozesses und ist nicht unabhängig von der Textproduktion zu betrachten. Während hinsichtlich der Rechtschreibung die Befunde uneindeutig sind, besteht aufgrund

von empirischen Studien kein Zweifel mehr daran, dass eine Störung graphomotorischer Fertigkeiten mit einer Störung der Textproduktion einhergeht. Dies kann dadurch erklärt werden, dass Defizite in den basalen Schreibfertigkeiten einen Teufelskreis aus Vermeidungsverhalten und Verfestigung der Defizite auslösen. Dieser Teufelskreis bewirkt schließlich, dass die mangelnden schreibbezogenen Erfahrungen und Fertigkeiten die Produktion von Texten erschweren. Folglich müssen Interventionen einerseits die Rechtschreibfähigkeiten erfassen und diese bei der Interpretation des Interventionserfolgs heranziehen und andererseits erst dann ansetzen, wenn die graphomotorischen Fertigkeiten ausreichend ausgebildet sind, wovon man in der vierten Klasse ausgehen kann. Darüber hinaus muss den Schülern verdeutlicht werden, dass die Qualität der Texte unabhängig von der Leserlichkeit der Handschrift bewertet wird.

Das Schreibentwicklungsmodell von Bereiter und Scardamalia (1987) bzw. dessen Erweiterung durch Kellogg (2008) betonen den progressiven Charakter des Übergangs von *Knowledge Telling* über *Knowledge Transforming* zu *Knowledge Crafting*, welches mit einem erhöhten Bewusstsein für die Ziele des Autors, die Aufgabenstellung und die Leserschaft einhergeht. Eine Zusammenfassung aller bisherigen Erkenntnisse ist Berninger und Winn (2006) mit dem *Not so simple view of writing-Modell* gelungen (siehe Abbildung 7).

Im Hinblick auf die Ziele des Dissertationsprojekts ergibt sich die Frage, ob bzw. wie durch Instruktion und Training aus einem Schreibnovizen ein Schreibexperte werden kann. Bezug nehmend auf die hier diskutierten Modelle und Theorien lassen sich einige Ansätze skizzieren, so dass die Frage, ob eine Einflussnahme auf die Entwicklung möglich ist, positiv beantwortet werden kann (Alamargot & Chanquoy, 2001):

- (1) Vermittlung inhalts- (z.B. spezifisches Wissen über das Thema) und sprachbezogenen (z.B. Wissen über die Textart, Strategien) **Wissens**;
- (2) Adäquate Abrufbarkeit und Verarbeitung des Wissens aus dem **Langzeitgedächtnis** durch intensives Üben, spezielle Instruktionen (z.B. kognitive Lehre) und Mnemotechniken;
- (3) Ausreichende Gedächtnisspanne und exekutive Kontrolle der am Schreiben beteiligten Prozesse durch das **Arbeitsgedächtnis**.

Im Gegensatz zu den ersten beiden Punkten, die bereits relativ gut in der schulischen Praxis umgesetzt werden, stellt die Realisierung des dritten Punktes im Unterricht und in Trainings eine große Herausforderung für die aktuelle Schreibforschung dar (Alamargot & Chanquoy, 2001; Berninger et al., 2009; Kellogg, 2008).

2.2 Selbstregulation beim Produzieren eines Textes

In den 80er Jahren wurden Ansätze zur Beschreibung der beim akademischen Lernen ablaufenden Prozesse, welche in der Regel außerhalb der Schule stattfinden, prominenter (Zimmerman, 1998). In diesem Zusammenhang gewann das Konzept der Selbstregulation immer mehr an Bedeutung und wurde schließlich als zentrale Kompetenz für erfolgreiches akademisches Lernen angesehen (Boekaerts, 1999). Im weiteren Verlauf untersuchten Studien die Rolle der Selbstregulation beim Schreiben. Es zeigte sich, dass die Anwendung selbstregulatorischer Strategien zu höheren Schreibleistungen führte (Graham & Harris, 2000; Zimmerman & Risemberg, 1997). Im Gegensatz dazu benutzten Schreibnovizen kaum selbstregulatorische Strategien, was auf eine unzureichende metakognitive Kontrolle der am Schreiben beteiligten Prozesse zurückzuführen ist (McCutchen, 1988).

In diesem Abschnitt wird zunächst das Konzept der Selbstregulation näher betrachtet und schließlich hinsichtlich der Relevanz für ein Trainingsprogramm zur Förderung der Schreibkompetenz bewertet. Dazu werden die Eigenschaften eines selbstregulierten Schreibers diskutiert.

2.2.1 Selbstreguliertes Lernen

Das Konzept der Selbstregulation umfasst mehrere unterschiedliche Prozesse und wird von verschiedenen Forschungsdisziplinen betrachtet, so dass in Abhängigkeit von der jeweiligen theoretischen Perspektive unterschiedliche Definitionen und Modelle entwickelt wurden. Verstärkt wurde dies durch das öffentliche und politische Interesse an der Operationalisierung und Förderung selbstregulierten Lernens (Boekaerts, 1999). Auch in Deutschland flossen Begriffe wie „Selbstständigkeit“ sowie „selbstorganisiertes und selbstgesteuertes Lernen“ als Schlüsselqualifikationen in die Bildungsstandards der einzelnen Länder ein (Drieschner, 2009).

Unter Selbstregulation wird nicht eine einzelne bestimmte Fertigkeit oder Fähigkeit verstanden, sondern eine Vielzahl an selbstgenerierten Gedanken, Gefühlen und Verhaltensweisen, die selbstgesteuert werden und auf ein bestimmtes Ziel ausgerichtet sind (Zimmerman, 1998). Folglich initiieren und lenken selbstgesteuerte Lerner ihren Lernprozess und verlassen sich weniger auf Lehrer, Eltern oder Freunde (Zimmerman, 1998). Sie sind sich ihrer eigenen Stärken und Schwächen bewusst. Zum Beispiel handelt ein Schüler selbstreguliert, wenn dieser in der Schule fehlt und sich selbst darum kümmert, die verpassten Inhalte nachzulernen. Dieser Schüler wird nicht nur mit höherer Wahrscheinlichkeit akademisch erfolgreich sein, sondern besitzt generell eine positivere Einstellung zum Lernen

und zur Zukunft, was für das spätere Berufsleben ebenfalls vorteilhaft ist (Drieschner, 2009; Zimmerman, 2002).

Ein prominentes Modell selbstregulierten Lernens wurde von Boekaerts (1999) entwickelt. Selbstregulatorisches Lernen wird dabei als ein Zusammenspiel mehrerer reziprok und interaktiv ablaufender Prozesse auf folgenden drei Regulationsebenen angesehen: (1) das Selbst, (2) der Lernprozess und (3) das Informationsverarbeitungssystem (siehe Abbildung 8).

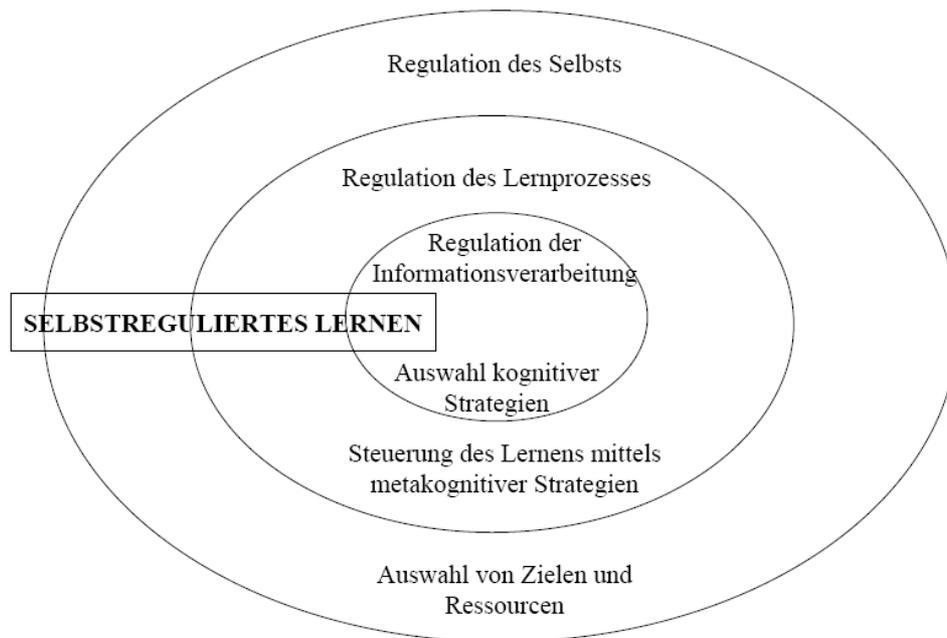


Abbildung 8: Drei-Schichten-Modell der Selbstregulation, adaptiert von Boekaerts (1999, S. 449).

Die meisten selbstregulatorischen Modelle unterscheiden drei verschiedene Arten von Strategien, über die selbstregulierte Lerner verfügen sollten (Boekaerts, 1999; Brunstein & Spörer, 2006; Pintrich, 1999):

(1) **kognitive Strategien:** Ein großes Repertoire an Lernstrategien gehört in allen Modellen der Selbstregulation zu den wesentlichen Charakteristika erfolgreicher Lerner. In Boekaerts Drei-Schichten-Modell ist die Auswahl geeigneter kognitiver Strategien der Kern selbstregulierten Lernens (siehe Abbildung 8). Selbstregulierte Personen können effektiv geeignete Strategien auswählen, kombinieren und koordinieren. Entscheidend ist dabei, dass sich der Lerner multiple kognitive Strategien aneignet, so dass in Abhängigkeit von der jeweiligen Aufgabe und den Rahmenbedingungen die optimale Strategie ausgesucht werden kann. In einer Studie von Zimmerman und Martinez-Pons (1988) konnten 14 verschiedene Lernstrategien den Lernerfolg sowie die von Lehrern eingeschätzte Anwendung der Lernstrategien über verschiedene Kontexte hinweg vorhersagen. Lernstrategien, die speziell für das Schreiben relevant sind, konnten ebenfalls anhand von Experteninterviews identifiziert werden (Zimmerman & Risemberg, 1997).

(2) **metakognitive Strategien:** Des Weiteren spielen die Verfügbarkeit und der effektive Einsatz von metakognitiven Strategien für erfolgreiches Lernen eine zentrale Rolle. Die selbstständige Kontrolle des Lernprozesses wird als eine Schlüsselkompetenz selbstregulierten Lernens angesehen (Zimmerman, 1999). Dazu zählt vor allem die Selbstüberwachung, aber auch die Planung und Bewertung des eigenen Lernprozesses. Boekaerts (1999) unterscheidet zwischen internaler, externaler und geteilter Regulation des Lernprozesses. Während die interne Regulation eine völlig selbstständige und von Lehrern, Eltern und Peers unabhängige Kontrolle des eigenen Lernens zur Erreichung der selbst gesetzten Ziele impliziert, ist bei der externalen Regulation zielgerichtetes Lernen nur durch Lenkung von außen möglich. Eine Mischform zwischen internaler und externaler Regulation stellt die geteilte Regulation dar. Die Regulation des Lernprozesses ist wiederum in Bezug zu setzen zu der inneren Schicht des Modells von Boekaerts, der Regulation der Informationsverarbeitung. Eine interne Regulation geht mit einer tiefen Informationsverarbeitung (z.B. strukturieren, Beziehungen herstellen) und eine externe Regulation mit einer oberflächlichen Informationsverarbeitung (z.B. erinnern, wiederholen) einher (Boekaerts, 1999; Rozendall, Minnaert, & Boekaerts, 2005). Die Regulation des Lernprozesses ist zudem kontextspezifisch und kann in Abhängigkeit von dem kognitiven und metakognitiven Wissen in spezifischen Domains variieren.

Kritisch ist jedoch bei Boekaerts' (1999) Unterscheidung zwischen internaler und externaler Kontrolle anzumerken, dass selbstregulierte Lerner ihr Verhalten anpassen und ihre Umwelt umstrukturieren, um optimal ihre Ziele zu erreichen. Dazu zählt jedoch ebenfalls die Heranziehung von externaler Hilfe, die von selbstregulierten Lernern jedoch nicht als Beschränkung, sondern vielmehr als Bereicherung des eigenen Lernens angesehen wird (Zimmerman, 1999).

(3) **motivationale Strategien:** Abgesehen von den Strategien, die selbstregulierte Lerner einsetzen und den Prozessen zur Regulation dieser Strategien ist die Betrachtung des *Selbsts* notwendig, um die Initiierung und Aufrechterhaltung von Lernaktivitäten zu erklären. Aus diesem Grund fügte Boekaerts in ihrem Modell die äußere Schicht des *Selbst* dazu (siehe Abbildung 8). Ein Lerner verfolgt zu einem Zeitpunkt immer eine Vielzahl von Zielen, die er abwägen und hinsichtlich der Priorität sortieren muss. Gemäß Boekaerts (1999) benötigt der selbstregulierte Lerner Zeit, um sich über die verschiedenen Ziele Gedanken zu machen sowie Wissen darüber, welche Anstrengungen notwendig sind, um die Ziele zu erreichen. Weitere Kennzeichen eines selbstregulierten Lerners sind eine hohe Selbstwirksamkeitserwartung, die adaptive Bewertung von Lernergebnissen sowie eine ausgeprägte intrinsische Motivation

(Brunstein & Spörer, 2006; Zimmerman, 2002). Die Beziehung ist insofern reziprok, da selbstregulierte Lerner von ihren eigenen akademischen Leistungen überzeugt sind und gleichzeitig die Förderung selbstregulatorischer Prozesse (z.B. Selbstüberwachung) die intrinsische Motivation steigert (Zimmerman, 2002). In Pfadanalysen konnten die indirekten und direkten Effekte von schulischen Selbstwirksamkeitserwartungen über Selbstregulationsprozesse auf die schulische Leistung nachgewiesen werden (Zimmerman, Bandura, & Martinez-Pons, 1992; Zimmerman & Bandura, 1994).

Insgesamt ergibt sich ein komplexes Bild an Strategien, Verhaltensmerkmalen, Einstellungen und Umweltfaktoren, die für das selbstregulierte Lernen von Bedeutung sind. Darüber hinaus sind die dargestellten selbstregulatorischen Prozesse nicht unabhängig voneinander zu sehen. Wie am Beispiel der Selbstwirksamkeit erläutert, beeinflussen sich die Prozesse gegenseitig. Folglich zeichnet sich ein erfolgreicher selbstregulierter Lerner durch ein vielschichtiges Zusammenspiel kognitiver, metakognitiver und motivationaler Strategien aus, wobei zu klären ist, wie die einzelnen Facetten selbstregulatorischen Lernens miteinander verwoben sind und wie die Interaktion zwischen den einzelnen Prozessen funktioniert. Aktuell gehen die meisten Forscher von einem zyklischen Zusammenhang der beim selbstregulierten Lernen involvierten Prozesse aus. Zimmerman (2002) entwickelte daraufhin ein zyklisches Phasenmodell der Selbstregulation (siehe Abbildung 9).

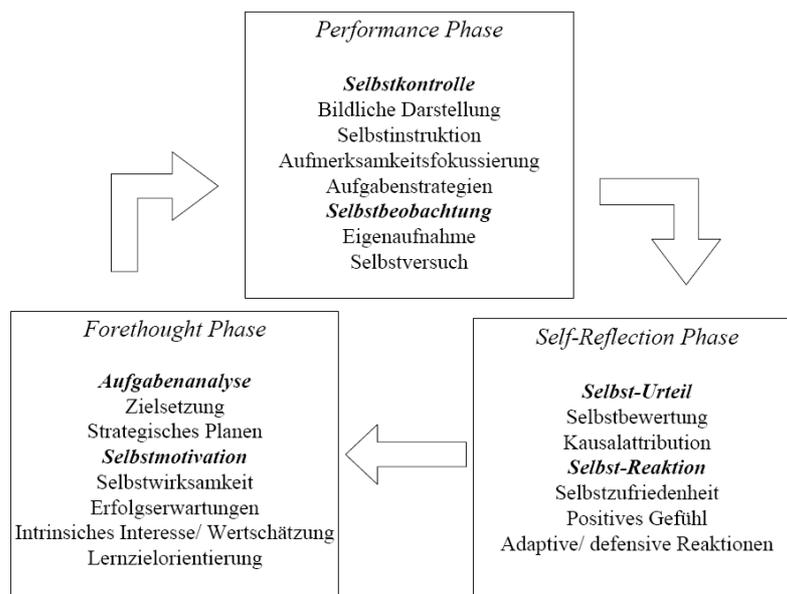


Abbildung 9: Phasen und Subprozesse der Selbstregulation, adaptiert von Zimmerman (2002, S. 67).

Wie in Abbildung 9 ersichtlich, unterscheidet Zimmerman (2002) die *Forethought Phase* vor dem Lernen, die *Performance Phase* während des Lernens sowie die *Self-Reflection Phase*, welche nach dem Lernen einsetzt.

Vor Beginn des Lernens analysiert der Lerner zunächst die Aufgabe, indem er sich Ziele setzt und die Erreichung der Ziele strategisch plant. Hierbei spielen selbstmotivierende Gedanken eine wichtige Rolle für den Lernerfolg. Durch eine hohe Selbstwirksamkeits- und Erfolgserwartung sowie einer Lernzielorientierung und intrinsischem Interesse sind die Schüler motiviert, selbstregulatorisch zu handeln.

Während der Ausführung des Lernens, in der *Performance Phase*, zeigt der Lerner ein Set an Verhaltensmerkmalen, die in zwei Kategorien eingeordnet werden können: Selbstkontrolle und Selbstbeobachtung. Die Selbstkontrolle dient dabei dem effektiven Einsatz der in der *Forethought Phase* ausgewählten Strategien. Dazu zählen Verhaltensweisen wie z.B. bildliche Darstellung, Selbstinstruktion und Aufmerksamkeitsfokussierung. Des Weiteren findet in der *Performance Phase* eine Selbstbeobachtung statt, welche es dem selbstregulierten Lerner ermöglicht, Störfaktoren zu identifizieren. Eine Form der Selbstbeobachtung ist z.B. der Selbstversuch, bei dem der Lerner die Aufgabe unter Ausblendung potenzieller Störfaktoren bearbeitet. Eine weitere Möglichkeit ist die Aufnahme des Lernverhaltens, indem z.B. die Zeit für das Lösen einer Aufgabe unter verschiedenen Bedingungen aufgezeichnet wird.

Im Anschluss an die *Performance Phase* erfolgt die *Self-Reflection Phase*, welche in zwei Subprozesse untergliedert wird: Selbst-Urteil und Selbst-Reaktion. Zunächst bildet sich der Lerner ein Urteil über das Lernergebnis. Dies geschieht u.a. durch Selbstbewertung, was einen Vergleich mit vorherigen Leistungen, einem Leistungsstandard und den in der *Forethought Phase* festgelegten Zielen impliziert. Eine internale oder externale Ursachenzuschreibung für Erfolg bzw. Misserfolg beeinflusst zudem zukünftige Lernprozesse. Wenn die Ursache für Misserfolg auf eine geringe Lernfähigkeit zurückgeführt wird, verringert dies die Anstrengungsbereitschaft, während bei einer externalen Attribuierung, z.B. ungünstige Lernumgebung, die Bereitschaft nicht reduziert wird. Abgesehen von der Urteilsbildung reagiert der Lerner überdies auf die erbrachte Leistung, was ebenfalls förderlich oder weniger förderlich im Sinne der Selbstregulation ausfallen kann. Selbstregulierte Lerner sind mit dem Ergebnis zufrieden und zeigen adaptive anstatt defensive Verhaltensmuster, um die eingesetzten Strategien zu verbessern.

Gemäß Zimmerman (2002) verlaufen die drei beschriebenen Phasen zyklisch in dem Sinne, dass die *Self-Reflection Phase* eine neue *Forethought Phase* initiieren kann, was

wiederum zu einer neuen *Performance Phase* und *Self-Reflection Phase* führt. Empirisch konnte die zyklische Modellannahme bestätigt werden, da sich hohe positive Korrelationen zwischen Selbst-Reaktion (d.h. Zufriedenheit mit dem Lernergebnis) und Selbstwirksamkeit (d.h. eigene Überzeugung, die Fertigkeit ausführen zu können) sowohl bei dem Erlernen einer motorischen Fertigkeit (Zimmerman & Kitsantas, 1997) als auch beim Schreiben (Zimmerman & Kitsantas, 1999) zeigten.

2.2.2 Förderung der Selbstregulation

Selbstregulatorische Fähigkeiten stellen eine Schlüsselkompetenz für akademischen Erfolg dar, so dass eine Förderung der Selbstregulation ein zentraler Bestandteil im Unterricht sein sollte. Da jedoch bereits die Definition des Begriffs Selbstregulation aufgrund seiner Vielschichtigkeit schwierig war, sind Förderkonzepte hierzu ebenfalls eine Herausforderung.

Selbstregulatorische Prozesse steigen mit zunehmendem Alter an, so dass bei der Förderung der Selbstregulation Entwicklungsprozesse berücksichtigt werden müssen (Graham & Harris, 2000; Paris & Paris, 2001). Allgemein kann Selbstregulation auf drei Wegen erhöht werden: (1) indirekt durch Reifung; (2) direkt durch Instruktionen und (3) durch Übung und adäquates Feedback (Paris & Paris, 2001). Selbstregulation sollte dabei nicht nur das Ziel, sondern ebenfalls der Weg sein (Brunstein & Spörer, 2006; Paris & Paris, 2001). Besonders vielversprechende Instruktionmethoden stellen dabei die Prinzipien der Beobachtung und der Nachahmung dar. Sowohl im Bereich Sport als auch beim Schreiben übernahmen Schüler erfolgreich die von den Lehrern modellierten selbstregulatorischen Techniken (Kitsantas, Zimmerman, & Cleary, 2000; Zimmerman & Kitsantas, 2002).

Es lassen sich drei Ziele für die Instruktion von selbstregulatorischen Techniken identifizieren (Cleary & Zimmerman, 2004): (1) Erhöhung der Selbstständigkeit der Schüler, Strategiedefizite zu entdecken und adäquate Strategien zu implementieren; (2) Erweiterung des Repertoires an Lernstrategien; (3) Befähigung der Schüler selbstregulatorische Strategien zyklisch anzuwenden. Auf Grundlage seines Phasenmodells der Selbstregulation (siehe Abbildung 9) entwickelte Zimmerman (1998) ein zyklisches Trainingsmodell der Selbstregulation, welches die von Cleary und Zimmerman (2004) postulierten Ziele verfolgt. In Abbildung 10 ist das Modell illustriert.

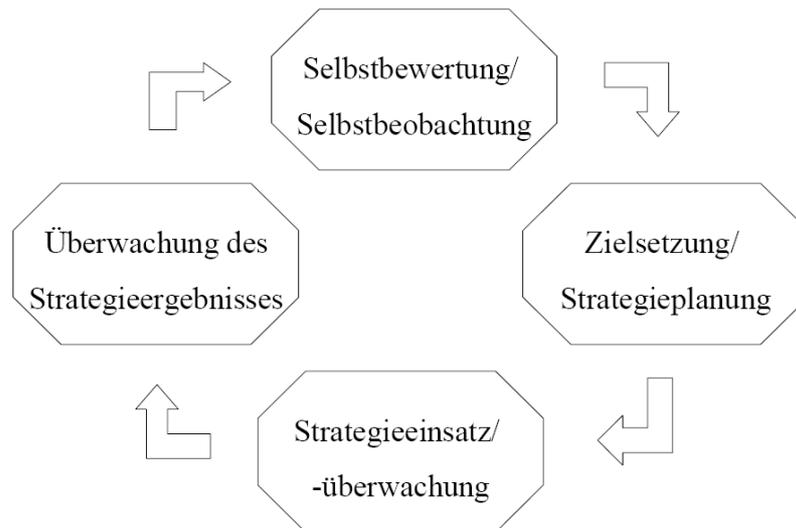


Abbildung 10: Zyklisches Trainingsmodell der Selbstregulation, adaptiert von Zimmerman (1998, S. 83).

Das zyklische Trainingsmodell von Zimmerman (1998) besteht aus vier Komponenten, deren inhaltliche Umsetzung auf die jeweilige zu fördernde Fertigkeit abgestimmt ist. Jeder Schritt stellt einen Ausgangspunkt für den nächsten Schritt dar und beeinflusst somit dessen Inhalt und Umfang. Erst das erfolgreiche Zusammenspiel der vier Komponenten macht selbstreguliertes Lernen möglich. Im Einzelnen handelt es sich um folgende Komponenten:

- *Selbstbewertung und Überwachung:* Als erstes wird die Effektivität der gegenwärtigen Lernstrategien bewertet. Lerndefizite werden festgestellt.
- *Zielsetzung und Strategieplanung:* Um die ermittelten Defizite zu beseitigen, werden Ziele gesetzt und Strategien ausgewählt, um die neuen Ziele zu erreichen. Hierbei ist die Verfügbarkeit von ausreichend vielen Lernstrategien sowie von sozialer Unterstützung zur Erklärung neuer Strategien entscheidend, damit der Lerner die für ihn und seine Ziele adäquatesten Strategien auswählen kann.
- *Strategieeinsatz und Überwachung:* Im weiteren Verlauf setzt der Lerner die ausgewählte Strategie ein und überwacht die Genauigkeit der Strategieimplementierung. Novizen benötigen bei diesem Schritt Hilfe, um eine neu erlernte Strategie korrekt einzusetzen.
- *Überwachung des Strategieergebnisses:* Am Ende des Lernzyklus bewertet ein selbstregulierter Lerner die Effektivität des Strategieeinsatzes, indem dieser selbst überprüft, ob das gewünschte Lernergebnis eingetreten ist.

Das von Zimmerman (1998) beschriebene Trainingsmodell zur Förderung der Schreibkompetenz ist zyklisch, da ein neuer Lernzyklus startet, sobald der Lerner mit dem erreichten Ergebnis nicht zufrieden ist und erneut damit beginnt, durch Selbstbewertung und

Selbstbeobachtung die Gründe für das Lernergebnis zu ermitteln. Der Kreislauf wird so lange durchgeführt bis eine erfolgreiche Lernstrategie identifiziert und erfolgreich umgesetzt wurde. Nach Erreichung eines Ziels setzt sich ein selbstgesteuerter Schreiber neue und anspruchsvollere Ziele.

Da Selbstregulation sowohl Ziel als auch Weg erfolgreichen Lernens darstellt, erfolgt die Förderung der Selbstregulation immer in Kombination mit fachspezifischen Fertigkeiten. Die Förderung der Selbstregulation zeigte sich bereits erfolgreich in Kombination mit einem Training zur Förderung der Problemlösekompetenz (Perels, Gürtler, & Schmitz, 2005), in Kombination mit einem Elterntaining zur Förderung der Hausaufgabenbetreuung (Bruder, Perels, & Schmitz, 2004) und im naturwissenschaftlichen Unterricht (Labuhn, Bögeholz, & Hasselhorn, 2008). Auch der sehr erfolgreiche *Reciprocal Teaching*-Ansatz von Palincsar und Brown (1984) greift auf Konzepte der Selbstregulation wie Selbstüberwachung zurück. In Bezug zum Schreiben führte ein selbstregulatorisches Training zu größeren Zuwächsen in den Planungsaktivitäten und der Textqualität als traditioneller Unterricht (Torrance, Fidalgo, & García, 2007). Das vorgestellte Trainingsmodell der Selbstregulation (Zimmerman, 1998) setzten Glaser und Kollegen explizit zur Förderung der Schreibkompetenz ein und konnten dabei ebenfalls die Effektivität des Ansatzes nachweisen (Glaser, 2005; Glaser & Brunstein, 2007b; Glaser et al., 2009).

Übersicht zu den Befunden: Hattie, Biggs und Purdie (1996) und Dignath, Büttner und Langfeldt (2008) führten jeweils eine Metaanalyse durch, in der sie den Effekt von Strategieinstruktion auf die Leistung untersuchten. In der Metaanalyse von Hattie et al. (1996) wurden insgesamt 270 Effektstärken aus 51 Studien untersucht. Dignath et al. (2008) führten eine Metaanalyse mit 263 Effektstärken aus 48 Studien durch, wobei sie speziell selbstregulatorisches Lernen in der Grundschule fokussierten. Die Ergebnisse sind der nachfolgenden Tabelle 3 zu entnehmen.

Insgesamt zeigte sich ein moderater bis großer Effekt von Strategieinstruktion. Die Effektstärken der Metaanalyse von Dignath et al. (2008), welche nur selbstregulatorische Strategien untersuchte, fielen dabei höher aus.

Tabelle 3: Effektstärken zur Beurteilung des Einflusses von Strategieinstruktion auf Leistung, Lernstrategien und Affekt.

	Leistung	Lernstrategien	Affekt/ Motivation	Insgesamt
Hattie et al. (1996)	0.57	0.16	0.48	0.45
Dignath et al. (2008)	0.62	0.73	0.76	0.69

Darüber hinaus zeigte sich in der Metaanalyse von Hattie et al. (1996), dass jüngere Schüler mit durchschnittlichen schulischen Leistungen sowie Underachiever, d.h. Schüler deren Leistungen erheblich von dem Niveau abweichen, das nach ihrer Intelligenz zu erwarten wäre, am meisten von Förderprogrammen zum strategischen Denken und Lernen profitierten. Auch Brunstein und Spörer (2006) weisen auf den potenziellen Erfolg selbstregulatorischer Trainings für Underachiever hin. Dignath et al. (2008) konnte nur für die Strategieranwendung einen Vorteil von jüngeren gegenüber älteren Schülern feststellen.

Des Weiteren identifizierten Brunstein und Spörer (2006) folgende zentrale Instruktionsprinzipien für die Förderung strategischen Lernens:

- Es sollten nicht nur effektive kognitive Strategien, sondern zusätzlich motivationale und metakognitive Strategien vermittelt werden, wobei dabei ein spezifisches Aufgabenfeld mit einer geringen Anzahl an Strategien abgebildet werden sollte.
- Der Einsatz der Strategien sollte explizit begründet und modelliert werden, so dass intensive Lehrer-Schüler-Interaktionen sowie Schüler-Schüler-Interaktionen den Aufbau selbstregulatorischer Kompetenzen erleichtern.
- Die neuen Strategien sollten weiterhin intensiv geübt werden. Ferner sollten die Schüler über die eigenen Fortschritte in der Anwendung der Strategien durch kontinuierliches Feedback unterrichtet werden. Die Bereitstellung von anregendem Material erhöht dabei die Lernbereitschaft und Motivation der Schüler.

Da die Förderung selbstregulatorischer Kompetenzen in Kombination mit einer spezifischen Fertigkeit erfolgen sollte, wird im nächsten Abschnitt das Konzept des selbstregulierten Schreibens genauer betrachtet.

2.2.3 Selbstreguliertes Schreiben

Aufgrund der komplexen Anforderungen, die die Textproduktion an einen Schreiber stellt, sind kognitive und metakognitive Prozesse, wie sie in den diskutierten Schreibmodellen fokussiert werden, zur Erklärung der Schreibkompetenz nicht ausreichend. Je intensiver der Schreibprozess analysiert wird und die sozialen Einflüsse auf das Schreiben identifiziert werden, desto komplexer wird das Bild zur Erklärung der am Schreiben beteiligten Prozesse (Hull & Rose, 1989). Um eine Schreibaufgabe bewältigen zu können, muss der Schreiber nicht nur über genügend kognitive Schreibstrategien verfügen und diese regulieren können, sondern vielmehr auch entscheiden, wann er schreiben möchte, welche Informationen er einholt und wie er seine Lernumwelt strukturiert. Folglich benötigt ein erfolgreicher Schreiber

selbstregulatorische Kompetenz, die nach Brunstein und Spörer (2006) kognitive, metakognitive und motivationale Strategien umfasst.

Zimmerman leistete einen wesentlichen Beitrag für das Verständnis selbstregulierten Schreibens. Er untersuchte das Vorgehen von Experten aus den Bereichen Schreiben, Sport, Musik sowie bei erfolgreichen Studenten. Durch Interviews mit Schreibexperten identifizierte er für das Schreiben folgende selbstregulatorische Prozesse (siehe Tabelle 4), die er hinsichtlich der Umwelt, des Verhaltens und des Selbst gliederte (Zimmerman, 1998; Zimmerman & Risemberg, 1997):

Tabelle 4: Selbstregulatorische Prozesse erfolgreicher Schriftsteller, adaptiert von Zimmerman und Risemberg (1997, S. 79).

Selbstregulatorischer Prozess	Verhaltensmerkmale
Prozesse bezogen auf die <i>Umwelt</i>	
Strukturierung der Umwelt	Kontrolle der Rahmenbedingungen durch Auswahl und Herstellung effektiver Schreibumgebungen
Hilfesuche	Heranziehen von Feedback, Suche nach sozialen Quellen für schreibbezogenes Wissen
Prozesse bezogen auf das eigene <i>Verhalten</i>	
Selbstüberwachung	Protokollierung der täglichen Schreibergebnisse
Selbstbelohnung	Hinauszögern von angenehmen Aktivitäten bis die Schreibaufgabe erledigt ist
Selbstinstruktion	Verbalisierung des zu schreibenden Textes
Prozesse bezogen auf das <i>Selbst</i> (<i>versteckte Prozesse</i>)	
Zeitmanagement	Festsetzen täglicher Schreibphasen, z.B. am Morgen
Zielsetzung	tägliche Zielsetzung in Form von Wörtern oder Seitenzahlen
Selbstbewertung	Formulierung von Kriterien zur Bewertung der eigenen Schreibleistung
Kognitive Aufgabenstrategien	Planungs- und Revisionsstrategien, z.B. das Produzieren von Stichwörtern und das Erstellen eines roten Fadens
Bildliche Vorstellung	Bildliche Vorstellung der Handlung, der Charaktere und des Handlungsortes

Insgesamt ist in dem Modell von Zimmerman und Risemberg (1997) die Kontrolle des Schreibprozesses auf Einflüsse der Umwelt, des Verhaltens und des Selbst zurückzuführen.

Umweltprozesse beziehen sich dabei auf alle außerhalb des Schreibers befindlichen sozialen und physischen Einflüsse, die offen und beobachtbar sind. Ein erfolgreicher Schreiber reguliert die Gestaltung und Optimierung seiner Lernumgebung selbst, indem er das

Wann?, *Wo?* und *Mit/von wem?* bestimmt und ausgehend von den verfügbaren Ressourcen sich die effektivsten auswählt (Brunstein & Spörer, 2006).

Verhaltensprozesse sind ebenfalls offen und beobachtbar und werden motorisch ausgeführt. Zimmerman und Risemberg (1997) zählen dazu Methoden der Selbstüberwachung, der Selbstbelohnung und der Selbstinstruktion. In Abhängigkeit von dem Erfolg der Methoden werden diese adaptiert oder neue werden ausgewählt und eingesetzt. Gemäß Brunstein und Spörer (2006) fallen in diesen Bereich Strategien, die kontrollieren, *was* gelernt werden soll. Der gesamte Lernprozess wird dabei beobachtet. Die erbrachten Leistungen werden hinsichtlich der gesetzten Ziele bewertet und der eigene Lernfortschritt wird stetig überwacht.

Im Gegensatz dazu sind Prozesse des Selbst oder der Person nicht beobachtbar. Zimmerman und Risemberg (1997) fassen darunter kognitive und affektive Strategien. Mit kognitiven Strategien ist die Auswahl geeigneter Lernstrategien, d.h. die Frage nach dem *Wie?* gemeint (Brunstein & Spörer, 2006). Erfolgreiche Lerner gehen planvoll bei der Bearbeitung einer Schreibaufgabe vor und können aus einem großen Repertoire an Methoden die jeweils effektivste auswählen. Folglich setzen sie sich anspruchsvolle, aber realistische Ziele und fragen sich nach dem *Warum?* ihres Lernens (Brunstein & Spörer, 2006). Aufgrund der persönlichen Relevanz der Ziele gehen sie selbst-motiviert an die Erreichung der Ziele heran.

In dem Modell von Zimmerman und Risemberg (1997) beeinflussen sich die drei Prozesskategorien Umwelt, Verhalten und Selbst gegenseitig in Form von Feedbackschleifen. Während des gesamten Schreibprozesses überwacht, reguliert und verändert der Schreiber den Erfolg der eingesetzten Strategien. Aufgrund des Zusammenspiels zwischen Umwelt, Verhalten und Selbst ist der Schreiber sensitiv für die Ursachen ineffektiven Lernens und kann adaptiv auf Schwierigkeiten in den drei Prozesskategorien reagieren. Gleichzeitig verändern die Feedbackschleifen die Selbstwirksamkeitserwartung des Schreibers. Bei andauerndem negativem Feedback aufgrund von schlechten Ergebnissen beginnt der Schreiber, an seinen eigenen Schreibfähigkeiten zu zweifeln, so dass seine Selbstwirksamkeitserwartung reduziert wird. Selbstwirksamkeit beim Schreiben umfasst einerseits den Glauben an sich selbst und seine Fähigkeiten zur Regulation der eigenen Schreibaktivität und andererseits die Befähigung zur Bewältigung einer vorliegenden Schreibaufgabe (Zimmerman & Bandura, 1994). Demzufolge ist die Selbstwirksamkeit in dem Modell von Zimmerman und Risemberg (1997) eine entscheidende Einflussgröße, da diese die Motivation, die einzelnen Prozesse auf den Ebenen der Umwelt, des Verhaltens und

des Selbst auszuführen, steuert. Eine verstärkte Anwendung von selbstregulatorischen Techniken führt wiederum zu einer höheren Selbstwirksamkeitserwartung, so dass die schreibbezogene Selbstwirksamkeitserwartung wiederum den Schreiberfolg vorhersagen kann (Zimmerman & Risemberg, 1997).

Die Rolle der Selbstwirksamkeit als Prädiktor und Mediator für akademischen Erfolg in unterschiedlichen Fächern konnte in Metaanalysen nachgewiesen werden (Stajkovic & Luthans, 1998). Obwohl in der Forschung die Selbstwirksamkeit bei der Betrachtung von Schreibprozessen zunächst vernachlässigt wurde, sind mittlerweile hinreichende Belege dafür vorhanden, dass eine schreibbezogene Selbstwirksamkeitserwartung einen unabhängigen Beitrag zur Vorhersage der Schreibkompetenz leistet (Pajares, 2003). Die Einflussnahme auf die Kompetenz ist jedoch vielschichtig. Zunächst reduziert eine hohe Selbstwirksamkeitserwartung die Angst vor dem Schreiben (Pajares, 2003). Gleichzeitig steigt das eigene Vertrauen in die Schreibfertigkeiten, was durch reguläres Feedback und durch Prozessziele von Lehrern unterstützt werden kann (Schunk & Schwartz, 1993). Insgesamt ist es laut Pajares (2003) zwingend notwendig, dass Lehrer die schreibbezogene Selbstwirksamkeit von Schülern durch Lob und Ermunterung steigern, da der Glaube an die eigene Schreibkompetenz für den Lernerfolg zwingend erforderlich ist. Dies ist insbesondere bei Schreibanfängern notwendig, da diese aufgrund von mangelnden Erfahrungen, Wissen und Feedback noch keine oder wenig Möglichkeiten zur Bildung einer adäquaten Selbstwirksamkeitserwartung hatten.

Abschließend ist festzustellen, dass erfolgreiche Schriftsteller eine Vielzahl an selbstregulatorischen Prozessen einsetzen (Zimmerman & Risemberg, 1997). Mit Ausblick auf die Konzeption eines Trainings zur Förderung der Schreibkompetenz ist es als nächstes erforderlich zu erörtern, welche Komponenten für die Effektivität eines Trainingsprogramms zwingend notwendig sind.

2.2.3.1 Wissenschaftlich fundierte Empfehlungen für den Schreibunterricht

Als Ausgangslage für die Identifikation von notwendigen selbstregulatorischen Techniken werden wissenschaftlich fundierte Empfehlungen für den Schreibunterricht in den Klassen vier bis zwölf herangezogen, die sowohl für durchschnittliche als auch für lernbehinderte Schüler gelten (Graham & Harris, 2009; Graham & Perin, 2007b; Graham et al., 2009).

Die Empfehlungen basieren größtenteils auf der Metaanalyse *Writing Next* (Graham & Perin, 2007b), welche explizit das Ziel verfolgte, als Leitfaden für die Unterrichtsgestaltung zu fungieren und dadurch den Schreibunterricht in den USA deutlich zu verbessern. Die dargestellten Empfehlungen wurden jedoch durch die Hinzuziehung weiterer Metaanalysen

optimiert, so dass die hier präsentierten wissenschaftlich fundierten Ergebnisse auf experimentellen und qualitativen Studien sowie Einzelfalluntersuchungen aus dem Bereich Schreiben basieren (Graham & Harris, 2009; Graham et al., 2009). Auch wenn die Höhe der berichteten Effektstärken variiert und die positiven Effekte auf die Schreibleistung nicht in jeder Metaanalyse nachgewiesen wurden, sind sie hilfreiche Ansatzpunkte für die Konzeption eines Trainingsprogramms. In der nachfolgenden Tabelle 5 werden die 13 Empfehlungen vorgestellt, wobei diese nach der Höhe der berichteten Effektstärke sortiert sind.

Tabelle 5: Wissenschaftlich fundierte Empfehlungen für den Aufsatzunterricht.

Empfehlungen aus <i>Writing Next</i> (Graham & Harris, 2009; Graham & Perin, 2007b; Graham et al., 2009)			
1	Instruktion von Planungs-, Überarbeitungs- und Editierungsstrategien	ES ⁴ = .82	<ul style="list-style-type: none"> ▪ Lehrer modelliert die Anwendung der Strategien. Anschließend üben die Schüler die neuen Strategien. Sie werden dabei vom Lehrer unterstützt.
2	Explizite Vermittlung von Zusammenfassungsstrategien	ES= .82	<ul style="list-style-type: none"> ▪ Effekt ist genrespezifisch. ▪ Es wurde nicht die Schreibkompetenz allgemein verbessert, sondern die Fertigkeit, eine eigene Zusammenfassung zu schreiben. ▪ Durch das Training wussten die Schüler besser, was eine gute Zusammenfassung ausmacht.
3	Schaffung einer Lernumgebung, in der Schüler kollaborativ zusammenarbeiten	ES= .75	<ul style="list-style-type: none"> ▪ Mehrere Schüler führten gemeinsam einen der drei Schreibprozesse (Planen, Schreiben, Überarbeiten) im Rahmen eines vorstrukturierten Prozesses gemeinsam aus. ▪ Peers stellen eine wichtige Referenzgruppe dar, anhand derer ein Schüler seine eigene Leistung durch Feedback beurteilen kann. ▪ Feedback ist vor allem bei jüngeren Schülern notwendig, um Kriterien zur Bewertung der eigenen Schreibleistung zu entwickeln.
4	Klare und spezifische Zielsetzung für das zu schreibende Textprodukt	ES= .70	<ul style="list-style-type: none"> ▪ Dazu zählen die Identifikation der Schreibaufgabe und die Festlegung der gewünschten Eigenschaften des zu schreibenden Textes.

⁴ ES = Effektstärke, berechnet nach Cohen's *d*.

5	Schaffung von Möglichkeiten zur Nutzung von Wortverarbeitungsprogrammen zur Produktion von Texten	ES= .55	<ul style="list-style-type: none"> ▪ Insbesondere bei Schülern mit Lernbehinderungen kann der Einsatz von Wortverarbeitungsprogrammen die Textqualität und die Lesbarkeit verbessern sowie die Anzahl der Rechtschreibfehler reduzieren. ▪ Es fehlen jedoch Belege dafür, dass durchschnittliche Schüler ebenfalls von einem Wortverarbeitungsprogramm profitieren.
6	Unterstützung der Schüler beim Schreiben komplexer Sätze	ES= .50	<ul style="list-style-type: none"> ▪ Effektive Strategien zur Formulierung adäquater Sätze entlasten das Arbeitsgedächtnis.
7	Anwendung des Schreibprozessansatzes	ES= .32	<ul style="list-style-type: none"> ▪ Der Schreibprozess besteht hierbei aus Ideengenerierung, Planen, Skizzieren, Überarbeiten und Veröffentlichen. ▪ Der Ansatz bietet den Schülern eine Vielzahl an Methoden, Schreibhilfen und Tutoren⁵, so dass die Lernumwelt optimal und hinsichtlich der eigenen Bedürfnisse gestaltet werden kann. ▪ Siehe Kapitel 2.3.1.1.
8	Förderung von Aktivitäten, die die Fertigkeit, Fragen zu stellen, erhöhen	ES= .32	<ul style="list-style-type: none"> ▪ Empfehlung trifft nur auf ältere Schüler der Klassen sieben bis zwölf zu. ▪ Das Fragenstellen führte zu einem erhöhten Bewusstsein für die Schreibaufgabe und die Bedürfnisse der Leserschaft.
9	Unterstützung der Schüler bei schreibvorbereitenden Aktivitäten wie das Generieren und Organisieren von Ideen	ES= .32	<ul style="list-style-type: none"> ▪ Darunter werden die Generierung von Ideen aus dem Langzeitgedächtnis, die Organisation der abgerufenen Informationen und die Formulierung von Kriterien zur Evaluation des geschriebenen Textes (Zielsetzung) gefasst.
10	Bereitstellung von guten Textbeispielen für jede Textsorte	ES= .25	<ul style="list-style-type: none"> ▪ Beispiele für besonders gelungene Texte versorgen den Schreiber mit Kriterien zur Bewertung der eigenen Schreibleistung.
11	Vermeidung von traditionellen Methoden des Grammatikunterrichts und der Satzlehre	ES= -.32	<ul style="list-style-type: none"> ▪ Traditionelle Methoden des Grammatikunterrichts wie z.B. die Analyse der einzelnen Satzteile wirkten sich ungünstig für die Textqualität aus.

⁵ Die männliche Form „Tutor“ schließt Tutorinnen mit ein.

Zusätzliche Empfehlungen (Graham & Harris, 2009; Graham et al., 2009)			
12	Anleiten der Schüler, Schreibleistung und Verhalten zu überwachen	PND ⁶ = 60%	<ul style="list-style-type: none"> ▪ Vor allem schreibschwache Schüler profitierten von dem Einsatz selbstüberwachender Methoden.
13	Bereitstellung von genügend Schreibzeit	-	<ul style="list-style-type: none"> ▪ Gute Lehrer und Schulen geben ihren Schülern viele verschiedene Schreibmöglichkeiten. ▪ Schreiben wird als ein integraler Bestandteil jeglichen Unterrichts ansehen.

Die höchsten Effektstärken von jeweils .82 konnten für die Instruktion von Planungs- und Überarbeitungsfertigkeiten sowie für die Vermittlung von Zusammenfassungsstrategien ermittelt werden (Graham & Harris, 2009; Graham & Perin, 2007b; Graham et al., 2009). Auch ein Trainingsprogramm zur Verbesserung der Fertigkeit, komplexe Sätze zu konstruieren, welches z.B. von Saddler und Graham (2005) entwickelt wurde, zeigten einen moderaten Effekt auf die Schreibleistung (Graham & Harris, 2009; Graham & Perin, 2007b; Graham et al., 2009). Folglich ist die selbstregulatorische Technik der *kognitiven Aufgabenstrategien* eine effektive Maßnahme zur Verbesserung der Textqualität (Zimmerman & Martinez-Pons, 1988; Zimmerman & Risemberg, 1997). Es zeigte sich überdies, dass der Einsatz von grafischen Planern zur Strukturierung von Planungsnotizen das Arbeitsgedächtnis entlastet und dadurch die Schreibleistung positiv beeinflusst (Englert, Okolo, & Mariage, 2009; Kellogg, 2008).

Darüber hinaus implizieren die meisten der dargestellten Empfehlungen selbstregulatorische Methoden. Zum Beispiel umfasst die dritte Empfehlung, kollaborative Lernumgebungen, und die zehnte Empfehlung, Textbeispiele, die Möglichkeit, die Methode der *Selbstbewertung* zu erlernen. Weiter bezieht sich die vierte Empfehlung explizit auf die Zielsetzung, welche ebenfalls eine selbstregulatorische Technik darstellt, deren positiver Effekt auf die Verbesserung der Schreibleistung sogar bei Schülern mit Lernbehinderungen nachgewiesen werden konnte (Graham, MacArthur, & Schwartz, 1995).

Graham, Olinghouse und Harris (2009) erörtern weitere Empfehlungen, die auf Studien mit lernbehinderten Schülern basieren. Sie betonen insbesondere die Wichtigkeit

⁶ PND (percentage of non-overlapping data): Beim PND-Wert handelt es sich im Gegensatz zur Effektstärke um ein nichtparametrisches Maß, welches die Prozentzahl der Daten der Treatment-Gruppe, die über dem höchsten Wert des Ausgangsniveaus liegen, angibt. Nach Rogers und Graham (2008) ist ein PND größer als 90% ein großer Effekt, ein PND zwischen 70% und 90% ein moderater Effekt und ein PND zwischen 50% und 70% ein kleiner Effekt. PND-Werte kleiner als 50% zeigen keinen Effekt an. Der PND-Wert kann als Effektstärke-Maß für Einzelfalluntersuchungen herangezogen werden.

motivationaler Komponenten wie die enthusiastische Vermittlung der Strategien und die Bildung einer anregenden und motivierenden Umwelt sowie Überlegungen, die den Grad der Unterstützung und Hilfestellung berücksichtigen. Hinsichtlich positiver Verstärkung berichten sie einen kleinen positiven Effekt (PND= 56%) auf die Schreibleistung (Graham et al., 2009). Sie legen dabei besonderen Wert darauf, dass die Belohnung sich auf das Erreichen gesetzter Ziele beziehen sollte.

Weiter ist anzumerken, dass die Empfehlungen auf bereits eingesetzte Trainings zurückgreifen, die nicht immer den neusten Stand der Forschung widerspiegeln. Folglich sollten die Empfehlungen zwar bei der Konzeption eines neuen Trainingsprogramms berücksichtigt werden, dennoch stellen sie nicht ein optimales Schreibprogramm für sämtliche Schüler der Jahrgangsstufen vier bis zwölf dar. Ein Trainingsprogramm der Klasse vier sollte sich inhaltlich von einem Trainingsprogramm der Klasse zwölf unterscheiden und spezifisch auf die Bedürfnisse und Lernvoraussetzungen von Schreibanfängern eingehen. Darüber hinaus handelt es sich bei den dargestellten Empfehlungen um eine Vielzahl an Strategien, die nicht alle in einem Programm integriert werden können.

Schließlich wurden in den Metaanalysen, auf denen die vorgenannten Empfehlungen basieren, Studien vermischt, in denen sowohl Lehrer als auch Trainer die jeweilige Intervention umsetzten. Die Stabilität und Generalisierbarkeit eines Trainings variiert jedoch je nachdem, wie und von wem die Trainingsmethode implementiert wurde, und sollte deswegen als Einflussgröße berücksichtigt werden (Graham & Harris, 2003).

2.2.3.2 Unterschiede zwischen Schreibnovizen und Schreibexperten

Schreibnovizen und Schreibexperten unterscheiden sich hinsichtlich mehrerer Aspekte voneinander. Graham und Harris (2003) identifizierten folgende Eigenschaften von Schreibnovizen:

- (1) *Schreibnovizen planen ihre Texte nicht.* Sogar nach expliziter Aufforderung zu planen, erstellten Schreibnovizen nur unzureichende Notizen in Form von ganzen Sätzen, die sie unreflektiert während des Schreibens übernahmen (Graham & Harris, 2000; Graham & Harris, 2003; Scardamalia et al., 1984). Im Gegensatz dazu planten Schreibexperten länger und produzierten qualitativ bessere Notizen als Schreibnovizen (Graham & Harris, 2003; Zimmerman & Risemberg, 1997). Darüber hinaus formulierten Schreibexperten mehrere Ziele sowie Zwischenziele und konnten flexibel auf die Ziele zugreifen (Harris & Graham, 2009).
- (2) *Schreibnovizen überarbeiten ihre Texte nicht.* Trotz ausreichendem Revisionswissen sowie Revisionsfertigkeiten konnten unerfahrene Schreiber den

Überarbeitungsprozess nicht ausführen (Fitzgerald, 1987). Die Vermittlung expliziter Revisionsstrategien zur Erhöhung der exekutiven Kontrolle während des Überarbeitens wirkte sich positiv auf die Anzahl und die Qualität der Revisionen aus (De La Paz, Swanson, & Graham, 1998; Graham, 1997). Ältere Schüler der Klassen fünf, sechs und acht erlernten die CDO-Strategie (*compare, diagnose, organize*), die es den Schülern ermöglichte, die vielfältigen Überarbeitungsprozesse zu koordinieren und zu regulieren.

Generell überarbeiten Schreibnovizen nur sehr wenig und sehr oberflächlich, indem sie die Lesbarkeit der Wörter erhöhen, Rechtschreibfehler verbessern und Punktierungen einfügen. Insgesamt betrafen weniger als 20% der Revisionen tatsächlich Verbesserungen des Textinhalts (Graham & Harris, 2003). Ferner berücksichtigen Schreibnovizen die potenzielle Leserschaft nicht (Graham, 1997). Im Gegensatz dazu überarbeiten Schreibexperten die Aussagen des Textes und verbessern ganze Sätze sowie Themen (Fitzgerald, 1987).

- (3) *Schreibnovizen besitzen wenig kognitives und metakognitives Strategiewissen.* Adäquate Strategien entlasten jedoch das Arbeitsgedächtnis, so dass die Vermittlung solcher Strategien vor allem bei Schreibanfängern erfolgversprechend ist (Kellogg, 2008).
- (4) *Der produzierte Text ist eine Aneinanderreihung von themenrelevanten Ideen ohne Zusammenhang.* Im Gegensatz zum Schreibexperten verwenden Schreibnovizen ausschließlich die *Knowledge Telling-Strategie*.
- (5) *Während der Schreibproduktion sind Schreibnovizen vornehmlich mit der Rechtschreibung, Grammatik und der Zeichensetzung beschäftigt und denken wenig über den Textinhalt nach.* Zudem sind die verfassten Texte sehr kurz. Während des Schreibens ist das Arbeitsgedächtnis aufgrund von noch nicht automatisierten graphomotorischen und orthografischen Prozessen überlastet. Die Schwierigkeiten mit der mechanischen Ausführung des Schreibens entmutigen Schreibnovizen schnell, so dass sie eher den Schreibprozess beenden und aufgeben (Graham, 1990). Generell zeichnet sich das Schreibprodukt von Schreibnovizen durch wenig Kohärenz, schlechtem Ausdruck, einen geringen Adressatenbezug und einen geringen Umfang aus (Graham, Harris, MacArthur, & Schwartz, 1991).
- (6) *Schreibnovizen zeigen erhebliche Schwierigkeiten sich kontinuierlich zu motivieren und überschätzen ihre eigene Schreibleistung.* Zu Beginn des Schreibenlernens sind die Schüler sehr von ihren eigenen Schreibfertigkeiten überzeugt (Graham &

Harris, 2003; Pajares, 2003). Mit fortschreitender Schulbildung entwickeln viele Schüler jedoch aufgrund von ausbleibenden Erfolgserlebnissen eine negative Einstellung zum Schreiben (Harris & Graham, 2009).

Einen weiteren Unterschied zwischen Experten und Novizen konnten Cleary und Zimmerman (2001) sowie Kitsantas und Zimmerman (2002) feststellen. In Untersuchungen mit Sportlern zeigte sich, dass Unterschiede zwischen Experten und Nicht-Experten auf einen unterschiedlichen Gebrauch an selbstregulatorischen Prinzipien während der *Forethought*, *Performance* und *Self-Reflection Phasen* zurückzuführen sind, wobei Erfahrungen (in Anzahl an Erfahrungen) sowie das bereichsspezifische Wissen keine Differenzierung ermöglichte (Cleary & Zimmerman, 2001; Kitsantas & Zimmerman, 2002). Folglich spielt das Wissen im Vergleich zu selbstregulatorischen Prozeduren eine untergeordnete Rolle.

In einer zusammenfassenden Darstellung der Ergebnisse zum Schreiben und zur Selbstregulation bündeln Graham und Harris (2000) die Erkenntnisse in vier wissenschaftlich untermauerten Thesen:

- (1) Erfahrene Schreiber sind selbstregulierter als unerfahrene Schreiber.
- (2) Schreibanfänger werden mit dem Alter und fortgeschrittener Beschulung selbstregulierter.
- (3) Individuelle Unterschiede hinsichtlich der Selbstregulation sagen individuelle Unterschiede im Schreiben voraus.
- (4) Die Vermittlung von selbstregulatorischen Strategien verbessert die Schreibkompetenz.

Auch wenn noch nicht für jede der vier Thesen ausreichend viele Studien als Beleg angeführt werden konnten, scheint das Erlernen von selbstregulatorischen Prozeduren den Übergang vom Schreibnovizen zu einem Schreibexperten zu erleichtern.

2.2.3.3 Selbstregulatorische Prozesse während des Planens, Schreibens und Überarbeitens

Aufgrund der Vielschichtigkeit und Komplexität des Konzepts der Selbstregulation ergibt sich eine große Menge an sinnvollen Instruktionmethoden. Daher sollen nun zusammenfassend die wichtigsten selbstregulatorischen Prozesse, die während der drei Schreibphasen des Planens, Schreibens und Überarbeitens notwendig sind, skizziert werden. Die Beschreibung der drei Phasen erfolgt in Anlehnung an das kognitive Schreibmodell von Hayes und Flower (1980a, 1980b).

Planungsprozesse: Drei der wissenschaftlich fundierten Empfehlungen beziehen sich explizit auf die Planungsphase, speziell auf die inhaltliche Planung des zu schreibenden Textes. Selbstregulatorische Techniken wie *Zielsetzung* und *kognitive Aufgabenstrategien* spielen dabei eine wichtige Rolle. Darüber hinaus erfordert der Planungsprozess eine Planung der Vorgehensweise (Hayes & Flower, 1980a; Hayes & Flower, 1980b). Hierbei sind selbstregulatorische Prozesse wie *Strukturierung der Lernumgebung* und *Zeitmanagement* für den Erfolg des Planens entscheidend. Des Weiteren muss der Planungsprozess überwacht werden (*Selbstüberwachung*), damit die generierten und organisierten Ideen während der Übersetzungs- und Überarbeitungsprozesse abgerufen werden können. Anknüpfend an das zyklische Selbstregulationsmodell von Zimmerman (2002) entspricht die Planungsphase der *Forethought Phase*, in der in Abhängigkeit von der gestellten Aufgabe Ziele geplant werden sowie sich selbst motiviert wird, um den gesamten Schreibprozess effizient durchführen zu können.

Übersetzungsprozesse: Eine zentrale Aufgabe des Übersetzungsprozesses ist die Transformation von Wissen in Sprache (Hayes & Flower, 1980a; Hayes & Flower, 1980b). Schwierigkeiten beim Übersetzungsprozess können durch den Einsatz selbstregulatorischer Techniken wie *Selbstüberwachung* und *Selbstinstruktion* bewältigt werden.

Des Weiteren spielt ein umfangreiches Repertoire an *kognitiven Aufgabenstrategien* eine wichtige Rolle bei der Ausführung des Schreibprozesses. Nachdem in der Planungsphase Ziele formuliert wurden, müssen während des Schreibens Strategien ausgewählt und eingesetzt werden, wobei selbstregulatorische Strategien der *Selbstkontrolle* und *Selbstbeobachtung* entscheidend sind (siehe *Performance Phase*).

Überarbeitungsprozesse: Der Überarbeitungsprozess unterteilt sich in einen Subprozess des analytischen Lesens, in dessen Verlauf der Text gescannt wird und einen Subprozess des Bearbeitens, bei dem der verfasste Text korrigiert wird (Hayes & Flower, 1980a; Hayes & Flower, 1980b). Während des gesamten Überarbeitungsprozesses reguliert ein kognitiver Monitor oder das Arbeitsgedächtnis die Ausführung, indem u.a. der produzierte Text mit den in der Planungsphase formulierten Zielen verglichen wird, was anhand der *Selbstbewertung* erfolgt. Schüler benötigen Standards, anhand derer sie ihre Revisionen ausführen können und die sie später internalisieren (Zimmerman & Risemberg, 1997).

Abgesehen davon ist die Reaktion auf das Schreibprodukt für die weiteren Schreibanstrengungen entscheidend (siehe *Self-Reflection Phase*). Eine erhöhte Zufriedenheit mit sich und dem Schreibprodukt geht mit einer erhöhten Selbstwirksamkeitserwartung

einher, was sich wiederum positiv auf die Anwendung selbstregulatorischer Techniken und die Schreibkompetenz auswirkt.

2.2.4 Zusammenfassung

Modelle der Selbstregulation gehen davon aus, dass Selbstregulation für den akademischen Erfolg zentral ist und eine Vielzahl an kognitiven, metakognitiven und motivationalen Strategien umfasst (Boekaerts, 1999; Brunstein & Spörer, 2006; Pintrich, 1999). Zimmerman (2002) erweiterte diese Modelle, indem er von einem zyklischen Charakter der beim Lernen ablaufenden selbstregulatorischen Prozesse ausging.

Auch für den Erwerb der Schreibkompetenz sind selbstregulatorische Prozesse wichtig, um die beim Schreiben beteiligten Prozesse zu regulieren und dadurch das Arbeitsgedächtnis zu entlasten. Die von Schriftstellern eingesetzten selbstregulatorischen Methoden umfassen die Regulation der Umwelt, des Verhaltens und des Selbst (Zimmerman, 1998; Zimmerman & Risemberg, 1997). Weiter führten metaanalytische Befunde zu der Formulierung von 13 wissenschaftlich fundierten Empfehlungen, die ebenfalls größtenteils selbstregulatorische Strategien implizierten. Auch der Vergleich von Schreibnovizen mit Schreibexperten führte zu der Schlussfolgerung, dass die Vermittlung von selbstregulatorischen Strategien die Entwicklung zum Schreibexperten erleichtert.

In dem kognitiven Schreibmodell von Hayes und Flower (1980a, 1980b) regulierten und steuerten Schreibexperten im Gegensatz zu Schreibnovizen die Schreibprozesse des Planens, Schreibens und Überarbeitens anhand eines kognitiven Monitors. In jüngeren Ansätzen wird vornehmlich von dem Arbeitsgedächtnis als die für die Koordination und Steuerung der Schreibprozesse verantwortliche Instanz gesprochen (Kellogg, 2008).

Insgesamt ergibt sich aus den dargestellten Modellen und Befunden die Konsequenz, dass ein Trainingsprogramm zur Förderung der Schreibkompetenz eine kombinierte Vermittlung von Schreibstrategien einerseits und selbstregulatorischen Strategien andererseits umfassen sollte. Zur Instruktion selbstregulatorischer Methoden sollte dabei auf das von Zimmerman (1998) entwickelte zyklische Trainingsmodell zurückgegriffen werden. Dieses umfasst die Schritte Selbstbewertung/Selbstbeobachtung, Zielsetzung/Strategieplanung, Strategieeinsatz/Überwachung und Überwachung des Strategieergebnisses. Neben der Instruktion effektiver, kognitiver, motivationaler und metakognitiver Strategien sollte ein erfolgreiches selbstregulatorisches Training zudem ausreichend Gelegenheiten für Übung, Interaktionen zwischen Lehrern und Schülern sowie Feedback geben.

2.3 Förderung der Schreibkompetenz

Auch wenn bereits hinreichend viele empirische Belege für die Ursachen geringer Schreibkompetenz vorliegen, fehlt bislang der Transfer theoretisch fundierter Modelle in die Praxis des Schreibunterrichts, was sich in dem schlechten Abschneiden amerikanischer Schüler bei national repräsentativen Erhebungen (NAEP) widerspiegelt (Graham & Perin, 2007a; Graham & Perin, 2007b; Persky et al., 2003; Rogers & Graham, 2008). Dort zeigten nur 28% der Viertklässler, 31% der Achtklässler und 24% der Zwölfklässler ein für das Alter angemessenes Kompetenzniveau (Persky et al., 2003). Auch wenn in Deutschland eine solche regelmäßige standardisierte Erfassung der Schreibleistungen fehlt, ist von ähnlichen Ergebnissen auszugehen.

Diese unbefriedigende Situation ist zum einen dadurch zu erklären, dass zunächst die Grundlagen für Schreibschwierigkeiten identifiziert werden müssen, aber zum anderen auch dadurch, dass experimentelle Interventionsstudien deutlich umfangreicher und für Störfaktoren anfälliger sind als Studien im Bereich der Grundlagenforschung. Dennoch haben sich bereits einige Programme zur Förderung der Schreibkompetenz herauskristallisiert, die sich sowohl historisch als auch inhaltlich in zwei Bereiche unterteilen lassen:

- (1) *Schreibprozessansatz*: Schüler lernen die Textproduktion durch die Beschäftigung mit authentischen und für sie persönlich wichtigen Schreibanlässen. Der Schüler sowie die Interaktion des Schülers mit dem Lehrer und den Mitschülern stehen im Vordergrund. Das Lernen erfolgt implizit mit wenig direkter Instruktion.
- (2) *Kognitive Strategieinstruktion beim Schreiben*: Im Gegensatz dazu erfolgt das Lernen beim zweiten Ansatz durch explizite Vermittlung von kognitiven Schreibstrategien. Anstatt der Schüler ist hier der Lehrer zentral, welcher durch Modellverhalten, Feedback und Scaffolding die Schüler bei der Anwendung der erlernten Strategien unterstützt. Als Beispiel werden Schreibprogramme, die die explizite Kombination von selbstregulatorischen und schreibbezogenen Strategien umfassen, vorgestellt.

2.3.1 Schreibprozessansatz zur Förderung der Schreibkompetenz

Als Reaktion auf die Schreibkrise in den 70er Jahren und der Kritik an den traditionellen Methoden des Schreibunterrichts, welche das Schreibprodukt als akademischen Erkenntnisgewinn akzentuierten, forderten Autoren der *National Assessment of Educational Progress* (NAEP) im Jahr 1983 die Schülerzentrierung und die gleichzeitige Zurücknahme

der Lehrerdominanz und des Frontalunterrichts (Hillocks, 1984). Dieser Perspektivwechsel basierte auf dem Schreibprozessansatz.

Der Kerngedanke des Schreibprozessansatzes beinhaltet, dass die Ausgangslage jeglicher Schreibintention eine Selbstdarstellung ist und auf eine bestimmte Leserschaft abzielt (Lipson, Mosenthal, Daniels, & Woodside-Jiron, 2000). Der traditionelle Fokus auf das Schreibprodukt und auf explizite Strategien zur Verbesserung der Formulierung eines Satzes oder eines Paragraphen verschob sich auf die verschiedenen am Schreiben beteiligten Subprozesse. Generell gehen Vertreter des Schreibprozessansatzes von folgenden Prozessen aus (Fletcher & Portalupi, 2008):

- (1) *Prewriting*: Darunter fallen Strategien der Ideengenerierung wie Brainstorming;
- (2) *Rough Draft*: Erstellung eines ersten Schreibentwurfs;
- (3) *Revise*: Überarbeitung des Textes, indem einzelne Sätze und Abschnitte umformuliert werden;
- (4) *Edit*: Korrektur von Grammatik- und Rechtschreibfehlern;
- (5) *Publish*: Veröffentlichung des Schreibprodukts.

Den Erkenntnissen der kognitiven Schreibforschung folgend, wird der Schreibprozess als eine rekursive Abfolge von verschiedenen Problemlöseaufgaben angesehen (Pritchard & Honeycutt, 2006). Während die ursprünglichen Verfechter des Schreibprozessansatzes jegliche explizite Strategievermittlung mit einer starken Lenkung durch den Lehrer ablehnten, herrscht heute generell der Konsens, dass einzelne kurze Instruktionen in Form von Minilektionen notwendig sind, um die Schreibkompetenz nachweislich zu verbessern.

Historisch ist der Schreibprozessansatz als sehr einflussreich zu bewerten. Mit seinem Beginn in den 70er Jahren, als sich mehrere Lehrer in San Francisco zusammantaten, um ihre Schreibprodukte auszutauschen, den damaligen Schreibunterricht mit dem Vorgehen professioneller Autoren verglichen und ihre neue Instruktionmethode *Process Approach* nannten, verbreitete sich der Schreibprozessansatz rasant in den amerikanischen Schulen, so dass dieser in den 80er Jahren als die wichtigste Unterrichtsmethode in den USA galt (Pritchard & Honeycutt, 2006). Auch heute wird der Schreibprozessansatz von vielen Lehrern in den USA eingesetzt. In einer Befragung der *National Assessment of Educational Progress* (NAEP) gaben sieben von zehn Lehrern an, einen Schreibprozessansatz in ihrem Unterricht zu verwenden, wobei jedoch die Definition und Umsetzung des Ansatzes stark variierten (Troia et al., 2009). Aus den von der NAEP festgestellten Defiziten in den

Schreibkompetenzen von Viert-, Acht- und Zwölftklässlern lassen sich jedoch Bedenken hinsichtlich der Effektivität des Schreibprozessansatzes ableiten (Persky et al., 2003).

2.3.1.1 Writer's Workshop

Als prominenteste Instruktionmethode des Schreibprozessansatzes entwickelte sich der *Writing Workshop* oder *Writer's Workshop*. Die Realisierung des *Writer's Workshops* ist stark von den persönlichen Erfahrungen und Einstellungen der Lehrer sowie den schulischen Rahmenbedingungen abhängig. Dennoch können folgende Schlüsseleigenschaften hervorgehoben werden (Fletcher & Portalupi, 2008; Lipson et al., 2000; Troia et al., 2009):

- (1) *Minilektionen*: Fünf- bis zehnmündige Vermittlung von Schreibstrategien, Rechtschreibung und Grammatik.
- (2) *Ausreichend Schreibzeit*: Einführung täglicher Schreibübungen für 20 bis 30 Minuten, wobei der Schreibanlass für die Schüler persönlich bedeutsam ist.
- (3) *Schreibkonferenzen*: Schüler- und lehrergeleitete Konferenzen über die verfassten Schreibpläne und Texte.
- (4) *Austausch mit anderen*: Entwicklung einer Schreibgemeinschaft in der Klasse und Sensibilisierung der Schüler für den Leser, indem die eigenen Texte vorgelesen und veröffentlicht werden.

Abgesehen davon implizieren die meisten *Writer's Workshops* individuell zugeschnittene Hilfestellungen durch den Lehrer, eine Förderung des Verantwortungsbewusstseins der Schüler für die eigenen Schreibprojekte und die Schaffung einer Lernumgebung, die die Kreativität der Schüler anregt (Fletcher & Portalupi, 2008; Gillet & Beverly, 2001; Johnson, 2004). Die Schüler sollen das Schreiben nicht nur als eine für den akademischen Erfolg relevante Kompetenz wahrnehmen, sondern zudem lernen, das Schreiben persönlich wertzuschätzen. Demnach intendiert der *Writer's Workshop* neben einer Verbesserung der Schreibleistung zusätzlich eine explizite Förderung der intrinsischen Motivation. Den Schülern soll der Schreibprozess Spaß machen, was durch ein positives Lehrermodell und durch die selbstständige Auswahl von Schreibanlässen verstärkt wird (Gillet & Beverly, 2001).

Effektivität des *Writer's Workshops*: Tatsächlich erfüllt der *Writer's Workshop* nur eingeschränkt die gewünschten Absichten. Im vorangegangenen Abschnitt wurden evidenzbasierte Prinzipien für den Schreibunterricht vorgestellt (siehe Tabelle 5). Eines der Prinzipien betrifft den Schreibprozessansatz. Es wird hierfür eine durchschnittliche Effektstärke von .32 angegeben, so dass generell davon ausgegangen werden kann, dass Instruktionmethoden wie der *Writer's Workshop* einen kleinen positiven Effekt auf die

Schreibleistung ausüben (Graham & Harris, 2009; Graham et al., 2009). Graham und Perin (2007a) berichten jedoch in ihrer Metaanalyse, dass der positive Effekt des Prozessansatzes auf die Schreibqualität durch professionelle Lehrerfortbildungen moderiert wird. In Studien, in denen keine professionellen Schulungen der Lehrer stattfanden, verschwand der positive Effekt. Nichtsdestotrotz erwiesen sich in einer Metaanalyse von Hillocks (1984) schülerorientierte Methoden als viermal so effektiv wie lehrerorientierte. Gleichzeitig kritisierte Hillocks (1984) den ursprünglichen *Writer's Workshop*, der jegliche Instruktionen durch den Lehrer ablehnte. Basierend auf den Ergebnissen seiner Metaanalyse schlussfolgerte er, dass ein Schreibtraining nur dann erfolgreich ist, wenn auch explizit Schreibstrategien instruiert werden und der Lehrer nicht nur unterstützend dabei ist.

Des Weiteren fehlen bislang fundierte wissenschaftliche Belege dafür, dass der *Writer's Workshop* seine explizite Intention, die Motivation der Schüler zu erhöhen, erfüllt (Pollington, Wilcox, & Morrison, 2001). In einer Studie von Graham, Harris und Mason (2005) zeigten Drittklässler, die an einem *Writer's Workshop* teilnahmen, im Gegensatz zu Drittklässlern, die alternativ ein kognitives Strategietraining durchliefen, innerhalb von fünf Monaten nur geringe Verbesserungen der Schreibkompetenz und keine Veränderungen der Motivation. Dabei ist jedoch zu berücksichtigen, dass das Strategietraining von geschulten Lerntrainern und im Gegensatz dazu der *Writer's Workshop* von regulären Klassenlehrern durchgeführt wurde, so dass nicht ausgeschlossen werden kann, dass die unterschiedlichen Effekte in dem Versuchsdesign begründet liegen.

Die Überprüfung der Effektivität des *Writer's Workshops* gestaltet sich schwierig, da aufgrund seiner rapiden Verbreitung in den Schulen der USA eine einheitliche wissenschaftlich überprüfbare Theorie als Grundlage für den *Writer's Workshop* fehlt (Lipson et al., 2000; Pritchard & Honeycutt, 2006). Folglich entwickelte sich eine Vielzahl an *Writer's Workshops*, die sich in Abhängigkeit von dem Wissen und den Erfahrungen der Lehrer unterscheiden, was einen Vergleich des *Writer's Workshops* mit anderen Unterrichtsmethoden erschwert. Aus diesem Grund sind die meisten Studien, die sich mit der Überprüfung des *Writer's Workshops* beschäftigten, Fallstudien. Eine häufig zitierte Fallstudie wurde von Lipson und Kollegen mit elf Lehrern der fünften Klasse durchgeführt (Lipson et al., 2000). Die Lehrereinstellungen zum Lernen und zum eigenen Unterricht standen dabei im Vordergrund. Es zeigte sich eine große Heterogenität bezüglich der im Unterricht realisierten *Writer's Workshops*. Auch wenn alle elf Lehrer angaben, einen Schreibprozessansatz in ihrem Unterricht einzusetzen, unterschieden sich diese maßgeblich im Grad der Unterrichtsstrukturierung, der Schülerzentrierung und in der Rolle des Lehrers. Da nur

Interviews mit den Lehrern durchgeführt wurden und keine Schülerdaten erhoben wurden, konnten keine Aussagen über die Schreibkompetenz der Schüler getroffen werden.

Im Gegensatz dazu überprüften Troia und Kollegen die Schreibkompetenz und die Motivation der Schüler in sechs vierten Klassen vor (Oktober 2002) und nach (Mai 2003) der Einführung eines *Writer's Workshops* (Troia, Lin, Monroe, & Cohen, 2009). Alle Lehrer waren von einer Schule, deren Kollegium kollektiv an einer umfangreichen Fortbildungsmaßnahme teilnahm, die u.a. den *Writer's Workshop* umfasste. Dementsprechend setzten alle Lehrer eine einheitliche Form mit festgelegten Inhalten eines *Writer's Workshops* um, wobei überdies Unterrichtsbeobachtungen stattfanden. Die Studie basiert auf einer kleinen Stichprobe von zehn überdurchschnittlichen Schülern, elf durchschnittlichen Schülern und zehn unterdurchschnittlichen Schülern. Es zeigten sich am Ende des Schuljahrs nur geringfügige Verbesserungen der Schreibkompetenz, wobei leistungsstarke Schüler deutlich mehr profitierten. Bei leistungsschwachen Schülern konnte kein signifikanter Anstieg der Leistungen festgestellt werden. In Bezug zur Motivation veränderten sich die schreibbezogenen Einstellungen und die Selbstwirksamkeitserwartungen nicht. Jedoch konnten signifikante, aber geringe Anstiege in den schreibbezogenen Zielorientierungen verzeichnet werden. Auch wenn die Studie erste Hinweise dafür liefert, dass ein *Writer's Workshop* nur zu minimalen Veränderungen der Schreibleistungen führte, können aufgrund einer fehlenden Kontrollgruppe keine kausalen Schlussfolgerungen gezogen werden.

Eine mögliche Erklärung für die geringen Verbesserungen der Schreibkompetenz könnte die fehlende Instruktion von selbstregulatorischen Fertigkeiten sein. Dies würde das schlechte Abschneiden schwacher Schüler begründen. Diese besitzen im Vergleich zu starken Schülern besonders große Schwierigkeiten, die während des Schreibens ablaufenden Prozesse zu kontrollieren und zu überwachen. Überdies weisen sie größere Probleme mit Orthographie, Rechtschreibung und Grammatik auf, so dass weniger Arbeitsgedächtniskapazitäten für die Generierung und Organisation des Textinhalts zur Verfügung stehen. Folglich vergrößerte sich der Leistungsunterschied zwischen starken und schwachen Schülern am Ende des Schuljahrs noch mehr (Troia et al., 2009).

Der dargestellte Schreibprozessansatz bzw. der *Writer's Workshop* beinhaltet einige für die Förderung der Schreibkompetenz wichtige Konzepte. Positive Entwicklungen in dem Schreibunterricht sind dem *Writer's Workshop* zu verdanken. Vor der Einführung des Schreibprozessansatzes wurden den Schülern kaum Gelegenheiten für die Planung und Revision von Texten eingeräumt. Der erste Entwurf war demnach häufig auch der letzte. Dies änderte sich zunehmend durch die Orientierung am Schreibprozess, so dass es heute in den

USA und in Deutschland gängige Praxis in den Schulen ist, explizite Zeit für die Planung, das Schreiben und das Überarbeiten zur Verfügung zu stellen.

Dennoch erfüllt der *Writer's Workshop* nicht die von ihm intendierten Ziele. Ein entscheidender Nachteil des *Writer's Workshops* besteht darin, dass dieser durch die Reduktion expliziter Instruktionen auf das Mindestniveau den Schülern nicht ausreichend viele Strategien anbietet, so dass insbesondere schwache Schüler häufig zwar wissen, was eine gute Geschichte ausmacht, dieses Wissen jedoch nicht in ihren eigenen Geschichten umsetzen können. Neuropsychologische Arbeiten, die erst durch Einführung entsprechender Methoden in den letzten Jahren möglich wurden, belegen dies (Berninger & Winn, 2006). Schreibprobleme sind demnach vor allem durch eine Überlastung des Arbeitsgedächtnisses zu erklären.

2.3.2 Kognitive Strategieinstruktion beim Schreiben

Die nachfolgenden Trainingsprogramme der kognitiven Strategieinstruktion sind nicht als ein Gegenpol zum *Writer's Workshop* zu verstehen. Vielmehr wird dieser durch weitere Elemente ergänzt, was zu einer Erhöhung der Effektivität führen soll. Allgemein werden unter kognitiven Strategien kognitive Prozesse verstanden, die von dem Lerner freiwillig zum Erreichen eines Lernziels eingesetzt werden, wobei metakognitive Strategien ebenfalls darunter subsumiert werden (Mayer, 2001).

Generell lassen sich folgende Unterschiede zwischen dem *Writer's Workshop* und dem kognitiven Strategietraining feststellen (Harris, Graham, & Mason, 2006):

- Für die Instruktion von Schreibstrategien und selbstregulatorischen Strategien wird mehr Zeit als im *Writer's Workshop* eingeplant. Darüber hinaus erfolgt die Instruktion systematisch und explizit, vor allem durch kognitives Modellieren.
- Während der *Writer's Workshop* sehr schülerzentriert ist, spielt der Lehrer bei der kognitiven Strategieinstruktion ebenfalls eine entscheidende Rolle. Die Instruktion setzt beim Ausgangsniveau der Schüler an, dennoch geht die Initiative zum Lernen und Schreiben größtenteils vom Lehrer aus. Zum Beispiel wählt in der Regel der Lehrer, nicht der Schüler, die Schreibanlässe aus.
- Darüber hinaus wird bei kognitiven Strategieinstruktionen der Schreibprozess und das Schreibprodukt hervorgehoben. Schüler erhalten ausführliche Rückmeldungen sowohl zum Schreibprodukt als auch zum Schreibprozess. Zusätzlich setzen sich die Schüler prozess- und produktbezogene Ziele.

Überdies wird den Schülern in Programmen der kognitiven Strategieinstruktion ebenfalls genügend Planungs-, Schreib- und Überarbeitungszeit zur Verfügung gestellt und der Austausch zwischen den Schülern gefördert.

Gemäß Gersten und Baker (2001) sollte eine Schreibstrategieinstruktion die explizite Vermittlung der Planungs-, Schreib- und Überarbeitungsschritte sowie eine regelmäßige Rückmeldeprozedur durch Lehrer oder Schüler beinhalten. Die Metaanalyse von Hattie et al. (1996) bestätigte diese Befunde: ein reines Lernstrategietraining erwies sich als weniger wirkungsvoll als metakognitive Interventionen. Folglich sollte ein Strategietraining nicht nur kognitive, sondern ebenfalls motivationale und metakognitive Komponenten umfassen, so dass die Schüler lernen, was, wann, wie und warum sie die Schreibstrategien einsetzen. Es wird davon ausgegangen, dass das Hinzufügen zusätzlicher Komponenten zu einem reinen Schreibstrategietraining das Arbeitsgedächtnis entlastet und freie Kapazitäten für das Anwenden der erlernten Strategien schafft.

De La Paz (2007) unterscheidet folgende typische Eigenschaften einer kognitiven Strategieinstruktion: (1) Die Lehrer verbalisieren ihre handlungsleitenden Gedanken laut; (2) die Instruktion ist zunächst lehrergesteuert, wobei diese Lenkung nach dem Scaffolding-Prinzip graduell zurückgenommen wird; (3) Schüler arbeiten anfangs in Gruppen und anschließend alleine; (4) die Regulation der Strategieanwendung wird zu Beginn von dem Lehrer übernommen und im weiteren Verlauf von den Schülern selbst durchgeführt; (5) Mnemotechniken werden eingesetzt, die die Anwendung kognitiver und metakognitiver Strategien anregen.

2.3.2.1 Pionierarbeiten in der Forschung der kognitiven Strategieinstruktion

Erste Arbeiten zur Förderung der Schreibkompetenz lieferten bereits Bereiter und Scardamalia (1987), indem sie durch *Procedural Facilitation* die Schreibleistung von Schülern und Studenten verbesserten. *Procedural Facilitation* umfasst Fragen oder Hinweiswörter zu den einzelnen Schreibphasen, welche die Schüler anregen, die bereits verfassten Notizen, Sätze und Wörter hinsichtlich der eigenen Ziele und den Erwartungen des Adressaten zu hinterfragen (Collins, Brown, & Newman, 1989). Auf diese Weise wird schreibrelevantes Wissen nicht nur abgerufen und niedergeschrieben (*Knowledge Telling*), sondern gleichzeitig transformiert und weiterentwickelt (*Knowledge Transforming*). Bereiter und Scardamalia (1987) setzten erstmals die Methoden des kognitiven Modellierens und des Scaffoldings, d.h. der zunehmenden Ausblendung von Hilfestellungen, zur Vermittlung der Schreibstrategien ein. Die Forscher entwickelten Hinweiswörter sowohl zur Förderung der Planungs- (Scardamalia et al., 1984) als auch der Revisionsfertigkeiten von Schülern (Bereiter

& Scardamalia, 1987). Beide Trainingsprogramme erwiesen sich als effektiv. Der *Procedural Facilitation*-Ansatz leistete wichtige Erkenntnisse bezüglich der Rolle der Selbstregulation beim Schreiben, da er zum einen die Überwachung der eigenen Schreibleistung und zum anderen die Regulation der Schreibziele implizierte.

Aufbauend auf dem *Procedural Facilitation*-Ansatz konzipierten Englert und Kollegen ein Interventionsprogramm zur Förderung der Schreibkompetenz für lernschwache Viert- und Fünftklässler, welches sie *Cognitive Strategy Instruction for Writing* (CSWI) nannten (Englert, Raphael, Anderson, Anthony, & Stevens, 1991; Englert, Raphael, & Anderson 1992). Sie unterscheiden vier grundlegende Komponenten des CSWI's, die den Kern des Schreibtrainings ausmachen (Englert, Raphael, & Anderson, 1992):

- *Entwicklung von deklarativem, prozeduralem und konditionalem Schreibstrategiewissen:* Die Vermittlung von deklarativem Strategiewissen wird durch metakognitive Komponenten ergänzt. Die Schüler lernen, wann das Wissen angewandt wird, warum die Anwendung der Strategien wichtig ist und wie die Strategien korrekt eingesetzt werden.
- *Betonung des „inneren Dialogs“:* Während der Lehrer eine Strategie anwendet, verbalisiert dieser alle seine Gedanken laut und macht dadurch die bei ihm ablaufenden kognitiven Prozesse für den Schüler sichtbar. Dadurch kann der Schüler das Vorgehen beim Anwenden einer Strategie nachempfinden und diese ebenfalls für sich übernehmen und internalisieren.
- *Scaffolding:* Das graduelle Ausblenden von Hilfestellungen bezieht sich zum einen auf die Unterstützung durch den Lehrer und zum anderen auf den Einsatz der Materialien. Das *Procedural Facilitation*-Prinzip wurde von Englert und Kollegen übernommen. Sie verwendeten ebenfalls Protokolle, die die Schüler zur Selbstbefragung anregen sollten.
- *Schreibkonferenzen:* Schreibkonferenzen von Schülern bzw. Schülern und Lehrern sollten die Zusammenarbeit der Schüler untereinander verstärken und die Möglichkeit geben, regelmäßig Feedbacks zu den Texten zu erhalten.

Darüber hinaus impliziert die CSWI-Instruktion die Anwendung von Mnemotechniken, speziell der POWER-Strategie, welche die Schüler an folgende Schreibprozesse erinnert: *Plan, Organize, Write, Edit, Revise*. Für jeden der fünf Prozesse werden den Schülern sogenannte *Think Sheets*, d.h. Denkprotokolle mit Fragen im Sinne des *Procedural Facilitation*-Ansatzes gegeben, die die Schüler anregen, während des gesamten Schreibprozesses Textinhalte zu generieren und zu organisieren (Wong, Harris, Graham, &

Butler, 2003). Wie auch bei Bereiter und Scardamalia (1987) erfolgt die Instruktion mittels des kognitiven Modellierens.

Die Effektivität der CSWI-Methode konnte in zwei Studien nachgewiesen werden. Zum einen konnte der Nachweis erbracht werden, dass nach der CSWI-Intervention im Gegensatz zum normalen Unterricht die Kluft zwischen lernschwachen und lernstarken Schülern bezüglich der Schreibkompetenz reduziert wurde (Englert et al., 1991). Zum anderen führte eine CSWI-Instruktion zu einem Anstieg metakognitiven Wissens, welches wiederum positiv mit der Schreibkompetenz sowie mit der Lesekompetenz korrelierte (Englert et al., 1992).

In ihren Arbeiten betonen Englert und Kollegen die Wichtigkeit des kognitiven Modellierens für den Erfolg von Interventionsmaßnahmen bei Schülern mit Lernschwierigkeiten (Englert et al., 1991; Englert et al., 1992). Durch die Verbalisierung aller handlungsleitenden Gedanken entwickeln Lehrer und Schüler eine gemeinsame Sprache. Insbesondere Schüler mit Lernschwierigkeiten verstehen häufig die Instruktionen der Lehrer nicht. Durch die Verbalisierung des „inneren Dialogs“ kann der Schüler das Vorgehen des Lehrers nachempfinden und somit für sich selbst einsetzen. Folglich sind die Schüler weniger auf externe Regulationen durch den Lehrer angewiesen und können eigene selbstregulierende Fertigkeiten entwickeln (Englert et al., 1992).

Die dargestellte Instruktionmethode des *Procedural Facilitation's* und der Interventionsansatz CSWI sind als sehr einflussreich einzustufen und wurden zudem bereits im Grundschulunterricht eingesetzt und erprobt. Darüber hinaus wurden noch weitere Trainingsprogramme entwickelt, die sich speziell mit der Förderung der Schreibkompetenz von Mittelstufenschülern und Erwachsenen beschäftigten. Da das intendierte Trainingsprogramm für Grundschüler gedacht ist, soll an dieser Stelle nur kurz auf zwei weitere Ansätze hingewiesen werden:

- (1) Seit Mitte der 70er Jahre erarbeiten Forscher an dem *Institute of Research in Learning Disabilities at the University of Kansas* (KU-IRLD) mehrere Ansätze der kognitiven Strategieinstruktion für verschiedene schulische Bereiche (z.B. Leseverständnis, Schreiben von Paragrafen und Sätzen, Fehlerüberwachung), die sie im Schulalltag erfolgreich überprüften, so dass diese noch heute im Sekundarstufenunterricht in den USA eingesetzt werden (Wong et al., 2003). Inhaltlich gibt es viele Überschneidungen zwischen KU-IRLD und CSWI, wie z.B. das kognitive Modellieren, lautes Denken, Selbstinstruktionen und die Verwendung von Mnemotechniken.

- (2) Ebenfalls aufbauend auf CSWI konzipierten Wong, Butler, Ficzere und Kuperis (1996) genrespezifische Schreibstrategietrainings. Sie entwickelten Programme für die High School, die das Verfassen eines Berichts, einer Stellungnahme und einer Gegenüberstellung förderten (Wong, 1997). Neben der Vermittlung von deklarativem und prozeduralem Strategiewissen fügte Wong (1997) affektive (insbesondere Selbstwirksamkeit) und metakognitive (insbesondere Bewusstsein für den Leser) Komponenten in die Schreibtrainings mit ein. Darüber hinaus ließ Wong (1997) die Schüler die Texte nicht handschriftlich, sondern mit einer Tastatur verfassen, so dass jedem Schreibtraining ein Tastaturtraining vorgeschaltet wurde. Insgesamt konnten positive Effekte der genrespezifischen Programme sowohl auf die Qualität als auch auf die Quantität der Schreibprodukte festgestellt werden (Wong et al., 2003).

Kritik an den Pionierarbeiten: Die hier skizzierten Pionierarbeiten im Bereich der kognitiven Strategieinstruktion beeinflussten spätere Schreibtrainings und wurden teilweise in den amerikanischen Schulalltag integriert. Dennoch müssen einige Punkte kritisiert werden. Zunächst basieren die berichteten positiven Befunde auf sehr kleinen Stichproben. Dies ist darauf zurückzuführen, dass der Fokus bei den Pionierforschern auf Schülern mit Lernschwierigkeiten lag, so dass nicht eine gesamte Klasse, sondern nur spezielle Schüler untersucht und mit durchschnittlichen Schülern verglichen wurden. Dieses Vorgehen birgt jedoch erhebliche Nachteile, da nicht nur Schüler mit Lernschwierigkeiten, sondern auch Schreibanfänger, wie sie in der Grundschule vorzufinden sind, unter Schwierigkeiten mit der Textproduktion leiden und erheblich von einem Trainingsprogramm profitieren könnten. Des Weiteren fehlen in den berichteten Studien adäquate Kontrollgruppen. Wenn überhaupt eine Vergleichsgruppe herangezogen wurde, war dies der normale Unterricht in einer anderen Klasse. Demnach sind die berichteten Ergebnisse nur eingeschränkt interpretierbar, da bereits die Anwesenheit eines Trainers, die Bereitstellung von neuem und anregendem Lernmaterial sowie das Bewusstsein, an einer wissenschaftlichen Studie teilzunehmen, die Befunde verzerren.

Die dargestellten Einschränkungen führten dazu, dass *Procedural Facilitation*, CSWI sowie die Trainingsansätze von KU-IRLD und Wong (1997) nicht im Mittelpunkt der vorliegenden Arbeit stehen.

2.3.2.2 Self-Regulated Strategy Development (SRSD)

Der von Harris und Graham (1996) entwickelte *Self-Regulated Strategy Development*-Ansatz, oder kurz SRSD, ist sowohl bei Forschern als auch bei Lehrern sehr beliebt, so dass seit 1985

mittlerweile eine Fülle von SRSD-Studien existiert. Die Effektivität von SRSD wurde zudem in mehreren Metaanalysen nachgewiesen. Zunächst werden jedoch die Ziele und Inhalte des Programms dargestellt.

Grundlagen und Ziele von SRSD: SRSD stellt nicht nur einen empirisch, sondern ebenfalls theoretisch gut fundierten Trainingsansatz zur Förderung der Schreibkompetenz dar und integriert dabei verhaltensorientierte sowie sozial-kognitive Interventionen. SRSD greift vor allem auf Arbeiten von Meichenbaum (Selbstinstruktionen), Vygotsky (Zone der proximalen Entwicklung), Brown und Campione (*Informed Instruction*), Pressley (*Good Information Processor*) und Zimmerman (Selbstregulation) zurück (Harris & Graham, 2009).

Aufbauend auf den genannten theoretischen Modellen verfolgt SRSD folgende Ziele (Graham, Harris, & Troia, 1998; Harris & Graham, 2009):

- Befähigung der Schüler, vor allem aber der Schüler mit Schreibproblemen, die komplexen kognitiven Anforderungen beim Schreiben zu bewältigen;
- Vermittlung, Förderung und Übung von effektiven Schreib- und Selbstregulationsstrategien, so dass die Ausführung der Strategien automatisiert wird;
- Qualifizierung der Schüler, sich selbst notwendige Strategien anzueignen und den eigenen Schreibprozess zu überwachen;
- Erweiterung des Repertoires an schreibbezogenem Wissen;
- Förderung einer positiven SchreibEinstellung sowie einer positiven Selbstwahrnehmung als Schreiber.

Auch wenn SRSD bereits mit durchschnittlich begabten Schülern durchgeführt wurde (Danoff, Harris, & Graham, 1993), wurden in der Mehrzahl der Studien Schüler mit Schreibschwierigkeiten oder mit Lernschwierigkeiten der Klassenstufen vier bis acht adressiert.

Prinzipien von SRSD: In den meisten Studien diente SRSD der Förderung der Schreibkompetenz. Die im Folgenden erläuterten SRSD-Prinzipien der Instruktionsgestaltung beziehen sich speziell auf das Schreiben, wobei diese auch für andere schulische Bereiche wie Mathe und Lesen adaptiert werden können (Reid & Lienemann, 2006). Insgesamt fünf Eigenschaften von SRSD lassen sich identifizieren (Graham & Harris, 2003; Harris & Graham, 2009):

- (1) *Explizite Vermittlung von Schreibstrategien in Kombination mit selbstregulatorischen Strategien:* Wie im vorangegangenen Abschnitt beschrieben, integrierten bereits andere Programme (z.B. CSWI) selbstregulatorische und

metakognitive Strategien in ihren Schreibprogrammen. SRSD dagegen vermittelt diese Strategien erstmals explizit den Schülern. Durch geeignete Instruktionmethoden wie kognitives Modellieren zeigt der Trainer oder Lehrer explizit die Anwendung selbstregulatorischer Strategien in Kombination mit Schreibstrategien.

- (2) *Interaktion zwischen Schüler und Lehrer*: Schüler werden als aktive Mitarbeiter und Mitgestalter in ihrer Entwicklung zum Schreibexperten angesehen.
- (3) *Individuelle Instruktion*: Die Inhalte der Instruktionseinheiten sind auf die jeweiligen Bedürfnisse und Kapazitäten der Schüler zugeschnitten. Folglich ist das SRSD-Programm nicht als ein rigider Katalog von festgelegten Themen zu verstehen, sondern eher als ein flexibles Programm, welches in Abhängigkeit von dem jeweiligen Schüler Freiräume für Adaptationen lässt.
- (4) *Kriteriumsorientierter Unterricht*: Der Ablauf der Strategieinstruktion erfolgt kriteriumsbezogen und orientiert sich an dem jeweiligen Kompetenzniveau der Schüler. Erst wenn der Schüler ein bestimmtes Kriterium, d.h. ein bestimmtes Leistungsniveau, erreicht hat, wird im Lernstoff vorangeschritten. Das kriteriumsorientierte Vorgehen steht im Gegensatz zu einem zeitabhängigen Vorgehen, bei dem der Unterricht nach Ablauf einer gewissen Zeitspanne fortgesetzt wird.
- (5) *SRSD als fortlaufender Prozess*: Die Instruktion neuer Strategien wird durch die Erweiterung bereits erlernter Strategien ergänzt. Zum Beispiel kann eine Strategie, die ursprünglich für narrative Texte eingeführt wurde, in abgewandelter Form auch für das Verfassen einer Erläuterung herangezogen werden.

Schreibstrategien von SRSD: Die Autoren von SRSD entwickelten eine Vielzahl an Schreibstrategien. Als erstes wird eine grundlegende Strategie, die *Three-Step-Strategy*, vorgestellt, die es den Schülern erleichtert, den gesamten Schreibprozess zu lenken und zu organisieren (Harris & Graham, 1996; Reid & Lienemann, 2006). Diese Strategie kann in Kombination mit genrespezifischen Planungs- und Revisionsstrategien, die als nächstes erläutert werden, eingesetzt werden. Die *Three-Step-Strategy* besteht aus folgenden Schritten: Im ersten Schritt, *Think*, werden allgemeine Schreibziele formuliert und über die Erwartungen des Lesers reflektiert. Während des zweiten Schrittes, *Plan what to say*, wird mit Hilfe einer weiteren Schreibstrategie ein inhaltliches Gerüst der wesentlichen Aussagen erstellt. Der letzte Schritt schließlich, *Write and say more*, beinhaltet das Verfassen eines Textes unter Berücksichtigung der Ziele und Pläne.

Abgesehen von der *Three-Step-Strategy* wurden zudem sowohl für unterschiedliche Textgenres als auch speziell für das Planen und Überarbeiten explizite Strategien konzipiert. Nachfolgend werden beispielhaft die wichtigsten erläutert.

Die wesentliche Strategie des SRSD-Programms für das Verfassen von narrativen Texten ist die *Story Grammar Strategy*, welche aus der Mnemotechnik *W-W-W*, *What=2*, *How=2* besteht (Harris & Graham, 1996). Die postulierten Geschichtenelemente gehen dabei auf Untersuchungen von Stein und Glenn (1979) zurück. Die Mnemotechnik erinnert die Schüler daran, sich selbst verschiedene Fragen zu stellen, um wichtige Informationen für den Text zu generieren. Die Inhalte der Strategie sind in Abbildung 11 aufgelistet.

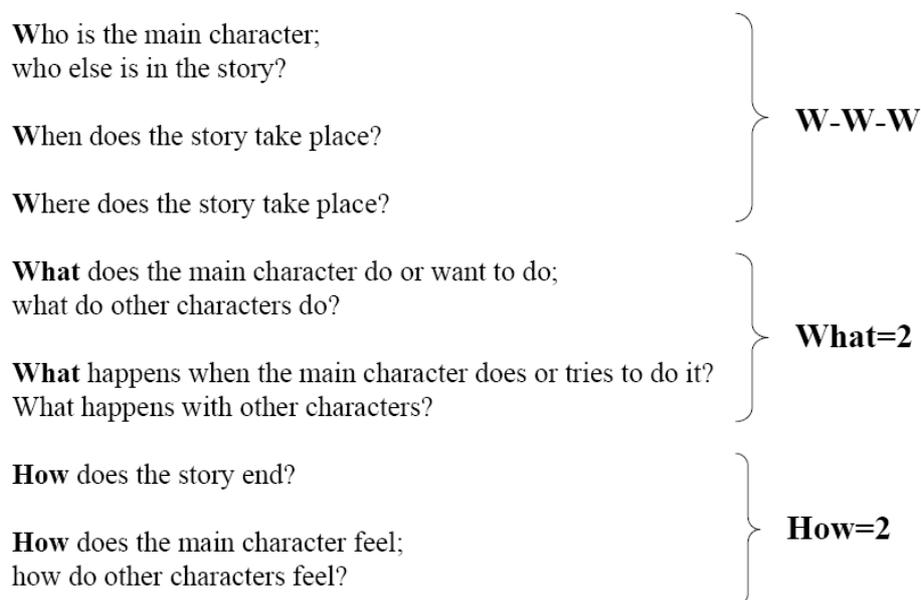


Abbildung 11: Story Grammar Strategy, adaptiert von Harris und Graham (1996, S. 49).

Es existieren Abwandlungen der Mnemotechnik für die *Story Grammar Strategy*, die sich aus anderen Fragen zusammensetzen, wobei jedoch die Inhalte identisch sind (Reid & Lienemann, 2006). Auch wenn die *Story Grammar Strategy* vornehmlich eine Planungsstrategie für das Schreiben ist, wird sie auch zur Förderung des Leseverständnisses eingesetzt.

Beliebte Planungsstrategien zum Verfassen einer Stellungnahme zu einem bestimmten Thema sind die Strategien *TREE* und *POW*. Während *POW* den Rahmen für das allgemeine Vorgehen beim Schreiben einer Stellungnahme darstellt, unterstützt *TREE* Schüler beim Generieren und Organisieren der Argumente. *POW* steht für die drei Schritte (1) *Pick my idea*, (2) *Organize my notes using TREE*, (3) *Write and say more* (Harris & Graham, 1996; Reid & Lienemann, 2006). *TREE* dagegen beinhaltet die klassischen Schritte zum Verfassen

einer Stellungnahme: *Topic sentence, Reasons, Examples, Ending* (Harris & Graham, 1996; Reid & Lienemann, 2006).

In der jahrelangen Arbeit mit SRSD entstanden weitere Planungsstrategien, welche sich auf andere Textgenres beziehen. Darüber hinaus wurden jedoch ebenfalls Revisionsstrategien entworfen.

Die *SCAN*-Strategie wurde ursprünglich für das Überarbeiten einer Stellungnahme mit einem Textverarbeitungsprogramm am Computer konzipiert. Sie wurde jedoch bereits ohne Computer eingesetzt. Anhand der Mnemotechnik *SCAN* wird jeder Satz einzeln überprüft. Die Buchstaben in dem Wort *SCAN* erinnern den Schüler daran, zunächst zu überprüfen, ob der Satz einen Sinn ergibt (*Does it make sense?*), anschließend zu überlegen, ob der Satz adäquat die vertretene Position repräsentiert (*Is it connected to my belief?*), sich zu fragen, ob noch weitere Informationen dazugefügt werden können (*Can I add more?*) und schließlich Rechtschreib- und Grammatikfehler zu identifizieren (*Note errors.*) (Harris & Graham, 1996; Reid & Lienemann, 2006).

Selbstregulatorische Strategien von SRSD: Bezüglich selbstregulatorischer Strategien beinhaltet das SRSD-Programm vor allem Strategien der Selbstüberwachung, der Selbstinstruktion, der Zielsetzung und der Selbstbelohnung. Der Einsatz sowie die Operationalisierung der selbstregulatorischen Strategien hängen erheblich von der instruierten Schreibstrategie und dem Textgenre ab. Der Kerngedanke von SRSD besteht jedoch darin, dass die kombinierte Vermittlung von Schreibstrategien mit selbstregulatorischen Strategien zu größeren Effekten führt als die alleinige Instruktion von Schreibstrategien oder von Selbstregulation. Der zusätzliche Gewinn durch selbstregulatorische Strategien kann in vielfältiger Weise erfolgen (Reid & Lienemann, 2006):

- Schüler erinnern sich leichter an die Inhalte einer bestimmten Schreibstrategie und lernen, wann sie diese einsetzen können.
- Schüler überwachen den korrekten Einsatz der Schreibstrategien, so dass diese korrekt angewendet werden.
- Schüler können Störfaktoren ausblenden und ihre Motivation und Anstrengungsbereitschaft aufrecht erhalten.

Instruktionsprinzipien von SRSD: Beim SRSD-Ansatz werden sechs Instruktionsstufen zur Vermittlung der beschriebenen Schreib- und selbstregulatorischen Strategien unterschieden (Harris & Graham, 1996; Reid & Lienemann, 2006). Die Abfolge der Stufen ist nicht als feststehend zu verstehen. Vielmehr können in Abhängigkeit vom Schüler Stufen übersprungen werden sowie rekursiv zwischen den Stufen gewechselt werden.

Stufe 1: Entwickeln und Aktivieren von Hintergrundwissen: Zu Beginn muss überprüft werden, ob die Schüler die für das Erlernen einer neuen Strategie benötigten Vorläuferfertigkeiten besitzen. Bei der Textproduktion muss u.a. gewährleistet werden, dass die Schüler über ausreichende Transkriptions-, Rechtschreib- und Grammatikfertigkeiten verfügen. Generell besteht die erste Stufe aus zwei Aufgaben: Definition der benötigten Fertigkeiten sowie Erfassung des vorhandenen Wissens bzw. der Fertigkeiten der Schüler.

Stufe 2: Diskutieren: Während der nächsten Stufe wird die zu erlernende Strategie vorgestellt. Dabei ist es wichtig, dass der Lehrer ein positives Beispiel darstellt und die Strategie engagiert und motiviert erklärt, so dass die Schüler eine positive Einstellung entwickeln und motiviert sind, die Strategie zu erlernen. Dazu erklärt der Lehrer die Nützlichkeit der neuen Strategie, indem er erläutert, wie bisherige Misserfolge durch den Einsatz der neuen Strategie vermieden werden können.

Stufe 3: Modellieren: Essenzieller Bestandteil der SRSD-Instruktion ist das kognitive Modellieren der Strategie. Während der Lehrer die Strategie anwendet, verbalisiert er alle handlungsleitenden Gedanken laut. Dadurch erlangen die Schüler metakognitives Wissen über den Strategieeinsatz. Der Lehrer erklärt, wie er die Strategie einsetzt, aber auch warum er die Strategie verwendet und wie er den Einsatz der Strategie überwacht und reguliert. Die vom Lehrer formulierten Selbstinstruktionen können vom Schüler übernommen und adaptiert werden.

Stufe 4: Memorieren: Durch den Einsatz von Lernhilfen, Mnemotechniken und Gedächtnisstützen wird die Strategie sowie die einzelnen Stufen der Strategieanwendung auswendig gelernt, so dass diese schließlich schnell und automatisiert abgerufen werden können.

Stufe 5: Unterstützen: Dem Scaffolding-Prinzip folgend, erhalten die Schüler zunächst sehr viel Unterstützung, die mit zunehmender Beherrschung der Strategie graduell ausgeblendet wird. Die Unterstützung bezieht sich dabei nicht nur auf Anleitungen und Hilfestellungen durch den Lehrer, sondern ebenfalls auf den Schweregrad der Aufgaben und auf das zur Verfügung gestellte Material. Kooperative Lernformen können in dieser Phase ebenfalls eingesetzt werden. Darüber hinaus erhalten Schüler kontinuierlich Feedback über den Lernfortschritt.

Stufe 6: Unabhängige Leistung: In der letzten Stufe führen die Schüler die Strategie alleine und ohne jegliche Form von Hilfestellungen aus. Der Lehrer überwacht jedoch weiterhin den Lernfortschritt.

Die Dauer der Strategieinstruktion variiert in Abhängigkeit von dem Leistungsniveau der Schüler (durchschnittliche Schüler oder Schüler mit Lernschwierigkeiten), dem Alter der Schüler und der jeweiligen Strategie. Für die Grundschule werden jedoch acht bis zwölf Unterrichtseinheiten (à 30 bis 40 Minuten) für die Durchführung der sechs Stufen vorgeschlagen (Harris, Graham, & Mason, 2003). Idealerweise sollte dabei die Instruktion mehrmals pro Woche erfolgen.

Effektivität von SRSD: Insgesamt erzielte SRSD Verbesserungen in folgenden Bereichen (Harris & Graham, 2009): Qualität des Textes, Integration genrespezifischer Geschichtenelemente in den Text, Schreibwissen, Selbstwirksamkeit, Schreibeinstellung, Planungs- und Revisionsfertigkeiten. Die Stabilität und Generalisierbarkeit der Effekte konnte zudem nachgewiesen werden (Harris & Graham, 2009). SRSD zeigte sich sowohl für Schüler mit Lernschwierigkeiten als auch für durchschnittliche Schüler in den Klassen vier bis acht als effektiv (Graham et al., 1998). Vier Metaanalysen wurden durchgeführt, um die Überlegenheit des SRSD-Ansatzes im Vergleich zu anderen Strategieinstruktionen nachzuweisen.

Die erste Metaanalyse wurde von Graham und Harris (2003) mit 18 SRSD-Studien in den Klassen zwei bis acht durchgeführt. Es wurden sowohl experimentelle (sieben Studien) als auch Einzelfalluntersuchungen (elf Studien) berücksichtigt. Das Ziel der Metaanalyse bestand darin, die Größe des Einflusses von SRSD auf verschiedene schreibbezogene Maße zu untersuchen. Insgesamt ergaben sich starke Effekte von SRSD auf die Aufsatzqualität (Posttest: $ES= 1.47$; Follow-Up: $ES= .74$), Integration der Geschichtenelemente (Posttest: $ES= 1.87$; Follow-Up: $ES= 1.60$) und Länge der Texte (Posttest: $ES= 2.07$; Follow-Up: $ES= .88$). In der Mehrheit der Studien wurden Schüler mit Lernschwierigkeiten untersucht, für die große Effekte nachgewiesen wurden. Für durchschnittliche und gute Schreiber lagen nur sehr wenige Studien vor, wobei die Ergebnisse ebenfalls zuversichtlich waren. Bezüglich der Stabilität der Effekte sanken die Effektstärken zum Follow-Up hin ab. Sie blieben aber in der Regel moderat bis groß. Darüber hinaus verglichen Graham und Harris (2003) die Effektivität von SRSD, wenn geschulte Trainer im Gegensatz zu Lehrern das Programm durchführten. Direkt im Anschluss an das Training zeigten sich positive Effekte zu Gunsten von Trainern. Hinsichtlich der Stabilität und der Generalisierbarkeit der Effekte, z.B. auf andere Textsorten, übertrafen jedoch die Lehrer die Trainer.

Graham (2006) rechnet eine weitere Metaanalyse mit 39 Studien, von denen 20 experimentelle und 19 Einzelfalluntersuchungen der Klassen zwei bis zwölf waren, um den Effekt von Strategieinstruktion allgemein auf die Schreibleistung zu untersuchen. Insgesamt

ergab sich eine mittlere Effektstärke für die Strategieinstruktion von 1.15, so dass sämtliche Strategieinstruktionen einen großen Effekt auf die Schreibleistungen ausübten. SRSD war die am häufigsten eingesetzte Unterrichtsmethode (45% der experimentellen und 68% der Einzelfalluntersuchungen). Vergleiche der Effektstärken von SRSD-Studien (Posttest: ES= 1.57, PND= 88%; Follow-Up: ES= 1.15, PND= 93%) mit Nicht-SRSD-Studien (Posttest: ES= .89, PND= 91%; Follow-Up: ES= 3.20, PND= 100%) zeigten beim Posttest für experimentelle Studien signifikante Unterschiede ($p < .02$) und für Einzelfallstudien nicht signifikante Unterschiede ($p = .35$). Vergleiche für den Follow-Up waren aufgrund von mangelnden Follow-Up-Erhebungen bei den Nicht-SRSD-Studien nicht durchführbar. Ein Vergleich von Lehrern und Trainern erwies sich beim Posttest als nicht signifikant ($p = .06$), bei den Einzelfallstudien wurden jedoch größere PND's berichtet, wenn Lehrer das Training durchführten ($p < .05$). Von den experimentellen Studien mit Prä-Post-Follow-Up-Erhebungen waren jedoch nur in sechs Fällen Lehrer für die Durchführung der Instruktion verantwortlich. Bei Trainern ergaben sich zudem größere Standardabweichungen bei der Berechnung der Effektstärken als bei Lehrern, was darauf schließen lässt, dass Lehrern bei der Instruktion eine bessere Binnendifferenzierung gelang.

Die umfangreichste und methodisch aufwendigste Metaanalyse wurde von Graham und Perin (2007a) mit 123 Studien der Klassenstufen vier bis zwölf durchgeführt, wobei nur experimentelle und quasi-experimentelle Studien aufgenommen wurden. Für die Analyse wurden alle Effektstärken neu berechnet und es wurde nur das Schreibmaß Textqualität berücksichtigt. Das primäre Ziel der Forscher war die Identifikation von effektiven Unterrichtsmethoden für den Schreibunterricht, so dass basierend auf den Ergebnissen der Metaanalyse die 13 wissenschaftlichen Empfehlungen für den Schreibunterricht entstanden (siehe Kapitel 2.2.3.1). Es ergab sich eine gewichtete Effektstärke von 1.14 für SRSD- und .62 für Nicht-SRSD-Studien. Dieser Unterschied erwies sich als signifikant ($p < .001$). Darüber hinaus untersuchte die Metaanalyse den Einfluss der methodischen Qualität der Studien (z.B. Randomisierung der Versuchspersonen, alternatives Kontrolltraining, Implementationsgenauigkeit) auf die Höhe der Effektstärken. Dabei zeigte sich kein Zusammenhang ($r = .07$).

Die letzte Metaanalyse von Rogers und Graham (2008) umfasste 88 Einzelfallstudien und diente wie auch die Analyse von Graham und Perin (2007a) der Identifikation effektiver Unterrichtsmaßnahmen. Die Instruktion von Planungsstrategien erzielte für die Geschichtenqualität einen durchschnittlichen PND-Wert von 99%. Da alle untersuchten

Strategieinstruktionsstudien den SRSD-Ansatz verwendeten, kann folglich geschlussfolgert werden, dass sich SRSD in Einzelfalluntersuchungen ebenfalls als hoch effektiv erwies.

Eine zusammenfassende Darstellung der Effektstärken (experimentelle Studien) und PND's (Einzelfalluntersuchungen) aller vier Metaanalysen ist in Tabelle 6 aufgelistet.

Tabelle 6: Effektstärken und PND's zur Beurteilung des Einflusses von SRSD auf die Aufsatzqualität.

	Posttest		Follow-Up		Insgesamt	
	ES	PND	ES	PND	ES	PND
Graham & Harris (2003)	1.47	97%	0.74	100%	-	-
Graham (2006)	1.57	88%	1.15	93%	-	-
Graham & Perin (2007)	-	-	-	-	1.14 ⁷	-
Rogers & Graham (2008)	-	99%	-	- ⁸	-	-

Anmerkungen: ES = Effektstärke für experimentelle Studien, berechnet nach Cohen's d ; PND = Percentage of non-overlapping data, Effektstärkemaß für Einzelfalluntersuchungen, siehe Fußnote 6.

Allgemein gelten Effektstärken von .80 als groß, von .50 als mittel und von .20 als klein (Graham, 2006; Graham & Harris, 2003; Graham & Perin, 2007a). Für den PND-Wert werden PND's über 90% als sehr effektiv, zwischen 70 und 90% als effektiv und unter 50% als ineffektiv angesehen (Graham, 2006; Graham & Harris, 2003; Rogers & Graham, 2008). Insgesamt ergibt sich für die Effektstärken ein sehr positives Bild, so dass der Effekt von SRSD auf die Aufsatzqualität im Vergleich zu Nicht-SRSD-Studien als groß zu bewerten ist. Zum Follow-Up hin scheint der Effekt jedoch leicht abzufallen. Hinsichtlich der PND's lässt sich ebenfalls die Schlussfolgerung ziehen, dass SRSD in Einzelfallstudien sehr effektiv ist.

Ein wesentlicher Kritikpunkt, der bei der Diskussion der Effektivität von SRSD aufkommt, betrifft die Unabhängigkeit der Analysen. Alle berichteten Metaanalysen sowie ein Großteil der in den Metaanalysen untersuchten Studien wurden von Graham und Harris sowie deren Mitarbeitern durchgeführt. Es existieren somit kaum Studien, die das SRSD-Programm unabhängig von Graham und Harris einsetzten und evaluierten. Weiter sind die meisten der Studien, auf die die Metaanalysen zurückgreifen, Einzelfalluntersuchungen. Einzelfalluntersuchungen bestehen in der Regel aus sehr kleinen Stichproben und befassen sich mit speziellen Subgruppen (z.B. Schülern mit Lernschwierigkeiten). Darüber hinaus ist die Stabilität der durch SRSD-Studien erzielten Effekte unklar, da diese häufig beim

⁷ Im Gegensatz zu den anderen Metaanalysen wurde in der Metaanalyse von Graham und Perin (2007) ein gewichtetes Feste-Effekte-Modell gerechnet, bei dem durch die gepoolte Standardabweichung der Experimental- und Kontrollbedingung dividiert wurde.

⁸ Für den Follow-Up kann kein durchschnittlicher PND-Wert angegeben werden, da weniger als vier Studien zur Berechnung vorlagen.

Follow-Up absanken. In der Gesamtheit ergibt sich eine Forderung nach weiteren methodisch sauberen experimentellen und quasi-experimentellen Studien, die nicht von Graham und seinen Mitarbeitern durchgeführt werden, die eine große leistungsheterogene Stichprobe umfassen und die die Stabilität der Effekte überprüfen.

2.3.2.3 Selbstregulatorisches Aufsatztraining (SAT)

Das im vorherigen Abschnitt erläuterte SRSD-Programm ist ein Schreibtraining aus den USA. Folglich ist es speziell auf das Verfassen englischer Texte ausgerichtet. Darüber hinaus sind sämtliche Strategien, inklusive Mnemotechniken und Übungsmaterialien, in englischer Sprache abgefasst.

Glaser (2005) adaptierte das SRSD-Programm, so dass dieses ebenfalls mit deutschen Schülern eingesetzt werden konnte und nannte es *Selbstregulatorisches Aufsatztraining* (SAT). Grundlegender Gedanke von SAT ist wie bei SRSD die Vermittlung von genrespezifischen Schreibstrategien mit selbstregulatorischen Prozeduren zur Überwachung und Steuerung des Schreibprozesses. Die Instruktion erfolgt ebenfalls nach dem Scaffolding-Prinzip, bei dem eine neue Strategie zunächst kognitiv modelliert, anschließend mit umfassender Hilfeleistung von den Schülern eingesetzt und zuletzt nach Zurücknahme sämtlicher Unterstützung von den Schülern unabhängig ausgeführt wird.

Schreibstrategien von SAT: Die im SAT vermittelten Schreibstrategien lehnen sich eng an die SRSD-Schreibstrategien an. Zunächst erlernten die Schüler eine allgemeine *Three-Step-Strategy*, die auf das kognitive Schreibmodell von Hayes und Flower (1980a, 1980b) zurückgreift und die drei Schritte Planen, Schreiben und Überarbeiten umfasst. Analog wird diese Strategie die 3-Schritte-Technik genannt. Sie diente als Grundlage für die weiteren genrespezifischen Strategien, die während des Planens, Schreibens und Überarbeitens eingesetzt wurden.

Auch wenn Glaser und Kollegen ebenfalls Trainingsprogramme zur Förderung der Revisionsfertigkeiten entwickelten (Glaser & Brunstein, 2007a), werden im Folgenden die Planungsaktivitäten fokussiert.

Aufbauend auf der *Story Grammar Strategy* (*W-W-W, What=2, How=2*) entwickelte Glaser (2005) eine Planungsstrategie zum Verfassen narrativer Texte. Im SAT-Programm erlernten Schüler die Mnemonik AHA-7W. Der erste Teil bezieht sich auf die drei Abschnitte einer Geschichte: Anfang, Hauptteil, Abschluss. Der zweite Teil leitet sich explizit von der von Harris und Graham (1996) postulierten *Story Grammar Strategy* sowie von den Bestandteilen einer guten Geschichte nach Stein und Glenn (1979) ab. Demnach besteht eine inhaltlich vollständige Geschichte aus den sieben Geschichtenelementen Hauptperson, Ziel,

Ort, Zeit, Handlungsschritte, Höhepunkt, Ende. Die entsprechenden 7W-Fragen sind in Abbildung 12 aufgelistet.

Anfang	Hauptteil	Abschluss
<p>Wer (... ist die Hauptperson?)</p> <p>Was (... ist ihr Anliegen?)</p> <p>Wo (... spielt die Geschichte?)</p> <p>Wann (... spielt die Geschichte?)</p>	<p>Was (... tut die Hauptperson nacheinander?)</p> <p>Was (... ist der Höhepunkt in der Geschichte?)</p>	<p>Wie (... endet die Geschichte?)</p>

Abbildung 12: AHA-7W-Strategie des SAT-Programms, adaptiert von Glaser (2005, S. 88).

Abgesehen von der Planung des Inhalts umfasst das SAT-Programm zudem Strategien zur Verbesserung des Ausdrucks einer Geschichte, indem interessante Verben und Adjektive sowie wörtliche Rede in die Geschichten integriert werden (Glaser & Brunstein, 2007b).

Selbstregulatorische Strategien von SAT: Das zweite Set an Strategien umfasst neben den Schreibstrategien die selbstregulatorischen Strategien, die explizit mit den Schreibstrategien verknüpft wurden. Dabei handelt es sich um folgende Punkte (Glaser & Brunstein, 2007b):

- *Selbstüberwachung der Planungsaktivitäten:* Es wurde ein Geschichtenplan eingeführt, der die einzelnen Komponenten der AHA-7W-Strategie abbildet. Zu jedem der drei Abschnitte (AHA) sowie zu den 7W-Fragen standen freie Zeilen für Notizen zur Verfügung, so dass der Schüler schnell feststellen konnte, ob zu jedem Element der Strategie Planungsnotizen formuliert wurden.
- *Selbstbewertung:* Anhand einer Checkliste lernten Schüler in den eigenen Geschichten, fehlende 7W-Fragen zu identifizieren. Auf einer Skala mit 0 (nicht vorhanden), 1 (vorhanden) sowie 2 (vorhanden und ausgestaltet) schätzten Schüler die einzelnen 7W-Fragen ein. Die Summe ergab ein Maß für die Gesamtqualität der Geschichte.
- *Selbstüberwachung der Revisionsaktivitäten:* Im Anschluss an die Selbstbewertung überarbeiteten die Schüler ihre Texte. Als Grundlage für die Überarbeitung diente dabei eine Checkliste, welche anzeigte, welche 7W-Fragen fehlten bzw. noch nicht ausgestaltet wurden.
- *Prozess- und produktbezogene Zielsetzung:* Zu Beginn einer neuen Schreibphase überlegten sich Schüler eine Punktzahl von maximal 14 Punkten, die sie in der

nächsten Geschichte zu erreichen anstreben. Die Schüler verglichen die erreichte Punktzahl mit der anvisierten, so dass die Zielerreichung von ihnen überwacht wurde. Abgesehen von den produktbezogenen Zielen überlegten sich Schüler zudem vor jeder Geschichte eine Strategie, auf die sie sich in der nächsten Geschichte insbesondere konzentrieren wollten.

Effektivität von SAT: Insgesamt liegen im Vergleich zum SRSD-Programm erst sehr wenige Studien vor, die die Effektivität von SAT untersuchten. Bislang sind jedoch die Ergebnisse sehr vielversprechend, da sich positive Effekte von SAT sowohl auf strategienahe und –ferne Maße der Aufsatzqualität als auch auf weitere schreibbezogene Maße (z.B. Wissen, Planungsaktivitäten, Selbstwirksamkeitserwartung) und Transfermaße ergaben (Glaser, 2005; Glaser & Brunstein, 2007b; Glaser et al., 2009). Ein Großteil der Effekte zeigte sich zudem fünf Wochen später als stabil.

2.3.3 Komponentenanalyse

Bereits Zimmerman und Risemberg (1997) machten auf das Problem aufmerksam, dass aufgrund der vielfältigen Strategien des SRSD-Programms die Identifikation essenzieller Komponenten schwierig ist. Insbesondere der inkrementelle Effekt der Selbstregulation im Vergleich zu einem reinen Schreibstrategietraining gilt als noch nicht ausreichend wissenschaftlich belegt (De La Paz, 2007). Dies bedeutet, dass weiterhin die Frage bestehen bleibt, welche Strategien für ein erfolgreiches Training zur Förderung der Schreibkompetenz notwendig sind (Harris & Graham, 2009). Da bislang erst wenige Forscher der Frage nachgegangen sind, welche Komponenten von SRSD für die Effektivität verantwortlich sind, berichtet De La Paz (2007) von vier Artikeln, die Strategieinstruktionen mit und ohne Selbstregulation gegenüberstellten.

Day (1986) verglich die Effektivität von vier verschiedenen Methoden zur Förderung der Fähigkeit, Zusammenfassungen zu schreiben, wobei sich die vier Bedingungen hinsichtlich des Umfangs an selbstregulatorischen Strategien unterschieden. Überprüft wurden die Methoden anhand von 62 College-Studenten, die entweder durchschnittliche oder schwache Leistungen im Schreiben aufwiesen. Explizit handelt es sich um folgende vier Bedingungen: (1) reines Selbstmanagement-Training, (2) reines Strategietraining (insbesondere das Lernen bestimmter Regeln), (3) Strategie- plus Selbstmanagement-Training, (4) Integration von Strategie- und Selbstmanagement-Training. Der wesentliche Unterschied zwischen den Bedingungen drei und vier bestand darin, dass in der vierten Bedingung, jede Strategie bzw. Regel explizit mit den selbstregulatorischen Strategien

verknüpft wurde. Dagegen stellt die Strategie- plus Selbstmanagement-Trainingsbedingung lediglich ein zweifaches Training bestehend aus den Bedingungen (1) und (2) dar.

Die Ergebnisse der Studie von Day (1986) lassen sich wie folgt zusammenfassen. Für die einzelnen instruierten Regeln ergaben sich unterschiedliche Ergebnisse. Der zusätzliche Nutzen von selbstregulatorischen Strategien erwies sich bei Regeln mit mittlerer Schwierigkeit am effektivsten. Bei leichten sowie bei schweren Regeln war der Gewinn durch Selbstmanagement-Strategien vergleichsweise gering. Des Weiteren führte ein reines Selbstmanagement-Training zu keinen Verbesserungen der Zusammenfassungsfertigkeiten. Die größten Zuwächse vom Prätest zum Posttest konnten für das integrierte Strategie- und Selbstmanagement-Training nachgewiesen werden. Das Strategie- plus Selbstmanagement-Training zeigte sich zudem als effektiver als das reine Selbstmanagement- und Strategietraining. Insgesamt profitierten durchschnittliche Schreiber mehr als schwache Schreiber und benötigten weniger explizite Instruktionen, um ein bestimmtes Kompetenzniveau zu erreichen.

Zwei weitere Studien, welche den inkrementellen Effekt der Selbstregulation zu einem reinen Strategietraining untersuchten, wurden von Graham und Harris (1989) sowie Sawyer, Graham und Harris (1992) durchgeführt. In der ersten Studie verglichen die Forscher 22 Schüler der fünften und sechsten Klasse mit Lernbehinderungen, die zufällig einer von zwei Bedingungen (1. Strategietraining, 2. Strategie- plus Selbstregulationstraining) zugeordnet wurden, mit elf durchschnittlichen Schülern, die kein Training durchliefen (Graham & Harris, 1989). In Anlehnung an die *Story Grammar Strategy* erhielten alle Schüler mit Lernbehinderungen ein Strategietraining zur Förderung der Schreibkompetenz beim Verfassen narrativer Texte. Darüber hinaus erhielten Schüler der Strategie- plus Selbstregulationsbedingung zusätzlich Instruktionen in folgenden Bereichen: Überwachung des Lernfortschritts, prozessbezogene Zielsetzung, Selbstbewertung. Die Strategievermittlung erfolgte in beiden Bedingungen nach dem SRSD-Ansatz anhand der sechs beschriebenen Instruktionsprinzipien (siehe Kapitel 2.3.2.2). Das Training wurde in Kleingruppen mit zwei bis drei Schülern zwei- bis dreimal pro Woche (à 45 Minuten) über einen Zeitraum von zwei bis drei Wochen durchgeführt.

Direkt nach dem Training ergab sich hinsichtlich der Integration von Geschichtenelementen kein signifikanter Unterschied zwischen den Schülern mit Lernbehinderungen und den durchschnittlichen Schreibern. Dennoch wurde am Ende des Trainings die Textqualität der durchschnittlichen Schreiber höher eingestuft, da diese u.a. längere Texte verfassten. Ein Vergleich der beiden Experimentalgruppen lieferte überdies

keinen Hinweis auf inkrementelle Effekte der Selbstregulation, da die beiden Gruppen sich hinsichtlich der Integration der Geschichtenelemente, der Geschichtenqualität und der Selbstwirksamkeitserwartung nicht unterschieden. Eine mögliche Erklärung für dieses Ergebnis liegt darin, dass eine hinreichend deutliche Abgrenzung zwischen den beiden Bedingungen nicht gelungen war, so dass in der reinen Strategiebedingung durch den Einsatz von Selbstinstruktionen und Merkkarten die Anwendung selbstregulatorischer Strategien ebenfalls suggeriert wurde.

Anknüpfend an die Erfahrungen der ersten Studie führten Sawyer, Graham und Harris (1992) eine zweite Studie mit 33 Schülern mit Lernbehinderungen durch, die sie zufällig einer von drei Bedingungen zuordneten und mit einer untrainierten Kontrollgruppe von zehn Schülern mit Lernbehinderungen sowie einer Kontrollgruppe mit 13 durchschnittlichen Schreibern verglichen. Die Studie fand ebenfalls in den Jahrgangsstufen fünf und sechs statt. Die Trainings wurden in Kleingruppen mit zwei bis drei Schülern durchgeführt und erstreckten sich über einen Zeitraum von drei Wochen, wobei dreimal wöchentlich (20 bis 56 Minuten) eine Sitzung stattfand. Zwei der drei Trainingsbedingungen waren identisch mit den beiden Experimentalgruppen der ersten Studie. Der einzige Unterschied zwischen der SRSD-Bedingung (ehemals Strategietraining) und der SRSD plus Selbstregulationsbedingung (ehemals Strategie- plus Selbstregulationstraining) bestand darin, dass nur Schüler der SRSD plus Selbstregulationsbedingung Feedback über die Prätestleistung erhielten, was die Zielsetzung erleichtern sollte. Die neue dritte Bedingung beinhaltete direkte Instruktion, welche die Vermittlung der Schreibstrategien ohne kognitives Modellieren, Selbstinstruktion, Selbstüberwachung und Zielsetzung umfasste.

Die Ergebnisse der Studie von Sawyer und Kollegen (1992) zeigen einerseits, dass durch SRSD bzw. SRSD plus Selbstregulation der Leistungsrückstand von lernbehinderten Schülern im Vergleich zu durchschnittlichen Schülern aufgehoben werden konnte, was durch einen normalen Förderunterricht (untrainierte Kontrollgruppe) und durch direkte Instruktion nicht erzielt wurde. Gleichzeitig liefern die Ergebnisse nicht genügend Hinweise darauf, dass selbstregulatorische Komponenten einen inkrementellen Effekt im Vergleich zu einem reinen Strategietraining erbrachten. Es zeigten sich Gruppenunterschiede bezüglich der Integration von Geschichtenelementen und nicht hinsichtlich eines holistischen Maßes der Geschichtenqualität zwischen den beiden SRSD-Bedingungen und der direkten Instruktionsbedingung. Dieser Unterschied war nur beim Posttest und nicht bei den Follow-Up-Erhebungen bedeutsam. Beim Follow-Up zeigten sich keine Gruppenunterschiede. Im Gegensatz dazu schnitt das SRSD plus Selbstregulation-Programm beim Transfermaß

(Situationstransfer) signifikant besser ab als die SRSD- und die direkte Instruktionsbedingung. Nur Schüler, die zusätzlich Feedback zu ihren Prätestergebnissen erhalten hatten, konnten das Erlernte auch bei einem anderen Lehrer zeigen.

Die vierte von De La Paz (1997) zitierte Studie wurde von Kernaghan und Woloshyn (1995) mit 35 durchschnittlich begabten Erstklässlern, die zufällig einer von drei Bedingungen zugeordnet wurden, durchgeführt. Im Gegensatz zu den vorherigen Studien ging es in der Studie von Kernaghan und Woloshyn (1995) um die Verbesserung der Rechtschreibung. Ein Rechtschreib- plus Selbstregulationstraining wurde einem reinen Rechtschreibtraining und traditionellen Rechtschreibübungen gegenübergestellt. Die selbstregulatorischen Prinzipien umfassten die Diskussion der Wichtigkeit der erlernten Rechtschreibstrategien sowie die Vermittlung von Wissen darüber, wie, wann und wo die Strategien angewendet werden können. In Bezug zu einem strategienahen Maß (Diktat der erlernten Wörter) zeigte sich die Rechtschreib- plus Selbstregulationsbedingung als signifikant besser als die reine Rechtschreibbedingung sowie der traditionelle Unterricht. Dieser Unterschied konnte jedoch für ein strategiefernes Maß (standardisierter Rechtschreibtest) nicht nachgewiesen werden. Außerdem zeigten im Anschluss an das Training alle Schüler bessere Leistungen bei der Textproduktion, so dass diesbezüglich ebenfalls keine Gruppenunterschiede festgestellt werden konnten.

Insgesamt lässt sich aus den bislang diskutierten Studien allenfalls ein mäßiger inkrementeller Effekt selbstregulatorischer Prozeduren ableiten. Dabei konnte dieser meistens nur für Transferaufgaben festgestellt werden. Abgesehen von der Studie von Sawyer, Graham und Harris (1992) weisen die übrigen Studien überdies erhebliche methodische Mängel auf. Zum Beispiel wurden keine Angaben zur Genauigkeit der Implementation (siehe Kapitel 2.4.2) gemacht. Darüber hinaus waren alle Interventionen sehr kurz und wurden nur mit einer kleinen Stichprobe durchgeführt. Folglich können die bisherigen Ergebnisse zur Analyse essenzieller Interventionskomponenten lediglich Ansatzpunkte für zukünftige Forschungsarbeiten sein. Weitere Untersuchungen in diesem Bereich sind wünschenswert, da insbesondere Studien, die den positiven Einfluss von Zielsetzung auf den Erfolg von Schreibtrainings nachwiesen (Ferretti, MacArthur, & Dowdy, 2000; Page-Voth & Graham, 1999), vermuten lassen, dass die bisher diskutierten Befunde die Rolle der Selbstregulation für die Effektivität von Schreibtrainings unterschätzten. Da Zielsetzung eine Form der Selbstregulation darstellt, müssten Schreibtrainings, die zusätzlich zur Zielsetzung noch weitere selbstregulatorische Komponenten umfassen, dementsprechend effektiver sein.

Aus diesem Grund führten Glaser und Kollegen zwei weitere Studien mit einer größeren methodischen Sorgfalt und ohne die beschriebenen Defizite der vorangegangenen Studien durch, um einen empirisch begründeten Nachweis zu erbringen, dass die kombinierte Vermittlung von Schreibstrategien mit selbstregulatorischen Strategien zu besseren Leistungen führt als ein reines Schreibtraining (Glaser, 2005; Glaser & Brunstein, 2007b; Glaser et al., 2009). In der ersten Studie verglich Glaser (2005; Glaser & Brunstein, 2007b) das zuvor beschriebene *Selbstregulatorische Aufsatztraining* (SAT) mit einem reinen Strategietraining sowie mit einer untrainierten Kontrollgruppe. Als Untersuchungsstichprobe dienten 113 Viertklässler aus sechs verschiedenen Klassen, die zufällig den drei Bedingungen zugeordnet wurden. Die untrainierte Kontrollgruppe erhielt regulären Deutschunterricht durch die Klassenlehrer. Die beiden Experimentalgruppen wurden in vier Trainingssitzungen (einmal pro Woche) in leistungsheterogenen Kleingruppen von vier bis sechs Schülern von geschulten Trainern unterrichtet. Hinsichtlich der abhängigen Maße Geschichtenelemente, Textlänge, Textqualität, Anzahl der Revisionen und eines Transfermaßes (Hörverständnis) übertrafen die Schüler der Strategie- plus Selbstregulationsbedingung die Schüler der beiden Vergleichsbedingungen. Beim Posttest produzierten Schüler der Strategie- plus Selbstregulationsbedingung zudem signifikant bessere Schreibpläne als die beiden Vergleichsbedingungen. Dieser Effekt blieb jedoch zum Follow-Up hin nicht stabil. Während die Schüler der untrainierten Kontrollbedingung keine Verbesserungen der Schreibleistungen beim Posttest und beim Follow-Up zeigten, verbesserten sich die Schüler der Strategiebedingung ebenfalls in ihren Leistungen, blieben mit ihren Leistungen jedoch deutlich unter dem Niveau der Schüler der Strategie- plus Selbstregulationsbedingung. Darüber hinaus fielen die Leistungen der Schüler der Strategiebedingung beim Follow-Up wieder ab, während Schüler der kombinierten Bedingung ihre Leistungen weiter aufrecht erhalten konnten. Obwohl Schüler der reinen Strategiebedingung und Schüler der Strategie- plus Selbstregulationsbedingung einen erheblichen Wissenszuwachs zeigten, konnten Schüler der kombinierten Bedingung ihr Wissen deutlich besser beim Planen, Schreiben und Überarbeiten anwenden als Schüler, die nur ein reines Strategietraining durchliefen.

In einer weiteren Studie mit 119 Viertklässlern konnten die berichteten Befunde bestätigt werden (Glaser et al., 2009). Die Schüler wurden wieder zufällig einer von zwei Bedingungen zugeordnet: reines Strategietraining versus Strategie- plus Selbstregulationstraining. Zusätzlich zu den Ergebnissen der vorherigen Studie konnten weitere positive Effekte zu Gunsten der Strategie- plus Selbstregulationsbedingung für die

Selbstwirksamkeitserwartung, das Fähigkeitsselbstkonzept und einer weiteren Transferaufgabe (Erlebniserzählung) nachgewiesen werden.

Diese Ergebnisse bestätigen und erweitern die bisherigen Erkenntnisse bezüglich des inkrementellen Effekts der Selbstregulation. Wie auch in den vorangegangenen Studien erleichterten selbstregulatorische Komponenten die Stabilität und Generalisierbarkeit der erlernten Strategien. Gleichzeitig gaben die Studien von Glaser und Kollegen einen deutlichen Hinweis darauf, dass selbstregulatorische Prozeduren die Effektivität eines reinen Strategietrainings hinsichtlich schreibbezogener Maße steigert.

Die durch das SAT-Programm erzielten Effektstärken für Schreibleistung, schreibbezogenes Wissen, Planen, Revidieren und Transfermaß Hörverständnis liegen in dem Bereich .61 und 4.48 (Glaser & Brunstein, 2007b) bzw. für Schreibleistung, Selbstwirksamkeitserwartung, Fähigkeitsselbstkonzept und Transfermaß Erlebniserzählungen in dem Bereich .32 und 1.13 (Glaser et al., 2009), wobei generell die Effektstärken beim Follow-Up leicht absanken. Insgesamt lässt sich somit ein moderater bis sehr großer Effekt des SAT-Programms auf die Schreibleistungen, schreibbezogenen Maße und auf motivationale Variablen feststellen.

Nichtsdestotrotz unterliegen die berichteten Effekte gewissen Einschränkungen. Bislang wurde das SAT-Programm nur von vorher intensiv geschulten Lerntrainern und in Kleingruppen von maximal sechs Schülern durchgeführt. Folglich fehlen Belege dafür, dass das SAT-Programm ebenfalls in gesamten Klassen und von Lehrern durchführbar ist. Dasselbe gilt für die Differenzierung zwischen Selbstregulation und Schreibstrategien. Die Trainer in den beiden dargestellten Studien wurden explizit in der Förderung selbstregulatorischer Kompetenzen sowie in den Instruktionmethoden der kognitiven Modellierung und des Scaffoldings geschult. Diese Prinzipien sind jedoch für Lehrer in deutschen Schulen ungewöhnlich. Folglich muss untersucht werden, ob Lehrer ebenfalls soweit vorbereitet werden können, um ein SAT-ähnliches Programm in den eigenen Klassen umzusetzen.

2.3.4 Zusammenfassung

Im vorangegangenen Abschnitt wurden Programme zur Förderung der Schreibkompetenz vorgestellt, wobei zwischen dem Schreibprozessansatz und der kognitiven Strategieinstruktion unterschieden wurde.

Der Schreibprozessansatz und dessen am häufigsten zitiertes Programm, der *Writer's Workshop*, fokussieren die am Schreiben beteiligten Prozesse. Eins der Hauptanliegen des *Writer's Workshops* besteht darin, den Schüler in den Mittelpunkt des Schreibunterrichts zu

stellen. Es erfolgen nur sehr wenige und kurze direkte Strategieinstruktionen durch den Lehrer, so dass die Schüler durch den Austausch mit anderen Schülern und durch die Freude am Schreibprozess das Schreiben erlernen.

Dagegen ist der zweite Ansatz, die kognitive Strategieinstruktion, als eine Erweiterung des *Writer's Workshops* anzusehen. Bei der kognitiven Strategieinstruktion wird zusätzlich zum Schreibprozess auch das Schreibprodukt betont. Folglich spielt der Lehrer eine zentrale Rolle. Dem Schüler wird durch explizite Strategievermittlung eine Vielzahl an unterschiedlichen Strategien zur Verfügung gestellt, so dass er selbst die für ihn adäquateste Strategie auswählen kann. Als Beispiele für den zweiten Ansatz wurden u.a. das SRSD- und das SAT-Programm vorgestellt, deren jeweilige Effektivität in mehreren Studien nachgewiesen wurde. Beiden ist dabei gemeinsam, dass die Vermittlung von Schreibstrategien explizit mit selbstregulatorischen Prozeduren verknüpft wird.

Ausgehend von der aktuellen Befundlage zeigte sich der zweite Ansatz sowohl bei durchschnittlichen als auch bei schwachen Schreibern als deutlich effektiver als der erste (Berninger et al., 2009; Troia, 2009a). Nichtsdestotrotz konnte bislang nicht hinreichend geklärt werden, welche der im „Gesamtpaket“ Schreibtraining enthaltenden kognitiven Strategien für die Effektivität verantwortlich sind. Die Studien von Glaser und Kollegen (Glaser, 2005; Glaser & Brunstein, 2007b; Glaser et al., 2009), die in Kleingruppen von Trainern durchgeführt wurden, erbrachten Hinweise für den inkrementellen Effekt von Selbstregulation. Bislang fehlen jedoch Studien, die die Überlegenheit selbstregulatorischer Schreibtrainings im Vergleich zu einem reinen Schreibstrategientraining unter den realen Bedingungen des Schulalltags nachweisen. Dies ist jedoch erforderlich, da die Metaanalyse von Graham und Harris (2003) zeigte, dass die Effekte einer Intervention in Abhängigkeit davon, ob Lehrer oder Trainer diese umsetzen, variieren.

2.4 Implementierung von Schreibprogrammen in den Regelunterricht

In den letzten Jahren ist in Deutschland die Qualität des Unterrichts und in diesem Zusammenhang die Professionalität der Lehrer immer stärker in den Fokus öffentlichen Interesses geraten. Diese Entwicklung ist auf vielfältige Veränderungen, wie die enttäuschenden Ergebnisse für Deutschland bei internationalen Schulleistungstudien (TIMMS, IGLU, PISA), die Konzeption von Bildungsstandards und von Instrumenten zur Qualitätssicherung in der Schule sowie die Einführung flächendeckender Vergleichsarbeiten (VERA) in den Klassenstufen drei, sechs und acht zurückzuführen (Helmke, 2009).

Dennoch existieren im Gegensatz zu den USA wenige Bemühungen, Interventionsprogramme, die sich in kontrollierten empirischen Studien als effektiv erwiesen haben, in den Unterrichtsalltag zu integrieren. Dies ist daran zu erkennen, dass die in den letzten Abschnitten diskutierten Programme fast ausschließlich in den USA durchgeführt wurden. Helmke (2009) erklärt die aktuelle unbefriedigende Situation damit, dass die Pädagogik in Deutschland traditionell geisteswissenschaftlich orientiert ist, der Praxisbezug in der Lehrerausbildung fehlt und die Schulen in Deutschland ein großes Maß an Autonomie besitzen, so dass die Verbreitung effektiver Fördermaßnahmen auf das Interesse der Schulen angewiesen ist.

Nichtsdestotrotz wurden spätestens seit TIMMS und PISA immer mehr Forderungen laut, die Qualität von Schulen und Unterricht zu optimieren, um dadurch die Schulleistungen der Schüler zu verbessern. Helmke (2009) stellt ein Rahmenmodell der Unterrichtsverbesserung vor, welches die Wirkmechanismen für den Erfolg aus Sicht des Lehrers skizziert (siehe Abbildung 13). Diese Perspektive ist bei der Implementierung neuer Unterrichtsmethoden wesentlich, da die Entscheidung, den eigenen Unterricht zu verbessern, fast ausschließlich durch die Lehrer selbst angeregt wird und selten von außen verordnet wird. Darüber hinaus ist der Erfolg einer Implementierung wahrscheinlicher, wenn der Lehrer sich bewusst für eine Verbesserung des Unterrichts entscheidet. Dem Modell von Helmke (2009) zufolge löst die Feststellung eines Lehrers, dass der aktuelle Unterricht nicht den gewünschten Kriterien guten Unterrichts entspricht, eine Sequenz von Denkprozessen und Verhaltensweisen aus, die erheblich von individuellen und externen Bedingungsfaktoren beeinflusst wird (siehe Abbildung 13).

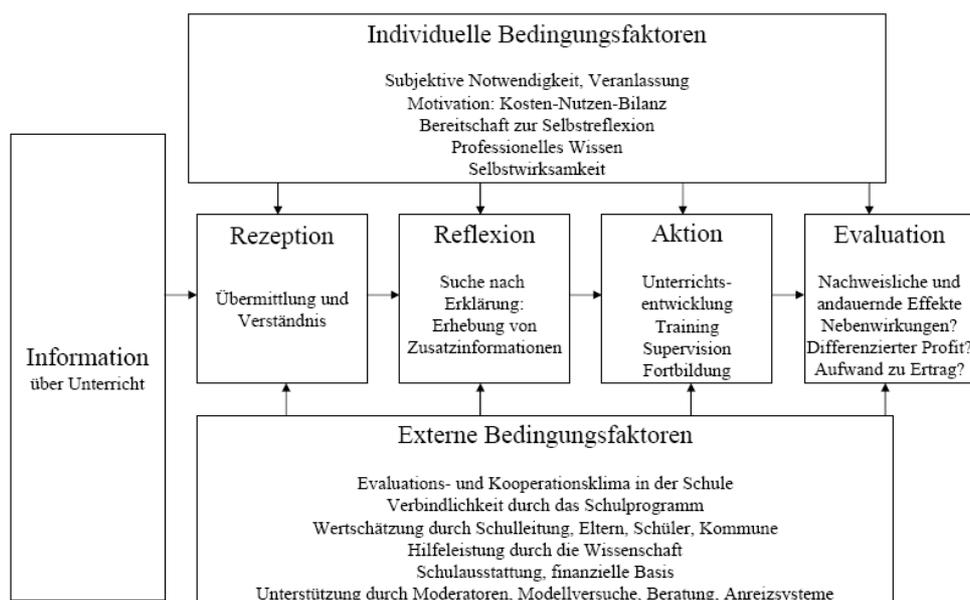


Abbildung 13: Rahmenmodell der Verbesserung des Unterrichts, adaptiert von Helmke (2009, S. 309).

Aus Sicht der Forschung ergeben sich zwei Ansatzpunkte, auf die Unterrichtsverbesserung einzuwirken. Eine Möglichkeit besteht in der direkten Einflussnahme auf den Prozess der *Aktion* durch Erweiterung des Repertoires der Lehrer an effektiven Unterrichtsmethoden. Die zweite Möglichkeit stellt eine indirekte Wirkung dar, indem die *externen und individuellen Bedingungsfaktoren* verändert werden. Dies kann z.B. die Veränderung des Schulklimas durch Veröffentlichungen von Ergebnissen von Schulleistungsvergleichen sein.

2.4.1 Implementationsforschung

Allgemein wird unter Implementation „die Umsetzung von wissenschaftlichen Erkenntnissen in die gesellschaftliche Praxis“ verstanden (Euler & Sloane, 1998, S. 312). Der Begriff der Implementation, wie er im Folgenden verwendet wird, schließt zudem die Vorbereitung des potenziellen Anwenders dieser Erkenntnisse, d.h. des Lehrers, mit ein. Es geht somit um die Frage, wie ein theoretisches Konzept in praktisches Handeln, ausgeführt durch Lehrer, transferiert werden kann.

In der Schule findet Implementation regelmäßig in der Form statt, dass curriculare Vorgaben in den Unterricht integriert werden. Jede Implementation erfolgt dabei nach einer bestimmten Strategie, welche wiederum in zwei verschiedenen Richtungen stattfinden kann. Zum einen können zunächst curriculare Inhalte konzipiert und im Anschluss daran so didaktisch gestaltet werden, dass sie in der Praxis umsetzbar sind. Diese Strategie wird als Top-down-Strategie bezeichnet. Zum anderen bedeutet eine symbiotische Strategie, dass ausgehend von den Anforderungen und Bedürfnissen der Praxis Konzepte entwickelt werden, die diesen gerecht werden.

Diese beiden Implementierungsstrategien werden in diesem Abschnitt vorgestellt. Zuvor wird jedoch ein neuer Forschungsansatz, *Educational Design Research* (EDR), vorgestellt, welcher sowohl Aspekte der Top-down- als auch der symbiotischen Strategie integriert: ausgehend von einer Theorie wird diese in die natürliche Umwelt integriert und gleichzeitig verändert. Der Wert einer bestimmten Theorie wird folglich durch die Konsequenzen in der natürlichen Umwelt beurteilt (Barab & Squire, 2004).

2.4.1.1 Educational Design Research (EDR)

Der EDR-Ansatz stellt eine Möglichkeit zur Entwicklung und Implementierung von Förderprogrammen in den Schulalltag dar. In den letzten Jahren gewann EDR eine zunehmende Bedeutung in der pädagogischen Forschung. Im Gegensatz zur symbiotischen und der Top-down-Strategie (siehe Kapitel 2.4.1.2) geht dieser Ansatz sowohl über die

alleinige Entwicklung einer neuen Intervention als auch über das alleinige Testen bestehender Interventionen hinaus (Design-Based Research Collective, 2003).

EDR ist keine einzelne Strategie, sondern vielmehr ein Set an Ansätzen mit teilweise unterschiedlichen Zielsetzungen. Gemeinsame Intention aller Studien, die auf einem EDR-Ansatz basieren, ist die Entwicklung neuer Theorien und Praktiken, welche das Lernen und Unterrichten in lebensnahen Situationen erklären und gravierend beeinflussen können (Barab & Squire, 2004). Insgesamt ergeben sich folgende charakteristische Eigenschaften des EDR-Ansatzes (Barab & Squire, 2004; van den Akker, Gravemeijer, McKenney, & Nieveen, 2006):

- *Interventionsorientiert*: Das Hauptziel des Ansatzes ist die Entwicklung einer Interventionsmethode, die in der natürlichen Umgebung (d.h. Klassenzimmer) umsetzbar ist.
- *Iterativ*: Der Ansatz umfasst einen Zyklus von Design, Evaluation und Überarbeitung, so dass immer wieder Anpassungen aufgrund von Erfordernissen des Kontexts vorgenommen werden.
- *Prozessorientiert*: Die Verbesserung von Interventionen und das Verständnis für Implementierungshindernisse stehen im Vordergrund.
- *Orientiert am Nutzen*: Der durch das Design gewonnene pädagogische Wert wird anhand der Nützlichkeit für den Anwender in realen Kontexten bewertet.
- *Theorieorientiert*: Das Design basiert auf theoretischen Konstrukten und leistet gleichzeitig einen Beitrag zur Erweiterung und Entwicklung neuer Theorien.

In Bezug zur Implementierung von Curricula zielt der EDR-Ansatz auf drei Ergebnisse ab (McKenney, Nieveen, & van den Akker, 2006):

- (1) Identifikation von *Design Principles*, womit neue Theorien, Heuristiken und Wissen gemeint sind;
- (2) Entwicklung von *Curricular Products*, d.h. von Unterrichtsprogrammen sowie
- (3) Verbesserung der *Professional Development*, speziell der Lehrerausbildung und -fortbildung.

Diese Ziele liegen in der ursprünglichen Motivation der *Design-Based Research*-Forscher begründet. Diese kritisieren die bisherige Bildungsforschung dahingehend, dass diese zu keinen nennenswerten großen Erfolgen, die sich im Schulunterricht widerspiegeln, geführt hat (Walker, 2006). Somit intendiert EDR die Etablierung eines alternativen Forschungsansatzes, welcher gegensätzlich zu traditionellen Forschungsmethoden ist. In Tabelle 7 sind die Unterschiede zwischen psychologischen Experimenten und der Forschung basierend auf dem *Design-Based Research* gegenübergestellt.

Tabelle 7: Vergleich psychologischer Experimente mit dem Design-Based Research-Ansatz, adaptiert von Barab und Squire (2004, S. 4).

Kategorien	Psychologische Experimente	Design-Based Research
Ort der Untersuchung	Durchgeführt in Laboren	Durchgeführt in realen lebensechten Situationen
Komplexität der Variablen	Wenige abhängige Variablen	Multiple abhängige Variablen, einschließlich Klimavariablen (z.B. Mitarbeit der Schüler), Erfolgsvariablen (z.B. Lerninhalt, Transfer), Systemvariablen (z.B. Verbreitung, Nachhaltigkeit)
Fokus der Forschung	Identifikation einiger essenzieller Variablen und Konstanthaltung dieser Variablen	Charakterisierung der Komplexität der Situation
Eingesetzte Prozeduren	Vorher festgelegte Prozeduren	Flexible Designüberarbeitungen; das anfängliche Design wird in Abhängigkeit von dessen Erfolg revidiert
Ausmaß sozialer Interaktionen	Isolierung der Lerner, um Interaktionen zu kontrollieren	Komplexe soziale Interaktionen mit Teilnehmern, Austausch von Ideen, etc.
Eigenschaften der Befunde	Hypothesentesten	Beschreibung der Eigenschaften des in der Praxis umgesetzten Designs
Rolle der Teilnehmer	Teilnehmer werden als Untersuchungsobjekte angesehen	Unterschiedliche Teilnehmer mit unterschiedlicher Expertise werden in der Entwicklung und Analyse des Designs mit eingebunden

Die Besonderheiten des EDR-Ansatzes stellen gleichzeitig dessen Hindernisse dar. EDR unterliegt mehreren Dilemmata (Barab & Squire, 2004; Design-Based Research Collective, 2003; McKenney et al., 2006):

Reale Lernumgebungen, in denen eine Vielzahl an nicht kontrollierbaren Störfaktoren auftreten, sind hoch komplex, so dass neue Interventionsmaßnahmen schwerer durchführbar sind. Auch wenn eine lebensnahe Situation am besten die Realität abbildet und folglich die Validität der Aussage über die Effektivität einer Interventionsmaßnahme steigert, ist die Replizierbarkeit und Generalisierbarkeit der Befunde auf andere Kontexte schwierig. Darüber hinaus werden viele der EDR-Studien über einen langen Zeitraum von mehreren Jahren durchgeführt, so dass die Aufrechterhaltung der Motivation zur Mitarbeit auf Seiten der

Lehrer sowie ebenfalls logistische Schwierigkeiten erschwert sind. In der Schule wie auch in der Wissenschaft werden überzeugende, schnell erzielbare Ergebnisse erwartet, so dass eine Verpflichtung zu einer langjährigen Mitarbeit zunächst auf Skepsis trifft. Gleichzeitig ist die Kooperation zwischen Akteuren der Schule und der Wissenschaft für den Erfolg des Designs essenziell.

Daran anknüpfend ergibt sich ein weiteres Dilemma, welches in der proklamierten Adaptationsvielfalt liegt. Ein sich ständig veränderndes Design ist nicht nur schwach in seiner Aussagekraft, sondern erfordert zugleich vom Forscher ein hohes Maß an Flexibilität, Organisationstalent und kommunikativen Fertigkeiten.

Die Rolle des Forschers ist schließlich der letzte Punkt, der zu erheblicher Kritik am EDR-Ansatz führt. In der Regel ist der Forscher nicht nur derjenige, der das Design entwirft, sondern er implementiert und beurteilt gleichzeitig das Design. Der Einfluss der Anwesenheit des Forschers in der Untersuchungssituation kann zu vielfältigen Verzerrungen führen. Dazu zählen der Hawthorne-Effekt (Beteiligung an einer Studie kann den Teilnehmern das Gefühl geben, etwas Besonderes zu sein, und somit deren Verhalten verändern), das Hypothesenerraten (Teilnehmer versuchen die Erwartungen des Forschers vorherzusagen und reagieren entsprechend) sowie Diffusion (Wissen über das Design beeinflusst andere Teilnehmer). Auch wenn es Vorteile mit sich bringt, wenn der Designer ebenfalls die Intervention durchführt und somit ein tieferes Verständnis über die Schwierigkeiten und Stärken des Designs erlangt, sind Beurteilereffekte nicht ausgeschlossen. Darüber hinaus wird der Designer, wahrscheinlich unbewusst, weniger offen für Kritik sein. Diese unerwünschten Effekte könnten größtenteils dadurch gelöst werden, dass die Durchführung einer neuen Maßnahme nicht von dem Designer selbst, sondern von unabhängigen Personen erfolgt.

Zusammenfassend ist EDR als ein interessanter Forschungsansatz zu bewerten, der jedoch noch in den Anfängen steckt. Abgesehen von den geschilderten Kritikpunkten sind die unternommenen Versuche, methodische Leitlinien und Kriterien zur Überprüfung der Qualität eines implementierten EDR-Ansatzes zu entwickeln, sehr unpräzise und heterogen (McKenney et al., 2006; Walker, 2006).

Dennoch kann EDR ein hilfreiches Instrument sein, effektive und umsetzbare Innovationen für den Unterrichtsalltag zu generieren. Für das vorliegende Dissertationsprojekt ist dies von besonderem Interesse, da durch EDR konzipierte Unterrichtsmethoden mit theoretisch und empirisch fundierten Interventionsprogrammen, wie dem *Selbstregulatorischen Aufsatztraining*, verglichen werden könnten. Walker (2006) vertritt die umstrittene Auffassung, dass die Notwendigkeit für EDR vor allem deswegen besteht, weil

bislang Studien, die die Effektivität von wissenschaftlich basierten Unterrichtsmethoden mit konventionellen Methoden verglichen, keinen signifikanten Unterschied zeigten. Gleichzeitig erfordert die Entwicklung einer Unterrichtsmethode nach Prinzipien der EDR einen großen zeitlichen und personellen Aufwand, da diese Projekte über mehrere Jahre hinweg angelegt sind. Für die vorliegende Arbeit scheinen deswegen die symbiotische und die Top-down-Strategie effizientere Methoden zur Implementierung des intendierten Aufsatztrainings zu sein.

2.4.1.2 Implementierungsstrategien

Ende der 60er bzw. Anfang der 70er Jahre verbreitete sich unter Forschern die Erkenntnis, dass ausbleibende Erfolge bei Interventionen nicht unbedingt auf die Ineffektivität der Maßnahme schließen lassen, sondern auch auf eine unzureichende Implementierung zurückzuführen sind (Snyder, Bolin, & Zumwalt, 1992). Befunde, die dies bestätigten, bewirkten die Entwicklung von Forschungsarbeiten in dem Bereich Curriculum Implementierung (Fullan & Pomfret, 1977; Snyder et al., 1992).

Im weiteren Verlauf kristallisierten sich zwei gegensätzliche Implementierungsansätze heraus. Snyder (1992) unterschied zwischen *Fidelity Perspective* und *Curriculum Enactment*. Die meisten der damaligen Interventionsstudien verfolgten eine *Fidelity Perspective*, bei der es darum ging, wie genau eine von außen entwickelte Innovation wie vorgesehen implementiert wurde und welche Faktoren die Implementierung erleichtern bzw. erschweren. Dagegen war das Forschungsfeld im Bereich *Curriculum Enactment* noch sehr klein. Hierbei entwickelten Lehrer und Schüler gemeinsam mit Experten ein Unterrichtsprogramm, wobei sie auf ihre eigenen Erfahrungen und Bedürfnisse zurückgriffen. Ursprünglich unterschied Snyder (1992) noch eine dritte Variante, die *Mutual Adaptation*, die zwischen der *Fidelity Perspective* und *Curriculum Enactment* angesiedelt wird und die Weiterentwicklung von wissenschaftsbasierten Interventionen durch Lehrer beinhaltet.

Gräsel und Parchmann (2004) griffen auf die Taxonomie von Snyder (1992) zurück und entwickelten diese weiter. Sie differenzierten zwischen der symbiotischen Strategie (siehe *Curriculum Enactment*) und der Top-down-Strategie (siehe *Fidelity Perspective*). Im Folgenden werden die Charakteristika sowie wichtige Einflussfaktoren der beiden Strategien vorgestellt.

Symbiotische Strategie: Mit dem Begriff symbiotische Strategie ist gemeint, dass Experten und Lehrer gemeinsam Lösungsansätze für die in der Unterrichtspraxis identifizierten und als relevant beurteilten Probleme entwickeln und erproben (Gräsel & Parchmann, 2004).

Die symbiotische Implementierungsstrategie verfolgt drei Intentionen: (1) Verbesserung der als defizitär bewerteten Ausgangslage; (2) Etablierung von langfristigen Kooperationsstrukturen zwischen Lehrern, Schulen und Wissenschaftlern; (3) Verbreitung der neuen Methoden unter weiteren Schulen und Lehrern (Dissemination) (Gräsel & Parchmann, 2004). Der Erfolg der Implementation wird demnach nach dem Ausmaß, wie gut die drei Ziele erreicht wurden, bewertet. Darüber hinaus strebt die symbiotische Strategie zudem die Vermittlung von Handlungskompetenzen an, die die Lehrer befähigt, sich selbst weiterzuentwickeln und adaptiv auf neue Problemlagen zu reagieren.

Gleichzeitig gehen verschiedene Schwierigkeiten mit den Zielen einher. Die Rolle des Wissenschaftlers und des Lehrers ist gleichrangig und das jeweilige Wissen wird für den Erfolg der Implementation als gleich wichtig angesehen (Gräsel & Parchmann, 2004). Aus diesem Grund setzt die symbiotische Strategie ein erhebliches Fachwissen und didaktische Kenntnisse von Seiten der Lehrer sowie die Fähigkeit, den eigenen Unterricht zu analysieren, voraus. Die Eigenschaften und Fähigkeiten des Lehrers stellen somit eine wichtige Einflussgröße dar. Infolgedessen empfiehlt es sich, der eigentlichen Implementierungsphase eine Qualifizierungsphase voranzustellen, in der Lehrern Informationen praxisnah vermittelt werden (Herzmann, Sparka, & Gräsel, 2006; Souvignier & Trenk-Hinterberger, 2007).

Des Weiteren können sich Kommunikationsschwierigkeiten zwischen Wissenschaftlern und Lehrern ergeben, die durch ein unterschiedliches Vokabular, aber auch durch unterschiedliche Zielvorstellungen, Rollenverständnisse und Überzeugungen entstehen (Gräsel & Parchmann, 2004). Für die Durchsetzung der symbiotischen Strategie ist die Akzeptanz der neuen Maßnahmen auf Seiten der Lehrer entscheidend, so dass diese davon überzeugt werden müssen, dass die Maßnahme praktikabel und wirksam ist (Souvignier & Trenk-Hinterberger, 2007).

Darüber hinaus könnte es schwierig sein, eine symbiotisch entwickelte Innovation in allgemeine Bildungsreformen zu integrieren und dadurch insgesamt die Kooperation und Kommunikation zwischen Schulen zu verbessern (Gräsel & Parchmann, 2004). Die in der Praxis identifizierten Probleme können von den Zielvorstellungen der Entscheidungsträger in der Bildungspolitik abweichen. Auch ist die Veränderung einer bestehenden Schulkultur erst dann als erfolgversprechend einzustufen, wenn mehrere Lehrer einer Schule engagiert an der Implementation beteiligt sind.

Als Hauptkritik an der symbiotischen Strategie ist jedoch die Frage nach der tatsächlichen Verbesserung der Unterrichtsqualität anzuführen. Konzepte, die von Lehrern und Wissenschaftlern gemeinsam ausgearbeitet wurden, müssen nicht zwangsläufig effektiv

sein und können aufgrund der fehlenden theoretischen Fundierung sogar negative Auswirkungen mit sich bringen. Einzelne qualitative Studien lieferten entsprechende widersprüchliche Ergebnisse. In einigen Studien wurden von positiven Bewertungen der Strategie durch Lehrer berichtet, während andere Gefühle der Überforderung und der zusätzlichen Belastung auf Seiten der Lehrer schilderten (Gräsel & Parchmann, 2004). Zentral für die Zufriedenheit der Lehrer sowie für die erfolgreiche Umsetzung der Programminhalte sind eine ausreichende Unterstützung der Lehrer durch die Wissenschaftler sowie eine Vermittlung theoretischer Inhalte, die für die Schulpraxis relevant sind (Souvignier & Trenk-Hinterberger, 2007).

Ein Beispiel für ein Projekt, in dem zusammen mit Lehrern erfolgreich eine neue Unterrichtsmethode implementiert wurde, ist die von der *Cognition and Technology Group at Vanderbilt* (CTGV) entwickelte Jasper-Serie, die problemorientiertes Lernen im Schulunterricht ermöglicht. Die Jasper-Serie umfasst zwölf Videos mit unterschiedlichen Hauptpersonen, die einem konkreten und realistischen Problem ausgesetzt sind, für das Schüler ab der fünften Klasse unter Heranziehung von mathematischen Fertigkeiten eine Lösung finden müssen. Das Hauptanliegen der Vanderbilt-Gruppe bestand darin, das Problem des trägen Wissens, welches häufig in der Schule produziert wird, zu überwinden (Cognition and Technology Group at Vanderbilt, 1990). Das Jasper-Projekt weist Parallelen zur symbiotischen Strategie auf, da von Beginn des Projekts an Lehrer aus der Praxis in die Entwicklung mit einbezogen wurden (Reinmann-Rothmeier & Mandl, 1998). Darüber hinaus war das Projekt explizit daran interessiert, den Erfordernissen des Schulalltags gerecht zu werden. Weiter war die Etablierung einer *Community of Learners* ein besonderes Ziel der Forschergruppe. Im Jahr 1990 wurde Jasper in 16 verschiedenen Schulen in neun Bundesländern der USA implementiert. Dabei durchliefen die teilnehmenden Lehrer zunächst ein zweiwöchiges Training, bevor sie im Anschluss Jasper in ihren eigenen Klassen umsetzten. Eine der wichtigsten Erfahrungen der CTGV war es, dass eine kontinuierliche Begleitung der Lehrer und eine regelmäßige Kommunikation zwischen Wissenschaftlern und Forschern für den Erfolg der Implementierung notwendig sind (Reinmann-Rothmeier & Mandl, 1998). Insgesamt erwiesen sich die Jasper-Serien als erfolgreich in der Verbesserung der mathematischen Fertigkeiten und der mathematischen Einstellungen sowie hinsichtlich der subjektiven Zufriedenheit der Lehrer mit dem Projekt (Cognition and Technology Group at Vanderbilt, 1992). Die positiven Effekte von Jasper auf die Matheleistung konnten in einer weiteren Studie von Hickey, Moore und Pellegrino (2001) bestätigt werden. Ferner berichten

die Autoren, dass die Materialien von den Lehrern häufig nicht so wie geplant eingesetzt wurden.

Ein weiteres Projekt, welches explizit die Verfolgung der symbiotischen Strategie anstrebte, wurde von Herzmann, Sparka und Gräsel (2006) durchgeführt. Dabei handelt es sich um ein mehrjähriges Projekt, in dem von Lehrern und Wissenschaftlern gemeinsam ein Lesetraining für die Hauptschule konzipiert wurde. Es wurden sowohl quantitative (Erfassung der Leseleistung der Schüler) als auch qualitative Daten (Lehrer-Interviews) erhoben. Nach einem halben Jahr zeigten sich jedoch keine Unterschiede zwischen Schülern, die die neue Unterrichtsmethode erhielten und Schülern, die kein Training durchliefen. Die Autoren erklären ihre Ergebnisse dahingehend, dass sich erst langfristig Verbesserungen zeigen würden (Herzmann et al., 2006).

Insgesamt ist festzustellen, dass abgesehen von den beiden diskutierten Projekten nur wenige empirische Studien vorliegen, welche die Effektivität symbiotischer Implementierungsstrategien evaluierten. Nichtsdestotrotz kann die symbiotische Strategie eine sinnvolle Methode zur Entwicklung von Innovationen darstellen, wenn mehrere Lösungsansätze sinnvoll erscheinen und diese stark kontextabhängig sind (Gräsel & Parchmann, 2004)

Top-down-Strategie: Die Top-down-Strategie impliziert, dass die Implementierung neuer Unterrichtskonzepte oder anderer Innovationen durch einen externen Experten erfolgt. Demnach werden die Ziele, Inhalte und Methoden des Unterrichtskonzepts sowie die Kriterien für den Erfolg von einem Experten erarbeitet. Die eigentliche Umsetzung der Innovation erfolgt jedoch durch Lehrer oder sonstige Adressaten der Maßnahme, die in entsprechenden Fortbildungen in der Durchführung geschult werden. Somit besteht sowohl eine personelle als auch zeitliche Abgrenzung zwischen der Konzeption und der Umsetzung der Innovation (Gräsel & Parchmann, 2004). Damit eine Top-down-Strategie Aussicht auf Erfolg hat, sollte diese in Voruntersuchungen ausreichend in verschiedenen Kontexten und bei Schülern mit unterschiedlichen Leistungsniveaus empirisch überprüft worden sowie theoretisch fundiert sein (Gräsel & Parchmann, 2004).

Die Top-down-Strategie, d.h. die Einführung von neuen Methoden von oben nach unten, bedeutet ferner, dass der Implementierungserfolg mit der Evaluation der Effekte der Innovation einhergeht. Folglich ist die Gewährung der Implementationsgenauigkeit notwendig, um aussagekräftige Schlussfolgerungen aus den Ergebnissen ziehen zu können. Der Erfolg der Umsetzung wird dabei von drei grundlegenden Einflussvariablen bestimmt: Lehrer, Schule und Experte.

Zunächst spielen die Einstellungen und Erwartungen des Lehrers eine wesentliche Rolle bei der Implementierung neuer Unterrichtsmethoden. In der Regel verfügen Lehrer über Handlungsrountinen und Überzeugungssysteme, die seit vielen Jahren bestehen und für ihre Unterrichtspraxis handlungsleitend sind. Folglich können Innovationen zunächst zu Ablehnungen von Seiten der Lehrer führen, vor allem wenn diese die Bedürfnisse und Rahmenbedingungen der Praxis nicht widerspiegeln. Zum Beispiel würden zeitlich und finanziell aufwendige Projekte zu Themengebieten, die aus Sicht der Lehrer bereits durch das bestehende Material gut abgedeckt werden, auf Widerwillen und Kritik stoßen. Infolgedessen erhöht sich der Erfolg einer Implementation entscheidend, wenn die Lehrer die neue Unterrichtsmethode als „nützlich, sinnvoll, realisierbar, wichtig usw.“ beurteilen (Gräsel & Parchmann, 2004, S. 203). Um dies sicherzustellen, entwickelten Blumenfeld, Fishman, Krajcik und Marx (2000) das Rahmenmodell CERA als Leitfaden für die Implementierung von Unterrichtskonzepten: *Collaborative construction of understanding, Enactment of new practices in classrooms, Reflection on practice, Adaption of materials and practices*. Demnach sollte nicht nur bei der symbiotischen, sondern auch bei der Top-down-Strategie den Lehrern Freiräume für Adaptationen gelassen und Mitspracherechte eingeräumt werden. Darüber hinaus ist es sinnvoll, wenn sich durch die geplanten Innovationen zwar für die Lehrer spürbare Veränderungen ergeben, diese jedoch nicht völlig von den gewohnten Routinen und Erfahrungen der Lehrer abweichen (Gräsel & Parchmann, 2004). Subjektive Theorien sowie die Selbstwirksamkeitserwartung der Lehrer werden ebenfalls als wichtige Einflussgrößen für die Umsetzung von neuen Unterrichtsmethoden diskutiert und folglich in vielen Implementationsmaßnahmen integriert (Timperley & Phillips, 2003).

Gemäß dem *Concern-Based Adaptation Model* (CBAM) äußern Lehrer in verschiedenen Phasen der Implementierung Befürchtungen, die zu Widerständen führen und die jeweilige Umsetzung der Methoden behindern können. Diese Befürchtungen betreffen am Anfang die eigene Person (z.B. „*Kann ich die Methoden umsetzen?*“) und wechseln im weiteren Verlauf der Implementierung zu Befürchtungen, die die Aufgabe (z.B. „*Lohnt sich die investierte Zeit für den Nutzen, den ich aus der Methode ziehen werde?*“) und andere Bereiche wie die Zusammenarbeit mit Kollegen, Eltern etc. (z.B. „*Werden meine Kollegen die Veränderungen mittragen?*“) betreffen (van den Berg & Ros, 1999).

Abgesehen von dem Lehrer ergeben sich außerdem Umsetzungshindernisse von Seiten der Schule als sozialer Kontext, in dem die Innovationen stattfinden. Neben dem Einfluss auf organisatorische Rahmenbedingungen wie die Einteilung der Lehrer zu bestimmten Klassen und die Erstellung des Stundenplans beeinflusst die Schule bzw. der Schulleiter als

Transformational Leadership die Anstrengungen und Bereitschaft von Lehrern, Innovationen umzusetzen (Geijsel, Slegers, Leithwood, & Jantzi, 2002; Geijsel, Slegers, van den Berg, & Kelchtermans, 2001). In Large-Scale-Untersuchungen konnten Geijsel und Mitarbeiter (2001, 2002) vor allem den Einfluss von *Vision Building* (wahrgenommene Einbindung bei der Etablierung von Zielen und Visionen für die Schule) und *Intellectual Stimulation* (wahrgenommene Unterstützung durch den Schulleiter bei der eigenen fachlichen Ausbildung) auf die Zustimmung der Lehrer zur Innovation sowie auf die tatsächliche Umsetzung in den eigenen Unterricht nachweisen (Geijsel et al., 2002; Geijsel et al., 2001).

Schließlich spielt der Experte, der die Innovation konzipiert und die Implementierung plant, für den Erfolg einer Implementation eine entscheidende Rolle. Vor allem ist ein genaues methodisches Vorgehen vor, nach und während der Implementierungsphase erforderlich. Vor der Implementierung sind die Entwicklung von geeigneten Hilfsmaterialien als Hilfestellung für die Lehrer sowie die Erfassung der Bedürfnisse der Lehrer und die Gegebenheiten der jeweiligen Schulen notwendig. Während der Implementierung muss eine ausreichende Schulung und Unterstützung der Lehrer gewährleistet werden. Weiter sind hier verschiedene Kontrollsysteme zur Sicherung der Implementationsgenauigkeit, d.h. die Übereinstimmung zwischen der extern geplanten und der umgesetzten Unterrichtsmethode, entscheidend. Dieser Punkt wird im nächsten Abschnitt genauer diskutiert. Um einen langfristigen Erfolg der Implementierung zu garantieren, muss nach der Implementierung sichergestellt werden, dass die Lehrer ausreichend vorbereitet sind, um ohne jegliche Unterstützung und Begleitung, die Unterrichtsmethode weiter einzusetzen. Somit ist eine Begleitung der Lehrer über die Einstiegsphase hinaus wünschenswert (Blumenfeld et al., 2000).

2.4.2 Standards zur Sicherung der Qualität der Implementierung

Bereits Campbell (1969) beschäftigte sich mit der Aussagefähigkeit von Studien, die die Evaluation von Programmen untersuchten und forderte einen „*truly experimental approach to social reform*“ (Campbell, 1969, S. 409). In diesem Zusammenhang verlangte er die Berücksichtigung der internen und externen Validität von Implementierungsmaßnahmen und arbeitete verschiedene Bedrohungen der Validität heraus. Die Einhaltung der Validität stellt eine wichtige Voraussetzung dar, um die Auswirkungen einer Maßnahme interpretieren zu können. Dazu müssen jedoch zunächst die Effekte, d.h. die Wirksamkeit einer Maßnahme, überprüft werden. Wie in dem vorangegangenen Abschnitt dargestellt, geht bei der Top-down-Strategie der Erfolg einer Implementierung, gemessen an der Effektivität des

Programms, mit der Qualität der Implementierung einher. Aus diesem Grund werden diese beiden Aspekte ebenfalls in dem nächsten Abschnitt diskutiert.

2.4.2.1 Interne und externe Validität der Implementierung

Mit interner Validität sind Gütekriterien der empirischen Untersuchung gemeint, die eine verlässliche Schlussfolgerung darüber zulassen, dass die hervorgerufenen Veränderungen in den abhängigen Variablen tatsächlich durch das eingesetzte Treatment, d.h. durch die Variation der unabhängigen Variablen und nicht durch andere Faktoren der Untersuchung verursacht wurden (Campbell, 1969; Häcker & Stapf, 2004). Im Gegensatz dazu bezieht sich die externe Validität auf die Generalisierbarkeit der Ergebnisse auf andere Situationen, Personen, Zeiten und Operationalisierungen (Campbell, 1969; Häcker & Stapf, 2004). Auch wenn sowohl die interne als auch die externe Validität niemals vollständig kontrolliert werden können, ist es durch verschiedene Maßnahmen möglich, eine hohe interne und externe Validität der Untersuchung zu garantieren. Diese Maßnahmen werden im Folgenden separat für interne und externe Validität dargestellt.

Interne Validität: Zur Gewährleistung einer hohen internen Validität sollten in experimentellen Studien potenzielle Störfaktoren kontrolliert werden. Dies geschieht in der Regel zum einen durch Randomisierung und zum anderen durch eine geeignete Kontrollgruppe (Campbell, 1969).

Mit Randomisierung ist die zufällige Zuordnung der Versuchspersonen zu den verschiedenen Bedingungen gemeint. Erfolgt eine vollständig randomisierte Zuteilung, wird von einem „wahrhaft“ experimentellen Design gesprochen. Es kann jedoch vorkommen, dass eine Randomisierung ethisch nicht vertretbar oder praktisch nicht umsetzbar ist. Vor allem wenn mit natürlich vorkommenden Stichproben, z.B. Schulen und Klassen, gearbeitet wird, ist eine zufällige Verteilung der Probanden auf Experimental- und Kontrollbedingung nicht möglich. Schüler aus verschiedenen Klassen können nicht neu zusammengesetzt werden, da dieses Vorgehen neue Störfaktoren produzieren würde und organisatorisch sowie ethisch nicht akzeptabel wäre. In diesem Fall empfiehlt sich ein Matching der Klassen hinsichtlich relevanter Eigenschaften wie z.B. Rechtschreibfertigkeiten, Alter etc. Darüber hinaus können retrospektiv etwaige Unterschiede zwischen den Bedingungen in den Prätestleistungen statistisch kontrolliert werden.

Campbell (1957, 1969) berichtet von Gefahren für die interne Validität, die z.B. auf natürliche Reifungsprozesse der Stichprobe, Regressionsartefakte aufgrund von extremen Merkmalsausprägungen, Sterblichkeit der Stichprobe sowie unreliablen Messungen zustande kommen. Aus diesem Grund betont er, dass die Erhebung von Prätest-Werten sowie die

Auspartialisierung des Einflusses der Prätest-Werte für eine hohe interne Validität unabdingbar sind. Insbesondere die Vergleichbarkeit der Interventions- und Kontrollgruppe hinsichtlich der Prätest-Werte sollte gewährleistet werden.

Die Konzeption einer geeigneten Kontrollgruppe stellt eine weitere wichtige Maßnahme zur Garantierung einer hohen internen Validität dar. Eine Randomisierung schließt nämlich eine Variablenkonfundierung nicht aus (Hager, Patry, & Brezing, 2000). Wird z.B. keine Kontrollgruppe oder eine untrainierte Kontrollgruppe verwendet, sind die Ergebnisse nicht alleinig durch die unabhängigen Variablen erklärbar, sondern können zudem durch normale Reifungsprozesse oder durch nicht-programmgebundene Faktoren (z.B. Zuwendung) zustande gekommen sein. Bei der Wahl einer geeigneten Kontrollgruppe ist es dabei wichtig, alle Faktoren, die für die Hypothese relevant sind, zwischen den Experimental- und Kontrollbedingungen zu variieren, während potenzielle Störfaktoren kontrolliert werden müssen (Hager et al., 2000). Eine angemessene Kontrollgruppe sollte außerdem so geplant werden, dass die Wünsche und Bedürfnisse der beteiligten Schulen berücksichtigt werden (Jaycox, McCaffrey, Ocampo, Shelley, Blake, Peterson et al., 2006). Aus diesem Grund ist eine untrainierte Kontrollgruppe oder eine Placebo-Gruppe aus Sicht der Schulen nicht vertretbar, wenn die Experimentalgruppe als ein Gewinn für die Schüler angesehen wird. Eine Wartelistekontrollgruppe erschwert jedoch wiederum die Erhebung von Langzeiteffekten (Jaycox et al., 2006).

Externe Validität: Ursprünglich wurde die Sicherung der internen Validität über die Sicherung der externen Validität gestellt, wobei diese Ansicht insbesondere im Hinblick auf die Implementierung und Dissemination von Trainingsprogrammen revidiert werden muss (Shadish, 2002). Die weitläufige Verbreitung eines Trainingsprogramms in mehreren Schulen und Staaten wird nur dann erfolgen, wenn sich das theoretisch fundierte Programm ebenfalls in mehreren unterschiedlichen Schulen, mit unterschiedlichen Schülergruppen und in verschiedenen Orten als effektiv erwiesen hat (O'Donnell, 2008).

Die Generalisierbarkeit der Ergebnisse bezieht sich hauptsächlich auf die Person und auf die Situation. Hinsichtlich der Person sollten Studien eine große Bandbreite an Personen mit unterschiedlichen Eigenschaften, Fertigkeiten und Einstellungen erfassen. Aus diesem Grund sollten aus der Gesamtstichprobe Subgruppen gebildet werden und die statistischen Analysen für die Subgruppen separat durchgeführt werden, um zu prüfen, ob die Effekte einer Maßnahme für sämtliche Mitglieder einer Stichprobe zutreffen. Darüber hinaus melden sich Teilnehmer einer Untersuchung in der Regel freiwillig dazu bereit, an einer Untersuchung

teilzunehmen, so dass häufig die Ergebnisse nur auf die speziell untersuchte Gruppe und nicht zwangsläufig auf nicht-freiwillige Personen generalisiert werden können.

Bezüglich der Situation ergeben sich ähnliche Schlussfolgerungen. Auch hier sollten variierende Situationen und Orte berücksichtigt werden. Eine Erhöhung der externen Validität wird dadurch erreicht, dass die Untersuchung in realen Situationen und unter normalen Bedingungen durchgeführt wird, so dass die so erzielten Ergebnisse leichter in den Alltag generalisierbar sind als die Ergebnisse von Laborexperimenten (Jaycox et al., 2006).

Insgesamt ermöglichen erst viele unterschiedliche Studien die Generalisierbarkeit von Ergebnissen. Metaanalysen stellen dabei ein hilfreiches Mittel dar, verallgemeinernde kausale Inferenzen aus multiplen Experimenten zu ziehen (Shadish, 2002).

Soziale Validität: Abgesehen von der internen und externen Validität heben Deshler und Schumaker (1994) zudem die soziale Validität einer Untersuchung hervor. Damit ist das Ausmaß, wie sehr die Unterrichtsmethode von Lehrern und Schülern als akzeptabel, hilfreich und praktikabel eingeschätzt wird, gemeint. Die Forscher schlagen aus diesem Grund die Durchführung von Interviews oder Fragebogenerhebungen am Ende einer Implementierung vor.

2.4.2.2 Überprüfung der Wirksamkeit des implementierten Programms

Die Frage, ob ein Trainingsprogramm unter realen Umständen wirksam ist oder nicht, ist das zentrale Anliegen jeglicher Implementationsforschung. Allgemein gelten statistische Signifikanztests sowie Effektgrößen als Wirksamkeitskriterien (Hager, 2000b).

Hinsichtlich der Evaluation der Wirksamkeit von Programmen differenziert Hager (2000a) zwischen der Wirksamkeitshypothese und der Wirksamkeitsunterschiedshypothese. Während die Wirksamkeitshypothese der Frage nachgeht, ob ein Programm generell wirksam ist und die vorher postulierten Ziele erfüllt, versucht die Wirksamkeitsunterschiedshypothese die Überlegenheit eines Programms im Vergleich zu einem Alternativprogramm nachzuweisen. Daraus ergeben sich unterschiedliche Versuchsdesigns zur Überprüfung der Hypothesen: isolierte Evaluation zur Prüfung von Wirksamkeitshypothesen und vergleichende Evaluation zur Überprüfung der Wirksamkeitsunterschiedshypothese.

Wirksamkeitshypothese: Die isolierte Evaluation zur Prüfung der Wirksamkeitshypothese beinhaltet eine Experimentalgruppe und eine Kontrollbedingung, die abgesehen von den zu prüfenden programmspezifischen Trainingsinhalten mit der Experimentalbedingung identisch ist (Hager, 2000a). Der Vergleich der Experimental- und der Kontrollbedingung sollte in einem Netto-Gewinn zu Gunsten der Experimentalbedingung resultieren, welcher unabhängig von normalen Reifungsprozessen oder der

Interventionssituation selbst ist. Deshalb ist die Konstanzhaltung jeglicher Randbedingungen für die Validität der Aussage über die Wirksamkeit wichtig. Zu den Randbedingungen gehören die gleiche Anzahl an Interventionssitzungen, die gleiche Dauer der Sitzungen, die gleichen Eigenschaften der Mitglieder der Versuchsgruppen, die gleichen Anreizsysteme etc. Ferner sollten die beiden Bedingungen von denselben eingangs geschulten Trainern durchgeführt werden, um auszuschließen, dass etwaige Trainerunterschiede für die Unterschiede zwischen den Bedingungen verantwortlich sind. Die Ziele der Experimental- und der Kontrollgruppe variieren. Als Kriteriumsmaße müssen jedoch nur die Ziele der Experimentalgruppe erhoben werden. Die Ziele der Kontrollgruppe können jedoch ebenfalls erfasst werden, um zu untersuchen, ob das Kontrollgruppentraining ebenfalls den intendierten Nutzen erbracht hat.

Wirksamkeitsunterschiedshypothese: Eine vergleichende Evaluation zur Prüfung der Wirksamkeitsunterschiedshypothese erfolgt erst dann, wenn im Rahmen der isolierten Evaluation die Wirksamkeit eines Trainingsprogramms nachgewiesen wurde. Das Experimentaltraining wird mit einem Alternativtraining verglichen, um zu prüfen, ob das Experimentaltraining wirksamer ist als das Alternativtraining (Hager, 2000a). Im Gegensatz zur isolierten Evaluation unterscheiden sich die beiden Versuchsgruppen hinsichtlich der Randbedingungen und sind bezüglich der Zielsetzung identisch. Die Differenzen in Bezug zu den Randbedingungen sind dadurch zu erklären, dass die Effektivität des Alternativprogramms unter festgelegten Randbedingungen (z.B. bestimmte Anzahl an Trainingssitzungen) entwickelt und überprüft wurde. Ein Vergleich der Effektivität des Alternativtrainings mit dem Experimentaltraining ist somit nur dann aussagekräftig, wenn die Trainings wie vorgesehen implementiert wurden. Die Ziele der beiden Gruppen sind gleich, da die Gegenüberstellung anhand der von außen oder programmimplizierten Ziele erfolgt. Somit werden Kriteriumsmaße erhoben, die sowohl für die Experimental- als auch für die Kontrollbedingung relevant sind. Weiter muss die Durchführung der Trainings nicht von demselben Trainer erfolgen, wobei jedoch die Trainer für das jeweilige Experimental- oder Alternativtraining explizit geschult sein sollten. Pressley und Harris (1994) betonen, dass bei der Auswahl und bei der Durchführung der Alternativprogramme genau und am besten mit Experten des Alternativprogramms gearbeitet werden sollte, da der Versuchsleiter zu Gunsten des Experimentaltrainings voreingenommen ist.

Unterscheidung zwischen Wirksamkeit und Effektivität: O'Donnell (2008) unterscheidet im Bereich der Implementationsforschung Wirksamkeit (*Efficacy*) und Effektivität (*Effectiveness*). Während Studien zur Überprüfung der Wirksamkeit unter idealen

und wissenschaftlich kontrollierten Bedingungen durchgeführt werden und der internen Validität des Trainings dienen, werden Studien zur Überprüfung der Effektivität unter den tatsächlichen und realen Bedingungen umgesetzt und geben Hinweise auf die externe Validität der Maßnahme. Wirksamkeitsstudien sollten Effektivitätsstudien vorgeschaltet werden. Darüber hinaus können aufgrund der starken Kontrolle von möglichen Störfaktoren in Wirksamkeitsstudien Trainingskomponenten analysiert werden, die für den Erfolg einer Intervention essenziell bzw. redundant sind. Effektivitätsstudien dagegen intendieren die Generalisierbarkeit der Befunde auf alltägliche Situationen und Personen wie z.B. Lehrer (O'Donnell, 2008). Erst wenn ein extern entwickeltes Programm sich in mehreren Situationen als effektiv erwiesen hat, sollte dieses von Lehrern umgesetzt und für diese nutzbar gemacht werden.

2.4.2.3 Überprüfung der Qualität der Implementierung

Bei der Top-down-Strategie geht der Erfolg einer Implementierung mit der Qualität der Umsetzung einher. Aus diesem Grund sind nur Aussagen über die Wirksamkeit eines Trainings, wie sie im vorangegangenen Abschnitt beschrieben wurden, zulässig, wenn die Qualität der Implementierung berücksichtigt wird. In diesem Abschnitt werden daher verschiedene Sicherungsmaßnahmen zur Gewährleistung der Implementationsgenauigkeit, insbesondere der inhaltlichen Vollständigkeit und der Unterrichtsqualität, diskutiert.

Implementationsgenauigkeit: Der Begriff Implementationsgenauigkeit bzw. *Treatment Integrity* bezeichnet die Genauigkeit, mit der die unabhängige Variable oder die Intervention so wie geplant implementiert wurde, so dass gültige Aussagen über die Beziehung zwischen unabhängigen und abhängigen Variablen getroffen werden können (Griffith, Duppong Hurley, & Hagaman, 2009). In diesem Sinne ist die Implementationsgenauigkeit sowohl für die Gewährung der internen als auch der externen Validität entscheidend (Griffith et al., 2009; O'Donnell, 2008). Die Implementationsgenauigkeit ist ein wichtiger Faktor für die Interpretierbarkeit der Ergebnisse, da bei mangelnder Implementationsgenauigkeit ausbleibende Effekte nicht notwendigerweise auf die Ineffektivität der Intervention schließen lassen (Griffith et al., 2009; Snyder et al., 1992).

O'Donnell (2008) unterscheidet fünf Kriterien zur Beurteilung der Implementationsgenauigkeit:

- (1) *Inhaltliche Vollständigkeit* der vorgesehenen Komponenten der Intervention;
- (2) *Dauer*, d.h. die Anzahl, Länge oder Frequenz der Trainingseinheiten;

- (3) *Qualität der Umsetzung* durch den Trainer bzw. den Lehrer unter Verwendung der vorgesehenen Techniken, Prozesse und Methoden;
- (4) *Reaktion der Teilnehmer*, d.h. die Beteiligung der Teilnehmer an den geplanten Aktivitäten und Inhalte der Intervention;
- (5) *Programmdifferenzierung*, d.h. das Ausmaß, in dem wichtige Faktoren zur Unterscheidung der Kontrollbedingung während der Implementierung vorhanden sind.

Insgesamt lassen sich die fünf Kriterien hinsichtlich der Struktur und des Prozesses gliedern. Die Struktur (d.h. Vollständigkeit, Dauer) umfasst den äußeren Rahmen, in dem die Implementierung stattfindet, während der Prozess (d.h. Qualität der Umsetzung, Programmdifferenzierung) die Art und Weise, wie die Umsetzung der Implementierung erfolgt, mit einschließt (Mowbray, Holter, Teague, & Bybee, 2003). Die Reaktion der Teilnehmer integriert sowohl Elemente der Struktur als auch des Prozesses.

Während im Gesundheitswesen die Berücksichtigung der Implementationsgenauigkeit schon seit vielen Jahren weit verbreitet ist, wurde sie bislang in der Evaluation pädagogischer Maßnahmen vernachlässigt. Nur in 5% der untersuchten Interventionsstudien im Bereich der Pädagogischen Psychologie wurden Angaben zur Implementationsgenauigkeit gemacht (Hsieh, Acee, Chung, Hsieh, Kim, Thomas, et al., 2005). Diesem Abfall an methodischer Qualität in der pädagogischen Interventionsforschung wird heute entgegengewirkt, indem die Analyseeinheiten den Bedingungen randomisiert zugeteilt werden, die Implementationsgenauigkeit erfasst wird und standardisierte Messinstrumente verwendet werden (O'Donnell, 2008).

Inhaltliche Vollständigkeit: Als Strukturbeispiel wird die inhaltliche Vollständigkeit erläutert, da diese in der Mehrzahl der Studien mit der Implementationsgenauigkeit gleichgesetzt wird (O'Donnell, 2008). Dies kann dadurch begründet werden, dass die Überprüfung der inhaltlichen Vollständigkeit durch den Einsatz von Checklisten einfach durch die teilnehmenden Lehrer operationalisierbar ist und nicht unbedingt externe Beobachter erfordert. Griffith und Kollegen berichten, dass in etwa der Hälfte der von ihnen untersuchten Interventionsstudien in dem Bereich Verhaltensstörungen Angaben zur inhaltlichen Vollständigkeit gemacht wurden (Griffith et al., 2009).

Dennoch war das Bewusstsein für die Notwendigkeit der Erfassung der inhaltlichen Vollständigkeit lange nicht in der pädagogischen Interventionsforschung verankert. Gräsel und Parchmann (2004) berichten von einer Studie aus dem Jahr 1971, die die Genauigkeit der Umsetzung einer Unterrichtsmethode mit Beobachtungsbögen überprüfte. Es zeigte sich, dass

die involvierten Lehrer nur einen geringen Anteil der Inhalte der Intervention wie vorgegeben umsetzen und den Großteil nach eigenem Ermessen adaptierten oder weglassen. Auch Souvignier und Trenk-Hinterberger (2007) heben hervor, dass die alleinige Bereitstellung von Unterrichtsmaterialien sowie Manualen die korrekte Umsetzung der inhaltlichen Vollständigkeit nicht garantieren. Sie argumentieren, dass Lehrer aufgrund von mangelnder Akzeptanz und Verinnerlichung der neuen Konzepte diese nicht wie geplant in ihren eigenen Unterricht integrieren.

Die Protokollierung der durchgeführten Inhalte einer Intervention durch die Lehrer oder Trainer ist für die Replizierbarkeit der Studien und zum Nachvollziehen der Interventionen zwingend erforderlich (Deshler & Schumaker, 1994; Pressley & Harris, 1994). Wenn die Intervention von unterschiedlichen Lehrern und Trainern umgesetzt wurde, helfen Checklisten, unterschiedliche Ergebnisse hinsichtlich der Genauigkeit der Durchführung zu erklären. Darüber hinaus sind Protokolle bei der Identifikation von schwierig umsetzbaren Trainingspassagen, die infolgedessen überarbeitet werden können, nützlich.

Qualitätsmerkmale wirksamen Unterrichts: Die Qualität des Unterrichts bzw. der Umsetzung einer neuen Unterrichtsmethode stellt ein Prozesskriterium zur Beurteilung der Implementationsgenauigkeit dar. In ihrer *Educational Reform Rating Rubric* zur Evaluation von groß angelegten Schulreformen listen Bessell, Burke, Plaza, Lee und Schumm (2008) sieben Erfolgsindikatoren auf, die auf eigene Vorarbeiten sowie Literaturrecherchen basieren. Diese sieben Indikatoren sind kollektive Teilnahme, Führungsstil des Schulleiters, Beteiligung der Eltern, Leistungen der Schüler, Projektunterstützung, Schulklima sowie die Instruktionsqualitäten der Lehrer. Die Noten und Testergebnisse der Schüler verbesserten sich dabei in Abhängigkeit von den durchgeführten Aktivitäten zur Verbesserung des eigenen Unterrichts und dem Verhältnis der Lehrer zu den Schülern und Eltern.

Die Bedeutsamkeit der Instruktionsqualität für das schulische Lernen wurde in einem Überblicksartikel von Wang, Haertel und Walberg (1993) bestätigt. Sie stellten fest, dass distale Faktoren wie der Schulleiter oder die Schulstruktur die Lernergebnisse kaum beeinflussten. Im Gegensatz dazu kristallisierten sich proximale Faktoren, wie die kognitiven Kompetenzen des Schülers, die Unterrichtsqualität und das häusliche Umfeld für erfolgreiches schulisches Lernen als wesentlich heraus. Bezüglich der Unterrichtsqualität werden vor allem eine effiziente Klassenführung und ein kompetenter Umgang mit Störungen als zentral angesehen (Wang et al., 1993).

Einflussreiche Arbeiten zur Unterrichtsqualität wurden von Helmke (2009) vorgelegt. Dieser versteht Unterricht als „didaktisch geplante und deshalb sowohl thematisch

abgrenzbare als auch zeitlich hinreichend umfassende Sequenzen des Lehrens und Lernens im Kontext pädagogischer Institutionen“ und Qualität als „die Beschaffenheit eines Sachverhaltes, die negativ oder positiv sein kann“ (Helmke, 2006, S. 812).

Basierend auf empirischen Studien, Monographien und Sammelbänden aus dem Bereich Bildungsforschung sowie Metaanalysen, unterscheidet Helmke (2009) zehn verschiedene fächerübergreifende Merkmale zur Erfassung der Unterrichtsqualität:

- (1) *Klassenführung*: Unter dem Begriff Klassenführung ist die Vorbeugung von Unterrichtsstörungen durch präventive, proaktive und reaktive Elemente gemeint. Insbesondere das Aufstellen und Durchsetzen von Regeln und Routinen ist für eine effektive Klassenführung zentral. Auf diese Weise werden Voraussetzungen für effektive Lehr- und Lernleistungen sowie genügend produktive Lernzeit geschaffen. Für die Unterrichtsqualität ist die Klassenführung extrem relevant, da sich einerseits hohe positive Korrelationen mit dem Lernfortschritt der Schüler und andererseits hohe negative Korrelationen mit dem Burn-Out von Lehrern zeigten (Helmke, 2009).
- (2) *Klarheit und Strukturiertheit*: Ein weiterer wichtiger Faktor der Unterrichtsqualität ist die klare, verständliche und strukturierte Vermittlung von Aufgaben, Wissen und Strategien, so dass bei den Schülern Lernprozesse angeregt werden.
- (3) *Konsolidierung und Sicherung*: Das Wiederholen und Üben von vermittelten Lerninhalten ist notwendig, damit durch die Aufnahme von neuen Informationen und Wissen nachhaltiges Lernen möglich wird. Lern- und gedächtnispsychologische Erkenntnisse sollten dabei berücksichtigt werden.
- (4) *Aktivierung*: Die Aktivierung der Schüler im Unterricht umfasst die kognitive Aktivierung (z.B. Einsatz von Lernstrategien), soziale Aktivierung (z.B. Mitarbeit in Gruppen), aktive Teilhabe an der Planung und Durchführung des Unterrichts (z.B. Unterrichtsbeteiligung) und körperliche Aktivierung (z.B. vielfältige Unterrichtsformen).
- (5) *Motivierung*: Die Motivation erleichtert die Initiierung und Aufrechterhaltung von Lernprozessen. Die Bereitstellung von Aufgaben mittlerer Schwierigkeit, die das Interesse und die Neugier der Schüler anregen sowie die Schaffung von Möglichkeiten, die eigene Fantasie sowie eigene Wünsche in die Aufgabengestaltung mit einfließen zu lassen, stellen Beispiele für Strategien zur Förderung der Motivation im Unterricht dar.

- (6) *Lernförderliches Klima*: Mit lernförderlichem Klima ist eine positive Beeinflussung des Lernens der Schüler durch die Lernumgebung gemeint. Insbesondere ein angemessener Umgang mit Fehlern, ein angemessenes Unterrichtstempo, eine entspannte Lernatmosphäre durch ein vertrauensvolles Lehrer-Schüler-Verhältnis sowie der Abbau von Leistungsangst spielen dabei für ein lernförderliches Unterrichtsklima eine entscheidende Rolle.
- (7) *Schülerorientierung*: Die Wertschätzung des Schülers als Person sowie ein fürsorglicher und als Ansprechpartner bereit stehender Lehrer sind Kennzeichen einer Schülerorientierung, was sich in einer vertrauensvollen Lehrer-Schüler-Beziehung widerspiegelt, so dass sich der Schüler im Unterricht wohlfühlt.
- (8) *Kompetenzorientierung*: Ein kompetenzorientierter Unterricht ist charakterisiert durch eine bewusste und geplante Ausrichtung des Unterrichts auf den Erwerb von Kompetenzen, wie sie in den Bildungsstandards formuliert wurden.
- (9) *Umgang mit Heterogenität*: Eine Schulklasse stellt eine heterogene Gruppe von Schülern mit unterschiedlichen Lernvoraussetzungen und Leistungsniveaus dar. Qualitativ guter Unterricht muss eine Binnendifferenzierung ermöglichen, so dass adaptiv auf die Leistungsunterschiede im Unterricht eingegangen wird. Dies erfolgt durch Individualisierung, was auf Seiten des Lehrers ein hohes Maß an diagnostischer Kompetenz, Fachwissen und einem großen Repertoire an Unterrichtsmethoden voraussetzt.
- (10) *Angebotsvielfalt*: Mit Angebotsvielfalt ist nicht nur eine Variation an Methoden (z.B. Frontalunterricht, offener Unterricht, verschiedene Sozialformen), sondern ebenfalls der Einsatz von unterschiedlichen Medien, Aufgabentypen, Textsorten, Lautstärken, Lernorten, Sinnesmodalitäten und Lern- bzw. Entspannungsphasen gemeint.

Abgesehen von den dargestellten Merkmalen von Helmke (2009) gibt es noch andere Klassifikationssysteme der Unterrichtsqualität wie z.B. das Quait-Modell (*Quality, Appropriateness, Incentives, Time*) von Slavin. Die Anzahl der Merkmale wurde von Helmke (2009) relativ willkürlich gewählt. Bei genauerer Betrachtung ergeben sich inhaltliche Überschneidungen der zehn Konzepte, so dass eine Zusammenfassung der Merkmale zu globaleren Dimensionen ebenfalls legitim ist.

Die Aufgliederung der Unterrichtsqualität in die zehn genannten Dimensionen stellt jedoch ein hilfreiches Gerüst dar, die Qualität des Unterrichts bzw. der Umsetzung einer neuen Unterrichtsmethode durch externe, geschulte Beobachter zu erfassen. Auf diese Weise

kann die Unterrichtsqualität als Erklärung für potenzielle Unterschiede in den Schülerleistungen herangezogen werden.

2.4.3 Implementationsforschung im Bereich Schreiben

In den bereits dargestellten Metaanalysen zu SRSD von Graham und Kollegen (Graham, 2006; Graham & Harris, 2003; Graham & Perin, 2007a; Rogers & Graham, 2008) wurde die praktische Bedeutsamkeit erfolgreicher Schreibprogramme für Lehrer stark betont. Auch wenn es bereits einige wenige experimentelle Studien gibt, die Lehrer anstatt Trainer für die Durchführung der Strategieinstruktionen heranzogen, ist die Qualität dieser Studien insgesamt als gering einzustufen, da z.B. nur ein geringer Anteil der Studien Follow-Up-Erhebungen umfasste und es sich größtenteils um Einzelfalluntersuchungen handelte (Graham, 2006). Gleichzeitig sind jedoch in den Metaanalysen Hinweise zu finden, dass die Effektivität der Trainings in Abhängigkeit davon, ob Lehrer oder Trainer diese durchführten, variieren (Graham, 2006; Graham & Harris, 2003). Generell zeigten sich Trainer effektiver bzw. genauso effektiv wie Lehrer, wenn es um kurzfristige Effekte direkt im Anschluss an das Training geht (Graham, 2006; Graham & Harris, 2003). Hinsichtlich der Stabilität und Generalisierbarkeit erbrachten Lehrer jedoch bessere Ergebnisse (Graham & Harris, 2003). In Einzelfalluntersuchungen schnitten Lehrer ebenfalls besser ab (Graham, 2006).

Folglich ist zu schlussfolgern, dass Studien mit Trainern anderen Einflussfaktoren ausgesetzt sind als Studien mit Lehrern. Ein Trainingsprogramm, welches unter stark kontrollierten Versuchsbedingungen von Trainern effektiv umgesetzt wurde, muss nicht zwangsläufig effektiv sein, wenn es von Lehrern unter normalen Schul- und Klassenbedingungen implementiert wird. Auf diese Problematik weisen Graham und Kollegen ausdrücklich hin und fordern mehr Forschungsarbeiten, die sich speziell mit Lehrern beschäftigen und explizit die Anforderungen des Schulalltags berücksichtigen (Graham & Perin, 2007a; Rogers & Graham, 2008).

Eine Implementierung von Interventionsmaßnahmen durch Lehrer impliziert eine Gefährdung der internen Validität der Untersuchung, da mehr Störfaktoren als bei Interventionen in Kleingruppen durch geschulte Trainer die Ergebnisse verzerren können. Aus diesem Grund verwundert es nicht, dass die bisherige Forschung in diesem Gebiet erst in den Anfängen steckt. Dies trifft insbesondere auf die Schreibforschung zu. In der Leseforschung wurden dagegen bereits erste Implementierungen vorgenommen.

Einer der ersten Forscher, der sich explizit mit der Schulung von Lehrern zur Implementierung von Lesestrategien und den dabei ablaufenden Veränderungsprozessen bei Lehrern beschäftigte, war Duffy (1993). Um die Gründe für das mangelhafte lesebezogene

Strategiewissen von schwachen Schülern herauszufinden, beobachtete Duffy (1993) vier Lehrer über einen Zeitraum von acht Monaten jeweils zweimal im Monat und führte Interviews mit Lehrern und Schülern durch. Zuvor durchliefen die Lehrer einen Vorbereitungskurs zum strategischen Lesen. Die Art, Form und Dauer der Umsetzung der erlernten Methoden wurde den Lehrern frei überlassen und in keiner Weise kontrolliert. Am Ende des Schuljahrs zeigten sich zwar Verbesserungen des strategischen Wissens bei den Schülern, diese Entwicklung vollzog sich jedoch sehr langsam und schwierig. Die Lehrer hatten große Probleme damit, die in den Kursen vermittelten Inhalte in ihrem Unterricht umzusetzen, obwohl sie selbst Unterrichtsmanuale erarbeitet hatten und zur Verfügung gestellt bekamen. Deswegen schlussfolgert Duffy (1993), dass traditionelle Lehrerfortbildungen, die in der Regel die Bereitstellung von Informationen, Unterrichtsskripten und Materialien umfassen, nicht ausreichen, damit Lehrer neue Unterrichtskonzepte verinnerlichen und dadurch adaptiv in ihren eigenen Unterricht integrieren.

Basierend auf diesen Erfahrungen konzipierten Souvignier und Trenk-Hinterberger (2007) ein Drei-Stufen-Modell zur Implementierung des Textdetektive-Programms zur Förderung des Leseverständnisses von Fünft- und Sechsklässlern. Das Textdetektive-Programm beinhaltet kognitive und metakognitive Lesestrategien, die mit selbstregulatorischen Strategien kombiniert werden. Es handelt sich dabei um folgende drei Stufen:

- Einführende Lehrerfortbildung mit praktischen und theoretischen Elementen;
- Textdetektive-Programm als ausgearbeitetes Unterrichtsmaterial für Lehrer;
- Wiederholungseinheit als Auffrischung der Strategieinhalte.

In einer Studie mit 22 Lehrern, die das dreistufige Implementierungsmodell durchliefen und fünf Lehrern, die den regulären Deutschunterricht abhielten, zeigte sich das dreistufige Modell als erfolgreich hinsichtlich der Verbesserung der Lesekompetenz auf Schülerebene und hinsichtlich der Akzeptanz des Textdetektive-Programms auf Lehrerebene.

Darüber hinaus konnten in zwei weiteren Studien das Textdetektive-Programm erfolgreich von Lehrern umgesetzt werden (Mokhlesgerami, Souvignier, Rühl, K., & Gold, 2007; Souvignier & Mokhlesgerami, 2006). In beiden Studien wurden Maßnahmen zur Sicherstellung der Implementationsgenauigkeit ergriffen. Während Mokhlesgerami et al. (2007) nur Unterrichtsbeobachtungen durchführten, integrierten Souvignier und Mokhlesgerami (2006) zusätzlich Lehrer-Checklisten zur Überprüfung der inhaltlichen Vollständigkeit und zwei Informationsnachmittage als Vorbereitung für die Lehrer in ihre

Untersuchung. Die positiven Ergebnisse des Textdetektive-Programms sind insbesondere deswegen hervorzuheben, da Lehrer häufig Schwierigkeiten haben, vorgegebene Trainingsprogramme gemäß der Top-down-Strategie in den eigenen Unterricht zu implementieren (Hacker & Tenet, 2002).

Hacker und Tenet (2002) begleiteten und beobachteten über einen Zeitraum von drei Jahren 17 Lehrer bei der Implementierung von reziproken Unterrichtsstrategien (*Reciprocal Teaching*, RT) zur Förderung des Leseverständnisses. Am Ende der qualitativen Studie zeigte sich, dass die Lehrer die RT-Methoden über die Jahre hinweg immer wieder modifizierten und dass die Art der Umsetzung ebenfalls stark von den jeweiligen Lehrern abhing. Daraus kann geschlussfolgert werden, dass für zukünftige Implementationsstudien die Genauigkeit der Implementation gesichert werden muss, um die interne Validität zu gewährleisten.

Im Gegensatz zum Lesen sind bezüglich des Schreibens nur zwei Studien hervorzuheben, die explizit die Implementierung des SRSD-Programms durch Lehrer intendierten. Insgesamt wurden in den meisten SRSD-Studien geschulte Trainer anstatt Lehrer eingesetzt und Schüler mit Lernbehinderungen anstatt durchschnittliche Schüler untersucht (De La Paz & Graham, 2002).

De La Paz und Graham (2002) legten eine Studie vor, in der zehn Klassen des siebten und achten Jahrgangs zufällig einer von zwei Bedingungen zugeordnet wurden: (1) Experimentalgruppe: sechs Klassen erhielten die SRSD-Instruktion; (2) Kontrollgruppe: vier Klassen wurden nach traditionellen Methoden im Schreiben unterrichtet, erhielten jedoch denselben Umfang an Übungen und Zeit zum Schreiben. Jeder Lehrer unterrichtete sowohl eine Experimental- als auch eine Kontrollgruppe. Zu drei verschiedenen Messzeitpunkten (Prätest, Posttest, Follow-Up) planten und schrieben alle Schüler der zehn Klassen einen argumentativen Aufsatz. Eine repräsentative Stichprobe von 58 Schülern (30 aus der Experimental- und 28 aus der Kontrollbedingung) wurden zur genaueren Analyse ausgewählt.

Verschiedene Maßnahmen zur Gewährung der Implementationsgenauigkeit wurden von De La Paz und Graham (2002) umgesetzt. Zunächst wurden die Lehrer in einem zweitägigen Workshop in der Durchführung von SRSD geschult. Zudem erhielten sie ausführliche Anleitungen sowie Materialien, die sie als Grundlage für die Umsetzung von SRSD heranzogen. Des Weiteren überprüfte ein Student im Hauptstudium die inhaltliche Vollständigkeit anhand von 20% der Audioaufnahmen der Unterrichtseinheiten in beiden Bedingungen. Außerdem beobachtete ein weiterer Student im Hauptstudium jede Woche den Unterricht in beiden Bedingungen. Aufbauend auf seinen Beobachtungen erhielten die Lehrer schriftliches Feedback mit Verbesserungshinweisen, wenn die Durchführung nicht wie

vorgesehen erfolgte. Überdies fanden wöchentliche Treffen mit den beteiligten Lehrern statt, bei denen Probleme und Veränderungen besprochen wurden.

Direkt nach dem Training (Posttest) sowie einen Monat später (Follow-Up) schrieben Schüler der Experimentalbedingung qualitativ bessere und längere Aufsätze, verwendeten einen reichhaltigeren Wortschatz sowie planten besser als Schüler der Kontrollbedingung. Somit konnten De La Paz und Graham (2002) nachweisen, dass die Implementierung des SRSD-Programms in den regulären Unterricht durch Lehrer möglich ist und dass durchschnittliche Schüler genauso wie lernschwache Schüler von der expliziten Vermittlung von Schreibstrategien mit selbstregulatorischen Strategien profitieren. Gleichzeitig heben De La Paz und Graham (2002) die Notwendigkeit der Replikation der Befunde hervor.

Um jedoch ein unverzerrtes Bild zu erlangen, sollten dabei die Daten aller Schüler einer Klasse ausgewertet werden, wobei die Anzahl der Experimental- und der Kontrollklassen identisch sein sollte. Auch wenn sich Vorteile daraus ergeben, dass ein Lehrer sowohl die Experimental- als auch die Kontrollbedingung unterrichtet, sind damit auch gewisse Schwierigkeiten und Risiken verbunden. Zum einen ist es organisatorisch schwierig, eine ausreichende Anzahl an Lehrern zu gewinnen, die mehrere Deutschklassen eines Jahrgangs gleichzeitig unterrichten und ebenso bereit sind, unterschiedliche Unterrichtsmethoden in den einzelnen Klassen durchzuführen. Dies bringt zudem ethische Probleme mit sich, da die Lehrer zwangsläufig eines der Programme als besser empfinden werden. Folglich sind berechtigte Zweifel daran zu setzen, ob die Umsetzung der unterschiedlichen Programme tatsächlich so wie geplant von den Lehrern umgesetzt wurden und ob die Lehrer unbewusst Komponenten der verschiedenen Programme verwechselten.

Ähnlich wie De La Paz und Graham (2002) gingen Torrance, Fidalgo und Garcia (2007) vor. Sie implementierten ebenfalls das SRSD-Programm durch Lehrer in den gesamten Klassenkontext mit durchschnittlichen Schülern. Sie nannten ihr Programm *Cognitive Self-Regulation Instruction* (CSRI), da es im Gegensatz zu dem ursprünglichen SRSD-Programm explizit auf durchschnittlich begabte Schüler abzielte. Als Erweiterung zu De La Paz und Graham (2002) implementierten Torrance und Kollegen (2007) das CSRI-Programm erstmalig in einen nicht-nordamerikanischen Kulturkreis mit einem entsprechend anderem Bildungssystem und untersuchten zudem Schüler der sechsten Klasse anstatt der siebten und achten Klasse. Insgesamt 95 Schüler aus vier verschiedenen Klassen nahmen an der Untersuchung teil. Drei der vier Klassen ($n=71$) kamen von einer Schule und bildeten die Experimentalgruppe, die sowohl Schreibstrategien für argumentative Texte als auch selbstregulatorische Strategien in Form von Selbstinstruktionen und spezifischen Planungs-

und Überarbeitungsstrategien erlernten. Das CSRI-Programm wurde in allen drei Klassen von einem Spanischlehrer durchgeführt, welcher zuvor in der Durchführung von CSRI geschult wurde. Die vierte Klasse ($n=24$) stellte die Kontrollbedingung dar und erhielt traditionellen Schreibunterricht von einem anderen Spanischlehrer. Die Implementationsgenauigkeit wurde nur bezüglich der inhaltlichen Vollständigkeit und lediglich in der Experimentalgruppe kontrolliert. Zu drei verschiedenen Messzeitpunkten (Prätest, Posttest, Follow-Up) schrieben die Schüler in allen vier Klassen Gegenüberstellungen. Sowohl beim Posttest als auch beim Follow-Up zeigten Schüler der Experimentalgruppe signifikant bessere Leistungen bei der Textproduktion (Qualität, Kohärenz, Struktur) und bei der Planung (Zeit). Hinsichtlich der Textrevisionen ergaben sich jedoch keine Unterschiede zwischen den Gruppen.

Die Studie von Torrance und Kollegen (2007) ist ein weiterer Beleg dafür, dass von Experten entwickelte Schreibtrainings erfolgreich in den normalen Unterricht implementiert werden können. Gleichzeitig weist die Studie einige Schwächen auf, was die Interpretierbarkeit der Befunde einschränkt. Abgesehen von der ungenügenden Gewährleistung der Implementationsgenauigkeit, handelt es sich um eine Studie mit einer kleinen Stichprobe, in der nur jeweils ein Lehrer die Experimental- bzw. Kontrollbedingung unterrichtete. Die Anzahl der Klassen in der Experimental- und der Kontrollbedingung variierten zudem. Folglich sind lehrer-, klassen- und schulspezifische Effekte nicht auszuschließen. Weiter erhielt nur der Lehrer der Experimentalbedingung Hilfestellungen und Anleitungen, so dass Hawthorne-Effekte oder Hypothesenraten nicht auszuschließen sind.

2.4.4 Zusammenfassung

Bei der Implementierung von Innovationen in den Regelunterricht können grundsätzlich zwei Strategien unterschieden werden. Während die symbiotische Strategie eine gemeinsam von Lehrern und Experten entwickelte Unterrichtsmethode impliziert, wird bei der Top-down-Strategie die neue Unterrichtsmethode von Experten entwickelt und von Lehrern umgesetzt. In diesem Zusammenhang wurde zudem der EDR vorgestellt, der zwar ebenfalls eine enge Zusammenarbeit von Experten und Lehrern mit einschließt, jedoch im Gegensatz zur symbiotischen Strategie theoretisch fundiert ist und eine Erweiterung der bestehenden Theorie intendiert.

Außerdem wurden in dem letzten Abschnitt wichtige Standards zur Sicherung der Qualität der Implementierung vorgestellt. Dazu gehören Maßnahmen zur Gewährung der internen und externen Validität sowie Untersuchungsparadigmen zur Überprüfung der Wirksamkeit der Intervention und der Qualität der Implementierung. Zum letzteren gehört

insbesondere die Genauigkeit der Implementation, wozu hauptsächlich die inhaltliche Vollständigkeit und die Qualität des Unterrichts zählen.

Im letzten Abschnitt wurden schließlich Implementationsstudien im Bereich Lesen und Schreiben vorgestellt. Beim Lesen gibt es im Vergleich zum Schreiben deutlich mehr Studien, die erfolgreiche Trainingsprogramme in den Schulunterricht integrierten. Bezüglich des Schreibens konnten nur zwei Studien vorgestellt werden (De La Paz & Graham, 2002; Torrance et al., 2007). Bezüglich folgender Punkten weisen die beiden Studien jedoch Grenzen auf:

- Die zwei Implementationsstudien im Bereich Schreiben wurden nur in den USA und Spanien durchgeführt. Der deutsche Kulturkreis und das deutsche Bildungssystem blieben bislang in der Schreibimplementationsforschung unberücksichtigt.
- Die beiden dargestellten Studien (De La Paz & Graham, 2002; Torrance et al., 2007) beschäftigten sich mit dem argumentativen Schreiben. Das Textgenre „erzählendes Schreiben“ wurde in diesem Zusammenhang noch nicht erforscht.
- Darüber hinaus umfassten die Studien von De La Paz und Graham (2002) bzw. von Torrance et al. (2007) nur Schüler der Klassenstufen sechs bis acht. Die Implementierung von Aufsatztrainings für Schreibanfänger, wie sie in der vierten Klasse vorzufinden sind, wurde bislang nicht untersucht.
- Während die Sicherung der Implementationsgenauigkeit von De La Paz und Graham (2002) bereits sehr gut realisiert wurde, weist die Studie von Torrance und Kollegen (2007) einige methodische Mängel auf.

3 Fragestellung und Hypothesen

3.1 Zielsetzung

Die langfristige Intention jeglicher pädagogisch-psychologischer Forschung sollte darin bestehen, praktischen Nutzen denjenigen zu bringen, die täglich mit Schülern arbeiten. Folglich sollten effektive Interventionen so adaptiert werden, dass sie für Lehrer in ihrem Unterrichtsalltag einen Gewinn darstellen (Paris & Paris, 2001). Das *Selbstregulatorische Aufsatztraining* (SAT) von Glaser (2005) stellt ein effektives Trainingsprogramm dar, welches bislang nur in Kleingruppen und von geschulten Lerntrainern umgesetzt wurde. Des Weiteren ist es in Deutschland das einzige Schreibtraining, welches aufbauend auf dem *Self-Regulated Strategy Development-* (SRSD) Ansatz von Harris und Graham (1996) explizit Schreibstrategien mit selbstregulatorischen Strategien verknüpft. Das vorliegende Dissertationsprojekt erweitert bisherige Forschungsarbeiten, so dass sich insgesamt zwei Hauptziele definieren lassen:

(1) Implementierbarkeit in den Regelunterricht

Der SRSD-Ansatz fokussiert vornehmlich die Förderung von lernschwachen Schülern, während das SAT-Programm für durchschnittlich begabte Schüler der vierten Klasse geeignet ist und überdies für den deutschen Sprachraum adaptiert wurde. Gleichzeitig handelt es sich bei allen durchgeführten SAT-Untersuchungen um kontrollierte experimentelle Studien, in denen geschulte Lerntrainer das Training in Kleingruppen von maximal sechs Schülern umsetzten. Es fehlen jedoch Belege für die ökologische Validität von Kleingruppentrainings, so dass eine Generalisierung der bisherigen Befunde auf den realen Klassenkontext nicht zulässig ist. Der Nachweis, dass das SAT-Programm auch einer ganzen Klasse in dem natürlichen Schulumfeld vermittelt werden kann, wurde bislang nicht erbracht. Insgesamt waren Lehrer in den bisherigen SAT-Studien nicht involviert, so dass der Schluss zulässig ist, dass die Schreibimplementationsforschung bislang im deutschen Kulturkreis und im deutschen Bildungssystem unberücksichtigt blieb. Graham und Harris (2003) stellten bereits fest, dass die Effektivität eines Schreibtrainings in Abhängigkeit davon, ob Lehrer oder Trainer die Intervention durchführten, variiert. Sie lieferten ermutigende Hinweise dafür, dass Lehrern eine Stabilisierung und Generalisierung der Leistungen besser gelingt als Trainern.

Generalhypothese 1: Es wurde angenommen, dass ein SAT-ähnliches Programm erfolgreich in den Schulalltag der vierten Klasse implementierbar ist. Die Umsetzung sollte dabei sowohl von geschulten Trainern als auch Lehrern durchführbar sein. Ferner wurde

erwartet, dass sämtliche im Projekt beteiligten Trainer und Lehrer die jeweiligen Schreibtrainings mit einer hohen Implementationsgenauigkeit ausführen. Dies bedeutet, dass die vorgesehenen Inhalte qualitativ gut realisiert werden, wobei sich die verschiedenen Trainer und Lehrer diesbezüglich nicht unterscheiden.

(2) Überlegenheit eines selbstregulatorischen Aufsatztrainings

Die in Metaanalysen ermittelten Effektstärken (nach Cohen's d) und PND's (siehe Fußnote 6) lieferten einen großen Effekt von SRSD- im Vergleich zu Nicht-SRSD-Studien (Graham, 2006; Graham & Harris, 2003; Graham & Perin, 2007a). Die Effektstärken lagen zwischen .74 und 1.57 und die PND's zwischen 88% und 100%. Nichtsdestotrotz sind die Studien von Glaser (2005; Glaser & Brunstein, 2007b; Glaser et al., 2009) die einzigen, die den inkrementellen Effekt selbstregulatorischer Strategien im Vergleich zu einem reinen Strategietraining belegten. Es fehlen jedoch Beweise dafür, dass die Effekte von Glaser (2005) auch im gesamten Klassenkontext bestehen bleiben. Überdies besteht weiterhin Unklarheit darüber, welche Komponenten von SRSD bzw. SAT für die Wirksamkeit zwingend erforderlich sind (De La Paz, 2007; Zimmerman & Risemberg, 1997). Reid und Lienemann (2006) betonen, dass selbstregulatorische Strategien es den Schülern erleichtern, Inhalte der Schreibstrategien zu memorieren, den Einsatz der Strategien zu überwachen und die Motivation aufrecht zu erhalten. In dem vorliegenden Dissertationsprojekt wurden Selbstregulation und Schreibstrategien separat betrachtet, um den Nachweis zu erbringen, dass Selbstregulation für die Wirksamkeit eines Schreibtrainings essenziell ist.

Generalhypothese 2: Folglich wurde angenommen, dass Schüler, die ein selbstregulatorisches Aufsatztraining absolvieren, im Vergleich zu Schülern, die ein Aufsatztraining ohne selbstregulatorische Komponenten durchlaufen, längere und qualitativ bessere Aufsätze schreiben, die Schreibstrategien häufiger einsetzen, besser planen und revidieren sowie über mehr schreibbezogenes Wissen verfügen. Des Weiteren wurde erwartet, dass Selbstregulation den Transfer der Schreibstrategien auf ungeübte Textgenres vereinfacht. Überdies wurde davon ausgegangen, dass die Effekte sich sowohl direkt im Anschluss an das Training als auch sechs Wochen später als stabil erweisen.

3.2 Grundprinzipien des empirischen Vorgehens

Zur Erreichung der beschriebenen Ziele und zum Testen der konstatierten Hypothesen wurde ein *Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining* (USAT) entwickelt, welches auf dem SAT-Programm von Glaser (2005) aufbaut und speziell auf die Anforderungen und

Rahmenbedingungen des Schulalltags der vierten Klasse angepasst ist. Das USAT-Programm integriert explizit Strategien zur Förderung der Selbstregulation mit Strategien zur Förderung der Schreibkompetenz.

Förderung der Selbstregulation: Paris und Paris (2001) nennen Reifung, Instruktion und Übung mit Feedback als drei Wege zur Erhöhung der Selbstregulation. Um einen praktischen Nutzen für den aktuellen Schreibunterricht in der Grundschule zu leisten, werden in der vorliegenden Arbeit die Instruktion und die Übung fokussiert. Diese sind zudem vielversprechend, da bereits in der Vergangenheit durch die Instruktion selbstregulatorischer Lernstrategien wie Zielsetzung, Selbstüberwachung und Selbstevaluation positive Effekte auf die Motivation und die Leistung nachgewiesen werden konnten (Paris & Paris, 2001; Schunk, 1996; Zimmerman & Kitsantas, 1997; Zimmerman & Kitsantas, 1999). Das zyklische Trainingsmodell von Zimmerman (1998) mit den Schritten Selbstbeobachtung, Zielsetzung, Überwachung des Strategieeinsatzes und Feedback diene dabei als Rahmenmodell. Da Selbstregulation nicht nur das Ziel, sondern ebenfalls der Weg sein sollte (Brunstein & Spörer, 2006; Paris & Paris, 2001), wurde die Instruktion selbstregulatorischer Strategien mit der Instruktion von Schreibstrategien kombiniert. Dies bedeutet, dass die Schüler zunächst ihr eigenes Schreibverhalten beobachteten, sich anschließend auf das Schreibprodukt und die Anwendung der Schreibstrategien bezogene Ziele setzten, diese Strategien im nächsten Schritt einsetzten und anschließend die Ergebnisse anhand von Feedback selbst bewerteten.

Förderung der Schreibkompetenz: Für jede der drei Schreibphasen des Planens, des Schreibens und des Überarbeitens wurden effektive selbstregulatorische Techniken identifiziert (siehe Kapitel 2.2.3.3). Der Schwerpunkt des intendierten Trainingsprogramms liegt bei Schreibanfängern. Da die kognitiven Voraussetzungen für das Ausführen von Revisionsfertigkeiten (z.B. Perspektivenübernahme, abstrakte Repräsentation von Schreibzielen) erst im späten Jugendalter ausgebildet werden (Feilke, 1993), hatte im vorliegenden Projekt die Instruktion von Planungsstrategien Vorrang gegenüber der Instruktion von Revisionsstrategien.

Abgrenzung zwischen Selbstregulation und Schreibstrategien: Um erfolgreich zwischen der Selbstregulation und den Schreibstrategien zu differenzieren, ist die Wahl und Konzeption einer geeigneten Kontrollgruppe von zentraler Bedeutung (Graham & Harris, 1989). Einerseits soll die Kontrollgruppe effektiv Schreibstrategien vermittelt bekommen, jedoch gleichzeitig keine selbstregulatorischen Methoden erlernen. Im Kapitel 2.3 wurden zwei grundlegende Ansätze zur Förderung der Schreibkompetenz, der Schreibprozessansatz bzw. *Writer's Workshop* und die kognitive Strategieinstruktion, vorgestellt, wobei sich der

zweite Ansatz, zu dem auch das SRSD- und das SAT-Programm zählen, als effektiver erwies (Berninger et al., 2009; Troia, 2009a). Nichtsdestotrotz erreichte der *Writer's Workshop* eine durchschnittliche Effektstärke von .32 (Graham & Harris, 2009; Graham & Perin, 2007b; Graham et al., 2009). Gleichzeitig umfasst der *Writer's Workshop* keine selbstregulatorischen Strategien. Außerdem wurde er bislang noch nicht in Deutschland umgesetzt und stellt somit für die Schüler neue und interessante Materialien bereit. Folglich erscheint die Implementierung des USAT-Programms im Vergleich zu einer Implementierung des *Writer's Workshops* in den regulären Deutschunterricht ein vielversprechender Forschungsansatz zu sein. Da der *Writer's Workshop* zudem kein spezifisches Trainingsprogramm enthält, konnten die Zielsetzung und die Rahmenbedingungen so gestaltet werden, dass analog zu den Vorgaben von Hager (2000a) die Wirksamkeit des USAT-Programms im gesamten Klassenkontext optimal überprüft werden konnte. Um den Nachweis für die Überlegenheit selbstregulatorischer Komponenten zu erbringen, sollten dieselben Schreibstrategien in beiden Bedingungen vermittelt werden.

Da das USAT-Programm neu konzipiert wurde, musste zu Beginn gewährleistet werden, dass die Unterschiede in den abhängigen Variablen tatsächlich auf Unterschiede zwischen den beiden Treatment-Bedingungen, d.h. auf die Selbstregulation, zurückzuführen sind. Auch wenn das USAT-Programm auf dem SAT-Programm aufbaut, sind die positiven Effekte des SAT-Programms nicht einfach auf den Klassenkontext zu generalisieren, da Trainings im Klassenkontext unter wesentlich anderen Bedingungen stattfinden als im stark kontrollierten Kleingruppen-Setting. Zur Sicherung der internen Validität empfiehlt es sich deshalb, die Schreibtrainings zunächst durch Trainer anstatt Lehrer implementieren zu lassen. Diese konnten randomisiert den Treatment-Bedingungen zugeteilt werden. Darüber hinaus konnte ein Trainer sowohl eine Experimental- als auch eine Kontrollgruppe instruieren. Insgesamt lassen sich durch den Einsatz von Trainern potenzielle Störfaktoren kontrollieren.

Nach erfolgreichem Nachweis der Wirksamkeit erfolgt als nächstes die Wirksamkeitsunterschiedshypothese (Hager, 2000a). Als vergleichendes Alternativtraining bot sich hier der traditionelle Aufsatzunterricht in der vierten Klasse an. Zur Erhöhung der externen Validität sollten dabei Lehrer die Schreibtrainings in ihren normalen Unterricht integrieren. Sollte sich das USAT-Programm als effektiver als der normale Aufsatzunterricht erweisen, würde dies dafür sprechen, das USAT-Programm weitläufig und für mehr Lehrer zugänglich zu machen.

Implementierung der Schreibtrainings: Nach Gräsel und Parchmann (2004) lassen sich zwei Implementationsstrategien, die Top-down-Strategie und die symbiotische Strategie,

unterscheiden. Die Implementierung des USAT-Programms und des *Writer's Workshops* entsprechen einer Top-down-Strategie, da Experten die Schreibtrainings entwickelten. Im Gegensatz dazu impliziert die symbiotische Strategie, dass ausgehend von den Anforderungen und Bedürfnissen der Praxis Experten zusammen mit Lehrern eine neue Unterrichtsmethode entwickeln. Der *Educational Design Research-* (EDR) Ansatz ähnelt der symbiotischen Strategie. Er betont jedoch verstärkt die Integration und Prüfung von theoretischen Konzepten.

Für das Dissertationsprojekt sind die symbiotische Strategie und der EDR-Ansatz insofern interessant, als dass sie als Vergleich zu der Top-down-Strategie herangezogen werden können. Der symbiotischen Strategie wurde dabei gegenüber dem EDR-Ansatz der Vorrang gegeben, da sie einen geringeren zeitlichen und personellen Aufwand erfordert und folglich in Hinblick auf die Rahmenbedingungen eher mit der Implementierung des USAT-Programms vergleichbar war. Aus diesem Grund wurde die symbiotische Strategie in der vorliegenden Arbeit verwendet, um zusammen mit Lehrern ein Trainingsprogramm zu entwickeln, welches den traditionellen Aufsatzunterricht widerspiegeln sollte.

Bei der Implementierung von Schreibtrainings müssen sowohl für die Top-down- als auch für die symbiotische Strategie Maßnahmen zur Sicherung der Implementationsgenauigkeit integriert werden. Ansonsten können ausbleibende Effekte auf eine unzureichende Umsetzung der Maßnahme durch die Trainer bzw. die Lehrer zurückgeführt werden. Es sollte dabei die inhaltliche Vollständigkeit und die Qualität der Implementierung überprüft werden.

3.3 Übersicht über die Studien

Um die dargestellten Ziele zu erreichen, wurden eine Lehrerbefragung, eine Pilotstudie und zwei Hauptstudien durchgeführt. Da die Intention der Arbeit die Implementierung des USAT-Programms in den Regelunterricht ist, wurden die natürlich vorkommenden Klassen als Stichprobe herangezogen. Folglich war eine randomisierte Zuordnung der Schüler zu den Bedingungen nicht möglich. Eine Übersicht über die Studien ist in Tabelle 8 zu finden.

Tabelle 8: Übersicht über die Studien.

Lehrerbefragung	Pilotstudie	Studie I	Studie II
Erfassung der Rahmenbedingungen für die Implementierung des USAT-Programms in den regulären Deutschunterricht	Implementierung des USAT-Programms durch Trainer	Implementierung des USAT-Programms sowie eines Aufsatztrainings ohne selbstregulatorische Komponenten durch Trainer	Implementierung des USAT-Programms sowie eines Aufsatztrainings ohne selbstregulatorische Komponenten durch Lehrer
Strukturiertes Interview mit Lehrern	Eingruppenplan mit Vortest- und Nachtesterhebung	Vortest-Nachtest-Follow-Up-Vergleichsgruppen-Versuchsplan	Vortest-Nachtest-Follow-Up-Vergleichsgruppen-Versuchsplan
-	<i>n</i> =21 (Schüler)	<i>n</i> =156 (Schüler)	<i>n</i> =242 (Schüler)
<i>n</i> =12 (Klasse)	<i>n</i> =1 (Klasse)	<i>n</i> =8 (Klasse)	<i>n</i> =11 (Klasse)

Zu Beginn des Dissertationsprojekts und somit vor der Konzeption des USAT-Programms wurde eine Lehrerbefragung mit zwölf Lehrern durchgeführt, um deren Bedürfnisse und Interessen bezüglich des Aufsatzunterrichts sowie die zeitlichen und organisatorischen Rahmenbedingungen für die Durchführung einer Implementationsstudie festzustellen.

Analog der Standards von Schiffler und Hübner (2000) erfolgte die Evaluation des USAT-Programms in mehreren Schritten. Während die Pilotstudie der formativen Evaluation diente, sollten die Studien I und II Aufsatztrainings mit Selbstregulation und ohne Selbstregulation vergleichen.

In der Pilotstudie wurde das USAT-Programm zunächst in den Klassenkontext einer vierten Klasse integriert. Die Durchführung des Programms erfolgte durch vier geschulte Lerntainer, die abwechselnd das Training in der Klasse umsetzten. Um zu überprüfen, ob das USAT-Programm im Klassenkontext anwendbar war und ob sich die Schreibleistungen der Schüler verbessert haben, wurden vor sowie direkt nach dem Training die Schreibleistungen der Schüler erfasst. Abgesehen von der Überprüfung der Umsetzbarkeit und Effektivität des Trainings war die Optimierung des USAT-Programms das Ziel der Pilotstudie.

Im Gegensatz dazu wurden in der ersten und zweiten Hauptstudie dem USAT-Programm vergleichende Trainings gegenübergestellt, welche sich in beiden Studien im Wesentlichen dadurch von der USAT-Bedingung abgrenzten, dass keine selbstregulatorischen Methoden explizit mit den Schreibstrategien vermittelt wurden. Darüber hinaus fanden in den Hauptstudien zu drei Messzeitpunkten Testungen statt.

Da gemäß O'Donnell (2008) vor der Effektivitätsprüfung eine hinreichende Wirksamkeitsprüfung erfolgen muss, wurde analog zur Pilotstudie das USAT-Programm in der ersten Studie zunächst von vier geschulten Trainern durchgeführt. Dieselben vier Trainer setzten ebenfalls das Kontrolltraining, *Writer's Workshop*, in ganzen Klassen um, so dass jeder Trainer eine USAT-Bedingung und einen *Writer's Workshop* realisierte.

Nachdem sich das USAT-Programm in der ersten Hauptstudie als wirksam erwiesen hatte, wurde in der zweiten Studie die Effektivität des USAT-Programms durch Lehrer untersucht. Dazu wurden Lehrer ausgebildet, das USAT-Programm in den eigenen Klassen umzusetzen. Darüber hinaus erarbeitete eine weitere Gruppe von Lehrern im Sinne der symbiotischen Implementationsstrategie ein eigenes Schreibtraining, welches auf den Schreibstrategien der USAT-Bedingung und den eigenen Erfahrungen der Lehrer basiert. Auf diese Weise diente die zweite Hauptstudie nicht nur der Replizierbarkeit der Befunde der ersten Studie, sondern stellte gemäß O'Donnell (2008) einen weiteren Schritt zur Überprüfung der Effektivität von schulbasierten Interventionen dar, nämlich den Vergleich mit traditionellen Lernmethoden in der Schule.

In beiden Hauptstudien sollten durch die Erfassung der Schreibleistungen sowie weiterer schreibbezogener Maße die Wirksamkeit, Generalisierbarkeit und Nachhaltigkeit des USAT-Programms gegenüber der Kontrollgruppe nachgewiesen werden. Durch den Einsatz von Checklisten zur Überprüfung der inhaltlichen Vollständigkeit und Beobachtungen zur Kontrolle der Qualität der Implementierung konnte überdies eine hohe Implementationsgenauigkeit sichergestellt werden. Ferner strebte die vorliegende Arbeit eine hohe soziale Validität an, so dass nach der Pilotstudie und den beiden Hauptstudien die Lehrer bezüglich der Zufriedenheit mit den jeweiligen Trainingsprogrammen befragt wurden. Ein zeitlicher Überblick der genannten Studien befindet sich in Abbildung 14.

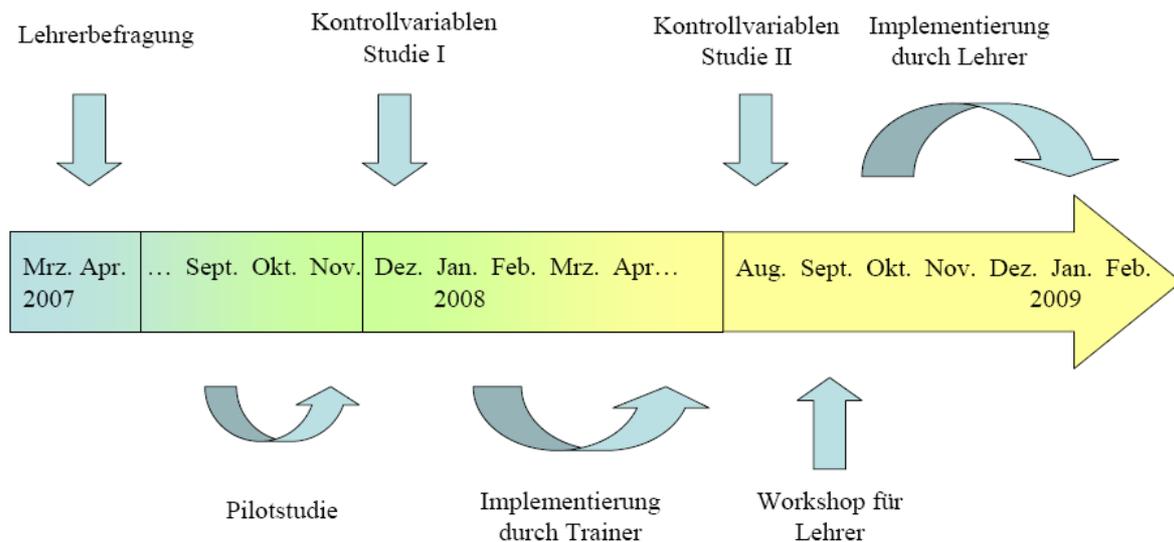


Abbildung 14: Zeitliche Reihenfolge der Studien.

3.4 Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (USAT)

Die für das USAT-Programm charakteristischen Komponenten sind genrespezifische Schreibstrategien, selbstregulatorische Prinzipien und Instruktionmethoden, welche im Folgenden erläutert werden. Eine ausführliche Darstellung der Trainingsinhalte erfolgt in den jeweiligen Kapiteln der Studien.

3.4.1 Schreibstrategien

3-Schritte-Technik: Als erstes erlernten die Schüler eine allgemeine Schreibstrategie zur Organisation und Strukturierung eines selbst konstruierten Textes. Die Strategie umfasst die drei Schreibprozesse des Planens, Schreibens und Überarbeitens von Texten und bezieht sich auf das Schreibmodell von Hayes und Flower (1980). Sie basiert auf der *Three-Step-Strategy* (Harris & Graham, 1996; Reid & Lienemann, 2006).

Beim Planen geht es um das Genieren und Organisieren von Ideen; das Schreiben umfasst Übersetzungsstrategien von Ideen in Sprache; und das Überarbeiten schließlich beinhaltet sowohl das analytische Lesen als auch die Korrektur des geschriebenen Textes.

In dem Dissertationsprojekt wurde die Textproduktion anhand von Bildergeschichten erlernt und überprüft. Bildergeschichten wurden aus dem Grund gewählt, da gemäß der Bildungsstandards für die vierte Klasse im Fach Deutsch nach Anregungen (Texte, Bilder, Musik) geschrieben werden soll (Kultusministerkonferenz, 2005).

Zusätzlich zu der allgemeinen Schreibstrategie, der 3-Schritte-Technik, erlernten die Schüler die für das erzählende Schreiben spezifischen Strategien, die sie während des Planens, Schreibens und Überarbeitens ausführten.

Inhalt: Die erste genrespezifische Strategie, AHA-7W, basiert auf der *Story Grammar Strategy* von Harris und Graham (1996) und wurde bereits in abgewandelter Form von Glaser (2005) eingesetzt (siehe Abbildung 12). Sie umfasst die wesentlichen inhaltlichen Elemente, die in einer vollständigen narrativen Geschichte enthalten sein müssen.

Ausdruck: Überdies lernten die Schüler Strategien zur Verbesserung des Ausdrucks einer Bildergeschichte. Eine wichtige Eigenschaft der SRSD-Instruktion ist die Anpassung alter Strategien an neue (Graham & Harris, 2003; Reid & Lienemann, 2006). Aus diesem Grund wurden die Strategien zur Verbesserung des Ausdrucks anhand der AHA-7W-Strategie erklärt. Die Schüler lernten, die 7W-Fragen durch Adjektive, abwechslungsreiche Verben, Gedanken und Gefühle anschaulich und bildlich zu beschreiben.

Zusammenhang: Das USAT-Programm wird komplettiert durch Strategien zur Verbesserung des Zusammenhangs einer Geschichte. Hierbei handelt es sich um die Berücksichtigung der Reihenfolge der Verben und die Verwendung unterschiedlicher Satzanfänge.

3.4.2 Selbstregulatorische Methoden

Ausgehend von dem Schreibmodell von Hayes und Flower (1980a) verwenden Schreibexperten einen kognitiven Monitor, um Schreibaufgaben erfolgreich zu bewältigen. Dieser kognitive Monitor übernimmt die Steuerung, Koordination und Überwachung der drei Schreibphasen: Planen, Schreiben, Überarbeiten. Es wird davon ausgegangen, dass die Instruktion selbstregulatorischer Prinzipien die Anwendung der Schreibstrategien erleichtert, so dass Schüler bessere Texte produzieren können. Zimmerman (1998) beschreibt einen Lerner als selbstreguliert, wenn dieser seine eigenen Anstrengungen selbst initiiert und lenkt, um Wissen und Fertigkeiten zu erwerben.

Im Rahmen des SRSD-Ansatzes von Harris und Graham (1996) wurden selbstregulatorische Prinzipien wie Zielsetzung, Selbstbewertung und Selbstüberwachung explizit mit den Schreibstrategien vermittelt. Analog zum SAT-Programm stellt dies eine wichtige Komponente des USAT-Programms dar.

Selbstbewertung: In dem USAT-Programm erlernten die Schüler ihre selbst konstruierten Texte anhand von Kriterien, die zuvor in der Klasse gemeinsam erarbeitet wurden, zu bewerten.

Zielsetzung: Darüber hinaus wurden die Schüler angeleitet, vor dem Verfassen eines neuen Textes sich eigene Ziele in Bezug auf die korrekte Anwendung der Schreibstrategien (prozessbezogene Zielsetzung) und in Bezug auf die Gesamtqualität des Textes (produktbezogene Zielsetzung) zu setzen. Schüler erhielten eine Rückmeldung über den Erfolg der angewandten Strategien und über die Qualität des Aufsatzes. Diese Rückmeldungen dienten den Schülern als Referenz für die neue Zielsetzung. Gleichzeitig kontrollierten die Schüler ihren eigenen Leistungsstand.

Selbstüberwachung: Als weitere wesentliche selbstregulatorische Methode lernten die Schüler während des Planens, Schreibens und Überarbeitens ihre Schreibaktivitäten zu überwachen. Dies erfolgte dadurch, dass die Schüler mit Hilfe von Schaubildern, Diagrammen und Checklisten die Ausführung der Schreibprozesse protokollierten.

3.4.3 Instruktionsprinzipien

Zur Vermittlung der Schreibstrategien und der selbstregulatorischen Methoden wurden folgende Instruktionstechniken eingesetzt: kognitives Modellieren, Fading, angeleitetes Üben, selbstständiges Üben, Selbstinstruktionen. Darüber hinaus umfasst die SRSD-Intervention von Harris und Graham (1996) eine sechsstufige Instruktion der Trainingsinhalte. Für das USAT-Programm wurde diese adaptiert und ebenfalls zur Vermittlung der Schreibstrategien und der selbstregulatorischen Methoden eingesetzt.

Stufe 1: Entwickeln und Aktivieren von Hintergrundwissen: Vor dem Beginn der Instruktion von neuen Schreibstrategien wurde das Vorwissen der Schüler in Bezug zu Schreibstrategien und Selbstregulation erfragt und aktiviert. Als erstes erzählten die Schüler, wie sie in der Regel beim Verfassen von Texten vorgehen. Auch ihr schreibbezogenes Wissen, d.h. der Aufbau und die Struktur eines guten Aufsatzes wurden erörtert. Gleichzeitig wurde in dem ersten Instruktionsschritt die Motivation der Schüler gefördert. Nachdem die Eigenschaften eines Schreibexperten erarbeitet worden waren, wurde den Schülern das Ziel in Aussicht gestellt, im Verlauf des Trainings Schreibexperte zu werden. Jedes Kind sollte sich dabei mit diesem Ziel identifizieren.

Stufe 2: Diskutieren: Als nächstes wurde die neue Strategie von dem Trainer bzw. dem Lehrer vorgestellt und erklärt. Zusammen mit den Schülern wurden die Vorteile und die Bedeutsamkeit, aber auch die Grenzen und Schwierigkeiten der Strategien herausgearbeitet. Auf einem Poster wurden die neuen Strategien festgehalten.

Stufe 3: Modellieren: Im Anschluss an das Diskutieren modellierte der Trainer bzw. der Lehrer den Gebrauch der Strategie. Er verwendete dabei Selbstinstruktionen, um laut alle

handlungsleitenden Gedanken zu verbalisieren und den Schülern den korrekten Einsatz der Strategien zu erleichtern.

Stufe 4: Memorieren: Zu Beginn jeder Trainingsstunde wurden die bereits erlernten Schreib- und selbstregulatorischen Strategien mit der gesamten Klasse wiederholt. Poster an den Wänden sowie die zuvor instruierten Mnemotechniken unterstützten die Wiedergabe der Strategien. Darüber hinaus wurden Übungen zur Wiederholung der Strategien, die von den Schülern in Gruppenarbeit oder alleine ausgeführt werden, durchgeführt. Weitere Lern- und Gedächtnisstützen in Form von Merkkarten, Blättern und Symbolen erleichterten zudem das Einprägen der Strategien.

Stufe 5: Unterstützen: Abgesehen von den genannten Hilfestellungen durch entsprechende Materialien unterstützte außerdem der Trainer bzw. der Lehrer die Schüler durch konkrete verbale Hilfestellungen, aber auch durch Lob und Anerkennung. Je nach Leistungsniveau der Schüler passte der Trainer bzw. der Lehrer die Unterstützungen an, so dass schreibschwache Schüler mehr gefördert wurden als schreibstarke Schüler.

Stufe 6: Unabhängige Leistung: Nachdem die Strategien ausreichend wiederholt und geübt worden waren, wurden die Hilfestellungen und die Unterstützung soweit zurückgenommen bis die Schüler selbstständig die Strategien anwenden konnten. Am Ende des Trainings sollten die Schüler alleine und ohne Hilfestellungen eine Geschichte planen, schreiben und überarbeiten, indem sie die erworbenen Schreibstrategien und selbstregulatorischen Techniken einsetzten.

4 Lehrerbefragung

4.1 Ziele der Lehrerbefragung

Bevor das *Unterrichtsintegrierte Selbstregulatorische Aufsatztraining* (USAT) im gesamten Klassenkontext erprobt wurde, musste zunächst festgestellt werden, welche Veränderungen im aktuellen Aufsatzunterricht der vierten Klasse notwendig sind. Wie bereits in Kapitel 2.4.1.2 beschrieben, setzen Lehrer neue Unterrichtsmethoden nur dann ein, wenn deren Bedürfnisse, Interessen und Erwartungen berücksichtigt werden (Blumenfeld et al., 2000; Zech, Gause-Vega, Bray, Secules, & Goldman, 2000). Aus diesem Grund wurde vor der Erprobung des USAT-Programms im gesamten Klassenkontext, eine Interviewstudie mit Lehrern der vierten Klasse durchgeführt. Die Ergebnisse der Befragung wurden bei der Konzeption des USAT-Programms und bei der Gestaltung der Lehrer-Workshops (siehe Kapitel 7) berücksichtigt. Folgende Fragen standen im Mittelpunkt der Erhebung:

- Wie sieht der traditionelle Aufsatzunterricht der vierten Klasse aus?
- Welche Strategien und Materialien werden im Aufsatzunterricht verwendet?
- Wie groß ist das Interesse der Lehrer an Fortbildungen im Bereich Aufsatzerziehung?

4.2 Methode

Stichprobe: An der Befragung nahmen zwölf Lehrer teil, die zum Zeitpunkt der Untersuchung eine vierte Klasse in Gießen oder im Gießener Umkreis unterrichteten. Alle beteiligten Lehrer waren weiblich. Die durchschnittliche Berufserfahrung lag bei 19,17 Jahren. Die Rekrutierung der Lehrerinnen erfolgte telefonisch und gestaltete sich als schwierig, da ein Großteil der angesprochenen Lehrerinnen aufgrund hoher Arbeitsbelastung und zu vielen Anfragen von Seiten der Universität absagte. Von den insgesamt zwölf Lehrerinnen erklärten sich neun zu einem persönlichen Gespräch und drei zu einem Telefoninterview bereit.

Ablauf der Befragung: Im Vergleich zum persönlichen Interview (siehe CD-Anhang A1) war das Telefoninterview (siehe CD-Anhang A1) um einige Fragen kürzer und konnte nicht mit einem Tonband aufgenommen werden. Während die Telefoninterviews zwischen zehn bis zwanzig Minuten dauerten, waren die persönlichen Kontakte mit den Lehrerinnen zeitlich umfangreicher (45 bis 60 Minuten). Die Treffen mit den Lehrerinnen fanden in den jeweiligen Klassenzimmern am Ende des Schultages statt, so dass die Lehrerinnen vor Ort Materialien zeigen konnten. Alle Interviews wurden am Ende des Schuljahres im März/April

2007 durchgeführt, d.h. zu einem Zeitpunkt, als ein Großteil des Unterrichtsstoffs der vierten Klasse bereits behandelt worden war. Die Befragung der zwölf Lehrerinnen erfolgte durch die Doktorandin.

Interviewleitfaden: Der für die Lehrerbefragung entwickelte Leitfaden (siehe CD-Anhang A1), der sowohl für das persönliche Gespräch als auch für das Telefongespräch als Grundlage diente, erfasst die Einstellungen der Lehrer zum Aufsatzunterricht in der vierten Klasse. Folgende Themen werden im Leitfaden angesprochen:

- Zeitpunkt sowie zeitlicher Umfang für die Instruktion, Wiederholung und Übung des erzählenden Schreibens;
- Schreibstrategien zum Planen, Schreiben und Überarbeiten von Texten;
- Selbstregulatorische Methoden und kooperative Lernformen;
- Materialien und Informationsquellen, die als Grundlage für die Unterrichtsplanung dienen;
- Fortbildungsmaßnahmen im Bereich Schreiben und Selbstregulation.

Kategoriensystem: In Anlehnung an Mayring (2008) wurden die Interviews anhand einer strukturierenden Inhaltsanalyse ausgewertet. Diese diente als Grundlage für weitere quantitative Analysen.

Zunächst wurden auf Grundlage des Interviewleitfadens Überkategorien gebildet (Gräsel, Fussangel, & Parchmann, 2006; Mayring, 2001). In einem weiteren Schritt wurden die Inhalte der Interviews diesen Kategorien zugeordnet. Anschließend wurden generelle Aussagen pro Kategorie getroffen. Die nach Mayring (2001) im Anschluss an eine qualitative Inhaltsanalyse mögliche quantitative Analyse diente der Ermittlung von Häufigkeiten der Kategorienbesetzung.

4.3 Ergebnisse

Insgesamt wurden folgende Überkategorien gebildet: zeitlicher Rahmen, Strategien, Materialien und Fortbildungen.

Zeitlicher Rahmen: Alle zwölf Lehrerinnen unterrichteten in der vierten Klasse mindestens vier bis maximal sechs Schulstunden Deutsch pro Woche. Der Großteil der Lehrerinnen (75%) schätzte, dass 25-30% des gesamten Deutschunterrichts das Geschichtschreiben umfasst und somit einen gleichwertigen Stellenwert mit Rechtschreibung, Grammatik und Lesen einnimmt. Ein kleinerer Anteil (25%) empfand das Geschichtschreiben als weniger wichtig als andere Bereiche des Deutschunterrichts und gaben an, sich weniger als 25% des Unterrichts mit dem Geschichtschreiben zu

beschäftigen. Keine der befragten Lehrerinnen konnte einen genauen Zeitpunkt für das Unterrichten des Geschichtschreibens nennen. Während in der dritten Klasse schwerpunktmäßig das erzählende Schreiben instruiert und geübt wird, passiert dies in der vierten Klasse eher „nebenher“ als Hausaufgabe oder „wie ein roter Faden“. Aus diesem Grund fiel es den meisten Lehrerinnen schwer anzugeben, wie viele Wochen sie benötigen, um das Geschichtschreiben einzuführen. Ein Drittel der befragten Lehrerinnen (33%) gab zwei bis drei Wochen an. Ein weiterer Teil (jeweils 17%) benötigte ein bis zwei bzw. vier bis fünf Wochen. 42% der Lehrerinnen begannen mit dem Aufsatzunterricht in der zweiten Klasse, weitere 33% in der dritten Klasse. Insgesamt konnten die Lehrerinnen weder den genauen Zeitpunkt, d.h. wann in der vierten Klasse das Schreiben thematisiert wird, angeben, noch die Dauer und den Umfang des Aufsatzunterrichts spezifizieren.

Strategien: Die Lehrerinnen wurden gefragt, ob sie den Kindern planungsspezifische Strategien sowie Strategien zum Überarbeiten von Texten vermitteln. Während 75% der Lehrerinnen keine, 17% ein paar und nur 8% (dies entsprach einer Lehrerin) planungsspezifische Strategien einsetzten, gaben 42% der Lehrerinnen an, ihren Schülern Revisionstechniken zu vermitteln. Von den übrigen Lehrerinnen instruierten 33% ein paar und nur 25% keine Revisionstechniken. Auf Nachfragen zeigte sich, dass es sich bei den Revisionstechniken um Schreibkonferenzen handelte, die in den letzten Jahren sehr beliebt geworden sind und in mehreren Grundschulzeitschriften vorgestellt wurden (Fix, 1999; Käferle, 2007; Simon, 1997). Des Weiteren berichteten 25% bzw. 42% der Lehrerinnen, manchmal bzw. regelmäßig kooperative Lernformen im Unterricht zu verwenden. Das kooperative Lernen fand größtenteils in den Schreibkonferenzen statt. Hinsichtlich selbstregulatorischer Methoden besaßen die befragten Lehrerinnen sehr wenig Wissen und kannten den Begriff größtenteils nicht. Nachdem dieser erklärt wurde, meinten 33% der Befragten ein paar selbstregulatorische Methoden, wie z.B. den Wochenplan, im Zusammenhang mit dem Geschichtschreiben zu benutzen. 67% der Lehrerinnen verwendeten keine selbstregulatorischen Methoden. Zusammenfassend ist festzustellen, dass bezüglich der Schreibstrategien und der Selbstregulation erhebliche Wissenslücken bei den befragten Lehrerinnen vorlagen.

Materialien: Die Lehrerinnen wurden danach gefragt, welche Materialien sie als Vorbereitung und für die Durchführung des Aufsatzunterrichts heranziehen. Mehrfachantworten waren möglich. 75% der Befragten gaben „Erfahrung“ und „Lehr- und Schulbücher“ als Grundlage für ihre Unterrichtsgestaltung an. Die verwendeten Schulbücher variierten jedoch. In der Regel wird eine Schulbuch-Serie von der gesamten Schule

verwendet. Als zweites nannten die Lehrerinnen „Fortbildungen“ (33%) und „Internet“ (17%). Des Weiteren wurde nach Kriterien für die Auswahl von Materialien gefragt. Ein Großteil gestaltete den Unterricht „passend zu den Kindern“ (42%) oder orientierte sich an dem Rahmenplan (33%). Weitere 17% suchten die Materialien „passend zu dem Gesamtunterricht“ aus. Eine Lehrerin (8%) berichtete, keine Kriterien bei der Auswahl zu berücksichtigen und das Material „willkürlich“ auszuwählen. Generell besaß keine der befragten Lehrerinnen einen strukturierten Unterrichtsablauf zum Thema erzählendes Schreiben, so dass jede nach eigenem Ermessen und aufbauend auf eigenen Erfahrungen bzw. Erfahrungen der Kollegen den Aufsatzunterricht gestalteten.

Fortbildungen: Als letzten inhaltlichen Schwerpunkt wurden den Lehrerinnen Fragen in dem Bereich Fortbildungen gestellt. Dabei schätzten 42% der Lehrerinnen den Schreibunterricht als schwerer bzw. 50% als etwas schwerer im Vergleich zu anderen Themengebieten ein. Schließlich äußerten 58% Interesse bzw. 33% etwas Interesse an Fortbildungsangeboten im Bereich Schreiben und Selbstregulation, wobei weitere Informationen in Bezug zu selbstregulatorischen Methoden explizit gewünscht wurden.

4.4 Diskussion

Die durchgeführte Lehrerbefragung gibt einen exemplarischen Überblick über den aktuellen Aufsatzunterricht in der vierten Klasse. Die dargestellten Ergebnisse geben wichtige Einblicke sowie Hinweise auf die Interessen und Bedürfnisse der Lehrerinnen, welche bei der Entwicklung eines unterrichtsintegrierten Aufsatztrainings berücksichtigt werden müssen.

Zusammenfassung der Ergebnisse: Zunächst ist festzustellen, dass das Aufsatzschreiben aus Sicht der Mehrheit der befragten Lehrerinnen ein wichtiger Bestandteil des Deutschunterrichts darstellt. Gleichzeitig verfügen die Lehrerinnen über unzureichendes Strategiewissen, über keine inhaltlich und zeitlich strukturierten Unterrichtseinheiten zum Thema Aufsatzschreiben und über keine wissenschaftlichen und empirisch fundierten Quellen für die Unterrichtsplanung. Folglich empfinden die meisten Lehrerinnen das Unterrichten von Schreiben als schwerer im Vergleich zu anderen Themen des Deutschunterrichts sowie im Vergleich zu anderen Unterrichtsfächern. Aus diesem Grund äußerte über die Hälfte der Lehrerinnen großes Interesse an Fortbildungen in dem Bereich Schreiben und Selbstregulation.

Die Ergebnisse der Untersuchungen stimmen teilweise mit den Ergebnissen der national repräsentativen Grundschullehrerbefragung von Cutler und Graham (2008) überein. Auch wenn in den USA bereits 16% des Schreibunterrichts für die Instruktion von Planungs-

und Revisionsfertigkeiten verwendet wird, befürworten die Autoren eine stärkere Fokussierung auf die Vermittlung von spezifischen Schreibprozessstrategien. In der dargestellten Studie besaß ein Viertel der befragten Lehrerinnen kein Wissen über planungsspezifische Strategien, wobei jedoch die meisten Lehrerinnen Revisionsstrategien vermittelten. In Deutschland scheint ebenfalls ein Bedarf an Weiterbildungen in diesem Bereich zu bestehen. Bezüglich der Schulung der Lehrerinnen fordern Cutler und Graham (2008) eine Verbesserung der bestehenden Angebote. Im Vergleich zu den oben dargestellten Ergebnissen sind die amerikanischen Lehrer jedoch deutlich besser vorbereitet und verfügen über ein größeres Repertoire an Unterrichtsmethoden für den Aufsatzunterricht. Auch wenn der Umfang stark variierte, setzten 90% der von Cutler und Graham (2008) befragten Lehrer drei der vier vorgegebenen Unterrichtspraktiken mindestens einmal im Jahr ein. Die meisten Lehrer (72%) setzten dabei einen Prozessansatz kombiniert mit traditioneller Strategievermittlung ein. Bei den in der vorliegenden Untersuchung befragten Lehrerinnen konnte keine einen spezifischen Unterrichtsansatz für das Aufsatzschreiben nennen. Dreiviertel verwendeten Erfahrung und Lehrbücher als Grundlage für ihre Unterrichtsplanung.

Schlussfolgerungen: Die berichteten Befunde sind für die Planung und Konzeption des USAT-Programms in mehreren Aspekten nützlich:

- Da kein fester Zeitpunkt für den Unterricht des erzählenden Schreibens vorzuliegen scheint, kann der Zeitpunkt der Durchführung sowohl in dem ersten als auch in dem zweiten Schulhalbjahr der vierten Klasse erfolgen.
- Aufgrund objektiver sowie subjektiv empfundener Wissens- und Fertigungsdefizite ist von einer hohen Bereitschaft und einem hohen Interesse an Fortbildungsangebote in dem Bereich Selbstregulation und Schreiben auszugehen. Eine wichtige Voraussetzung für den Erfolg von Implementationsmaßnahmen kann somit als gegeben angenommen werden (Blumenfeld et al., 2000; Gräsel & Parchmann, 2004; Zech et al., 2000).

Abgesehen davon weisen die dargestellten Befunde Grenzen und Einschränkungen auf. Zunächst handelt es sich um eine sehr kleine Stichprobe, die nur weibliche Lehrer aus dem Raum Gießen umfasst, und deren Teilnahme nicht zufällig erfolgte. Die Lehrerinnen meldeten sich freiwillig bereit, die Fragen zu beantworten. Mehrere Lehrer bzw. Schulen lehnten eine Teilnahme ab. Da die befragten Lehrerinnen sich bereits aufgrund ihrer Teilnahmebereitschaft von der restlichen Lehrerpoptation hervorheben könnten, sind die Ergebnisse zunächst nur für diese spezielle Gruppe zutreffend.

Diese Einschränkungen beeinflussen jedoch nicht die Schlussfolgerungen für die Konzeption des USAT-Programms. Die durchgeführte Interviewstudie intendierte eine rein deskriptive Darstellung der aktuellen Situation im Aufsatzunterricht der vierten Klasse und diente als Voruntersuchung, um möglichst früh die Meinungen und Bedürfnisse von Lehrern in die Konzeption des Trainings mit einfließen zu lassen. Die obigen Ergebnisse geben einen exemplarischen Überblick über die aktuelle Unterrichtssituation in der Region, in der die Implementationsstudien durchgeführt werden sollten. Eine Interviewstudie mit einer größeren, zufällig gezogenen und repräsentativeren Stichprobe, in der Unterrichtsbeobachtungen bzw. Videoaufnahmen durchgeführt werden, könnte globalere Aussagen treffen. Für die Zukunft ist dies ein interessantes Aufgabenfeld, um explizit die Implementationsforschung auf die Situation und somit die Bedürfnisse und Interessen der Lehrer und Schulen abzustimmen. Folglich wäre eine national repräsentative Erhebung in Deutschland wünschenswert, so wie sie von Cutler und Graham (2008) in den USA durchgeführt wurde.

5 Pilotstudie

5.1 Vorarbeiten

Aus der vorab durchgeführten Lehrerbefragung ergab sich, dass die Lehrerinnen größtenteils positiv neuen Lehrmethoden im Bereich Schreiben gegenüberstanden, aber auch, dass für Lehrer wichtige Rahmenbedingungen bei der Implementierung des USAT-Programms berücksichtigt werden müssen. Aktuelle Implementationsansätze betonen, dass neue Unterrichtsmethoden eher von Lehrern umgesetzt werden, wenn diesen Adaptations- und Mitsprachemöglichkeiten eingeräumt werden (Blumenfeld et al., 2000). Dazu zählen zum einen die frühe Einbindung von praktizierenden Lehrern in die Konzeptionsphase sowie zum anderen die Schaffung von Freiräumen bei der Gestaltung der zu implementierenden Innovation (Gräsel & Parchmann, 2004). Aus diesem Grund erfolgten die Vorarbeiten vor Beginn der Pilotstudie in Kooperation mit einer Grundschulreferendarin.

Konzeption des USAT-Programms: Die Grundschulreferendarin wurde bereits bei der Konzeption der Trainingsinhalte des USAT-Programms mit einbezogen. Sie unterrichtete seit einem Jahr an einer hessischen Grundschule in den Jahrgängen eins bis vier. Zudem hatte sie sich bereits im Rahmen ihrer Examensarbeit an der Justus-Liebig-Universität mit dem SAT-Programm befasst.

Die Konzeptionsphase gestaltete sich zunächst so, dass theoriegeleitet die Manuale für die einzelnen Trainingseinheiten sowie die dazugehörigen Unterrichtsmaterialien von der Doktorandin erstellt wurden. Anschließend wurden die einzelnen Trainingseinheiten mit der Grundschulreferendarin sowie den Trainern der Pilotstudie besprochen. Änderungsideen und Implementierungsschwierigkeiten wurden diskutiert und in die Manuale eingebaut. Die Ideen der Grundschulreferendarin waren in diesem Zusammenhang besonders bereichernd, da diese potenzielle Problemfelder im Unterrichtsalltag aufzeigte sowie die Sichtweise der Lehrer mit einbrachte. Die in der Pilotstudie eingesetzten Manuale sind im Anhang zu finden (siehe CD-Anhang A4.2).

Erprobung einzelner Trainingseinheiten: In einem weiteren Schritt zur Entwicklung des USAT-Programms führte die Grundschulreferendarin im September 2007 einzelne Elemente der drei ersten Unterrichtseinheiten, in denen der Großteil der Instruktionen erfolgte, in einer vierten Klasse in der Nähe von Hanau durch. Dies diente folgenden Zielen:

- Überprüfung der Verständlichkeit zentraler Instruktionen und Übungen für Schüler und Trainer;

- Überprüfung der Praktikabilität der Instruktionen und Übungen in dem gesamten Klassenkontext;
- Überprüfung, ob das Übungsmaterial für den Unterricht in der vierten Klasse adäquat schwierig ist;
- Vergleich der geplanten Zeit mit der tatsächlichen Dauer für Instruktionen und Übungen.

In der folgenden Tabelle 9 sind die wichtigsten Inhalte der drei von der Referendarin durchgeführten Trainingseinheiten aufgelistet. Eine Unterrichtsstunde dauerte 90 Minuten. Jede Unterrichtseinheit wurde von der Doktorandin beobachtet und im Anschluss mit der Referendarin reflektiert.

Tabelle 9: Inhalte der Trainingseinheiten der Vorarbeiten.

TE⁹ 1	<p>Lehrerin leitet zusammen mit den Schülern die AHA-7W-Fragen-Strategie her.</p> <p>Lehrerin modelliert das Planen und Schreiben einer Bildergeschichte anhand des Geschichtenplans.</p> <p>Schüler planen und schreiben eigene Geschichten.</p> <p>Lehrerin modelliert das Überarbeiten mit Hilfe einer Checkliste.</p> <p>Schüler überarbeiten ihre Geschichten.</p> <p>Lehrerin erklärt die Hausaufgabe (Übungen zur Wiederholung der 7W-Fragen).</p>
TE 2	<p>Lehrerin wiederholt die erlernten Strategien der letzten Einheit.</p> <p>Lehrerin erklärt die Gruppenarbeit. Ein Gruppenchef wird ernannt, welcher mit Hilfe von Fragen die Gruppenarbeit leitet.</p> <p>In Gruppen von vier bis sechs Kindern planen die Schüler gemeinsam eine Geschichte mit dem Geschichtenplan.</p> <p>Lehrerin gibt den Schülern strategie- und ergebnisbezogene Rückmeldungen zu der in der letzten Einheit geschriebenen Geschichte anhand von Protokollen.</p> <p>Lehrerin modelliert die Zielsetzung mit Hilfe der Protokolle. Anschließend setzen sich Schüler strategie- und ergebnisbezogene Ziele für die nächste Geschichte.</p> <p>Lehrerin erklärt die Hausaufgabe: die Schüler sollen eine Bildergeschichte planen, schreiben und überarbeiten.</p> <p>Lehrerin erklärt die zweite Gruppenarbeit: die Schüler sollen in Gruppen einen vorgegebenen Text mit Hilfe der Checkliste überprüfen und anschließend überarbeiten.</p>

⁹ TE = Trainingseinheit; Dauer: 90 Minuten.

TE 3	<p>Lehrerin wiederholt die erlernten Strategien der letzten Einheiten und gibt den Schülern strategie- und ergebnisbezogene Rückmeldungen zu der Bildergeschichte, welche die Schüler als Hausaufgabe geschrieben haben.</p> <p>Lehrerin modelliert die Anwendung der Ausdrucksstrategien.</p> <p>Lehrerin erklärt die Gruppenarbeit und ernennt einen Gruppenchef, der die Arbeit mit Hilfe von Fragen leitet. Die Schüler sollen in Gruppen eine Geschichte bildlich und anschaulich ausgestalten.</p> <p>Die Lehrerin modelliert die Verbesserung des Zusammenhangs in einer Geschichte. Anschließend sortieren die Schüler im Rahmen einer Gruppenarbeit einzelne Sätze, so dass eine zusammenhängende Geschichte entsteht.</p>
-------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Während die Instruktionen der Schreibstrategien in der Regel gut funktionierten und von den Schülern sowohl verstanden als auch richtig umgesetzt wurden, gestalteten sich die Gruppenarbeiten schwierig, da die Schüler sehr unruhig wurden und nicht konzentriert an der Aufgabe arbeiteten. Obwohl die Klassenlehrerin die Gruppenaufteilung vornahm, gab es teilweise Konflikte innerhalb der Gruppen. Für die folgenden Studien wurden deswegen Optimierungen der Gruppenarbeit vorgenommen, die hauptsächlich die Motivation zur Zusammenarbeit steigern sollten. Darüber hinaus wurden einzelne Abschnitte der Manuale sowie Einzelheiten der Materialien aufgrund der praktischen Erfahrungen optimiert. Die Veränderungen betrafen vornehmlich Formulierungen, aber auch Darstellungsformen, die den Schülern die Strategien verständlicher machen sollten.

Ferner wurden für die Studie I Tipps für Trainer zusammengestellt, welche sich insbesondere auf den Umgang mit Störungen bezogen. Es handelt sich dabei um folgende Punkte, die mit den Trainern besprochen wurden:

- Trainings in den ersten (erste und zweite Stunde) sowie in den letzten (fünfte und sechste Stunde) Unterrichtsstunden sollten vermieden werden, da die Kinder müde und unaufmerksam sind.
- Schüler, die mit dem Planen, Schreiben und Überarbeiten eher fertig sind, müssen Füllübungen zur Verfügung gestellt bekommen, da diese ansonsten andere Schüler stören.
- Während der Instruktionsphasen durch den Trainer sollen die Schüler sämtliche Materialien auf den Tischen zur Seite legen, damit sie nicht abgelenkt werden.
- Insbesondere beim Überarbeiten müssen die Schüler motiviert werden, gut mitzuarbeiten, da die Konzentration am Ende nachlässt.
- Gespräche der Schüler untereinander müssen sofort unterbunden werden.

- Vor dem Training sollte ein Gespräch mit den Klassenlehrern erfolgen, in der die Besonderheiten einzelner Schüler oder der gesamten Klasse besprochen werden. Auch Regeln und Routinen, die die Schüler gelernt haben, sollten erfragt werden.
- Die Trainer sollten sich einen Zeitplan mit genauen Zeitvorgaben für die einzelnen Trainingsinhalte erstellen. Verzögerungen sollten die Trainer so früh wie möglich erkennen und diesen entgegenwirken.
- Materialien sollten abgezählt werden, um das Austeilen zu beschleunigen.

5.2 Fragestellung und Hypothesen

Die deskriptiv orientierten Vorarbeiten zeigten bereits wichtige Punkte zur Verbesserung des USAT-Programms auf. Im Rahmen der Vorarbeiten führte jedoch eine im Unterrichten erfahrene Referendarin nur einzelne Ausschnitte des Trainings durch. In der Pilotstudie dagegen wurde das Training von geschulten Lerntrainern, die keine Erfahrungen im Unterrichten besaßen, über einen Zeitraum von sechs Wochen umgesetzt, so dass vielfältigere Probleme auftreten konnten. Des Weiteren fokussierten die Vorarbeiten vornehmlich die Überprüfung der Verständlichkeit sowie Durchführbarkeit der zentralen Instruktionen des USAT-Programms. Während die Pilotstudie dies ebenfalls nachweisen wollte, sollte sie zudem zeigen, dass das USAT-Programm tatsächlich die Schreibleistungen der Schüler nach dem Training signifikant verbesserte. Insgesamt ergaben sich folgende Ziele für die Pilotstudie:

- Überprüfung der Verständlichkeit zentraler Instruktionen und Übungen für Schüler und Trainer;
- Überprüfung der Praktikabilität der Instruktionen und Übungen in dem gesamten Klassenkontext;
- Überprüfung, ob das Übungsmaterial für den Unterricht in der vierten Klasse adäquat schwierig ist;
- Vergleich der geplanten Zeit mit der tatsächlichen Dauer für Instruktionen und Übungen;
- Überprüfung der Effektivität des USAT-Programms mit Hilfe eines Eingruppenplans mit Vortest-Nachtest-Erhebungen;
- Schulung und Vorbereitung der Trainer auf die Durchführung der Studie I.

Unter Anknüpfung an die im Kapitel 3.1 formulierten Ziele und Hypothesen intendierte die Pilotstudie vornehmlich die Überprüfung der Implementierbarkeit des USAT-Programms in den Regelunterricht. Folglich konnten folgende Hypothesen formuliert werden.

Hypothese 1a: Es wurde erwartet, dass am Ende des USAT-Programms die Schüler inhaltlich vollständigere, sprachlich besser ausgestaltete, umfangreichere und qualitativ bessere Aufsätze schreiben als vor dem Training.

Hypothese 1b: Ferner wurde angenommen, dass die Schüler im Anschluss an das USAT-Programm bessere Planungsnotizen anfertigen und mehr überarbeiten.

Um am Ende des Trainings aussagefähige Schlussfolgerungen ziehen zu können, wurden Maßnahmen zur Sicherung der Implementationsgenauigkeit ergriffen.

5.3 Methode

5.3.1 Untersuchungsdesign, Ablauf und Stichprobe

Der Pilotstudie lag ein Eingruppenplan mit Vortest- und Nachtesterhebungen zu Grunde. Das USAT-Programm wurde in einer vierten Klasse mit 21 Schülern (elf Mädchen, zehn Jungen) aus einer öffentlichen Grundschule aus einem Vorort von Gießen erprobt. Das Durchschnittsalter betrug 9.5 Jahre. 73% der Schüler gaben an, zu Hause vornehmlich Deutsch zu sprechen.

Die Rekrutierung erfolgte aufgrund eines persönlichen Kontakts zwischen einer Trainerin und der Klassenlehrerin. Vor Beginn der Testung und des Trainings wurden die Eltern der Schüler auf einem Elternabend und durch einen Brief von der Klassenlehrerin informiert.

Das USAT-Programm wurde von vier Trainerinnen in der gesamten Klasse umgesetzt und fand im regulären Deutschunterricht statt. Trainerin A war die Doktorandin. Die Trainerinnen B und C waren Psychologiestudentinnen im Hauptstudium. Trainerin D war eine Lehramtsstudentin (Grundschule) mit den Fächern Deutsch und Musik im Hauptstudium. Das USAT-Programm umfasste sechs Trainingseinheiten, wobei zwei der Trainerinnen (Trainerinnen A und B) eine Trainingseinheit (à 90 Minuten) durchführten und die übrigen zwei (Trainerinnen C und D) jeweils zwei Trainingseinheiten umsetzten. Alle Trainerinnen nahmen an allen Trainingseinheiten teil.

Da pro Woche eine bis zwei Trainingseinheiten stattfanden, dauerte das Training insgesamt vier Wochen. Darüber hinaus wurden an zwei verschiedenen Messzeitpunkten die Schreibleistungen erhoben. Der Prätest erfolgte eine Woche vor Beginn des Trainings und der Posttest eine Woche nach der letzten Trainingsstunde. Die gesamte Pilotstudie wurde im Zeitraum Oktober/November 2007 durchgeführt.

5.3.2 Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (USAT)

Im Kapitel Fragestellung wurde bereits der prinzipielle Aufbau des USAT-Programms vorgestellt. In dem folgenden Abschnitt werden nun die Trainingseinheiten und die eingesetzten Materialien der Pilotstudie genauer beschrieben.

5.3.2.1 Schreibstrategien

Im USAT-Programm wurden sowohl allgemeine Schreibstrategien wie die 3-Schritte-Technik als auch genrespezifische Schreibstrategien zur Verbesserung des Inhalts, des Ausdrucks und des Zusammenhangs einer Geschichte vermittelt.

3-Schritte-Technik: Als erstes lernten die Schüler eine Strategie, mit der sie eine Schreibaufgabe strukturieren und organisieren konnten: die 3-Schritte-Technik. Sie basiert auf dem kognitiven Schreibmodell von Hayes und Flower (1980).

Die Schüler wurden angeleitet, ihre Geschichten zu planen, zu schreiben und zu überarbeiten. Damit die Schüler jedoch effektiv planen, schreiben und überarbeiten konnten, wurden ihnen zusätzlich genrespezifische Strategien vermittelt, die sie während der drei Schreibphasen anwenden sollten. Da es sich beim USAT-Programm um das narrative Schreiben auf der Grundlage von Bildern handelt, erlernten die Schüler Strategien zum Verfassen von Bildergeschichten. Die 3-Schritte-Technik diente somit als Gerüst für sämtliche Schreib- und selbstregulatorischen Strategien. Als Erinnerungshilfe hing im Klassenzimmer ein Poster (siehe Abbildung 15). Jeder Buchstabe in den Wörtern „plane“, „schreibe“ und „überarbeite“ steht für eine Aktivität, die die Schüler jeweils beim Planen, Schreiben und Überarbeiten ausführen sollten. Vor der Einführung des Posters leiteten sich die Schüler die Mnemonik „3-Schritte-Technik“ im Rahmen einer Hausaufgabe her (siehe CD-Anhang A4.3).

Explizite Planungsstrategien erlernten die Schüler anhand einer Bildergeschichte (siehe CD-Anhang A7.4). Das Überarbeiten modellierte die Trainerin mit einer Beispielgeschichte (siehe CD-Anhang A7.5), die sie zunächst überprüfte und anschließend korrigierte. Sie verbalisierte dabei laut alle handlungsleitenden Gedanken, so dass die Schüler jeden Schritt nachvollziehen konnten.

3-SCHRITTE-TECHNIK

P-L-A-N-E

- P** = Pass auf das Thema auf und überlege, was du machen musst
- L** = Lass deiner Phantasie freien Lauf und stelle dir die Geschichte bildlich vor
- A** = Arbeite mit den 7-W-Fragen
- N** = Nicht durcheinander planen, sondern eine Reihenfolge überlegen
- E** = Einzelne Stichpunkte im Geschichtenplan aufschreiben

SCH-R-EI-B-E

- Sch** = Schreibe eine spannende Überschrift
- R** = Richte dich nach dem Geschichtenplan
- Ei** = Eine Reihe frei lassen, wenn du deine Geschichte schreibst
- B** = Beachte deine Ziele aus dem 7-W-Protokoll
- E** = Einen Haken machen, wenn du etwas aus deinem Geschichtenplan verwendet hast

ÜBER-A-R-B-EI-T-E

- Über** = Überlege, was deine Ziele sind
- A** = Alles noch einmal langsam durchlesen
- R** = Richte dich nach der Checkliste
- B** = Beachte die 7-W-Fragen
- Ei** = Einen Haken machen, wenn du eine der 7-W-Fragen entdeckst
- T** = Text in die freie Zeile einfügen und mit einem Pfeil markieren
- E** = Eine mene miste, die Geschichte ist in der Kiste

Abbildung 15: 3-Schritte-Technik-Poster der USAT-Bedingung.

Inhalt: Als erste genrespezifische Schreibstrategie, die die Schüler während des Planens, Schreibens und Überarbeitens einsetzen sollten, lernten sie die Merkhilfe AHA-7W zur Verbesserung des Inhalts einer Geschichte kennen. Die Erinnerungshilfe AHA-7W ist eine strukturelle und inhaltliche Orientierungshilfe während der drei Schreibphasen.

Der erste Teil der Mnemotechnik, AHA, erinnert dabei an den **A**nfang, **H**auptteil und den **A**bschluss einer guten Geschichte. Die Abkürzung 7W bezieht sich auf die 7W-Fragen, die in jeder inhaltlich vollständigen Geschichte vorkommen müssen. Der Anfang der Geschichte enthält den Namen der Hauptperson (**W**er ist die Hauptperson?), ihr Ziel (**W**as ist das Ziel der Hauptperson?), den Ort (**W**o spielt die Geschichte?) und die Zeit (**W**ann spielt die Geschichte?) der Geschichte. Der längste Teil der Geschichte ist der Hauptteil. Hier werden die Handlungsschritte (**W**as macht die Hauptperson Schritt für Schritt?) und der Höhepunkt (**W**as ist der Höhepunkt?) beschrieben. Schließlich endet die Geschichte mit dem Abschluss, in dem erläutert wird, **w**ie die Geschichte endet. Zur Erinnerung und zur Erleichterung des Einprägens wurde ein Poster mit den 7W-Fragen in das Klassenzimmer gehängt (siehe Abbildung 16).



Abbildung 16: 7W-Fragen-Poster.

Ausdruck: Als nächstes wurden Strategien zur Verbesserung des Ausdrucks einer Geschichte eingeführt. Die in diesem Zusammenhang verwendete Ausdruckskarte (siehe Abbildung 17) orientiert sich an den 7W-Fragen und enthält Beispiele für eine bildliche und anschauliche Beschreibung der einzelnen 7W-Fragen. Als Merkhilfe wurde den Schülern mitgeteilt, Adjektive, abwechslungsreiche Verben, Gedanken und Gefühle in die Geschichte zu integrieren. Während der Ausdrucksinstruktion generierte die Trainerin zusammen mit den Schülern Ausgestaltungsbeispiele für die 7W-Fragen und griff dabei auf die Erfahrungen und das Wissen der Schüler zurück.

 WIE STELLST DU DIR DEINE PERSON VOR?			
AUSSEHEN			
EIGENSCHAFTEN	<table border="0"> <tr> <td>hilfsbereit, fleißig, lustig, schlau, neugierig</td> <td>Sport, malen, rechnen, schreiben, musizieren</td> </tr> </table>	hilfsbereit, fleißig, lustig, schlau, neugierig	Sport, malen, rechnen, schreiben, musizieren
hilfsbereit, fleißig, lustig, schlau, neugierig	Sport, malen, rechnen, schreiben, musizieren		
 WARUM IST ES DEINER HAUPTPERSON WICHTIG, DASS SIE IHR ZIEL ERREICHT?			
 WIE KANNST DU DIE ZEIT (JAHRESZEIT/ TAGESZEIT) BILDLICH BESCHREIBEN?			
FRÜHLING	Bäume und Blumen blühen, Vögel zwitschern, Ostern		
SOMMER	Sonnenschein, strahlend blauer Himmel, warme Tage, Sommerferien		
HERBST	Blätter fallen von den Bäumen, Regen, Wind, Nebel		
WINTER	Schneeflocken, Eiskristalle, Frostig, Weihnachten		
MORGEN	Sonne geht auf, Hahn kräht, es wird langsam hell, kühler und nebliger Morgen		
ABEND	Sonne geht unter, es wird dunkel		
 WIE STELLST DU DIR DEINEN ORT BILDLICH VOR?			
GEBÄUDE	<table border="0"> <tr> <td>Haus, Schule, Schloss</td> <td>alt/neu, hoch/niedrig</td> </tr> </table>	Haus, Schule, Schloss	alt/neu, hoch/niedrig
Haus, Schule, Schloss	alt/neu, hoch/niedrig		
ZIMMER	Bilder an der Wand, hell/dunkel, sauber/schmutzig, klein/groß		
LANDSCHAFT	<table border="0"> <tr> <td>Wiese, Berg, Wald, See</td> <td>viele bunte Blumen, grünes Gras, dunkler Wald</td> </tr> </table>	Wiese, Berg, Wald, See	viele bunte Blumen, grünes Gras, dunkler Wald
Wiese, Berg, Wald, See	viele bunte Blumen, grünes Gras, dunkler Wald		
 WIE KANNST DU SPANNEND BESCHREIBEN, WAS DEINE HAUPTPERSON MACHT?			
ADJEKTIVE	schnell – langsam, laut – leise, fröhlich – traurig		
WIE KANNST DU BESCHREIBEN, WAS DEINE HAUPTPERSON DENKT UND FÜHLT?			
ANGST	blass werden, Atem anhalten, starr vor Schreck, am ganzen Körper zittern, Herz pocht ganz wild		
FREUDE	lachen, vor Freude in die Luft springen, leuchtende Augen		
WUT	hochroter Kopf, auf den Boden stampfen, Tür zuschlagen		
ÜBERRASCHUNG	aufgerissene Augen, staunen, verwundert, erschrocken		
 WIE KANNST DU DEN HÖHEPUNKT ANGSCHAULICH BESCHREIBEN?			
GEDANKEN	Traurig dachte sie „...“; Erleichtert rief er „...“; Entsetzt bemerkte sie „...“; Leise flüsterte er „...“		
GEFÜHLE	Angst Freude Wut Überraschung		
 WIE KANNST DU BESCHREIBEN, WIE SICH DEINE HAUPTPERSON AM ENDE FÜHLT?			
Angst Freude Wut Überraschung			
SATZANFÄNGE			
Zuerst..., Auf einmal..., Anschließend...,	Darauf..., Danach..., Plötzlich...,	Später..., Schließlich..., Endlich...,	Zuletzt...

Abbildung 17: Ausdruckskarte.

Zusammenhang: Die Zusammenhangsinstruktion erfolgte relativ spät im USAT-Programm. Der Schwerpunkt lag auf der korrekten Reihenfolge der Handlungsschritte und der Verwendung unterschiedlicher Satzanfänge. Die Ausdruckskarte enthält Beispiele für verschiedene Satzanfänge. Die Reihenfolge der Handlungsschritte sollte den Schülern durch die Einführung eines Zusammenhangssymbols (siehe Abbildung 18) und den Hinweis, bereits beim Planen darauf zu achten, erleichtert werden. Das ursprüngliche Symbol der Handlungsschritte wurde dafür um Ketten, die die einzelnen Treppenstufen verbinden, und das Höhepunkt-Symbol oben auf der Treppe ergänzt.

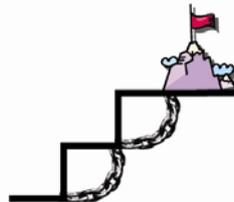


Abbildung 18: Zusammenhangssymbol.

5.3.2.2 Selbstregulatorische Methoden

Einer der zentralen Aspekte des SRSD-Ansatzes von Harris und Graham (1996) ist die explizite Kombination von selbstregulatorischen Techniken mit Schreibstrategien. Aus diesem Grund ist eine explizite Trennung zwischen den selbstregulatorischen Techniken einerseits und den Schreibstrategien andererseits nicht möglich. Zur besseren Verständlichkeit werden diese jedoch separat vorgestellt. Einen kurzen Überblick aller im Training vorkommenden selbstregulatorischen Strategien liefert Tabelle 10. Die eingesetzten Materialien dienten teilweise mehreren Funktionen und regten unterschiedliche selbstregulatorische Prozesse an, so dass es in der nachfolgenden Tabelle Überschneidungen gibt.

Tabelle 10: Selbstregulatorische Materialien und deren Funktionen.

Selbstregulatorische Funktion	Material
Zielsetzung	Leistungsprotokoll, 7W-Protokoll
Selbstüberwachung	Geschichtenplan, Ausdruckskarte
Selbstkontrolle	Checkliste, Leistungsprotokoll, 7W-Protokoll

Abgesehen von den Materialien spielte die Verbalisierung aller handlungsleitenden Gedanken während des Planens, Schreibens und Überarbeitens eine wichtige Rolle bei der Vermittlung selbstregulatorischer Techniken. Die Trainerin modellierte die Anwendung der Materialien und begründete dabei jeweils den Einsatz der Materialien. Dadurch sollten die Schüler die kognitiven Prozesse, die während der Selbstregulation ablaufen, verinnerlichen, so dass mit zunehmender Übung die Hilfestellungen durch Materialien nicht mehr notwendig waren und ausgeblendet werden konnten. Folglich stellten die Materialien nur eine temporäre Hilfe zum Erlernen der selbstregulatorischen Strategien dar.

Zielsetzung: In jeder Trainingseinheit planten, schrieben und überarbeiteten die Schüler eine neue Geschichte. Bevor die Schüler dies jedoch taten, setzten sie sich prozess- und produktbezogene Ziele für die neue Geschichte mit einem 7W- und einem Leistungsprotokoll. Die Zielsetzung ist nach Zimmerman (1998) eine wichtige selbstregulatorische Komponente.

Prozessbezogene Zielsetzung: Auf dem 7W-Protokoll (siehe Abbildung 19) sind die Symbole für die 7W-Fragen abgebildet. Es dient der prozessbezogenen Zielsetzung und gibt den Schülern Rückmeldungen darüber, ob die Schreibstrategien in der letzten Geschichte korrekt angewendet wurden. In der ersten Trainingseinheit erhielten die Schüler in der ersten

Zeile des Protokolls eine Rückmeldung zu der Bildergeschichte, die sie im Prätest geschrieben hatten. Basierend auf diesem Feedback, überlegten sich die Schüler Ziele für die neue Geschichte, die sie mit einer anderen Farbe ebenfalls in der ersten Zeile des 7W-Protokolls markierten. Zuvor wurde die Zielsetzung von der Trainerin am Overheadprojektor modelliert. Dabei wiederholte sie zudem den AHA-7W-Trick, was das Einprägen der Inhaltsstrategie förderte.

7-W-FRAGEN PROTOKOLL

							
7-W-Fragen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>				
7-W-Fragen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>				
7-W-Fragen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>				

Abbildung 19: 7W-Protokoll I.

Da im Verlauf des USAT-Programms weitere Strategien dazugelernt wurden und die Schüler sich folglich weitere strategiebezogene Ziele setzen konnten, veränderte sich das 7W-Protokoll zunehmend. Zunächst wurden im 7W-Protokoll die Ausdrucks- und schließlich die Zusammenhangsstrategien ergänzt, so dass das 7W-Protokoll nach Einführung aller Schreibstrategien wie in Abbildung 20 aussah.

Dieses 7W-Protokoll enthält zusätzlich zu dem ursprünglichen eine weitere Zeile sowie eine weitere Spalte pro Trainingseinheit. In der oberen Zeile wurden die 7W-Fragen markiert, die in der Geschichte genannt wurden (Inhalt) und in der unteren Zeile die 7W-Fragen, die darüber hinaus anschaulich beschrieben wurden (Ausdruck). Die zusätzliche Spalte enthält ein Kästchen, welches markiert wurde, wenn die Schüler die Reihenfolge der Verben sowie verschiedene Satzanfänge in ihren Geschichten verwendet hatten. Gleichzeitig konnten die Schüler mit dem neuen 7W-Protokoll nicht nur auf den Inhalt bezogene Ziele, sondern ebenfalls Ziele zur Verbesserung des Ausdrucks und des Zusammenhangs setzen.

7-W-FRAGEN PROTOKOLL

								
7-W-Fragen	<input type="checkbox"/>	<input type="checkbox"/>						
Adjektive Gedanken & Gefühle Verben	<input type="checkbox"/>	<input type="checkbox"/>						
7-W-Fragen	<input type="checkbox"/>	<input type="checkbox"/>						
Adjektive Gedanken & Gefühle Verben	<input type="checkbox"/>	<input type="checkbox"/>						
7-W-Fragen	<input type="checkbox"/>	<input type="checkbox"/>						
Adjektive Gedanken & Gefühle Verben	<input type="checkbox"/>	<input type="checkbox"/>						

Abbildung 20: 7W-Protokoll II.

Produktbezogene Zielsetzung: Abgesehen von prozessbezogenen Zielen und Rückmeldungen erhielten die Schüler ein auf die Gesamtqualität des Aufsatzes bezogenes Feedback mit Hilfe des Leistungsprotokolls (siehe Abbildung 21). Auf dem Leistungsprotokoll sind auf der x-Achse die Geschichten, die die Schüler im Verlauf des Trainings schrieben, aufgelistet. Mit Ausnahme von der Prätestgeschichte („vor dem Training“) enthält das Leistungsprotokoll eine Spalte für die Zielsetzung („Ziel“) und eine Spalte für die Rückmeldung („Punkte“). Auf der y-Achse sind die Kriterien für eine gute Geschichte abzulesen: Inhalt, Ausdruck, Zusammenhang. Insgesamt konnten die Schüler 21 Punkte erreichen, d.h. sieben Punkte pro Kriterium. In der ersten Trainingseinheit erhielten die Schüler eine Rückmeldung zur Prätestgeschichte. Anschließend modellierte die Trainerin die ergebnisbezogene Zielsetzung anhand des Leistungsprotokolls auf dem Overhead-Projektor. Schließlich überlegte sich jeder Schüler ein eigenes Ziel und markierte dies in dem Leistungsprotokoll in der Spalte „Ziel“.

LEISTUNGSPROTOKOLL

Name

Inhalt/Ausdruck/Zusammenhang	VOR		2		3		4		5		6		7		NACH	
	Ziel	Punkte	Punkte													
21																21
20																20
19																19
18																18
17																17
16																16
15																15
14																14
13																13
12																12
11																11
10																10
9																9
8																8
7																7
6																6
5																5
4																4
3																3
2																2
1																1

VOR DEM TRAINING
GESCHICHTE 2
GESCHICHTE 3
GESCHICHTE 4
GESCHICHTE 5
GESCHICHTE 6
7
NACH DEM TRAINING

Abbildung 21: Leistungsprotokoll.

Zusammenhang Schreibstrategien und Gesamtpunktzahl: Abgesehen von der prozess- und produktbezogenen Zielsetzung sowie Rückmeldung sollten die Schüler zudem den Zusammenhang zwischen der erfolgreichen Anwendung der Strategien und der höheren Punktzahl im Leistungsprotokoll erfahren. Ziel der Rückseite des Leistungsprotokolls war es, die Schüler zu motivieren, sich bei der Anwendung der Schreibstrategien anzustrengen. Aus diesem Grund wurde in der ersten Trainingseinheit zusammen mit dem Leistungsprotokoll die Rückseite des Leistungsprotokolls (siehe Abbildung 22) eingeführt. Auf der Rückseite konnten die Schüler die Gesamtpunktzahl aus dem Leistungsprotokoll neben der Anzahl korrekt angewandter Schreibstrategien aus dem 7W-Protokoll ablesen und folglich miteinander in Verbindung setzen. In der Spalte „Punkte“ wurden die Überschrift der Geschichte des Schülers sowie die Punktzahl des Leistungsprotokolls eingetragen. In der nächsten Spalte „7W-Fragen“ wurden die Kästchen der 7W-Fragen, die genannt wurden, ausgemalt bzw. ab der dritten Trainingseinheit umkreist, wenn die 7W-Fragen zusätzlich anschaulich beschrieben wurden. Die dritte Spalte „Zusammenhang“ wurde ab der fünften Trainingseinheit markiert, wenn der Zusammenhang in der Geschichte berücksichtigt wurde. Auf diese Weise erhielten die Schüler einen Überblick über ihre Verbesserungen und über ihren aktuellen Leistungsstand.

PUNKTE	7-W-FRAGEN							ZUSAMMENHANG
<input type="checkbox"/>								
<input type="checkbox"/>								
<input type="checkbox"/>								
<input type="checkbox"/>								
<input type="checkbox"/>								
<input type="checkbox"/>								
<input type="checkbox"/>								

Abbildung 22: Rückseite des Leistungsprotokolls.

Selbstüberwachung: Wie in der Tabelle 10 ersichtlich, erfolgte die Selbstüberwachung des Schreibstrategieeinsatzes anhand des Geschichtenplans und der Ausdruckskarte.

Geschichtenplan: Der Geschichtenplan (siehe Abbildung 23) entspricht einer grafischen Darstellung des AHA-7W-Tricks und wurde während des Planens und Schreibens eingesetzt. Auf der linken Seite des Geschichtenplans sind die 7W-Fragen aufgelistet. Als Erinnerungshilfe sind nicht nur die Symbole, sondern auch Hilfsfragen auf dem Geschichtenplan abgebildet. Daneben befinden sich freie Zeilen und ein Ankreuzkästchen. Während des Planens wurden Stichwörter zu den 7W-Fragen in die freien Zeilen geschrieben, welche beim Schreiben abgehakt wurden, sobald diese in der eigenen Geschichte integriert waren.

GESCHICHTENPLAN ✓

ANFANG

Person
Wer ist die Hauptperson?

Ziel
Was will die Hauptperson erreichen?

Zeit
Wann spielt die Geschichte?

Ort
Wo spielt die Geschichte?

HAUPTTEIL

Handlungsschritte
Was macht die Hauptperson Schritt für Schritt?

Höhepunkt
Was ist der Höhepunkt?

ABSCHLUSS

Ende
Wie endet die Geschichte?

Abbildung 23: Geschichtenplan.

Ausdruckskarte: Auch die bereits im vorangegangenen Abschnitt beschriebene Ausdruckskarte (siehe Abbildung 17) diente der Selbstüberwachung der Anwendung der Ausdrucksstrategien. Anhand der Ausdruckskarte konnten die Schüler Schritt für Schritt die einzelnen 7W-Fragen durchgehen und sich anschauliche und bildliche Beschreibungen für diese überlegen.

Selbstkontrolle: Die Selbstkontrolle, d.h. die eigenständige Überprüfung des korrekten Strategieeinsatzes, erfolgte vornehmlich mittels der Checkliste (siehe Abbildung 24). Gleichzeitig stellten das 7W- und das Leistungsprotokoll Hilfsmittel dar, die korrekte Umsetzung der Schreibstrategien in den eigenen Geschichten zu kontrollieren.

Checkliste: Die Checkliste (siehe Abbildung 24) besteht aus einer Auflistung der Schreibstrategien sowie Abhakkästchen für jede einzelne Strategie, die während des Überarbeitens von den Schülern markiert wurden, wenn die jeweilige Schreibstrategie in den eigenen Geschichten identifiziert wurde. Die Anwendung der Checkliste wurde von der Trainerin modelliert. Nach Einführung der Ausdrucks- und der Zusammenhangsstrategien wurde jeweils eine neue Checkliste eingeführt, die um die neuen Strategien ergänzt wurde. Die endgültige Checkliste ist ebenfalls in Abbildung 24 zu sehen. Für Inhalt, Ausdruck und Zusammenhang ist jeweils eine Spalte mit Abhakkästchen dargestellt.

CHECKLISTE

✓

ÜBERSCHRIFT

ANFANG

Person 

Ziel 

Zeit 

Ort 

HAUPTTEIL

Handlungsschritte 

Höhepunkt 

ABSCHLUSS

Ende 

CHECKLISTE

✓ ✓ ✓

ÜBERSCHRIFT

ANFANG

Person 

Ziel 

Zeit 

Ort 

HAUPTTEIL

Handlungsschritte  

Höhepunkt 

ABSCHLUSS

Ende 

Abbildung 24: Checklisten zur Überprüfung des Inhalts, Ausdrucks und Zusammenhangs einer Geschichte.

5.3.2.3 Unterrichtsmethoden

Die im USAT-Programm eingesetzten Unterrichtsmethoden basieren vornehmlich auf dem SRSD-Ansatz von Harris und Graham (1996), speziell auf der sechsstufigen Instruktion, welche bereits in der Fragestellung (siehe Kapitel 3.4.3) beschrieben wurde. Zudem waren kognitives Modellieren, Fading, angeleitetes Üben, selbstständiges Üben sowie Selbstinstruktionen zentrale Instruktionsprinzipien des USAT-Programms. Darüber hinaus wurden Gruppenarbeiten durchgeführt und verschiedene Aspekte zur Förderung der Motivation der Schüler eingesetzt.

Gruppenarbeit: Die Einteilung der Schüler in leistungsheterogene Gruppen erfolgte bereits vor Beginn des Trainings durch die Klassenlehrerin. Ab der zweiten Trainingseinheit fand in jeder Einheit eine Gruppenarbeit statt. Sie diente als Übung für die in der jeweils vorangegangenen Trainingseinheit erlernten Strategien. Jeder Gruppe wurde einer der folgenden Namen zugewiesen: Skater, Sprinter, Schwimmer, Reiter, Biker. Die Assoziation zum Sport sollte die Schüler motivieren, als Team zusammenzuarbeiten. Darüber hinaus wurde jede Woche einer aus der Gruppe zum Tisch-Chef ernannt. Dieser erhielt einen Stapel

mit Leitfragen (siehe CD-Anhang A4.1), welche ihm bei der Durchführung der Gruppenarbeit helfen sollte. Die Leitfragen entsprachen den von der Trainerin modellierten handlungsleitenden Gedanken und sollten zudem die Verinnerlichung der selbstregulatorischen Techniken fördern. Abgesehen von der Gruppenarbeit halfen die Tisch-Chefs zudem beim Austeilen von Blättern und sorgten dafür, dass ruhig und konzentriert gearbeitet wurde. Am Ende jeder Trainingsstunde wurde ein Gruppensieger benannt. Entsprechend der durchgeführten Übungen gab es Planungs-Champions, Überarbeitungs-Helden, Spannungs-Gewinner, Zusammenhangs-Sieger und Freundlichkeits-Meister, welche auf dem Gruppenposter (siehe Abbildung 25) eingetragen wurden.



Abbildung 25: Poster zur Würdigung der besten Teams.

Motivationsfördernde Komponenten: Brunstein und Spörer (2006) betonen, dass bei der Förderung der Selbstregulation kognitive, metakognitive sowie motivationale Komponenten berücksichtigt werden müssen. Deswegen wurde im Training ein besonderer Wert auf die Motivation gelegt. In dem vorherigen Abschnitt wurde bereits die Gruppenarbeit beschrieben. Die Ernennung des Tisch-Chefs sowie des Gruppenarbeitssiegers sollten nicht nur einen geregelten Ablauf der Arbeitsphasen garantieren, sondern zudem die Schüler motivieren, sich anzustrengen. Abgesehen davon wurden weitere motivierende Elemente in dem Training integriert. Zum Beispiel wurden sehr gute Geschichten zu verschiedenen Zeitpunkten im Training vorgelesen.

Erinnerungshilfen: Ein zentrales Prinzip der SRSD-Instruktion ist das Einprägen der neuen Strategien (Harris & Graham, 1996). Aus diesem Grund wurden in jeder Trainingseinheit sämtliche Schreib- und selbstregulatorischen Strategien wiederholt. Die Wiederholung umfasste die Verwendung von verschiedenen Materialien wie Poster an den Wänden, das mündliche Abfragen der Strategien und das Bearbeiten von Übungsblättern.

5.3.2.4 Ablauf des Trainings

Tabelle 11 vermittelt einen Überblick über den Ablauf der Pilotstudie.

Tabelle 11: Inhaltlicher Überblick der Pilotstudie.

Prätest	Erfassung der Schreibleistung sowie der Planungs- und Überarbeitungsaktivitäten	
	Strategien	Gruppenarbeit
TE 1	Inhalt (AHA-7W) Einführung des 7W- und des Leistungsprotokolls Planen und Schreiben einer Geschichte mit Hilfe des Geschichtenplans	
TE 2	Überarbeiten einer Geschichte mit Hilfe der Checkliste	Gruppe plant zusammen eine Bildergeschichte.
TE 3	Ausdruck	Gruppe überarbeitet einen Text mit der Checkliste.
TE 4	Wiederholung mit dem Schwerpunkt Überarbeiten und Ausdruck	Gruppe ergänzt bildliche Beschreibungen in einem Geschichtenplan.
TE 5	Zusammenhang	Gruppe findet in einer Geschichte die Satzanfänge/Verben und markiert diese.
TE 6	Wiederholung	Gruppe wiederholt alle Strategien anhand von Schülergeschichten.
Posttest	Erfassung der Schreibleistung sowie der Planungs- und Überarbeitungsaktivitäten	

In der ersten Trainingseinheit erfolgte die Einführung der Inhaltsstrategien und sämtlicher selbstregulatorischer Methoden. Im weiteren Verlauf wurden die Ausdrucks- und Zusammenhangsstrategien instruiert und die selbstregulatorischen Materialien um die neu erlernten Strategien ergänzt. In zwei weiteren Trainingseinheiten wurden schließlich Übungen zu sämtlichen Schreibstrategien durchgeführt.

Der Ablauf einer Trainingseinheit gestaltete sich immer sehr ähnlich und gliederte sich nach den Komponenten des zyklischen Trainingsmodells der Selbstregulation von Zimmerman (1998) (siehe Abbildung 10) wie folgt:

- *Rückmeldung:* Zu Beginn jeder Trainingseinheit erfolgte eine prozess- und produktbezogene Rückmeldung zu der in der letzten Trainingseinheit bzw. beim Prätest geschriebenen Geschichte anhand des 7W- und des Leistungsprotokolls. Die Rückmeldung fand jeweils im Stuhlkreis statt.
- *Motivierung:* Nach der Rückmeldung wurden die Schüler für den bisherigen Strategieeinsatz und für die verbesserten Geschichten gelobt.
- *Wiederholung:* Die Schüler wurden mündlich oder in Form von Übungen über die bereits erlernten Schreibstrategien abgefragt.
- *Strategieinstruktion:* Die Trainerin modellierte kognitiv den Einsatz neuer Schreibstrategien. Die Schüler wurden dabei mit einbezogen.
- *Zielsetzung:* Hinsichtlich aller erlernten Strategien setzten sich die Schüler prozess- und produktbezogene Ziele für die nächste Geschichte.
- *Geschichtenschreiben:* In jeder Trainingseinheit wurde eine neue Geschichte geschrieben, in der die Schüler die neuen bzw. bereits bekannten Strategien anwenden sollten. Sie hatten jeweils zehn Minuten Zeit eine Geschichte zu planen, zwanzig Minuten eine Geschichte zu schreiben und zehn Minuten die Geschichte zu überarbeiten. In den ersten Trainingseinheiten leitete die Trainerin die Schüler beim Verfassen der Geschichte an. Sie blendete jedoch im Verlauf des Trainings die Hilfestellungen immer weiter aus, so dass in der sechsten Trainingseinheit eine Geschichte ohne jegliche Hilfestellungen und Materialien geschrieben wurde.

Am Ende jeder Trainingseinheit wurden die Aufsätze der Schüler eingesammelt und von allen vier Trainerinnen gemeinsam anhand der 7W- und Leistungsprotokolle bewertet. Als nächstes werden die sechs Trainingseinheiten detailliert beschrieben. Jede Trainingseinheit dauerte 90 Minuten.

Trainingseinheit 1 (Schwerpunkt: Inhalt, Planen, Schreiben): In der ersten Trainingseinheit wurden den Schülern zunächst die groben Inhalte des Trainings erläutert, wobei an das Vorwissen der Schüler angeknüpft wurde. Den Schülern wurde erklärt, dass es um das Aufsatzschreiben geht und dass alle Schüler im Verlauf des Trainings Geschichtenexperten werden. Zu Anfang zeigte die Trainerin drei Schilder mit den Begriffen Inhalt, Ausdruck und Zusammenhang. Die Schüler wurden gebeten, sich zu überlegen, in welchem Zusammenhang diese Begriffe mit dem Geschichtenschreiben stehen könnten. Anschließend wurde den Schülern erläutert, dass ein Geschichtenexperte seine Geschichte plant, schreibt und überarbeitet (*3-Schritte-Technik*), dass er eine inhaltlich vollständige Geschichte schreibt (*Inhalt*), bildliche und anschauliche Beschreibungen verwendet

(*Ausdruck*) und dass die Geschichte des Geschichtenexperten zusammenhängt und einen Sinn ergibt (*Zusammenhang*). Daran anknüpfend zeigte die Trainerin das Leistungsprotokoll (siehe Abbildung 21) in Form eines Posters. Sie erklärte, dass der Inhalt, der Ausdruck und der Zusammenhang Kriterien für eine gute Geschichte sind und dass im Training die Geschichten anhand des Leistungsprotokolls bewertet werden. Anschließend teilte die Trainerin jedem Schüler eine Mappe mit einem Leistungsprotokoll aus, in dem die jeweilige Rückmeldung zur Prätestgeschichte markiert war.

Als nächstes erfolgte die Herleitung der AHA-7W-Strategie. Die Trainerin fragte zunächst, ob jemand die Teile einer guten Geschichte kennen würde. Zusammen mit den Schülern erarbeitete sie den ersten Teil der Merkhilfe: Anfang, Hauptteil, Abschluss. Anschließend las sie die Geschichte „Der schlaue Fuchs“ (siehe CD-Anhang A7.1), welche eine inhaltlich vollständige Geschichte repräsentiert, vor und leitete anhand dieser die 7W-Fragen her. Am Ende erklärte die Trainerin die Merkhilfe AHA-7W.

Darauf aufbauend zeigte die Trainerin das 7W-Protokoll (siehe Abbildung 19) und modellierte die prozess- und produktbezogene Zielsetzung mit dem 7W- und dem Leistungsprotokoll. Sie erläuterte überdies die Rückseite des Leistungsprotokolls (siehe Abbildung 22) und ermutigte die Schüler, sich Mühe zu geben, die Schreibstrategien einzusetzen, um mehr Punkte im Leistungsprotokoll zu erreichen.

Danach erklärte die Trainerin die 3-Schritte-Technik und erläuterte, wie selbstregulatorische Prinzipien beim Planen, Schreiben und Überarbeiten die Anwendung der AHA-7W-Strategie erleichtern. Anhand einer Bildergeschichte (siehe CD-Anhang A7.4) modellierte die Trainerin die Anwendung des Geschichtenplans (siehe Abbildung 23) zunächst beim Planen und anschließend beim Schreiben einer Geschichte.

Als erstes zeigte die Trainerin den Aufbau des Geschichtenplans und erklärte, dass das bloße Beschreiben der Bilder nicht für das Verfassen einer guten Bildergeschichte ausreicht. Sie erläuterte den Schülern, dass in der Regel beim Schreiben wichtige Informationen vergessen werden und es deswegen wichtig sei, sich beim Planen zu den einzelnen 7W-Fragen Stichwörter aufzuschreiben. Aus diesem Grund modellierte die Trainerin die Anwendung des Geschichtenplans und verbalisierte dabei alle handlungsleitenden Gedanken laut. Nachdem der Geschichtenplan ausgefüllt war, zeigte die Trainerin, wie sie mit Hilfe des Geschichtenplans Sätze formuliert und aufschreibt. Sie demonstrierte, wie sie die erfolgreich integrierten Stichwörter im Geschichtenplan abhakt und wie sie beim Schreiben eine Reihe frei lässt. Auch dieses Mal verbalisierte sie laut alle handlungsleitenden Gedanken, so dass die

Schüler ihr Vorgehen besser nachvollziehen konnten. Gleichzeitig wurden die Schüler mit einbezogen, so dass alle Schüler aufmerksam und motiviert mitarbeiteten.

Am Ende der ersten Trainingseinheit konnten die Schüler schließlich ihr neu gewonnenes Wissen einsetzen, indem sie eine eigene Bildergeschichte planten und schrieben. Sie gingen dabei genauso wie die Trainerin vor und verwendeten alle gezeigten Materialien. Während der Einzelarbeiten half die Trainerin schwachen Schülern und beantwortete Fragen. Teilweise remodellierte die Trainerin die Anwendung des Geschichtenplans oder erklärte erneut den AHA-7W-Trick.

Trainingseinheit 2 (Schwerpunkt: Überarbeiten): Am Anfang erhielten die Schüler die Mappen mit den neu ausgefüllten 7W- und Leistungsprotokollen zurück, so dass sie kontrollieren konnten, ob sie ihr selbst gesetztes Ziel der letzten Trainingseinheit erreicht hatten. Anschließend wurde der AHA-7W-Trick mündlich wiederholt.

Als nächstes führte die Trainerin die Gruppenarbeit wie bereits im vorangegangenen Abschnitt erläutert ein. Es wurde ein Tisch-Chef ernannt, welcher ausgestattet mit Leitfragen die Gruppenarbeit leitete. Die Gruppenaufgabe bestand darin, gemeinsam eine Bildergeschichte zu einem vorgegebenen Bild (siehe CD-Anhang A4.1) mit Hilfe des Geschichtenplans zu planen. Die ausgefüllten Gruppen-Geschichtenpläne wurden zum Schluss eingesammelt. Am Ende der Trainingseinheit wurde die beste Gruppe als „Planungs-Champion“ gekürt.

Danach setzten sich die Schüler für die nächste Geschichte Ziele im 7W- und Leistungsprotokoll. Anschließend planten und schrieben die Schüler unter Verwendung des Geschichtenplans eine Reizwortgeschichte (siehe CD-Anhang A3).

Schließlich erfolgte die Instruktion der Überarbeitungsstrategien mit der Checkliste (siehe Abbildung 24). Dabei ging die Trainerin ähnlich wie bei der Einführung des Geschichtenplans vor. Es wurde das selbstregulatorische Prinzip der Selbstbewertung modelliert und den Schülern erläutert. Anhand der Geschichte „Tim beim Einkaufen“ (siehe CD-Anhang A7.5) zeigte die Trainerin, wie sie selbst eine Geschichte überarbeitet. Zuerst las die Trainerin die Geschichte, die für alle Schüler sichtbar an der Tafel stand, laut vor. Anschließend modellierte sie die Anwendung der Checkliste, indem sie zunächst jeden Satz einzeln vorlas. Danach überprüfte sie, ob eine der 7W-Fragen in dem vorgelesenen Satz genannt wurde. Sie bezog die Schüler mit ein. Wenn die Schüler z.B. das Ziel in dem Satz entdeckten, hakte die Trainerin das 7W-Element „Ziel“ auf der Checkliste, welche auf einem Overhead-Projektor für alle Schüler sichtbar war, mit einem Folienstift ab. Nachdem der Anfang der Geschichte von der Trainerin und der Klasse hinsichtlich der 7W-Fragen

analysiert worden war, überprüfte die Trainerin, ob eine 7W-Frage des Anfangs in der Geschichte fehlte. In der Beispiel-Geschichte fehlten der Ort und die Zeit. Zusammen mit den Schülern generierte die Trainerin passende Orts- und Zeitangaben, die sie anschließend in die Geschichte an der Tafel mit einem farbigen Stück Kreide einfügte. Analog ging die Trainerin bei der Überarbeitung des Hauptteils und des Abschlusses vor, so dass schließlich die gesamte Beispiel-Geschichte revidiert wurde.

Am Ende der Trainingseinheit setzten die Schüler die neue Strategie selbstständig ein, indem sie ihre eigenen Geschichten mit der Checkliste überarbeiteten. Sie wurden dabei von der Trainerin unterstützt.

Trainingseinheit 3 (Schwerpunkt: Ausdruck): Gemäß des typischen Trainingsablaufs erfolgte zu Beginn der Trainingseinheit die Rückmeldung zur letzten Trainingsgeschichte im Stuhlkreis und die Wiederholung der bisher erlernten Schreibstrategien. Anschließend startete die Gruppenarbeit. Es wurde ein Tisch-Chef ernannt, der wieder Leitfragen erhielt. Die Trainerin erklärte die Aufgabe, die darin bestand, dass jede Gruppe eine vorgegebene Geschichte anhand der Checkliste überprüfte und anschließend überarbeitete (siehe CD-Anhang A4.1). Am Ende der Trainingseinheit wurde die beste Gruppe als „Überarbeitungs-Held“ gekürt.

Im Anschluss daran erläuterte die Trainerin die Bildergeschichte, zu der in der aktuellen Trainingseinheit eine Geschichte geschrieben werden sollte (siehe CD-Anhang A3). Danach füllten die Schüler den Geschichtenplan aus.

Als nächstes erfolgte die Herleitung der Ausdrucksstrategien. Analog zur Inhaltsinstruktion wurde zunächst das Hintergrundwissen der Schüler aktiviert. Die Trainerin begann mit dem ersten 7W-Element (Hauptperson) und fragte die Schüler nach Vorschlägen für die Ausgestaltung der Person. Anschließend wurde in der Klasse über die Vorschläge diskutiert. Schließlich zeigte die Trainerin, wie sie selbst die Person in einer Geschichte ausgestaltet, indem sie Beispiele aus der Ausdruckskarte (siehe Abbildung 17) ablas. Die Ausdruckskarte war dabei mit Hilfe eines Overhead-Projektors für alle Schüler sichtbar. Genauso ging die Trainerin mit den übrigen 7W-Elementen vor. Nachdem schließlich die Ausdrucksstrategien für alle 7W-Fragen hergeleitet und auf der Ausdruckskarte gezeigt wurden, teilte die Trainerin für jeden Schüler die Ausdruckskarte aus und forderte die Schüler auf, diese beim Planen, Schreiben und Überarbeiten einzusetzen.

Direkt im Anschluss setzten sich die Schüler neue Ziele für die nächste Geschichte in dem 7W- und dem Leistungsprotokoll. Die Trainerin teilte dabei ein neues 7W-Protokoll mit den Ausdrucksstrategien aus. Dadurch konnten die Schüler einerseits in den neuen

Protokollen ablesen, ob sie bereits den Ausdruck in ihren Geschichten berücksichtigt hatten und andererseits speziell auf den Ausdruck bezogene Ziele setzen, wozu die Trainerin sie maßgeblich ermutigte.

Daran anknüpfend konnten die Schüler das neu gewonnene Wissen einsetzen, indem sie den vorher bereits ausgefüllten Geschichtenplan mit der Ausdruckskarte überarbeiteten. Sie sollten anschauliche und bildliche Beschreibungen ergänzen. Mit Hilfe des Geschichtenplans schrieben sie anschließend die Geschichte. Nach zwanzig Minuten unterbrach die Trainerin das Schreiben und führte eine neue Checkliste, die um die Ausdrucksstrategien ergänzt wurde, ein. Danach überprüften die Schüler als Erstes die eigene Geschichte mit der neuen Checkliste, bevor sie abschließend diese mit einem farbigen Stift überarbeiteten. Während des gesamten Schreibprozesses erinnerte die Trainerin die Schüler an die Anwendung der Ausdruckskarte und unterstützte sie dabei.

Trainingseinheit 4 (Schwerpunkt: Wiederholung): In der vierten Trainingseinheit wurden keine neuen Strategien eingeführt. Dafür wurden die bereits erlernten Strategien, insbesondere die Ausdrucks- und Überarbeitungsstrategien, intensiv wiederholt. Wie jedoch in jeder Trainingseinheit erhielten die Schüler zunächst am Anfang die Rückmeldungen zur letzten Geschichte mit dem 7W- und dem Leistungsprotokoll. Da in der dritten Trainingseinheit der Ausdruck eingeführt wurde, bekamen die Schüler dieses Mal ebenfalls ein Feedback über den korrekten Einsatz der Ausdrucksstrategien.

Als nächstes las die Trainerin eine Geschichte vor (siehe CD-Anhang A4.4). Die Schüler sollten ein Gesamturteil zu der Geschichte abgeben, indem sie zwischen einem stark weinenden, einem weinenden, einem lachenden und einem stark lachenden Smiley wählen konnten. In der gesamten Klasse wurde diskutiert, was in der Geschichte nicht so gut gelungen war. Schließlich überprüften und verbesserten die Schüler die Geschichte mit Hilfe der Checkliste.

Die Wiederholung sämtlicher bereits erlernter Schreibstrategien erfolgte im nächsten Schritt anhand einer Schülergeschichte. Dies sollte zudem die Schüler motivieren, sich beim Geschichtenschreiben Mühe zu geben, da die Schüler darauf hingewiesen wurden, dass in der letzten Trainingseinheit abermals Schülergeschichten vorgelesen wurden.

Anschließend wurde der neue Tisch-Chef verkündet und mit Leitfragen ausgestattet. Die Gruppenarbeit bestand in der vierten Trainingseinheit darin, einen teilweise ausgefüllten Geschichtenplan hinsichtlich der Ausgestaltung zu ergänzen (siehe CD-Anhang A4.1). Der Gruppensieger war der „Spannungs-Gewinner“.

Im Anschluss an die Gruppenarbeit setzten sich die Schüler prozess- und produktbezogene Ziele mit dem 7W- und Leistungsprotokoll. Danach planten die Schüler zunächst eine Reizwortgeschichte (siehe CD-Anhang A3) mit dem Geschichtenplan, schrieben sie danach und überarbeiteten diese abschließend mit der Checkliste.

Trainingseinheit 5 (Schwerpunkt: Zusammenhang): Gemäß dem prototypischen Ablauf einer Trainingseinheit wurden zunächst die Bewertungen der letzten Trainingsgeschichte anhand des 7W- und Leistungsprotokolls den Schülern zurückgemeldet und mündlich sämtliche bereits erlernten Strategien wiederholt.

Anschließend begann die Herleitung der Zusammenhangsstrategien mit einer Partnerarbeit, bei der zwei Schüler gemeinsam Schnipsel mit Sätzen zu einer Geschichte „Tim und Tobi“ sortieren sollten (siehe CD-Anhang A7.3). Bei dieser Übung sollte bereits das Hintergrundwissen der Schüler aktiviert werden, da die Schüler durch das Identifizieren der 7W-Elemente in den Schnipseln den Anfang, Hauptteil und das Ende ermitteln konnten. Im Anschluss an die Partnerarbeit diskutierte die Trainerin mit den Schülern das optimale Vorgehen beim Lösen der Aufgabe.

Als nächstes zeigte die Trainerin anhand der von den Schülern sortierten Geschichte, wie sie selbst beim Überprüfen des Zusammenhangs vorgeht. Es wurde dabei der Schwerpunkt auf die Kontrolle der Reihenfolge der Verben und die Verwendung unterschiedlicher Satzanfänge gelegt. Die Trainerin führte ein Zusammenhangssymbol (siehe Abbildung 18) ein, welches die Schüler daran erinnern sollte, bei den Handlungsschritten die Reihenfolge der Verben zu berücksichtigen. Mit der gesamten Klasse identifizierte die Trainerin die Handlungsschritte und die Satzanfänge in der „Tim und Tobi“-Geschichte. Sie betonte, wie wichtig die richtige Reihenfolge der Handlungsschritte und die Verwendung von abwechslungsreichen Verben für das Verständnis einer Geschichte sind. Darüber hinaus wies die Trainerin die Schüler darauf hin, die auf der Ausdruckskarte (siehe Abbildung 17) aufgelisteten Satzanfänge in ihren Geschichten einzubauen.

Die darauf aufbauende Gruppenarbeit sollte die Zusammenhangsstrategien weiter vertiefen (siehe CD-Anhang A4.1). Sie ähnelte der Partnerarbeit, die zu Beginn der Trainingseinheit durchgeführt wurde. Nachdem ein neuer Tisch-Chef ernannt wurde und Leitfragen erhalten hatte, sortierte die Gruppe eine aus verschiedenen Abschnitten bestehende Geschichte, so dass eine vollständige und ausgestaltete Geschichte mit einer richtigen Reihenfolge entstand. Anschließend beantworteten die Schüler in der Gruppe verschiedene Verständnis- und Transferfragen zu der Geschichte. Am Ende der Stunde wurde der „Zusammenhangs-Sieger“ gekürt.

Zuvor wurde jedoch wieder von jedem Schüler eine neue Bildergeschichte verfasst (siehe CD-Anhang A3). Dazu setzte sich jeder Schüler zunächst neue Ziele in dem 7W- und dem Leistungsprotokoll. Die Trainerin verteilte ein neues 7W-Protokoll (siehe Abbildung 20), auf dem die Schüler ablesen konnten, ob sie bereits den Zusammenhang in ihrer Geschichte berücksichtigt hatten. Falls nicht, konnten sie das Zusammenhangs-Kästchen markieren und sich für die nächste Geschichte vornehmen, speziell darauf zu achten.

Während des Planens, Schreibens und Überarbeitens unterstützte die Trainerin die Schüler speziell bei der Anwendung der Zusammenhangsstrategien. Nach der Schreibphase führte sie eine neue Checkliste (siehe Abbildung 24) ein, mit der die Schüler nicht nur die Anwendung der Inhalts- und Ausdrucksstrategien, sondern zudem der Zusammenhangsstrategien kontrollieren konnten.

Trainingseinheit 6 (Schwerpunkt: Wiederholung): Nachdem die Schüler die Rückmeldungen zur letzten Trainingsgeschichte erhalten hatten, erfolgte eine ausführliche Wiederholungsphase anhand von drei Schülergeschichten. Damit die Schüler adäquates Feedback zu den Geschichten geben und gut zuhören konnten, wurden zuvor Regeln eingeführt und diskutiert. Während die Schüler beim Feedback darauf achten sollten, in der Ich-Form zu sprechen und mit etwas Positivem zu beginnen, wurden die Schüler beim Zuhören darauf hingewiesen, auf die 7W-Fragen sowie sämtliche im Verlauf des Trainings erlernten Strategien zu achten (siehe CD-Anhang A4.4).

Jede der drei von der Trainerin vorgelesenen Geschichte betonte eins der für eine gute Geschichte relevanten Kriterien des Inhalts, Ausdrucks und des Zusammenhangs. Die Schüler identifizierten die einzelnen Strategien in den Schülergeschichten und wiederholten mündlich alle Schreibstrategien.

Anschließend wurde der Tisch-Chef der sechsten Trainingseinheit ernannt. Die Gruppenarbeit bestand darin, in Abschnitten der von den Schülern verfassten Geschichten die 7W-Fragen bzw. die Ausgestaltungselemente der 7W-Fragen zu identifizieren. Der Tisch-Chef las dabei die jeweiligen Geschichtenabschnitte vor und sorgte dafür, dass die Regeln beim Zuhören und für das Feedback eingehalten wurden. Die beste Gruppe wurde zum „Freundlichkeits-Meister“ ernannt.

Danach erfolgte eine weitere Wiederholungsübung. Zu zweit bearbeiteten die Schüler Tandem-Karten, auf denen auf der einen Seite eine Frage und auf der anderen Seite die Antwort stand (siehe CD-Anhang A4.1). Im Wechsel lasen die Schüler die Fragen zu den erlernten Strategien vor und beantworteten diese.

Schließlich setzten sich die Schüler Ziele für die nächste Geschichte in dem 7W- und dem Leistungsprotokoll. Hinterher planten, schrieben und überarbeiteten die Schüler eine neue Bildergeschichte. Dabei wurden sämtliche Hilfsmittel ausgeblendet, d.h. die Schüler verwendeten keinen Geschichtenplan, keine Ausdruckskarte und keine Checkliste. Die Schüler sollten eigenständig die Schreibstrategien einsetzen. Auch die Trainerin half nur in Ausnahmefällen. Jegliche Erinnerungsstützen, einschließlich der Poster und der verbalen Hilfestellungen, wurden entfernt, so dass die Schüler unter denselben Bedingungen wie in den Testungen eine Geschichte planten, schrieben und überarbeiteten.

Zum Abschluss des Trainings erhielt jeder Schüler einen persönlichen Brief, in dem die individuellen Stärken und Schwächen in Bezug zum Geschichtschreiben berichtet wurden (siehe CD-Anhang 4.4). Dieser sollte die Schüler motivieren, sich bei der Posttestung besonders anzustrengen.

5.3.3 Genauigkeit der Implementation

In der Pilotstudie wurden verschiedene Maßnahmen zur Sicherstellung einer hohen Implementationsgenauigkeit eingesetzt.

Schulung der Trainerinnen: Ein wichtiger Faktor für den Erfolg einer Implementation ist die intensive Schulung der Trainerinnen. Wie im ersten Abschnitt des Kapitels beschrieben, waren die beteiligten Trainerinnen bereits sehr früh in der Konzeptions- und Planungsphase involviert. Des Weiteren erhielt jede Trainerin detaillierte Manuale für jede Trainingseinheit. Die Manuale waren ausformuliert und konnten von den Trainerinnen komplett auswendig gelernt werden (siehe CD-Anhang A4.2). Ferner erhielt jede Trainerin Checklisten, mit deren Hilfe sie nach der Durchführung der Trainingseinheit die inhaltliche Vollständigkeit überprüfen konnte (siehe CD-Anhang A5.3.2). Diese Checklisten stellten zudem eine kurze Zusammenfassung aller wichtigen Punkte der jeweiligen Trainingseinheiten dar.

Unmittelbar vor der Durchführung der jeweiligen Trainingseinheiten wurden die einzelnen Trainingseinheiten von der für die jeweilige Stunde zugeteilten Trainerin modelliert. Alle Trainerinnen waren bei diesen Treffen anwesend. Es wurde explizit auf Schwierigkeiten und Probleme hingewiesen, so dass die Trainerin für schwierige Instruktionen sensibilisiert wurde. Unklarheiten wurden überdies besprochen. Auch mögliche Störungen und Schwierigkeiten wurden diskutiert sowie Lösungsvorschläge gesammelt.

Unterrichtsbeobachtung: Als weitere Maßnahme zur Gewährleistung einer hohen Implementationsgenauigkeit beobachteten die übrigen Trainerinnen die Durchführung des Trainings im Unterricht mit standardisierten Beobachtungsbögen (siehe CD-Anhang A5.3.1).

Die Beobachtungsbögen enthielten die Kriterien Klarheit der Instruktion, Individualisierung, Motivierung und Klassenführung. Da alle Trainerinnen bei der Durchführung der Pilotstudie dabei waren, konnte unmittelbar eingegriffen werden, sobald die jeweilige Trainerin von dem Originalprogramm abwich oder etwas vergessen bzw. einen Fehler gemacht hatte. Aus diesem Grund wurden in der Pilotstudie 100% der geplanten Trainingseinheiten umgesetzt.

Trainertreffen: Darüber hinaus fand im Anschluss an jede Trainingseinheit ein Treffen mit allen Trainerinnen statt. Zum einen wurde gemeinsam über die Einheit reflektiert, wobei die Beobachtungsbögen als Orientierungsrahmen dienten. Zum anderen wurden die Aufsätze der Schüler korrigiert und die Leistungen der Schüler in die 7W- und Leistungsprotokolle eingetragen. Auf diese Weise konnte zudem garantiert werden, dass die Trainerinnen zur Bewertung der Schüleraufsätze ähnliche Standards heranzogen und dass sie das Ausfüllen der Protokolle übten.

5.3.4 Testungsmaterial

Zu zwei Messzeitpunkten, vor sowie nach dem Training, wurden die Schreibleistungen der Schüler im gesamten Klassenverband erhoben. Die Pilotstudie intendierte vor allem die Überprüfung der Wirksamkeit des USAT-Programms. Aus diesem Grund wurden neben den Schreibleistungen keine weiteren Maße erfasst.

Die Bildergeschichten wurden hinsichtlich strategienaher (Geschichtenelemente, Planungsaktivitäten, Überarbeitungsaktivitäten) und strategieferner Maße (Narrative Qualität/ Geschichtenqualität) ausgewertet. Die Texte, Stichpunkte und Revisionen wurden von zwei trainierten Personen gemeinsam ausgewertet. Darüber hinaus wurden die Wörter unabhängig von der korrekten Schreibweise und grammatikalischen Regeln gezählt (Wortanzahl).

Schreibanlass: Als Schreibanlass dienten drei verschiedene Bildergeschichten der Serie *Vater und Sohn* von Plauen (2003), welche bereits von Glaser und Kollegen eingesetzt wurden (Glaser & Brunstein, 2007a; Glaser, Keßler, & Brunstein, 2009). Die Bildergeschichten bestehen aus vier zusammenhängenden Bildmotiven. Die drei Bilderserien wurden von neun unabhängigen Beurteilern aus insgesamt zehn verschiedenen Bilderserien als ähnlich im Hinblick auf Schwierigkeit und Interesse eingestuft. Die Schüler erhielten zu den beiden Messzeitpunkten zufällig eine der drei Bildergeschichten, so dass jede Bildergeschichte von einer gleichen Anzahl von Schülern pro Messzeitpunkt bearbeitet wurde. Da in Studie I und II, welche jeweils drei Messzeitpunkte umfassten, dieselben Bildergeschichten verwendet werden sollten, wurden in der Pilotstudie ebenfalls drei verschiedene Bilderserien eingesetzt (siehe CD-Anhang A2). Zu jedem Messzeitpunkt erhielten die Schüler eine andere Bildergeschichte. Während die Testungsgeschichten den

Schülern nicht bekannt waren, stellten sie dennoch einen trainingsnahen Schreibanlass dar, da vier der sechs im Training geschriebenen Geschichten ebenfalls Bildergeschichten bestehend aus ein bis sechs Bildern waren.

Ablauf der Testungen: Die Prä- sowie Posttests wurden von denselben Trainerinnen, die das Training umsetzten, geleitet und dauerten jeweils 45 Minuten. Zuvor wurden die Trainerinnen explizit in der Durchführung der Testungen geschult. Zunächst wurden die Schüler gebeten, zu einer der drei Bildergeschichten eine Geschichte zu planen und sich Stichpunkte zu notieren (zehn Minuten), anhand der Stichpunkte eine Geschichte zu schreiben (zwanzig Minuten) und schließlich die Geschichte zu überarbeiten (zehn Minuten). Wie auch die Trainingsmanuale wurden die Instruktionen für die Trainerinnen ausformuliert (siehe CD-Anhang A2). Für das Planen und Schreiben der Bildergeschichte erhielten die Schüler Extra-Blätter. Die im Training eingesetzten Hilfsmaterialien entfielen in den Testungen.

Vor der Posttestung wurden die Schüler darüber informiert, dass die Klassenlehrerin eine Rückmeldung zu den Posttest-Leistungen der Schüler erhielt. Sowohl beim Prä- als auch beim Posttest fehlte ein Kind. Das bei der Prätestung fehlende Kind wurde von der Lehrerin nachgetestet. Das beim Posttest fehlende Kind entfiel bei der Auswertung, so dass sich eine Stichprobengröße von 21 Schülern ergab.

Geschichtenelemente: Zur Bewertung der inhaltlichen Vollständigkeit sowie der Ausgestaltung wurde die *Story Grammar Scale* von Harris und Graham (1996) herangezogen. Die adaptierte Version wurde bereits in verschiedenen Studien erprobt und eingesetzt (Glaser & Brunstein, 2007a; Glaser, 2005; Glaser & Brunstein, 2007b; Glaser et al., 2009). Die Schülergeschichten wurden in Bezug zu den 7W-Fragen (Person, Ziel, Zeit, Ort, Handlungsschritte, Höhepunkt, Ende) ausgewertet. Es wurde jeweils ein Punkt für das Nennen einer der 7W-Fragen vergeben. Ein weiterer Punkt konnte erreicht werden, wenn die 7W-Fragen noch zusätzlich anschaulich und sprachlich ausgestaltet wurden. Insgesamt ergab sich eine 14-stufige Skala.

Narrative Qualität/Geschichtenqualität: Um strategiefere Trainingseffekte zu erfassen, wurde die Gesamtqualität der Schüleraufsätze auf einer sechsstufigen Skala holistisch eingeschätzt. Die eingesetzte Skala basiert auf dem von Gentile (1992) entwickelten *National Scoring Guide*, welcher in den USA im Rahmen nationaler Schreibleistungserhebungen (NAEP) regelmäßig zur Beurteilung von Schülertexten verwendet wird. Für jede Stufe der Skala wurden spezifische Kriterien hinsichtlich des Inhalts, des Ausdrucks und des Zusammenhangs festgelegt, die für die Gesamtqualität einer Geschichte für wichtig erachtet wurden. Somit orientieren sich die Bewertungskriterien nicht

an den 7W-Fragen. Eine Geschichte wurde mit einem Punkt bewertet, wenn diese nur aus einer Aneinanderreihung mehrerer zusammenhangsloser Sätze bestand. Die Höchstpunktzahl mit sechs Punkten entsprach einer Geschichte, die überwiegend inhaltlich vollständig und sprachlich ausgestaltet war und dessen Gesamtstruktur aufgrund von hervorragenden Übergängen und Sinnzusammenhängen eine Einheit bildete.

Planen: In Anlehnung an die Auswertung der Geschichtenelemente wurden die Stichpunkte ebenfalls hinsichtlich der inhaltlichen Vollständigkeit und der Ausgestaltung der 7W-Fragen ausgewertet. Es wurde jeweils ein Punkt für das Nennen einer 7W-Frage und jeweils ein weiterer Punkt für die anschauliche und bildliche Beschreibung einer 7W-Frage vergeben, was ebenfalls in einer 14-stufigen Skala resultierte.

Revisionen: Jede die Geschichte verbessernde Einfügung wurde gewertet. Die Revisionen wurden kategorisiert in inhaltliche (z.B. eine 7W-Frage) und sprachliche (z.B. Adjektive, Verben, Gedanken und Gefühle) Einfügungen.

5.4 Ergebnisse

Die Daten wurden auf Schülerebene analysiert. Es wurden T-Tests für abhängige Stichproben durchgeführt. Die Ergebnisse sind in Tabelle 12 dargestellt.

Tabelle 12: Mittelwerte, Standardabweichungen und T-Werte für alle abhängigen Variablen getrennt nach Messzeitpunkt (Pilotstudie).

	Prätest		Posttest	
	M	SD	M	SD
Geschichtenelemente	3.74	1.59	7.84	2.39
T-Wert	7.60***			
Geschichtenqualität	2.10	0.81	3.48	0.89
T-Wert	6.87***			
Planen	0.62	0.74	6.05	1.50
T-Wert	13.78***			
Revisionen	0.38	0.67	0.52	0.75
T-Wert	0.72			
Wortanzahl	88.76	37.01	135.43	66.25
T-Wert	3.37**			

Anmerkungen: *** $p < .001$, signifikanter Unterschied zwischen Prä- und Posttest;

** $p < .01$, signifikanter Unterschied zwischen Prä- und Posttest.

Die Schreibleistungen in Bezug auf Geschichtenelemente, Wortanzahl und Geschichtenqualität sowie die Planungsaktivitäten verbesserten sich signifikant vom Prätest zum Posttest. Bezüglich des Überarbeitens ergaben sich keine signifikanten Unterschiede zum Ausgangsniveau.

Insgesamt schrieben die Schüler am Ende des Trainings Geschichten, die inhaltlich vollständiger, sprachlich besser ausgestaltet, vom Umfang her länger und von der Qualität hochwertiger als die Geschichten vor dem Training waren.

5.5 Diskussion

Die Ergebnisse der Pilotstudie zeigen, dass das USAT-Programm erfolgreich von Lerntrainerinnen in den regulären Deutschunterricht implementiert werden konnte. Da dem USAT-Programm keine Kontrollgruppe gegenübergestellt wurde, sind die Ergebnisse jedoch nur eingeschränkt gültig. Natürliche Reifungsprozesse können genauso wenig ausgeschlossen werden wie stichproben- und trainingspezifische Effekte. Zum einen kann die untersuchte Klasse besonders empfänglich für das USAT-Programm gewesen sein und zum anderen kann die alleinige Durchführung eines Trainings und die Verwendung von anregenden Materialien zu Verbesserungen der Leistungen geführt haben. Dennoch lassen sich folgende Aussagen bezüglich der Ergebnisse der Pilotstudie treffen.

Die Hypothese 1a konnte bestätigt werden, da die Schüler am Ende des Trainings ihre Schreibleistungen (Geschichtenelemente, Geschichtenqualität, Wortanzahl) signifikant steigern konnten. Dagegen konnte die Hypothese 1b nur bezüglich der Planungsaktivitäten erfüllt werden. Die Revisionen der Schüler verbesserten sich tendenziell, jedoch nicht signifikant von der Prätestung zur Posttestung.

Da das USAT-Programm schwerpunktmäßig Planungsstrategien und keine expliziten Revisionsstrategien vermittelt, ist dieses Ergebnis nicht überraschend. Glaser und Brunstein (2007a) betonen, dass Revisionsfertigkeiten explizit mit Methoden der Steuerung und Überwachung kombiniert werden müssen. Nichtsdestotrotz sollte in den folgenden beiden Hauptstudien das Überarbeiten stärker fokussiert werden. Dies kann zum einen dadurch erfolgen, dass die Schüler stärker als bislang motiviert werden, die Checkliste zu verwenden. Bei den Viertklässlern sank die Konzentration am Ende einer Doppelstunde erheblich ab, so dass die Motivation, in den letzten zehn Minuten die eigene Geschichte zu überarbeiten, gering war. Der Anreiz, die Geschichte zu überarbeiten, sollte deswegen z.B. durch verbale Ermutigungen oder durch Hinweise auf mehr Punkte im Leistungsprotokoll erhöht werden.

In der Pilotstudie waren alle Lerntrainerinnen während des gesamten Trainings anwesend und konnten somit bei Schwierigkeiten, z.B. wenn ein Schüler störte oder wenn ein Schüler eine Hilfestellung benötigte, eingreifen. In der Pilotstudie wurde deutlich, dass die Unterstützung durch weitere Personen notwendig ist. Da durch die Anwesenheit einer weiteren Person Störungen sowie Verzögerungen effizient behoben werden konnten, ohne dass die Durchführung des Trainings unterbrochen werden musste, sollte in der Studie I ein Tutor der Trainerin zur Seite gestellt werden. Des Weiteren waren die zeitlichen Vorgaben sehr eng bemessen, so dass ein Tutor dazu beitragen könnte, den Schülern alle relevanten Trainingsinhalte in der angesetzten Zeit zu vermitteln.

Bezüglich der Durchführung der Instruktionen funktionierte die Vermittlung der AHA-7W-Strategie (Inhalt), die Modellierung der selbstregulatorischen Methoden (Selbstüberwachung, Zielsetzung, Selbstkontrolle) und des Zusammenhangs sehr gut. Die Strategien zur Verbesserung des Ausdrucks einer Geschichte wurden ebenfalls von den Schülern verstanden und größtenteils eingesetzt. Die Ausdrucksinstruktion dauerte jedoch sehr lange und erforderte von den Schülern sehr viel Aufmerksamkeit. Da die Schüler der Pilotstudie sich größtenteils sehr gut konzentrieren konnten, könnte die lange Zuhörphase bei schwächeren Schülern und allgemein unruhigen Klassen ein Lernhindernis darstellen. Deswegen sollte die Ausdrucksinstruktion für die Studie I geändert werden. Anstatt die gesamte Instruktion in einer Trainingseinheit zu vermitteln, sollte diese in zwei Trainingseinheiten vollzogen werden. Weiter sollte die Instruktion für die Schüler ansprechender gestaltet werden. Dies kann durch die Herleitung der Ausdrucksstrategien anhand einer konkreten Geschichte ermöglicht werden. Darüber hinaus dauerte die Ausdrucksinstruktion sehr lange, da die Schüler sowohl vor als auch nach der Einführung der Strategien jeweils den Geschichtenplan ausfüllten. In der Studie I sollte das vorherige Ausfüllen des Geschichtenplans entfallen, um Zeit zu sparen. Die Erfahrungen der Pilotstudie zeigen, dass es genügt, wenn die Schüler nach der Ausdrucksinstruktion eine Geschichte planen, schreiben und überarbeiten.

Schließlich zeigten sich trotz der Verbesserungen im Anschluss an die Vorarbeiten Schwierigkeiten bei der Durchführung der Gruppenarbeiten. Während die Gruppenarbeiten den Schülern sowie der Klassenlehrerin subjektiv sehr gut gefielen, dauerten diese erheblich länger als vorgesehen, so dass sich die übrigen Trainingsinhalte (Strategieinstruktion, Planen, Schreiben und Überarbeiten einer Geschichte) verzögerten. Für die Studie I sollte aus diesem Grund die Gruppenarbeit zur Partnerarbeit umgestaltet werden. Da der Fokus der Gruppenarbeiten auf dem Üben der Strategien lag, wird erwartet, dass der Lerneffekt bei den

Partnerarbeiten größer ist, da keine Erklärungen für den Ablauf der Gruppenarbeiten notwendig sind und folglich mehr Übungszeit zur Verfügung steht. Dadurch entfällt jedoch gleichzeitig die regelmäßige Ernennung eines Tisch-Chefs, so dass ein neues motivierendes Element in der Studie I eingeführt werden sollte.

6 Studie I: Implementierung durch Trainer

6.1 Fragestellung und Hypothesen

Die Erkenntnisse aus der Pilotstudie führten zu Optimierungen des *Unterrichtsintegrierten Selbstregulatorischen Aufsatztrainings* (USAT)-Programms. In der Studie I wurde der USAT-Bedingung eine Kontrollbedingung gegenübergestellt, welche dieselben Schreibstrategien wie das USAT-Programm enthält, jedoch keine selbstregulatorischen Komponenten umfasst. Es handelt sich dabei um einen *Writer's Workshop* (WW), welcher in Anlehnung an den Schreibprozessansatz konzipiert wurde (siehe Kapitel 2.3.1.1). Dieselben Trainerinnen wie in der Pilotstudie führten das USAT- und das WW-Programm im gesamten Klassenkontext im regulären Deutschunterricht durch. Um die Wirksamkeit, Stabilität und Generalisierbarkeit des USAT-Programms im Vergleich zum WW-Programm zu analysieren, wurde direkt vor und nach der Durchführung der Trainings sowie sechs Wochen später folgende Variablen erhoben:

- Strategienahe Variablen (Geschichtenelemente, Planungsqualität, Anzahl der Revisionen, schreibbezogenes Wissen);
- Strategieferne Variablen (Geschichtenqualität, Wortanzahl, Transfermaß Erlebniserzählungen).

Die Richtung der zu erwartenden Effekte ergibt sich analog zu den in Kapitel 3.1 konstatierten Generalhypothesen. Insgesamt lassen sich folgende Fragestellungen und Hypothesen spezifizieren:

(1) Sind von Experten entwickelte Trainingsprogramme wie das *Unterrichtsintegrierte Selbstregulatorische Aufsatztraining* (USAT) und der *Writer's Workshops* (WW) in den Regelunterricht durch Trainer implementierbar?

Jede Lerntrainerin führte in jeweils einer Klasse das USAT- und das WW-Programm im gesamten Klassenkontext durch. Das USAT-Programm baut auf dem *Selbstregulatorischen Aufsatztraining* (SAT) von Glaser (2005) auf, welches wiederum auf dem *Self-Regulated Strategy Development* (SRSD)-Ansatz von Harris und Graham (1996) basiert. Das SAT-Programm wurde bislang nur in Kleingruppen durchgeführt. Im Gegensatz dazu konnte SRSD bereits erfolgreich im gesamten Klassenkontext implementiert werden (De La Paz & Graham, 2002; Torrance et al., 2007). Darüber hinaus ist der *Writer's Workshop* in den USA eine populäre Unterrichtsmethode und ist folglich von Lehrern im Unterricht einsetzbar (Troia et al., 2009).

Hypothese 1: Insgesamt wurde erwartet, dass sowohl das USAT- als auch das WW-Programm erfolgreich von Lerntainerinnen mit ganzen Klassen durchführbar sind. Ferner wurde angenommen, dass die Vollständigkeit und Qualität der Durchführung der beiden Trainings zwischen den Bedingungen und den Trainerinnen nicht variiert, so dass sowohl bei der Umsetzung des USAT- als auch des WW-Programms eine hohe Implementationsgenauigkeit unterstellt wird. Maßnahmen wie z.B. Unterrichtsbeobachtungen zur Sicherung der inhaltlichen Vollständigkeit und einer hohen Qualität der Implementierung wurden ergriffen.

(2) Ist das *Unterrichtsbegleitete Selbstregulatorische Aufsatztraining* (USAT) effektiver in der Förderung schreibbezogener Maße als der *Writer's Workshop* (WW)?

Das WW-Programm enthält dieselben Schreibstrategien wie das USAT-Programm. Des Weiteren wurde die Zeit für Übung, Wiederholung und Instruktion in beiden Bedingungen konstant gehalten. Mit Kleingruppen konnten Glaser und Kollegen die Überlegenheit eines selbstregulatorischen Aufsatztrainings bereits nachweisen (Glaser, 2005; Glaser & Brunstein, 2007b; Glaser et al., 2009). Zwingende Belege für den inkrementellen Effekt der Selbstregulation stehen jedoch noch aus (De La Paz, 2007).

Hypothese 2: Es wurde vermutet, dass bezüglich der strategienahen und –fernen Schreibmaße das USAT-Programm deutlich besser abschneidet als das WW-Programm. Das Hinzufügen von selbstregulatorischen Strategien sollte die Effektivität eines Schreibtrainings erhöhen, da erwartet wurde, dass sich die Schüler die Schreibstrategien besser merken und in ihre eigene Geschichten leichter umsetzen können. Darüber hinaus wurde angenommen, dass Selbstregulation die Motivation der Schüler erhöht, an ihren eigenen Aufsätzen zu arbeiten.

(3) Können die Effekte nach sechs Wochen ebenfalls nachgewiesen werden?

Direkt vor und direkt nach dem Training sowie sechs Wochen später wurden die strategienahen und –fernen Schreibleistungen erhoben, um einerseits den unmittelbaren Lernzuwachs als auch andererseits den zeitlichen Verlauf der Wirksamkeit der beiden Trainingsprogramme zu überprüfen.

Hypothese 3: Die explizite Vermittlung von selbstregulatorischen Strategien sollte dabei das Memorieren und den Einsatz der Schreibstrategien nachhaltiger fördern. Aus diesem Grund wurde postuliert, dass sich die Unterschiede zwischen der USAT- und der

WW-Bedingung direkt im Anschluss an die Trainings zeigen und dass die Leistungsunterschiede sechs Wochen später beim Follow-Up weiterhin stabil bleiben.

(4) Werden die erlernten Schreibstrategien auch bei ungeübten Textgenres angewendet?

Der Transfer der erlernten Fertigkeiten auf eine ungelernete Fertigkeit (Anforderungs-Transfer) wurde dadurch erfasst, dass die Schüler beider Bedingungen jeweils zum Prätest, Posttest und Follow-Up eine Erlebniserzählung schrieben.

Hypothese 4: Selbstregulatorische Strategien sollten es den Schülern erleichtern, die erlernten Schreibstrategien zu abstrahieren, so dass sie auch bei anderen Textgenres leichter die wesentlichen inhaltlichen Elemente identifizieren können. Folglich wurde bezüglich des Transfermaßes angenommen, dass Schüler, denen die Schreibstrategien kombiniert mit selbstregulatorischen Strategien vermittelt werden, das Erlernete besser auf ein ungeübtes Textgenre anwenden können als Schüler, die sich die Schreibstrategien ohne Methoden der Zielsetzung, Selbstbewertung und Selbstkontrolle aneignen.

Davon abgesehen wurden in der Studie I weitere Analysen durchgeführt, um die Effekte des USAT- und des WW-Programms genauer zu untersuchen.

(5) Zusätzliche Analysen

Um den Einfluss des gruppenrandomisierten Designs auf die Ergebnisse zu untersuchen, wurden Mixed Model-Analysen durchgeführt und Intra-Klassen-Korrelationen berechnet.

(6) Prüfung von Moderatoreffekten

Um die Ergebnisse für Schüler mit unterschiedlichen Lernvoraussetzungen zu generalisieren, wurden anhand der Kontrollvariablen Geschlecht, Rechtschreibung und Muttersprache Deutsch jeweils zwei Subgruppen gebildet und die Leistungen für die Subgruppen differenziert betrachtet. So wurde überprüft, ob der Einfluss des USAT- und des WW-Programms in gleicher Weise auf die einzelnen Subgruppen der Kontrollvariablen wirkt.

6.2 Methode

6.2.1 Untersuchungsdesign und Ablauf

Zur Überprüfung der oben beschriebenen Hypothesen wurde ein Vortest-Nachtest-Follow-Up-Vergleichsgruppen-Versuchsplan gewählt. Das USAT-Programm wurde mit dem WW-

Programm ohne selbstregulatorische Komponenten verglichen. Insgesamt führten vier geschulte Lerntainerinnen das USAT- und das WW-Programm im gesamten Klassenkontext durch. Die Klassen bzw. Trainer wurden dabei zufällig einer der beiden Bedingungen zugeordnet. Eine grafische Darstellung des Untersuchungsdesigns der Studie I zeigt Abbildung 26.

Die Rekrutierung der Klassen und Lehrer erfolgte telefonisch. Falls Lehrer am Telefon Interesse an der Teilnahme äußerten, erhielten sie per Fax ein Schreiben mit ausführlichen Informationen. Gleichzeitig konnten sich die Lehrer mit dem Schreiben für eine Informationsveranstaltung anmelden, mit der eine verbindliche Teilnahme an der Studie verbunden war. Da die Kapazitäten auf zwei Klassen pro Trainerin begrenzt waren, erhielten nur die ersten acht Lehrer, die sich für die Informationsveranstaltung anmeldeten, eine Zusage. Vier Lehrern musste abgesagt werden. Alle acht Lehrer waren weiblich und kamen zur Informationsveranstaltung. Dort wurden ihnen die theoretischen Grundlagen der Schreibtrainings vermittelt. Darüber hinaus trafen Trainerinnen und Lehrerinnen organisatorische Absprachen.

Die Zuordnung der Klassen zu den beiden Bedingungen erfolgte auf Klassenebene und anhand der Klassengröße und der Rechtschreibfertigkeiten, so dass in beiden Bedingungen eine nahezu gleiche Anzahl an Schülern mit durchschnittlich ähnlichen Rechtschreibleistungen vorlag. Die Zuteilung der Klassen zu den vier Trainerinnen erfolgte zufällig.

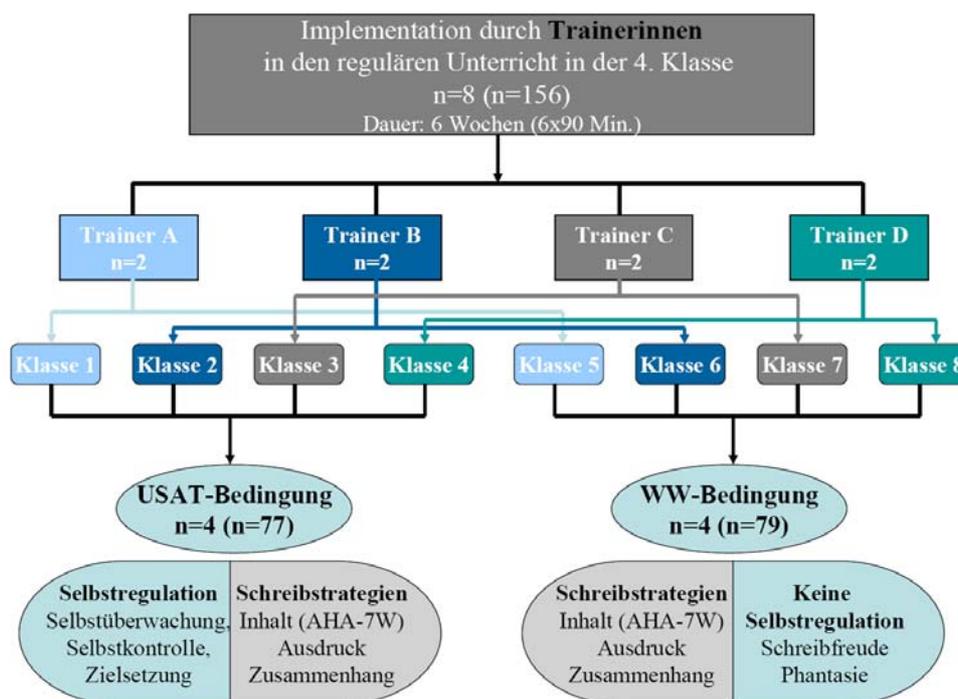


Abbildung 26: Untersuchungsdesign der Studie I.

Die beiden Aufsatzbedingungen wurden in kompletten Klassen von vier geschulten Lerntainerinnen durchgeführt, wobei jede Trainerin sowohl eine Experimental- als auch eine Kontrollbedingung instruierte. Die Trainerinnen der Pilotstudie und der Studie I waren identisch (siehe Kapitel 5.3.1). In beiden Bedingungen wurde über einen Zeitraum von sechs Wochen einmal wöchentlich eine 90minütige Trainingseinheit realisiert. Basierend auf den Erfahrungen aus der Pilotstudie wurde jeder Trainerin eine Tutorin zur Unterstützung zugeordnet. Drei der vier Tutorinnen studierten Psychologie im Grundstudium. Die vierte Tutorin studierte Sonderschullehramt. Gemeinsam mit den Lerntainerinnen nahmen die Tutorinnen vor Beginn des Trainings an mehreren Schulungen teil, in denen sie die Inhalte der USAT- und WW-Bedingung kennenlernten. Während der Durchführung der beiden Trainings beobachteten und bewerteten die Tutorinnen anhand von standardisierten Bögen die Implementierung der beiden Bedingungen durch die jeweilige Trainerin. Abgesehen davon assistierten die Tutorinnen den Trainerinnen bei der Durchführung der Trainings, indem sie z.B. beim Austeilen der Materialien mithalfen.

Die Testungen zu den drei verschiedenen Messzeitpunkten (Prä, Post, Follow-Up) wurden ebenfalls von den Trainerinnen in den jeweiligen Klassen durchgeführt. Zuvor wurden die Trainerinnen explizit in der Durchführung der Testungen geschult. Sie erhielten zudem ausformulierte Manuale für die Testungen (siehe CD-Anhang 4.2).

Als Kontrollmaß wurden die Rechtschreibleistungen der Schüler mit dem DRT 4 (Grund, Haug, & Naumann, 2004) erhoben. Wie in Abbildung 14 dargestellt ist, fanden die Testungen der Rechtschreibleistungen einige Monate vor Beginn des Trainings statt.

6.2.2 Stichprobe

An der Studie I waren insgesamt 156 Viertklässler in acht Klassen aus fünf verschiedenen Grundschulen der Landkreise Gießen und Vogelsberg beteiligt.

Die Größen der Klassen variierten in der USAT-Bedingung zwischen 16 und 21 und in der WW-Bedingung zwischen 14 und 23 Kindern. Im Mittel umfasste eine Klasse 19.5 Schüler (SD: 3.02). Die Gesamtstichprobe bestand zu 49% aus Mädchen und zu 51% aus Jungen. Das Durchschnittsalter betrug 9.92 Jahre (SD: .74). Die Mehrheit der Schüler war deutscher Herkunft (95%). In der Tabelle 13 sind die Merkmale zur Beschreibung der Stichprobe getrennt für die beiden Bedingungen aufgelistet.

Tabelle 13: Beschreibung der Stichprobe getrennt nach Bedingung (Studie I).

	USAT		WW	
<i>n</i>	77		79	
Geschlecht	w: 48.1%	m: 51.9%	w: 50.6%	m: 49.4%
Alter (in Jahren)	M: 9.87	SD: 0.81	M: 9.97	SD: 0.66
Nationalität Deutsch	93.5%		96.2%	
Andere Nationalität	6.5%		3.8%	
Deutsch als Muttersprache	76.6%		72.2%	
Andere Muttersprache	13.0%		12.7%	
Deutsch und andere Sprache	10.4%		15.2%	
Klassengröße	M: 19.25	SD: 2.22	M: 19.75	SD: 4.03
Rechtschreibung	M: 49.09	SD: 9.17	M: 46.78	SD: 8.62
Fehltermine	M: 0.40	SD: 0.54	M: 0.28	SD: 0.48

Hinsichtlich aller in Tabelle 13 berichteten Merkmale (Geschlecht, Alter, Nationalität, Muttersprache, Klassengröße, Rechtschreibung, Fehltermine) bestanden keine signifikanten Unterschiede zwischen den beiden Bedingungen ($ps > .05$).

6.2.3 Trainingsbedingungen

In der Studie I wurden zwei Bedingungen gegenübergestellt: das *Unterrichtsbasierte Selbstregulatorische Aufsatztraining* (USAT) und der *Writer's Workshop* (WW). Während das USAT-Programm in weiten Teilen mit dem in der Pilotstudie vorgestellten Training übereinstimmt, stellt das WW-Programm ein Aufsatztraining dar, welches dieselben Schreibstrategien wie das USAT-Programm enthält, diese jedoch ohne den Einsatz selbstregulatorischer Methoden vermittelt. Für die Konzeption der WW-Bedingung wurden Anregungen aus dem in den USA weit verbreiteten *Writer's Workshop* (Fletcher & Portalupi, 2008; Gillet & Beverly, 2001; Johnson, 2004) herangezogen, welcher bereits im Theorieteil dieser Arbeit vorgestellt wurde (siehe Kapitel 2.3.1.1). Die WW-Bedingung ist jedoch nicht komplett mit dem dort beschriebenen Ansatz identisch, da die Zeiten für Instruktion, Übung und Wiederholung in beiden Bedingungen konstant gehalten wurden.

6.2.3.1 Vergleich der USAT- und der WW-Bedingung

Zur Verdeutlichung der Unterschiede bzw. Gemeinsamkeiten der beiden Bedingungen sind in Tabelle 14 zunächst prototypische Einheiten der USAT- und der WW-Bedingung aufgelistet.

Tabelle 14: Prototypische Trainingsinhalte der USAT- und der WW-Bedingung.

USAT	WW
<p>Rückmeldung</p> <p>Prozess- (7W-Protokoll) und produktbezogene Rückmeldungen (Leistungsprotokoll)</p>	<p>Rückmeldung</p> <p>Keine prozess- oder produktorientierte Rückmeldungen, allgemeine Rückmeldungen, Betonung der Selbstständigkeit, Phantasie und Spaß</p>
<p>Motivierung</p> <p>Motivierung der Schüler, die Schreibstrategien anzuwenden (Strategieregulation) und Geschichtenexperte zu werden (Performanzregulation)</p> <p>Loben der Schüler</p> <p>Ernennung des „Besten Bergsteigers“ (siehe Abbildung 30)</p> <p>Berechnung der Punktedifferenz zwischen letzter und vorletzter Geschichte im Leistungsprotokoll</p>	<p>Motivierung</p> <p>Motivierung der Schüler, schöne Geschichten zu schreiben, Spaß am Schreiben zu entwickeln und die Phantasie zu verwenden</p> <p>Loben der Schüler</p> <p>Ernennung des „Größten Schreibfans“ (siehe Abbildung 30)</p>
Keine Übungen zur Förderung der Phantasie	Phantasiereise mit der Leitfigur „Balthasar“ oder zusätzliche Übungen zur Förderung der Phantasie
<p>Wiederholung</p> <p>Wiederholung der Schreibstrategien in Bezug zu den Geschichten der Schüler</p>	<p>Wiederholung</p> <p>Wiederholung der Schreibstrategien durch Rezitieren und Memorieren der Schreibstrategien</p>
<p>Partnerarbeit (TE¹⁰ 2 bis TE 6)</p> <p>Übungen zum Inhalt, Planen, Überarbeiten, Ausdruck und Zusammenhang sowie der Anwendung selbstregulatorischer Strategien</p>	<p>Gruppenarbeit (TE 2 bis TE 6)</p> <p>Übungen zum Inhalt, Planen, Überarbeiten, Ausdruck und Zusammenhang sowie der Phantasie</p>
<p>Strategieinstruktion (TE 1 bis TE 3, TE 5) oder Übungen (TE 4 und TE 6)</p> <p>Kognitives Modellieren der Schreibstrategien in Kombination mit den selbstregulatorischen Methoden</p>	<p>Strategieinstruktion (TE 1 bis TE 3, TE 5) oder Übungen (TE 4 und TE 6)</p> <p>Kognitives Modellieren der Schreibstrategien ohne selbstregulatorische Methoden</p>
Prozess- und produktbezogene Zielsetzung mit dem 7W- und Leistungsprotokoll	Keine Zielsetzung
<p>Geschichtenschreiben:</p> <p>Angeleitetes Planen mit Geschichtenplan</p> <p>Angeleitetes Schreiben mit Geschichtenplan</p> <p>Angeleitetes Überarbeiten mit Checkliste</p>	<p>Geschichtenschreiben:</p> <p>Angeleitetes Planen</p> <p>Angeleitetes Schreiben</p> <p>Angeleitetes Überarbeiten</p>

¹⁰ TE = Trainingseinheit; Dauer: 90 Minuten.

In Tabelle 14 wird deutlich, dass die meisten Trainingsinhalte sowohl in der USAT- als auch in der WW-Bedingung integriert wurden. Die Komponenten Rückmeldung, Motivierung, Wiederholung, Partnerarbeit, Instruktion von Schreibstrategien und das Planen, Schreiben, Überarbeiten einer Geschichte waren in beiden Bedingungen enthalten. Der einzige Unterschied besteht darin, dass in der USAT-Bedingung selbstregulatorische Strategien (Zielsetzung, Selbstüberwachung, Selbstkontrolle) vermittelt wurden und dass in der WW-Bedingung alternative Übungen zur Förderung der Phantasie (z.B. Phantasiereisen) und der Schreibfreude (z.B. Austausch der Geschichten mit dem Tischnachbarn) integriert wurden. Auf diese Weise wurde sichergestellt, dass die Instruktions- und Übungszeit in beiden Bedingungen konstant war.

Zur besseren Veranschaulichung sind in Abbildung 27 die Gemeinsamkeiten und Unterschiede der beiden Bedingungen noch einmal gegenübergestellt.

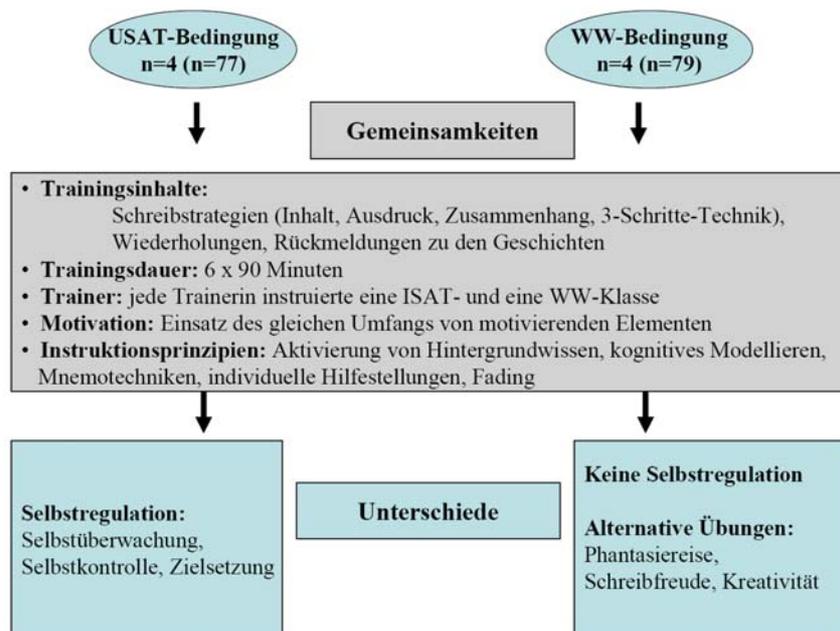


Abbildung 27: Gemeinsamkeiten und Unterschiede der USAT- und der WW-Bedingung.

Abbildung 27 macht deutlich, dass beim Vergleich der USAT- mit der WW-Bedingung die Gemeinsamkeiten überwiegen und kaum Unterschiede feststellbar sind. Es wurden dieselben Schreibstrategien, Instruktionsprinzipien und motivationalen Faktoren vermittelt, wobei in der USAT-Bedingung diese explizit mit der Selbstregulation verknüpft wurden. Des Weiteren wurden dieselben Schreibansätze in beiden Bedingungen eingesetzt.

Die konkrete Beschreibung der Trainingsinhalte erfolgt im nächsten Abschnitt. Dennoch werden in der nachfolgenden Tabelle 15 die wichtigsten Inhalte der USAT- und der WW-Bedingung aufgelistet.

Tabelle 15: Vergleich der USAT- und der WW-Bedingung.

	USAT	WW
Prätest	Erfassung der Schreibleistung sowie der Planungs- und Überarbeitungsaktivitäten und des schreibbezogenen Wissens Erfassung der Transferleistung (Erlebniserzählung) durch den jeweiligen Klassenlehrer im Anschluss an die Prätestung	
TE 1	Herleitung der AHA-7W-Strategie und der 3-Schritte-Technik Rückmeldungen zu den Prätestgeschichten Modellierung des Planens und Schreibens einer Bildergeschichte Angeleitetes Planen und Schreiben einer Bildergeschichte	
	Modellierung und angeleitetes Üben der Selbstüberwachung beim Planen und Schreiben mit dem Geschichtenplan Modellierung der Zielsetzung mit dem 7W-Protokoll und dem Leistungsprotokoll	Modellierung der Anwendung der Textlupe (siehe Abbildung 34) zur Bewertung von Geschichten
TE 2	Motivierung der Schüler durch anerkennendes Lob Rückmeldungen zu den Geschichten der letzten Trainingseinheit Wiederholung der AHA-7W-Strategie und der 3-Schritte-Technik Einführung der Partner-/Gruppenarbeit Angeleitetes Planen, Schreiben und Überarbeiten einer Reizwortgeschichte Modellierung des Überarbeitens einer Geschichte Ernennung eines Gruppensiegers am Ende der Trainingseinheit	
	Modellierung und angeleitetes Üben der Selbstbewertung und -kontrolle beim Berechnen der Punktedifferenz zwischen der Prätest- und der letzten Geschichte; Einkleben von entsprechenden Belohnungen in das Leistungsprotokoll Rückmeldung und Zielsetzung mit dem Leistungs- und dem 7W-Fragen-Protokoll Durchführung einer Partnerarbeit zum Planen mit Hilfe des Geschichtenplans Modellierung und angeleitetes Üben der Selbstbewertung und -kontrolle beim Überarbeiten mit der Checkliste	Durchführung einer Phantasiereise mit der Leitfigur „Balthasar“ Durchführung einer Gruppenarbeit zum Sammeln von Ideen für Abenteuer mit „Balthasar“

TE 3	<p>Einführung eines Belohnungsbuttons für den Schüler mit einer besonders guten Geschichte</p> <p>Motivierung der Schüler durch Lob und Ernennung des Gruppensiegers</p> <p>Rückmeldungen zu den Geschichten der letzten Trainingseinheit</p> <p>Wiederholung der erlernten Strategien</p> <p>Erarbeitung der Strategien zur Verbesserung des Ausdrucks einer Geschichte anhand der Ausdruckskarte</p> <p>Angeleitetes Planen, Schreiben und Überarbeiten einer Bildergeschichte</p>	
	<p>Rückmeldung und Zielsetzung mit dem Leistungs- und dem 7W-Fragen-Protokoll</p> <p>Einführung des Buttons „Bester Bergsteiger“ für den Schüler, der sich am meisten verbessert hat</p> <p>Durchführung einer Partnerarbeit zum Überarbeiten mit Hilfe der Checkliste</p> <p>Erarbeitung der Strategien zur Verbesserung des Ausdrucks Anfangs einer Geschichte</p> <p>Modellierung und angeleitetes Üben der Anwendung der Ausdruckskarte Anfang während des Planens, Schreibens und Überarbeitens</p>	<p>Einführung des Buttons „Größter Schreibanlass“ für den Schüler mit der schönsten und originellsten Geschichte</p> <p>Durchführung einer Gruppenarbeit zum Generieren eines Geschichtenanfangs und Hauptteils bei vorgegebenem Geschichtenabschluss (Förderung der Phantasie)</p> <p>Durchführung einer Partnerübung zum anschaulichen Beschreiben von Nomen</p> <p>Erarbeitung sämtlicher Strategien zur Verbesserung des Ausdrucks einer Geschichte</p>
TE 4	<p>Motivierung der Schüler mit Button, Lob und Gruppensieger</p> <p>Rückmeldungen zu den Geschichten der letzten Trainingseinheit</p> <p>Wiederholung der erlernten Strategien</p> <p>Begleitetes Planen, Schreiben und Überarbeiten einer Geschichte</p>	
	<p>Berechnen der Punktedifferenz zwischen der aktuellen Aufsatzbewertung und der Aufsatzbewertung der Prätestgeschichte; Einkleben von Belohnungen in das Leistungsprotokoll</p> <p>Rückmeldung und Zielsetzung mit dem Leistungs- und dem 7W-Fragen-Protokoll</p> <p>Durchführung einer Partnerarbeit zum Üben der Ausdrucksstrategien mit Hilfe der Ausdruckskarte</p> <p>Modellierung und angeleitetes Üben der Anwendung der Ausdruckskarte Hauptteil/ Ende während des Planens, Schreibens und Überarbeitens</p> <p>Reizwörter als Schreibanlass</p>	<p>Durchführung einer Partnerarbeit zum Überarbeiten mit der Textlupe</p> <p>Durchführung einer Phantasiereise mit „Balthasar“</p> <p>Phantasiereise als Schreibanlass</p>

TE 5	<p>Motivierung der Schüler durch Button, Lob und Gruppensieger Rückmeldungen zu den Geschichten der letzten Trainingseinheit Wiederholung der gelernten Strategien Durchführung einer Partnerübung zur Erarbeitung der Strategien zur Verbesserung des Zusammenhangs einer Geschichte Begleitetes Planen, Schreiben und Überarbeiten einer Bildergeschichte</p>	
	<p>Rückmeldung und Zielsetzung mit dem Leistungs- und dem 7W-Fragen-Protokoll</p> <p>Modellierung und angeleitetes Üben der Selbstkontrolle bei der Überprüfung des Zusammenhangs in der Geschichte</p> <p>Durchführung einer Partnerarbeit zur Selbstkontrolle der richtigen Anwendung der Zusammenhangsstrategien</p>	<p>Identifizierung der Zusammenhangsstrategien in einer Geschichte</p> <p>Durchführung einer Partnerübung zur Beschreibung von Handlungsabläufen</p>
TE 6	<p>Motivierung der Schüler mit Button, Lob und Gruppensieger Rückmeldungen zu den Geschichten der letzten Trainingseinheit Wiederholung der erlernten Strategien Selbstständiges Planen, Schreiben und Überarbeiten einer Bildergeschichte</p>	
	<p>Berechnen der Punktedifferenz zwischen der aktuellen Aufsatzbewertung und der Aufsatzbewertung der Prätestgeschichte; Einkleben von Belohnungen in das Leistungsprotokoll</p> <p>Rückmeldung und Zielsetzung mit dem Leistungs- und dem 7W-Fragen-Protokoll</p> <p>Wiederholung der gesamten Schreibstrategien anhand von drei Schülergeschichten</p> <p>Abgesehen von dem 7W- und Leistungsprotokoll werden keine selbstregulatorischen Materialien verwendet</p>	<p>Durchführung einer Phantasiereise mit „Balthasar“</p> <p>Verfassen einer Balthasar-Geschichte in einer Gruppe nach der „Sätze-Weiterschreiben“- Methode</p> <p>Gegenseitiges Feedback zu den Bildergeschichten</p>
Posttest	<p>Erfassung der Schreibleistung sowie der Planungs- und Überarbeitungsaktivitäten und des schreibbezogenen Wissens</p> <p>Erfassung der Transferleistung (Erlebniserzählung) durch den jeweiligen Klassenlehrer im Anschluss an die Posttestung</p>	
Follow-Up	<p>Erfassung der Schreibleistung sowie der Planungs- und Überarbeitungsaktivitäten und des schreibbezogenen Wissens</p> <p>Erfassung der Transferleistung (Erlebniserzählung) durch den jeweiligen Klassenlehrer im Anschluss an die Follow-Up-Erhebung</p>	

6.2.3.2 Schreibstrategien

Eine wesentliche Gemeinsamkeit zwischen der USAT- und der WW-Bedingung besteht darin, dass dieselben Schreibstrategien den Schülern vermittelt wurden (siehe Abbildung 27). Die Schreibstrategien sind dabei mit denen der Pilotstudie identisch (siehe Kapitel 5.3.2.1). Im Folgenden werden die 3-Schritte-Technik sowie die Strategien zur Verbesserung des Inhalts, des Ausdrucks und des Zusammenhangs einer Geschichte dargestellt, wobei insbesondere auf die Unterschiede zur Pilotstudie eingegangen wird.

Trickposter: Am Anfang des Trainings wurde das Trickposter in beiden Bedingungen eingefügt (siehe Abbildung 28). Für die USAT- und die WW-Bedingung wurden unterschiedliche Trickposter verwendet. Im Gegensatz zur USAT-Bedingung wurde nicht von einem „Geschichtenexperten“ gesprochen, sondern den Schülern erklärt, dass im Verlauf des Trainings Tricks für das Geschichtschreiben gelernt werden. Die Felder im Trickposter waren zunächst leer und wurden nach Einführung neuer Schreibstrategien mit den jeweiligen Strategien gefüllt. Das Vorgehen war in beiden Bedingungen identisch. Das Trickposter diente der Orientierung und verdeutlichte den Schülern, was sie bereits gelernt hatten und was sie noch lernen würden.



Abbildung 28: Trickposter der Studie I und II.

3-Schritte-Technik: Die Einführung der 3-Schritte-Technik (Planen, Schreiben, Überarbeiten) sowie der Planungs- und Überarbeitungsstrategien erfolgte wie in der Pilotstudie (siehe Kapitel 5.3.2.1).

Als Erinnerungshilfe wurde in beiden Bedingungen ebenfalls ein Poster in das Klassenzimmer gehängt, welches zuvor als Hausaufgabe hergeleitet wurde (siehe CD-Anhang A4.3). Jeder Buchstabe in „plane“, „schreibe“ bzw. „überarbeite“ repräsentiert dabei eine Aktivität, die beim jeweiligen Schreibprozess erfolgen sollte. Da die Schüler in der USAT-Bedingung selbstregulatorische Methoden beim Planen, Schreiben und Überarbeiten einsetzen sollten, unterschieden sich die beiden Poster. In der USAT-Bedingung wurde das Poster aus

der Pilotstudie eingesetzt (siehe Abbildung 15). Das 3-Schritte-Technik-Poster der WW-Bedingung ist in Abbildung 29 visualisiert. Anstatt selbstregulatorischer Methoden wurden allgemeine Hinweise wie z.B. „Beachte die 7W-Fragen.“ oder „Schreibe einzelne Stichpunkte auf.“ aufgelistet.



Abbildung 29: 3-Schritte-Technik-Poster der WW- und der LUG-Bedingung.

Inhalt: Die Schreibstrategie AHA-7W zur Verbesserung des Inhalts einer Geschichte wurde bereits in der Pilotstudie erfolgreich instruiert, so dass in der Studie I die Herleitung der 7W-Fragen analog zur Pilotstudie anhand der inhaltlich vollständigen Geschichte „Der schlaue Fuchs“ (siehe CD-Anhang A7.1) erfolgte. Es wurden auch dieselben Erinnerungsstützen wie das 7W-Poster (siehe Abbildung 16) verwendet.

Ausdruck: Die Ausdrucksstrategien dienen der sprachlichen Ausgestaltung der Geschichte und der Erzeugung von Spannung. Wie bereits in der Pilotstudie zeigte die Trainerin anhand der Ausdruckskarte (siehe Abbildung 17), auf der für jede 7W-Frage Ausgestaltungsbeispiele (Adjektive, Verben, Gedanken, Gefühle, wörtliche Rede) illustriert sind, wie die Schüler ihre eigenen Geschichten stilistisch optimieren können.

Im Gegensatz zur Pilotstudie erfolgte die Ausdrucksinstruktion jedoch anhand einer konkreten Beispielgeschichte (siehe CD-Anhang A7.2). Zu Beginn las die Trainerin eine

vollständige, aber nicht ausgestaltete Geschichte vor. Anschließend erarbeitete die Trainerin zusammen mit den Schülern heraus, dass in der vorgelesenen Geschichte zwar alle 7W-Fragen genannt wurden, die Geschichte sich aber trotzdem langweilig anhörte, weil keine Ausgestaltungsmittel enthalten waren. Daraufhin modellierte die Trainerin die Anwendung der Ausdruckskarte (siehe Abbildung 17). Sie zeigte den Schülern anhand der vorgelesenen Geschichte, wie die 7W-Fragen mit Hilfe der Ausdruckskarte anschaulich und bildlich beschrieben werden können. Am Ende las die Trainerin die anfängliche Geschichte noch einmal vor, wobei die Geschichte dieses Mal um Adjektive, Verben, Gedanken und Gefühle ergänzt war. Auf diese Weise konnten die Schüler die Verbesserungen der Geschichte durch die Verwendung von stilistischen Elementen nachvollziehen.

Die beschriebene Instruktion war in der USAT- und in der WW-Bedingung identisch. Im Gegensatz zur WW-Bedingung modellierte die Trainerin in der USAT-Bedingung zudem die Anwendung der Ausdruckskarte beim Planen und Schreiben einer Geschichte. Außerdem erhielt jeder Schüler der USAT-Bedingung eine eigene Ausdruckskarte und wurde angeleitet, diese beim Planen, Schreiben und Überarbeiten einzusetzen.

Zusammenhang: Als letzte genrespezifische Strategie erlernten Schüler Strategien zur Verbesserung des Zusammenhangs. Die Zusammenhangsinstruktion erfolgte wie bereits in der Pilotstudie beschrieben. In der Studie I standen jedoch die Inhalts- und Ausdrucksstrategien im Fokus, so dass der Zusammenhang als zusätzliche Strategie für besonders gute Schüler vorgesehen war. Schwächere Schüler wurden angewiesen, sich nur auf den Inhalt und den Ausdruck zu konzentrieren. Die Zusammenhangsinstruktion erfolgte in der zweitletzten Trainingseinheit, so dass die Schüler wenige Möglichkeiten zum Üben und Wiederholen der Strategie hatten. Da in der letzten Trainingseinheit keine Hilfsmittel eingesetzt wurden, wurden für den Zusammenhang in der fünften Trainingseinheit keine neuen Materialien (Checkliste, 7W-Protokoll) eingeführt, um zusätzliche Verwirrungen zu vermeiden. Aus diesem Grund wurde in der USAT-Bedingung zwar das Zusammenhangssymbol (Treppe mit Ketten, siehe Abbildung 18) eingeführt, aber nicht wie in der Pilotstudie auf der Checkliste zur Selbstkontrolle ergänzt. Vielmehr modellierte die Trainerin, wie man beim Überarbeiten durch Einkreisen der Satzanfänge und Unterstreichen der Handlungsschritte überprüfen kann, ob der Zusammenhang in der eigenen Geschichte berücksichtigt wurde. Da es sich hierbei um eine selbstregulatorische Methode handelt, entfiel das Einkreisen und Unterstreichen in der WW-Bedingung.

6.2.3.3 Unterrichtsmethoden

In der Abbildung 27, welche die Gemeinsamkeiten und Unterschiede zwischen der USAT- und der WW-Bedingung darstellt, ist erkennbar, dass in beiden Bedingungen dieselben Unterrichtsmethoden eingesetzt wurden.

Partnerarbeit: Die in der Pilotstudie durchgeführte Gruppenarbeit wurde in der Studie I zu einer Partnerarbeit umgestaltet. Einige Elemente aus der Pilotstudie wie Gruppennamen (Skater, Sprinter, Schwimmer, Reiter, Biker), Teamposter (siehe Abbildung 25) und Gruppengewinner (Planungs-Champion, Überarbeitungs-Held, Spannungs-Gewinner, Zusammenhangs-Sieger, Freundlichkeits-Meister) blieben jedoch erhalten, wobei zwei bis drei Paare immer eine Gruppe ausmachten und zusammen an einem Tisch saßen. Die Zusammenstellung von heterogenen Gruppen erfolgte im Vorhinein durch die jeweilige Lehrerin.

Hinsichtlich der Inhalte der Partnerarbeiten unterschieden sich die beiden Bedingungen. Während es in der USAT-Bedingung vornehmlich darum ging, die Anwendung der selbstregulatorischen Methoden in Kombination mit den Schreibstrategien zu schulen, fokussierten die Übungen der WW-Bedingung die Förderung der Schreibfreude und der Phantasie. Es wurde dabei auf Übungen aus Grundschulzeitschriften zurückgegriffen (Altenburg, 2006; Delp, Müller, & Stöveken, 2004; Menauer, 2006).

Instruktionsprinzipien: Die zur Instruktion der Schreibstrategien und selbstregulatorischen Methoden eingesetzten Prinzipien entsprechen denen der Pilotstudie und orientieren sich in beiden Bedingungen an den sechs Instruktionsstufen von Harris und Graham (1996).

Wiederholung: Die Zeit für Wiederholungen und Übungen wurde in beiden Bedingungen konstant gehalten. Die Schüler wurden zudem auf die Poster als Erinnerungstützen hingewiesen. Im Gegensatz zur USAT-Bedingung erfolgte in der WW-Bedingung die Wiederholung nicht anhand von konkreten Schülergeschichten.

Motivationsfördernde Komponenten: In der USAT- und in der WW-Bedingung wurde der gleiche Umfang an motivationsfördernden Elementen implementiert. Während die Motivierung in der USAT-Bedingung sich insbesondere auf den Einsatz der Schreibstrategien bezog, wurden die Schüler in der WW-Bedingung angeregt, Schreibfreude zu entwickeln und ihre Phantasie einzusetzen. Da der Tisch-Chef aus der Pilotstudie in der Studie I entfiel, wurde ab der dritten Trainingseinheit der „Beste Bergsteiger“ bzw. der „Größte Schreibfan“ eingeführt (siehe Abbildung 30).



Abbildung 30: Button Bester Bergsteiger (USAT-Bedingung) und Größter Schreibfan (WW-Bedingung).

Jede Woche wurde ein Kind für die zuletzt geschriebene Geschichte mit einem Button, welcher analog zu einem Wanderpokal jede Woche an ein anderes Kind weitergereicht wurde, belohnt. In der USAT-Bedingung wurde der Schüler, der sich im Vergleich zur vorherigen Woche um die meisten Punkte im Leistungsprotokoll gesteigert hatte, zum „Besten Bergsteiger“ ernannt. Da nicht das Kind mit den meisten Punkten oder mit der besten Geschichte belohnt wurde, sondern die Steigerung im Fokus stand, hatten insbesondere schwächere Schüler die Möglichkeit, die Auszeichnung zu bekommen. Im Gegensatz dazu erhielt in der WW-Bedingung das Kind, welches die schönste Geschichte geschrieben hatte, den Button „Größter Schreibfan“. Es wurde jedoch ebenfalls darauf geachtet, insbesondere schwächere Schüler zu motivieren

Hausaufgaben: Nach jeder Trainingsstunde erhielten die Schüler eine Hausaufgabe, in der sie die in der vorangegangenen Stunde eingeführten Strategien wiederholten. Die Hausaufgaben dienten dazu, dass schreibbezogene Wissen zu verfestigen, so dass in beiden Bedingungen dieselben Hausaufgaben durchgeführt wurden.

6.2.3.4 Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (USAT)

In der Pilotstudie wurde das *Unterrichtsintegrierte Selbstregulatorische Aufsatztraining* (USAT) bereits erprobt und hinsichtlich der Effektivität analysiert. Aufbauend auf den Ergebnissen der Pilotstudie wurden für die Studie I die selbstregulatorischen Materialien, die motivierenden Komponenten und die Lern- und Gedächtnisstützen optimiert.

Im folgenden Abschnitt werden insbesondere die Unterschiede zwischen der Pilotstudie und der Studie I hervorgehoben. Die Schreibstrategien und die Unterrichtsmethoden, die mit denen der WW-Bedingung identisch sind, wurden bereits im vorangegangenen Abschnitt beschrieben. Im Folgenden werden zunächst die

selbstregulatorischen Strategien erörtert und anschließend die einzelnen Trainingssitzungen skizziert.

6.2.3.4.1. Selbstregulatorische Strategien

In dem USAT-Programm wurden selbstregulatorische Strategien wie Zielsetzung, Selbstüberwachung und Selbstkontrolle explizit mit Schreibstrategien, die im obigen Abschnitt beschrieben wurden und mit dem WW-Programm identisch sind, kombiniert.

Geschichtenplan: Da in der Pilotstudie die Anwendung des Geschichtenplans zur Selbstüberwachung der Schreibstrategien sehr gut von den Schülern verstanden und erfolgreich umgesetzt wurde, waren keine Optimierungen des Geschichtenplans notwendig (siehe Abbildung 23).

Checkliste: Die Ergebnisse der Pilotstudie zeigten für Revisionen keine signifikanten Unterschiede zwischen den Prä- und Posttestwerten. Aus diesem Grund wurde die Checkliste zur Bewertung der eigenen Schreibleistung optimiert. Die Instruktion erfolgte wie in der Pilotstudie beschrieben anhand einer Beispielgeschichte (siehe CD-Anhang A7.5). Um das Einprägen und Identifizieren der Schreibstrategien in den eigenen Geschichten stärker zu fördern, wurde die Checkliste um zentrale Schlüsselwörter für die einzelnen 7W-Fragen ergänzt (siehe Abbildung 31).

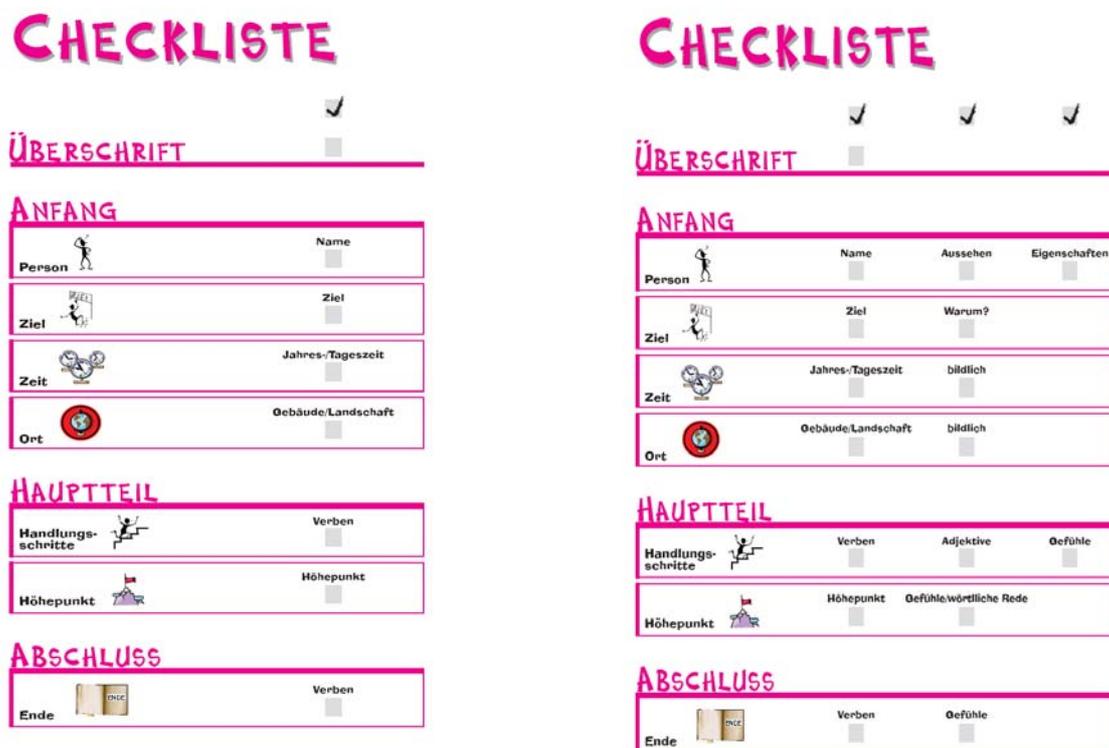


Abbildung 31: Optimierte Checklisten I und II der Studie I und II.

In der Pilotstudie wurde beobachtet, dass die Schüler beim Überarbeiten schnell die einzelnen Kästchen abhakten, ohne sich genau die einzelnen Schreibstrategien in Erinnerung zu rufen. Die Schlüsselwörter in den neuen Checklisten sollten dies verhindern und die Schüler anregen, ihre eigenen Geschichten genauer zu analysieren, um anschließend ihre Geschichten besser zu überarbeiten. Abgesehen davon wurden in der Studie I nur noch zwei anstatt drei verschiedene Checklisten eingesetzt. Eine separate Checkliste für die Kontrolle der Zusammenhangesstrategien entfiel.

Zielsetzung: Wie in der Pilotstudie setzten sich die Schüler sowohl prozess- als auch produktorientierte Ziele bevor sie mit dem Verfassen einer neuen Geschichte begannen. Die Zielsetzung funktionierte in der Pilotstudie bereits zufriedenstellend, so dass nur wenige Veränderungen vorgenommen wurden.

Prozessbezogene Zielsetzung: Hinsichtlich des 7W-Protokolls wurde analog zur Checkliste kein neues Protokoll für den Zusammenhang eingeführt. Das erste 7W-Protokoll mit den Inhaltsstrategien blieb im Vergleich zur Pilotstudie unverändert (siehe Abbildung 19). Das optimierte 7W-Protokoll II enthält sowohl die Strategien des Ausdrucks und des Zusammenhangs und ist in Abbildung 32 abgebildet. Da es bereits nach der Ausdrucksinstruktion in der dritten Trainingseinheit angewandt wurde, enthält es im Vergleich zum alten 7W-Protokoll eine Zeile mehr.

Name _____

7-W-FRAGEN PROTOKOLL

| 7-W-Fragen | <input type="checkbox"/> |
|-------------------------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| Adjektive
Gedanken & Gefühle
Verben | <input type="checkbox"/> |
| 7-W-Fragen | <input type="checkbox"/> |
| Adjektive
Gedanken & Gefühle
Verben | <input type="checkbox"/> |
| 7-W-Fragen | <input type="checkbox"/> |
| Adjektive
Gedanken & Gefühle
Verben | <input type="checkbox"/> |
| 7-W-Fragen | <input type="checkbox"/> |
| Adjektive
Gedanken & Gefühle
Verben | <input type="checkbox"/> |

Reihenfolge
Verben
Satzanfänge

Reihenfolge
Verben
Satzanfänge

Abbildung 32: Optimiertes 7W-Protokoll II der Studie I und II.

Produktbezogene Zielsetzung: Mit Hilfe des Leistungsprotokolls setzten sich die Schüler produktbezogene Ziele und erhielten Rückmeldungen zu der Gesamtqualität ihrer Aufsätze. Das USAT-Programm intendierte neben der Verbesserung der Aufsatzqualität auch

die damit verbundenen Veränderungen der Kognitionen und Einstellungen. Deswegen sollte in der Studie I stärker die Selbstwahrnehmung der Schüler als bessere Schreiber gefördert werden. Dies wurde dadurch angestrebt, dass sich die Schüler in jeder zweiten Trainingsstunde (Trainingsstunde 2, 4 und 6) die Differenz zwischen der aktuellen Aufsatzbewertung und der Aufsatzbewertung der Prättestgeschichte ausrechneten. Je nach Höhe der Differenz durften sich die Schüler daraufhin „guter Schreiber“, „prima Schreiber“, „super Schreiber“ oder „Meister-Schreiber“ nennen und sich einen entsprechenden Klebepunkt in das Leistungsprotokoll kleben (siehe CD-Anhang A5.1.3). Die Abbildung 33 illustriert das optimierte Leistungsprotokoll der Studie I, welches die Bewertungsabstufungen enthält.

Die in der Pilotstudie vorgestellte Rückseite des Leistungsprotokolls (siehe Abbildung 22) wurde unmodifiziert beibehalten. Sie sollte den Schülern den Zusammenhang zwischen der Anwendung der Schreibstrategien und der Punktzahl im Leistungsprotokoll verdeutlichen, so dass die Schüler sich stärker anstrengten, die Strategien einzusetzen und „gute Schreiber“ zu werden.

LEISTUNGSPROTOKOLL

Name _____

Inhalt/Ausdruck/Zusammenhang	VOR		2		3		4		5		6		7		NACH
	Ziel	Punkte	Ziel												
21															21
20															20
19															19
18															18
17															17
16															16
15															15
14															14
13															13
12															12
11															11
10															10
9															9
8															8
7															7
6															6
5															5
4															4
3															3
2															2
1															1

VOR DEM TRAINING
GESCHICHTE 2
GESCHICHTE 3
GESCHICHTE 4
GESCHICHTE 5
GESCHICHTE 6
7
NACH DEM TRAINING

Legende: Joker Guter Schreiber Prima Schreiber Super Schreiber Meister-Schreiber

Abbildung 33: Optimiertes Leistungsprotokoll der Studie I und II.

6.2.3.4.2. *Ablauf des Trainings*

Nachfolgend werden die Inhalte der sechs Trainingseinheiten, die jeweils 90 Minuten dauerten, skizziert. Es werden nur die Unterschiede zur Pilotstudie dargestellt. Eine ausführliche Darstellung der USAT-Trainingseinheiten ist in Kapitel 5.3.2.4 nachzulesen. Nach jeder Trainingseinheit wurden die Aufsätze der Schüler eingesammelt und von der jeweiligen Trainerin mit dem 7W- und Leistungsprotokoll bewertet. Ein kurzer Überblick der Trainingseinheiten bietet Tabelle 15.

Trainingseinheit 1 (Schwerpunkt: Inhalt, Planen, Schreiben): Die erste Trainingseinheit umfasste die Instruktion der 3-Schritte-Technik und der Inhaltsstrategien. Gleichzeitig wurden die meisten selbstregulatorischen Materialien und Techniken eingeführt. Die im weiteren Verlauf des Trainings vorkommenden Strategien stellen Erweiterungen und Adaptationen der in der ersten Trainingseinheit erlernten Techniken dar.

Wie auch in der Pilotstudie wurden den Schülern zunächst die groben Inhalte des Trainings erläutert, wobei an das Vorwissen der Schüler angeknüpft wurde. Den Schülern wurde erklärt, dass es um das Aufsatzschreiben geht und dass alle Schüler im Verlauf des Trainings Geschichtenexperte werden. Zur besseren Strukturierung der Schreibstrategien stellte die Trainerin das zunächst leere Trickposter „*Mit diesen Tricks wirst du ein Geschichtenexperte*“ (siehe Abbildung 28) vor. Sie klebte anschließend die drei Begriffe „Inhalt“, „Ausdruck“ und „Zusammenhang“ ein. Das Poster wurde in das Klassenzimmer gehängt. Im Verlauf des Trainings wurde das Poster immer weiter mit Schreibstrategien gefüllt. Sobald eine neue Strategie, z.B. der AHA-7W-Trick, gelernt worden war, wurde dieser in der dafür vorgesehenen Stelle im Poster geklebt. Auf diese Weise erhielten die Schüler schnell einen Überblick über die für das Schreibtraining relevanten Strategien.

Die daran anschließenden Trainingsinhalte waren mit denen der Pilotstudie identisch:

- Modellierung der Anwendung des Leistungsprotokolls und Verteilung der Trainingsmappen mit Rückmeldungen zur Prätest-Geschichte;
- Herleitung der AHA-7W-Strategie anhand des Textes „Der schlaue Fuchs“ (siehe CD-Anhang A7.1) und Aufhängen des 7W-Posters in der Klasse (siehe Abbildung 16);
- Einführung des 7W-Protokolls und Verteilung der 7W-Protokolle mit Rückmeldungen zur Prätest-Geschichte;
- Modellierung der Zielsetzung anhand des 7W- und des Leistungsprotokolls;
- Zielsetzung;
- Einführung der 3-Schritte-Technik;

- Modellierung der Anwendung des Geschichtenplans anhand einer Bildergeschichte (siehe CD-Anhang A7.4);
- Planen und Schreiben einer Bildergeschichte mit dem Geschichtenplan.

Trainingseinheit 2 (Schwerpunkt: Überarbeiten): Die zweite Trainingseinheit mit dem Schwerpunkt Überarbeiten unterscheidet sich von der Pilotstudie in der Weise, dass die Schüler zu Beginn die Punktedifferenz ausrechneten und dass Partner- anstatt Gruppenarbeiten stattfanden und somit kein Tisch-Chef ernannt wurde.

Nachdem die Schüler am Anfang die Rückmeldungen zur letzten Geschichte anhand des 7W- und Leistungsprotokolls erhielten und die Trainerin die Anstrengungen der Schüler gelobt hatte, zeigte sie, wie die Differenz zwischen der Prätestgeschichte und der letzten Geschichte ausgerechnet wird und welcher Klebepunkt ab welcher Differenz in das Leistungsprotokoll eingeklebt werden sollte. Auf diese Weise sollte den Schülern die bereits erzielte Verbesserung verdeutlicht werden. Im Anschluss berechnete jeder Schüler seine eigene Punktedifferenz und klebte einen Klebepunkt in das Leistungsprotokoll ein. Die Trainerin unterstützte, wenn nötig, die Schüler.

Nach der Wiederholung des AHA-7W-Tricks erklärte die Trainerin den allgemeinen Ablauf der Partnerarbeit und die Aufgabe der zweiten Trainingseinheit, die darin bestand, zu zweit zu einem vorgegebenen Bild eine Geschichte unter Verwendung des Geschichtenplans zu planen (siehe CD-Anhang A5.1.1). Analog zur Pilotstudie führte die Trainerin das Poster „Beste Teams“ (siehe Abbildung 25) ein, motivierte die Schüler, sich während der Partnerarbeit anzustrengen und kürte am Ende der Trainingseinheit die beste Gruppe als „Planungs-Champion“.

Die weiteren Inhalte der zweiten Trainingseinheit erfolgten wie in der Pilotstudie:

- Zielsetzung;
- Planen und Schreiben einer Reizwortgeschichte;
- Modellierung der Anwendung der Checkliste (siehe Abbildung 31) und des Einfügens fehlender 7W-Fragen anhand einer Beispielgeschichte (siehe CD-Anhang A7.5);
- Überarbeiten der Geschichten mit der Checkliste.

Trainingseinheit 3 (Schwerpunkt: Ausdruck Anfang): In der dritten Trainingseinheit wurde erstmalig der Button „Bester Bergsteiger“ (siehe Abbildung 30) eingeführt, welcher dem Schüler verliehen wurde, der sich von der vorletzten (TE 1) zur letzten (TE 2) Trainingseinheit am meisten gesteigert hatte. Jede Woche wurde der Button neu verliehen, so dass der Button ein zusätzlicher Anreiz war, sich anzustrengen. Die

Ernennung des „Besten Bergsteigers“ erfolgte nach der Rückgabe der Protokolle (prozess- und produktbezogene Rückmeldung zur letzten Trainingsgeschichte).

Während die Wiederholung der bereits erlernten Strategien und die Partnerarbeit (Vertiefung der Überarbeitungsstrategien) wie in der Pilotstudie durchgeführt wurden, variierte die Ausdrucksinstruktion. Zum einen erfolgte die Einführung der Ausdrucksstrategien wie im vorherigen Abschnitt beschrieben anhand einer vollständigen, jedoch nicht ausgestalteten Geschichte (siehe CD-Anhang A7.2) und zum anderen wurde in der dritten Trainingseinheit nur die Ausdrucksstrategien am Anfang einer Geschichte vermittelt.

Zu Beginn der Ausdrucksinstruktion las die Trainerin eine vollständige, jedoch nicht ausgestaltete Geschichte vor. Danach wurden mit Hilfe der Ausdruckskarte Ausgestaltungsmöglichkeiten für die 7W-Fragen am Anfang einer Geschichte generiert. Zudem zeigte die Trainerin, wie die Ausdruckskarte beim Ausfüllen des Geschichtenplans hilfreich sein kann. Am Ende las die Trainerin den überarbeiteten Geschichtenanfang vor. Des Weiteren erhielt jeder Schüler wie in der Pilotstudie eine eigene Ausdruckskarte (siehe Abbildung 17).

Die restlichen Inhalte der dritten Trainingseinheit (Zielsetzung, Planen/Schreiben/Überarbeiten) inklusive der Einführung der Checkliste II wichen nicht von dem Ablauf der Pilotstudie ab.

Trainingseinheit 4 (Schwerpunkt: Ausdruck Hauptteil/Ende): Die vierte Trainingseinheit unterschied sich in mehreren Punkten von der Pilotstudie und wird deswegen ausführlicher beschrieben. Wie in der zweiten Trainingseinheit erfolgte nach der regulären Rückmeldung mit dem 7W- und Leistungsprotokoll die Berechnung der Punktedifferenz zwischen der Prätestgeschichte und der letzten Geschichte. Darüber hinaus wurde ein neuer „Bester Bergsteiger“ ernannt. In der anschließenden Wiederholungsphase wurden explizit die Ausdrucksstrategien des Geschichtenanfangs wiederholt.

Anschließend las die Trainerin die Geschichte zur Herleitung der Ausdrucksstrategien, welche bereits in der dritten Trainingseinheit eingesetzt wurde, vor (siehe CD-Anhang A7.2). Unter Einbezug der Schüler wurden nun die Ausdrucksstrategien für den Hauptteil und den Abschluss einer Geschichte erarbeitet und auf der Ausdruckskarte gezeigt. Die Trainerin erklärte, dass die Beschreibung von Gedanken und Gefühlen und die Verwendung von wörtlicher Rede wichtige Ausdrucksstrategien darstellen. Anhand der Ausdruckskarte modellierte die Trainerin den Einsatz der neu erlernten Strategien beim Planen, Schreiben und Überarbeiten einer Geschichte, insbesondere bei der Verwendung des Geschichtenplans. Am

Ende der Ausdrucksinstruktion las die Trainerin die Geschichte vom Anfang noch einmal mit kompletter Ausgestaltung vor.

Um die Ausdrucksstrategien zu verfestigen, wurde im Anschluss eine Partnerarbeit durchgeführt. Die Aufgabe umfasste das Ergänzen von Ausgestaltungselementen in einem Geschichtenplan und war mit der Pilotstudie identisch.

Schließlich setzten sich die Schüler nach der Gruppenarbeit Ziele im 7W- und im Leistungsprotokoll, wobei sie insbesondere die Verbesserung des Ausdrucks anstreben sollten. Daran anknüpfend planten, schrieben und überarbeiteten die Schüler eine Reizwortgeschichte (siehe CD-Anhang A3).

Trainingseinheit 5 (Schwerpunkt: Zusammenhang): Nach der Rückgabe der 7W- und Leistungsprotokolle, der Ernennung des neuen „Besten Bergsteigers“ und der Wiederholung der erlernten Strategien wurde dieselbe Partnerarbeit wie in der Pilotstudie (Sortierung von Satz-Schnipseln zu einer Geschichte) eingesetzt, um zur Herleitung der Zusammenhangsstrategien überzuleiten.

In der zusammengefügt Geschichte identifizierte die Trainerin mit den Schülern die Satzanfänge und die Handlungsschritte. Im Gegensatz zur Pilotstudie führte die Trainerin keine neue Checkliste ein, sondern modellierte, wie Satzanfänge und Verben einkreist und unterstrichen werden, um zu kontrollieren, ob der Zusammenhang in der Geschichte gelungen war. Als Merkhilfe führte die Trainerin jedoch wieder das Zusammenhangssymbol ein (siehe Abbildung 18). Zusätzlich wies die Trainerin auf die Ausdruckskarte hin, die Beispiele für verschiedene Satzanfänge enthält.

Analog zur Pilotstudie wurden direkt im Anschluss an die Instruktion die Zusammenhangsstrategien in der Gruppenarbeit vertieft. Die Schüler sollten wieder Sätze zu einer sinnvollen Geschichte sortieren. Als neue Zusammenhangsübung sollten die Schüler zusätzlich zu zweit unter Heranziehung der Ausdruckskarte Satzanfänge ergänzen (siehe CD-Anhang A5.1.1).

Die übrigen Trainingsinhalte (Zielsetzung, Planen/Schreiben/Überarbeiten) waren wieder mit der Pilotstudie identisch.

Trainingseinheit 6 (Schwerpunkt: Wiederholung): Ziel der letzten Trainingseinheit war es, alle erlernten Strategien noch einmal zu wiederholen und sämtliche Hilfestellungen auszublenden, damit die Schüler eigenständig eine Geschichte planen, schreiben und überarbeiten können.

Der Ablauf der Trainingseinheit war den Schülern bereits gut bekannt. Am Anfang erhielten sie ihre Mappen mit den Rückmeldungen zurück und berechneten die

Punktedifferenz zu der Prätestgeschichte. In der letzten Stunde sollte der größte Teil der Schüler den Klebepunkt „super Schreiber“ oder „Meister-Schreiber“ in ihr Leistungsprotokoll einkleben können. Weiter wurde zum letzten Mal der „Beste Bergsteiger“ ernannt. Wie in der Pilotstudie erfolgte anschließend die Wiederholungsphase anhand von drei Schülergeschichten. Es wurde jeweils eine Geschichte zum Inhalt, Ausdruck und Zusammenhang vorgelesen. Die Schüler identifizierten die jeweiligen Strategien in den Geschichten. Im Gegensatz zur Pilotstudie wurden jedoch nicht explizit Feedback- oder Zuhör-Regeln eingeführt.

Die Gruppen- bzw. Partnerarbeit war ebenfalls mit der Pilotstudie identisch. Zu zweit bearbeiteten die Schüler Tandemkarten (siehe CD-Anhang A5.1.1). Auch der weitere Ablauf der letzten Trainingseinheit war mit der Pilotstudie gleich:

- Zielsetzung;
- Planen, Schreiben und Überarbeiten einer Bildergeschichte ohne Hilfsmittel;
- Verteilung von persönlichen Briefen zur Motivierung der Schüler für die Posttestung (siehe CD-Anhang A5.1.3).

Die Geschichten wurden eingesammelt und nach dem Posttest den Schülern bewertet -zusammen mit allen anderen Geschichten- zurückgegeben. Überdies erhielt jede Klasse eine Urkunde, die den Schülern bescheinigte, dass alle zu Schreibexperten geworden sind (siehe CD-Anhang A5.1.3).

6.2.3.5 Writer's Workshop (WW)

Es wurde bereits erläutert, dass sich die beiden Trainingsbedingungen, USAT und WW, nur hinsichtlich der Selbstregulation unterschieden. Um die Zeit für Training, Übung und Instruktion in beiden Bedingungen konstant zu halten, wurden in der WW-Bedingung Übungen zur Förderung der Selbstständigkeit, der Schreibfreude und der Phantasie durchgeführt. Diese, die Selbstregulation ersetzenden Elemente, werden im Folgenden beschrieben. Anschließend werden die Inhalte der einzelnen Trainingseinheiten erläutert.

6.2.3.5.1. Bedingungsspezifische Elemente

Textlupe: Schüler der WW-Bedingung erhielten wie auch Schüler der USAT-Bedingung am Anfang jeder Trainingseinheit eine Rückmeldung zu der zuletzt geschriebenen Geschichte. Um die beiden Bedingungen voneinander abzugrenzen, wurde in der WW-Bedingung auf Methoden zurückgegriffen, die in Grundschulzeitschriften vorgeschlagen und in dem Praxisalltag bereits erprobt wurden. In vielfältigen Artikeln wurden die Schreibkonferenz (Käferle, 2007; Simon, 1997) und die Textlupe (Kruse, 2006; Fix, 1999;

Uerdingen, 2006) als angemessene Überarbeitungsstrategien in der Grundschule vorgestellt. Die Textlupe wurde als Bewertungsinstrument der Schüleraufsätze der WW-Bedingung gewählt, da sie einerseits keine selbstregulatorischen Komponenten enthält und andererseits eine kindgerechte und motivierende Rückmeldung ermöglicht. Sie diente dabei sowohl für die Trainerinnen als auch für die Schüler als Leitfaden zur Bewertung von Geschichten.

Wie in der Abbildung 34 ersichtlich besteht die Textlupe aus drei Spalten: „*Das hat mir gut gefallen!*“, „*Das fällt mir auf!*“ und „*Hier habe ich Vorschläge!*“. Um das Ausfüllen der Textlupe den Trainerinnen zu erleichtern und zu gewährleisten, dass keine prozess- oder produktbezogenen Rückmeldungen gegeben wurden, erhielten die Trainerinnen eine Schulung in dem Ausfüllen der Spalten sowie verschiedene Formulierungsbeispiele (siehe CD-Anhang A5.2.4).

Abbildung 34: Textlupe (WW-Bedingung).

In die erste Spalte der Textlupe, die etwas Gelungenes hervorhebt, wurde z.B. „*Ich konnte mir deine Geschichte gut vorstellen.*“ eingetragen. Beim zweiten Punkt, der einen Kritikpunkt widerspiegelt, bekamen die Schüler Anmerkungen wie z.B. „*Mir ist aufgefallen, dass du mit deiner Geschichte nicht fertig geworden bist.*“. In der letzten Spalte schließlich erhielten die Schüler Verbesserungstipps, wie z.B. „*Versuche das nächste Mal, deine Geschichte spannend zu schreiben.*“.

Phantasiereise: In der WW-Bedingung wurde die Förderung der Phantasie stark betont. Darüber hinaus sollten die Schüler Spaß am Schreiben und Hören von Geschichten gewinnen. Deswegen war die Phantasiereise ein wichtiger Bestandteil der WW-Bedingung. In Anlehnung an Schumacher (2001) und Dießner (2003) wurden drei Phantasiereisen (siehe

CD-Anhang A5.2.2) erarbeitet. Jeweils eine der Geschichten wurde in der zweiten, vierten und sechsten Trainingseinheit den Schülern vorgelesen. Wie ein roter Faden zog sich die Hauptperson „Balthasar“ mit seinem fliegenden Teppich durch alle Geschichten (siehe CD-Anhang A5.2.2).

Die Phantasiereisen fanden jeweils im Stuhlkreis statt. Nach einer Einleitung, in der die Schüler angewiesen wurden, sich zu entspannen und die Augen zu schließen, las die Trainerin die jeweilige Phantasiereise vor. Nach der Phantasiereise wurden alle Schüler langsam wieder zurückgeholt. Im Stuhlkreis wurden anschließend die Ideen und Vorstellungen der Schüler besprochen. So fragte die Trainerin z.B., wie sich die Schüler den Balthasar vorgestellt haben. In diesem Zusammenhang nahm die Trainerin außerdem Bezug zu den Geschichten der Schüler und wies diese an, beim Verfassen der eigenen Geschichte sich ebenfalls alles genau vorzustellen und die Phantasie zu verwenden. Nach dem Erfahrungsaustausch kehrten alle Schüler zu ihren Plätzen zurück. Das Thema der Phantasiereise wurde jedoch im weiteren Verlauf der Trainingseinheit wieder aufgenommen. In der zweiten Trainingseinheit sollten die Schüler zu zweit eine weitere Phantasiereise mit Balthasar planen. In der vierten Trainingseinheit diente die Phantasiereise als Schreibstimuli, so dass die Schüler eine Fortsetzung der Reise verfassten. In der sechsten Trainingseinheit schrieben die Schüler als Gruppe gemeinsam eine Balthasar-Geschichte.

6.2.3.5.2. Ablauf des Trainings

Umfang und Dauer des WW-Trainings waren mit dem USAT-Programm identisch, so dass es ebenfalls sechs Trainingseinheiten à 90 Minuten umfasste. Wie auch im USAT-Programm sammelten die Trainerinnen die wöchentlich verfassten Schüleraufsätze ein und bewerteten diese. Eine Übersicht der Trainingsinhalte der WW-Bedingung ist in Tabelle 15 vorzufinden. Im Folgenden werden nur die Unterschiede zur USAT-Bedingung aufgelistet.

Trainingseinheit 1 (Schwerpunkt Inhalt, Planen, Schreiben):

- Keine Rückmeldung mit dem 7W- und Leistungsprotokoll;
- Einführung der Textlupe und Verteilung der Textlupen: jeder Schüler erhielt eine Textlupe mit Feedback zu der Prätestgeschichte;
- Modellierung des Planens und Schreibens einer Geschichte ohne Geschichtenplan.

Trainingseinheit 2 (Schwerpunkt: Überarbeiten):

- Vorstellung von Balthasar und Durchführung der ersten Phantasiereise (siehe CD-Anhang A5.2.2);
- Einführung der Gruppenarbeit: Planen einer Phantasiereise mit Balthasar (siehe CD-Anhang A5.2.1);

- Modellierung des Überarbeitens einer Geschichte ohne Checkliste.

Trainingseinheit 3 (Schwerpunkt: Ausdruck):

- Einführung des Buttons „Größter Schreibfan“ (Abbildung 30) für den Schüler mit der schönsten Geschichte;
- Durchführung der Gruppenarbeit „Fremder Planet“ (siehe CD-Anhang A5.2.1): Generierung eines Geschichtenanfangs und -hauptteils zu einem vorgegebenen Geschichtenabschluss und Vorlesen der Geschichten;
- Vermittlung sämtlicher Ausdrucksstrategien (Anfang, Hauptteil, Abschluss);
- Keine Modellierung der Anwendung der Ausdruckskarte beim Planen, Schreiben und Überarbeiten;
- Durchführung einer Gruppenarbeit zum Thema Ausdruck in Anlehnung an Menauer (2006): Die Schüler einer Gruppe sollten zu vorgegebenen Begriffen passende Verben, Adjektive und bedeutungsähnliche Wörter suchen und in eine Tabelle eintragen (siehe CD-Anhang A5.2.1).

Trainingseinheit 4 (Schwerpunkt: Phantasiereise, Wiederholung):

- Durchführung einer Gruppenarbeit zum Thema Überarbeiten: Schüler bewerteten zu zweit eine vorgegebene Geschichte mit der Textlupe (siehe CD-Anhang A5.2.1);
- Durchführung der zweiten Phantasiereise mit Balthasar (siehe CD-Anhang A5.2.2);
- Generierung von weiteren Balthasar-Abenteuern mit Schwerpunkt auf phantasievolle und bildliche Beschreibungen;
- Planen, Schreiben und Überarbeiten eines selbst erfundenen Balthasar-Abenteuers;
- Tischnachbarn tauschten Geschichten aus und bewerteten diese mit der Textlupe.

Trainingseinheit 5 (Schwerpunkt: Zusammenhang):

- Herleitung der Zusammenhangsstrategien, ohne die Handlungsschritte zu unterstreichen und die Satzanfänge einzukreisen;
- Durchführung einer Gruppenarbeit zum Thema Zusammenhang: Zu vorgegebenen Handlungen sollten die Schüler zu zweit folgende Fragen stichpunktartig beantworten: 1. *Was brauche ich?*, 2. *Warum brauche ich es?*, 3. *Was muss ich zuerst machen?*, 4. *Was muss ich am Ende machen?* (siehe CD-Anhang A5.2.1). Am Ende erzählten sich die einzelnen Gruppenmitglieder gegenseitig die unterschiedlichen Handlungen, die sie vorbereitet hatten.

Trainingseinheit 6 (Schwerpunkt: Wiederholung):

- Durchführung einer Phantasiereise mit Balthasar (siehe CD-Anhang A5.2.2);
- Durchführung einer Gruppenarbeit: Verfassen einer Balthasar-Geschichte in der Gruppe. Jedes Gruppenmitglied schrieb einen Satz, welcher in der Gruppe diskutiert wurde. Anschließend wurde die Geschichte weitergereicht. Am Ende lasen die Gruppen ihre Geschichten in der Klasse vor.
- Schüler tauschten Geschichten mit ihren Tischnachbarn und bewerteten diese mit Hilfe der Textlupe;
- Verteilung von persönlichen Briefen zur Motivierung für den Posttest (siehe CD-Anhang A5.2.4). Die Anmerkungen entsprachen den Formulierungen der Textlupe, d.h. fokussierten die Schreibfreude, Phantasie und Selbständigkeit, und nahmen keinen Bezug zu den Schreibstrategien.

Zum Schluss erhielt die Klasse eine Urkunde von der Trainerin, die allen Schülern zertifizierte, dass sie „Größte Schreibfans“ geworden waren.

6.2.4 Genauigkeit der Implementation

Um zu gewährleisten, dass jede Trainingseinheit von jeder der vier Trainerinnen so wie geplant durchgeführt wurde und dass etwaige Unterschiede in den abhängigen Variablen auf Unterschiede zwischen den beiden Bedingungen zurückzuführen sind, wurden verschiedene Maßnahmen zur Sicherung der Implementationsgenauigkeit ergriffen. Zu diesen Maßnahmen zählen die Schulung der Trainerinnen, die Beobachtung der Trainingseinheiten durch Tutorinnen, die Erfassung der inhaltlichen Vollständigkeit durch Checklisten und die Befragung der Lehrerinnen.

Schulung und Betreuung der Trainerinnen: Da die vier an der Studie I beteiligten Trainerinnen bereits aktiv an der Vorbereitung und Durchführung der Pilotstudie beteiligt waren, besaßen sie ausreichende Erfahrungen bezüglich des USAT-Programms. Nichtsdestotrotz war eine Schulung der Trainerinnen unmittelbar vor Durchführung der Studie I notwendig, da das USAT-Programm nach der Pilotstudie etwas verändert wurde und die Trainerinnen in der Durchführung der WW-Bedingung eingewiesen werden mussten.

An einem Wochenende wurden den Trainerinnen sowie den Tutorinnen im Rahmen eines Workshops das Design und die Inhalte der Studie I erklärt. Es wurde insbesondere auf die Unterschiede zwischen den beiden Bedingungen hingewiesen. Dazu erhielten die Trainerinnen eine Gegenüberstellung der beiden Bedingungen (siehe CD-Anhang A5.3.3), in der die wichtigsten inhaltlichen Unterschiede aufgeführt waren. Ziel des Workshops war es zudem, die Trainerinnen für Variationen beim Loben und in den verbalen Äußerungen in den

beiden Bedingungen zu sensibilisieren. Darüber hinaus wurde in dem Workshop die erste Trainingseinheit besprochen und von einer Trainerin modelliert. Abgesehen davon beschäftigte sich der Workshop inhaltlich mit der Durchführung der Testungen. Jede Trainerin führte, begleitet von der Tutorin, die Testungen in den von ihr betreuten Klassen durch. Die Durchführung der Testung wurde im Workshop von der Doktorandin modelliert. Jede Trainerin erhielt zudem ein ausformuliertes Testungsmanual (siehe CD-Anhang A2). Weiter wurden die einzelnen Testinstrumente vorgestellt. Am Ende fühlte sich jede Trainerin ausreichend auf die Durchführung der Testung vorbereitet.

Nach dem Workshop fanden jede Woche Einzeltreffen der Trainerinnen mit der Doktorandin statt. Es wurde dabei über die durchgeführte Trainingseinheit reflektiert. Die von den Trainerinnen berichteten Probleme betrafen größtenteils die Einhaltung der zeitlichen Vorgaben, Unterrichtsstörungen und die Förderung besonders schwacher Schüler. In den Nachtreffen wurde deswegen gemeinsam nach Lösungswegen gesucht. Durch die wöchentlichen Treffen war die Doktorandin kontinuierlich über den von den Trainern subjektiv empfundenen Erfolg der Implementierung informiert.

Außerdem dienten die Nachtreffen zur Vorbereitung auf die nächste Trainingseinheit. Die jeweilige Trainerin modellierte die wichtigsten Instruktionen der beiden Bedingungen. Im Anschluss bekam sie Verbesserungsvorschläge und Hinweise, auf die sie in dem Training besonders achten sollte. Die Hinweise betrafen insbesondere die Differenzierung zwischen den beiden Bedingungen.

Darüber hinaus erhielten die Trainerinnen und Tutorinnen ausführliche Manuale für jede Trainingssitzung sowohl für die USAT- (siehe CD-Anhang A5.1.2) als auch für die WW-Bedingung (siehe CD-Anhang A5.2.3). Auf der ersten Seite von jedem Manual ist eine Material-Checkliste aufgelistet. Getrennt für Trainerin und Schüler konnte dort die Vollständigkeit der Materialien überprüft sowie kontrolliert werden, ob alle Vorbereitungen für die Durchführung getroffen wurden. Zudem stand den Trainerinnen ein Ordner zur Verfügung, der sortiert nach Trainingseinheiten Kopiervorlagen sowie Folien und sonstige Materialien (z.B. Trickschilder, Klebepunkte, Bergsteiger-Button) enthielt.

Darüber hinaus waren am Anfang der Manuale für die jeweilige Trainingsstunde wichtige Durchführungshinweise und die Ziele der Trainingseinheiten aufgeführt. Weiter konnten die Trainerinnen nachlesen, welche Vorbereitungen für die Trainingseinheit notwendig waren. Außerdem umfasste das Manual eine komplett ausformulierte Anleitung, welche von den Trainerinnen teilweise wörtlich übernommen wurde. Dick gedruckte Sätze sollten sogar explizit wortgetreu übermittelt werden, da diese für die jeweilige Bedingung

zentral waren. Darüber hinaus enthielten die Manuale explizite Zeitangaben für die Durchführung der Trainingseinheiten, welche auf Zeitmessungen während der Vorarbeiten und der Pilotstudie basierten. Auf diese Weise konnten die für jede Trainingssitzung zur Verfügung stehenden 90 Minuten effizient genutzt werden.

Trainer-Checklisten: Direkt im Anschluss an die jeweiligen Trainingseinheiten füllten die Trainerinnen eine Checkliste aus (siehe CD-Anhang A5.3.2). Die Trainer-Checklisten dienten der Überprüfung der inhaltlichen Vollständigkeit der Implementierung. Die Checkliste enthielt stichpunktartig alle wesentlichen Inhalte der jeweiligen Trainingseinheiten. Die Trainerin hakte die Inhalte, die sie in der Stunde durchgeführt hatte, ab und bewertete zudem die Qualität der Durchführung mit 1, d.h. *guter Ablauf*, 0, d.h. *mittelmäßiger Ablauf* oder mit -1, d.h. *schlechter Ablauf*. Darüber hinaus konnten sich die Trainerinnen Anmerkungen zu den einzelnen Inhalten sowie zu der Trainingseinheit insgesamt machen. Die Trainer-Checkliste diente als Grundlage für die wöchentlichen Reflexionen mit der Doktorandin.

Insgesamt wurden von allen Trainerinnen 100% der geplanten Inhalte in den beiden Bedingungen realisiert. Keine Trainerin bewertete ihre Implementation mit -1, d.h. die Trainerinnen waren subjektiv mit der Durchführung der Trainings zufrieden.

Beobachterschulung: Vor Beginn der Trainings wurden die vier Tutorinnen in der Durchführung der Beobachtungen geschult. Eine Tutorin beobachtete eine Trainerin während sämtlicher Trainingseinheiten und in beiden Bedingungen. Aus organisatorischen Gründen war es nicht möglich, die Beobachter randomisiert den Trainingseinheiten und den Trainerinnen zuzuordnen.

Wie oben beschrieben, fand vor den Trainings eine Workshop-Schulung für Trainerinnen und Tutorinnen statt. Während die Trainerinnen die Trainingseinheiten modellierten, erprobten die Tutorinnen das Ausfüllen des Beobachtungsbogens (siehe CD-Anhang A5.3.1). Im Anschluss an die Modellierung verglichen die Tutorinnen ihre Einschätzungen, so dass sie auf diese Weise einen gemeinsamen Bewertungsmaßstab entwickelten. Darüber hinaus konnten während des Workshops Unklarheiten besprochen und Durchführungshinweise gegeben werden.

Zusätzlich fanden vor jeder Trainingseinheit separate Treffen mit den Tutorinnen statt, in denen die wichtigsten Inhalte sowie die Besonderheiten der jeweiligen Trainingseinheit besprochen wurden. Die Tutorinnen bekamen, wie die Trainerinnen auch, die ausformulierten Trainingsmanuale. Der Bedingungscheck (siehe CD-Anhang A5.3.3), welcher ebenfalls von

den Tutorinnen ausgefüllt wurde und zur Differenzierung zwischen den beiden Bedingungen diente, wurde ebenfalls vor Trainingsbeginn mit den Tutorinnen erprobt und diskutiert.

Unterrichtsbeobachtung: Ziel der Beobachtungen war die Überprüfung der Vollständigkeit und der Qualität der Implementierung der beiden Trainingsprogramme. Sämtliche Trainingseinheiten in beiden Bedingungen wurden von den Tutorinnen anhand von standardisierten Beobachtungsbögen (siehe CD-Anhang A5.3.1) eingeschätzt. Um eine differenzierte Einschätzung der Trainingsstunde zu gewährleisten, wurde jede Stunde in vier Einheiten eingeteilt, die von den Tutorinnen separat bewertet wurden. Inhaltlich ähnelten sich die Einheiten über alle Trainingsstunden hinweg. In der Regel umfassten die Beobachtungsbögen die Einschätzung folgender vier Unterrichtsschwerpunkte: 1. Rückmeldung und Zielsetzung, 2. Partnerarbeit, 3. Strategieinstruktion, 4. Anleitung zum Planen, Schreiben und Überarbeiten. Für jeden der vier Unterrichtsschwerpunkte schätzte die Tutorin die Qualität der Implementierung anhand von vier Faktoren separat ein. Diese vier Faktoren basieren auf den von Helmke (2009) postulierten Merkmalen guten Unterrichts und wurden jeweils auf einer fünfstufigen Skala von -2 (*sehr schlechte Umsetzung*) bis +2 (*sehr gute Umsetzung*) bewertet:

- Unter *Klarheit der Instruktion* sind folgende Eigenschaften gefasst: kurze, direkte und prägnante Ausdrucksweise, Betonung der Wichtigkeit bestimmter Aspekte, Überwachung der Verständlichkeit der Instruktion, Abstimmung der verbalen Instruktion mit den Visualisierungen und den Materialien.
- Mit *Individualisierung* sind die Anpassung der Erklärungen und Beispiele, das Eingehen auf Fragen und Verständnisprobleme, die Variation von Fragen sowie die Anpassung der verbalen Unterstützung auf individuelle Lernvoraussetzungen gemeint.
- *Motivierung* bedeutet folgendes: Blickkontakt, freundlich, ermutigend, anerkennend, Berücksichtigung möglichst vieler Kinder.
- Schließlich gehören zu dem Faktor *Klassenführung* Eigenschaften wie der effektive und ökonomische Umgang mit Störungen und Unterbrechungen, das Bemerkens und Signalisieren von Störungen, Nutzung der Trainingszeit für fachliche Ziele und Minimierung des Zeitaufwandes für außerfachliche Angelegenheiten.

Zusätzlich hatten die Tutorinnen die Möglichkeit, sich auf dem Beobachtungsbogen Notizen zu den vier Trainingsschwerpunkten zu machen. Zu den Notizen zählten u.a. Verbesserungsvorschläge oder Auffälligkeiten einzelner Schüler, die die Trainerin nicht

bemerkt hatte. Darüber hinaus machten die Tutorinnen die Trainerinnen bereits während des Trainings darauf aufmerksam, wenn Inhalte vergessen oder falsch umgesetzt wurden. Aus diesem Grund wurde die inhaltliche Vollständigkeit durch die Tutorinnen garantiert.

Darüber hinaus besprachen die Tutorinnen im Anschluss an das Training die Bedingungschecks und die Beobachtungsbögen mit der jeweiligen Trainerin. Auf diese Weise erhielten die Trainerinnen ein direktes Feedback zu der Qualität der Implementierung und erhielten zudem Hinweise für die Durchführung der anderen Trainingseinheiten. Außerdem wurden Verbesserungsvorschläge, die sich insbesondere auf die Motivierung der Schüler und die Vermeidung von Störungen bezogen, diskutiert.

Ergebnisse der Unterrichtsbeobachtungen: Die Ergebnisse der Beobachtungen zeigen, dass über alle Trainingszeitpunkte hinweg in beiden Bedingungen die Umsetzungen entweder als sehr gut oder als gut von den Tutorinnen beurteilt wurden. Eine schlechte Bewertung wurde kein einziges Mal vergeben. Da die Varianz der Antworten somit nahezu Null war, konnten keine Reliabilitäten als Quotient wahre durch beobachtete Varianz der Skalen berechnet werden.

Tabelle 16 zeigt die prozentualen Anteile der sehr guten Bewertungen für die vier Trainerinnen und die vier Faktoren getrennt nach Bedingung. Alle übrigen Trainingssitzungen wurden mit gut bewertet. Die prozentualen Anteile an sehr guten Bewertungen variierten nur minimal in den Bedingungen und zwischen den Trainerinnen.

Tabelle 16: Angabe der Anteile sehr guter Umsetzungen getrennt nach Faktoren, Trainer und Bedingung (Studie I).

	Klarheit		Individualisierung		Motivierung		Klassenführung	
	USAT	WW	USAT	WW	USAT	WW	USAT	WW
Trainerin 1	83.3%	83.3%	83.3%	83.3%	83.3%	100%	100%	100%
Trainerin 2	83.3%	100%	100%	100%	100%	100%	100%	66.7%
Trainerin 3	100%	100%	100%	100%	100%	100%	100%	100%
Trainerin 4	100%	100%	100%	100%	100%	100%	100%	83.3%

Da die Tutorinnen die Klasse und nicht einzelne Schüler beobachteten, wurden die nachfolgenden Analysen auf Klassenebene gerechnet. Für jeden der vier Faktoren Klarheit, Individualisierung, Motivierung und Klassenführung wurde eine Messwiederholungsanalyse durchgeführt.

Die Messwiederholungsanalysen mit den jeweiligen Faktoren als abhängige Variable, Messzeitpunkt (TE 1 bis TE6) als Innersubjektfaktor und Bedingung (USAT, WW) als Zwischensubjektfaktor lieferten für keinen Faktor einen signifikanten Bedingungseffekt ($p > .21$). Ferner ergaben Messwiederholungsanalysen mit den jeweiligen Faktoren als abhängige Variable, Messzeitpunkt (TE 1 bis TE6) als Innersubjektfaktor und Trainer (Trainerin 1 bis Trainerin 4) als Zwischensubjektfaktor für keinen Faktor einen signifikanten Trainereffekt ($p > .21$). Deswegen ist der Schluss zulässig, dass die vier Trainerinnen in den beiden Bedingungen hinsichtlich der vier Faktoren das Training gleich gut umsetzten.

Bedingungscheck: Abgesehen von den Beobachtungsbögen füllten die Tutorinnen zusätzlich für jede der beiden Bedingungen einen Bedingungscheck aus (siehe CD-Anhang A5.3.3), welcher der Überprüfung der Programmdifferenzierung diente. Gemäß O'Donnell (2008) ist dies ein prozessbezogenes Kriterium der Implementationsgenauigkeit.

Die Tutorinnen untersuchten mit Hilfe des Bedingungschecks, ob die Trainerinnen durch verbale Äußerungen zwischen den beiden Bedingungen adäquat differenzierten. Zum Beispiel sollten in der USAT-Bedingung Sätze wie „*Wenn ihr euch anstrengt und die 7W-Fragen in euren Geschichten verwendet, werdet ihr immer bessere Schreiber.*“ fallen. Sobald solche Aussagen von der Trainerin getroffen wurden, kreuzte die Tutorin dies in dem Bedingungscheck für die USAT-Bedingung ab. Dagegen sollten in der WW-Bedingung diese Sätze nicht von der Trainerin verwendet werden, was die Tutorin auf dem Bedingungscheck der WW-Bedingung markierte.

Falls die Trainerin in der USAT-Bedingung Aussagen vergessen hatte, erinnerte die Tutorin sie bereits während des Trainings daran, so dass diese von der Trainerin ergänzt wurden. Bezüglich der WW-Bedingung kam es in keiner Trainingseinheit und bei keiner Trainerin zu Äußerungen, die der USAT-Bedingung vorbehalten waren. Insgesamt lässt sich nach Durchsicht aller Bedingungschecks schließen, dass alle Trainerinnen sehr gut zwischen den beiden Bedingungen differenzierten.

Befragung der Lehrerinnen: Nach Abschluss der Trainings wurde mit jeder der acht Lehrerinnen ein strukturiertes Interview durchgeführt. Da in den beiden Bedingungen unterschiedliche Methoden und Materialien eingesetzt wurden, variierten die Inhalte der Interviews in den beiden Bedingungen. Die Anzahl der Items sowie die erfragten Faktoren waren jedoch identisch (siehe CD-Anhang A5.3.4). Die Lehrerinnen beider Bedingungen waren während der gesamten Trainings im Klassenraum anwesend und konnten sich dadurch ein eigenes Bild über die beiden Trainingsprogramme sowie über die subjektiv empfundene Effektivität des Programms machen.

Ziel der Lehrerbefragung war die Erfassung der Meinungen der Lehrerinnen bezüglich der Trainings und insbesondere die Einschätzungen des Inhalts, der Arbeitsmaterialien, der Lehr-/ Lernmethoden und der Neuheit für den Deutschunterricht in der Grundschule. Mit der Lehrerbefragung sollten hilfreiche Erkenntnisse für die Konzeption der Studie II gewonnen werden. Darüber hinaus betonen Deshler und Schumaker (1994) die Wichtigkeit der Erfassung der sozialen Validität einer Untersuchung. Dazu zählen die Autoren die Einschätzung der Nützlichkeit und Praktikabilität der Intervention.

Jedem der vier Faktoren Inhalt, Materialien, Methoden und Neuheit wurden drei Items zugeordnet, die die Lehrerinnen auf einer fünfstufigen Skala von 1 *gar nicht geeignet*, 2 *wenig geeignet*, 3 *etwas geeignet*, 4 *gut geeignet* und 5 *sehr gut geeignet* einschätzten. Für jeden Faktor ergab sich somit eine Maximalpunktzahl von 15. Die Reliabilitäten (Cronbach's α) der Skalen sind als mittelmäßig bis gut zu beurteilen: Inhalt: $\alpha = .64$, Materialien: $\alpha = .79$, Methoden: $\alpha = .75$, Neuheit: $\alpha = .90$. Die teilweise geringen Reliabilitäten sind auf die geringe Anzahl von Items zurückzuführen.

Tabelle 17: Mittelwerte, Standardabweichung und T-Werte für Inhalt, Materialien, Methoden und Neuheit getrennt nach Bedingung (Studie I).

	Inhalt		Materialien		Methoden		Neuheit	
	M	SD	M	SD	M	SD	M	SD
USAT-Gruppe	14.25	0.96	13.25	2.22	14.00	1.41	13.00	2.16
WW-Gruppe	12.75	0.96	13.00	1.41	13.00	1.83	10.25	4.86
T-Wert	2.22		0.19		0.87		1.04	

Die Daten wurden mittels T-Tests für unabhängige Stichproben ausgewertet. Alle in Tabelle 17 aufgeführten T-Werte sind nicht signifikant ($ps > .07$), so dass geschlussfolgert werden kann, dass die Lehrerinnen der USAT- und der WW-Bedingung hinsichtlich des Inhalts, der Materialien, der Methoden und der Neuheit die jeweiligen Trainings als gleich gut geeignet einschätzten. Insgesamt schätzten alle Lehrerinnen in beiden Bedingungen die Trainings hinsichtlich der erfassten Kriterien als sehr nützlich ein. Tendenziell waren die Bewertungen des USAT-Programms jedoch besser als die des WW-Programms.

Schließlich sollten die Lehrerinnen ein Gesamturteil für die Trainings abgeben. Im Mittel bewerteten sowohl die USAT- als auch die WW-Lehrerinnen ihr jeweiliges Training mit der Note 1.75, wobei 1 sehr gut und 6 ungenügend bedeutete.

6.2.5 Testungsmaterial

Zu jedem der drei Messzeitpunkte (Prätest, Posttest, Follow-Up) wurden dieselben abhängigen Maße erhoben. Jede Testung wurde im gesamten Klassenkontext von der jeweiligen Trainerin, die zuvor in der Durchführung der Testungen geschult und von einer Tutorin begleitet wurde, durchgeführt. Jeder Schüler arbeitete alleine an den einzelnen Aufgaben und durfte dabei keine Hilfsmittel einsetzen.

Schreibanlass: Als Schreibanlass dienten drei Bilderserien (siehe CD-Anhang A2), die aus jeweils vier zusammenhängenden Bildern bestehen und aus der Serie *Vater und Sohn* von Plauen (2003) stammen. Um mögliche Unterschiede in der Schwierigkeit der Schreibstimuli über alle Klassen hinweg auszugleichen, wurden die Bilderserien den Schülern so zugeordnet, dass zu jedem Messzeitpunkt ein gleicher Anteil von Schülern innerhalb einer Klasse die Bildergeschichte 1, 2 und 3 bearbeitete und dass jeder Schüler eine unterschiedliche Bildergeschichte zu jedem Messzeitpunkt erhielt. Folglich wurden in jeder Klasse alle drei Bilderserien zu allen drei Messzeitpunkten bearbeitet, wobei jeder Schüler sich mit denselben drei Bildergeschichten, jedoch zu unterschiedlichen Zeitpunkten beschäftigte. Die drei Bilderserien wurden bereits in der Pilotstudie erprobt.

Ablauf der Testung: Jede Testung dauerte 90 Minuten (eine Doppelstunde). Der Ablauf war in beiden Bedingungen identisch. Das ausführliche Testungsmanual ist im Anhang (siehe CD-Anhang A2) zu finden.

In der ersten Schulstunde (45 Minuten) wurden die Bildergeschichten verfasst. Nach einer kurzen Einführung und Einleitung wurden die Schüler aufgefordert, abgesehen von einem Stift, sämtliche Materialien von den Tischen zu entfernen. Anschließend wurden die Bildergeschichten und Schreibblätter verdeckt ausgeteilt. Den Schülern wurde erklärt, dass sie in den nächsten zehn Minuten die Geschichte planen, dann zwanzig Minuten schreiben und schließlich zehn Minuten überarbeiten sollten. Beim Planen wurden die Schüler darauf hingewiesen, sich auf einem leeren Blatt Notizen zu den Bildern zu machen, anhand derer sie später die Geschichte schreiben sollten. Beim Schreiben wiesen die Trainerinnen die Schüler an, sich anhand der Notizen, eine gute Geschichte zu überlegen, wobei sie darauf achten sollten, eine Reihe freizulassen. Wenn Schüler weniger als zwanzig Minuten für das Schreiben benötigten, erhielten sie ein Mandala. Auf diese Weise wurden Unruhen in der Klasse vermieden. Wenn Schüler in den zwanzig Minuten nicht mit dem Schreiben der Bildergeschichte fertig wurden, konnten sie an der Geschichte in der Überarbeitungsphase weiter schreiben, sobald sie ihren Text fertig überarbeitet hatten. Vor dem Überarbeiten wurde den Schülern erklärt, dass sie mit einem bunten Stift die Geschichte überarbeiten und

dabei Ergänzungen in die freien Zeilen notieren sollten. War ein Schüler vor Ablauf der zehn Minuten mit dem Überarbeiten fertig, beschäftigte sich der Schüler mit einem Mandala. Folglich arbeiteten alle Schüler in beiden Bedingungen gleich lang an der Bildergeschichte.

In der zweiten Stunde wurde der Wissenstest (siehe CD-Anhang A2) durchgeführt. Dazu las die Trainerin sämtliche Aufgaben des Fragebogens vor. Hinterher wurden Verständnisfragen der Schüler geklärt. Schließlich wurde den Schülern zehn Minuten Zeit zur Beantwortung der Fragen eingeräumt. Der Herkunftsfragebogen wurde im Anschluss von den Schülern ausgefüllt.

Als Transferaufgabe schrieben die Schüler eine Erlebniserzählung. Um die Ergebnisse nicht durch Ermüdungserscheinungen oder Motivationsverluste zu verzerren, fand die Erhebung der Erlebnisgeschichten einen Tag nach der Testung statt und wurde von der jeweiligen Klassenlehrerin durchgeführt. Die Lehrerinnen erhielten ausformulierte Anleitungen (siehe CD-Anhang A2) und wurden explizit darum gebeten, sich an diese zu halten sowie die vorgegebenen Zeitangaben nicht zu überschreiten. Die Instruktion ähnelte der für die Bildergeschichte. Die Schüler erhielten zehn Minuten Zeit für das Planen, zwanzig Minuten für das Schreiben und zehn Minuten für das Überarbeiten der Erlebnisgeschichte.

6.2.5.1 Abhängige Variablen

Als nächstes werden die im Einzelnen erhobenen abhängigen Maße ausführlich dargestellt.

Wortanzahl: Zur Beurteilung des Textumfangs wurden die geschriebenen Wörter unabhängig von der korrekten Anwendung der Rechtschreib- und Grammatikregeln gezählt.

Strategienahe Maße: Strategienahe Maße beziehen sich auf die im Training instruierten Strategien. In der Studie I wurden folgende strategienahen Maße erhoben: Geschichtenelemente als strategienahes Schreibleistungsmaß, Planungs- und Überarbeitungsaktivitäten sowie Wissen.

Geschichtenelemente: Wie auch in der Pilotstudie wurden die Geschichten in Anlehnung an die Studien von Glaser und Kollegen (Glaser, 2005; Glaser & Brunstein, 2007a; Glaser & Brunstein, 2007b; Glaser, Keßler, & Brunstein, 2009) sowie an die *Story Grammar Scale* von Harris und Graham (1996) analysiert.

Die Geschichten wurden hinsichtlich der Nennung (sieben Punkte) und der Ausgestaltung (sieben Punkte) der 7W-Fragen ausgewertet, so dass sich eine Maximalpunktzahl von 14 ergab. Eine ausführliche Beschreibung der Bewertungskriterien ist in Tabelle 18 aufgeführt.

Tabelle 18: Bewertungskriterien der Geschichtenelemente.

	Inhalt	Ausdruck
Person	Name der Hauptperson wurde genannt.	Eigenschaften und Aussehen der Hauptperson wurden beschrieben.
Ziel	Ziel der Hauptperson wurde unter Verwendung von Wörtern wie „will“, „möchte“, „nimmt sich vor“ genannt.	Ziel der Hauptperson wurde durch die Beschreibung der Wichtigkeit des Ziels für die Hauptperson begründet.
Zeit	Zeitpunkt wurde genannt.	Zeitpunkt wurde durch anschauliche Beschreibungen und Adjektive ausgestaltet.
Ort	Ort wurde genannt.	Ort wurde durch anschauliche Beschreibungen und Adjektiven ausgestaltet.
Handlungsschritte	Mehrere Handlungsschritte, die die Handlung der Hauptperson nachvollziehen lassen, wurden beschrieben.	Handlungsschritte wurden durch Adjektive, abwechslungsreiche Verben, Gedanken und Gefühle ausgestaltet.
Höhepunkt	Höhepunkt wurde aus Sicht der Hauptperson beschrieben.	Gedanken und Gefühle der Hauptperson beim Höhepunkt wurden beschrieben.
Ende	Ende wurde aus Sicht der Hauptperson beschrieben.	Gedanken und Gefühle der Hauptperson am Ende wurden beschrieben.

Planungsaktivitäten: Vor dem Schreiben der Geschichte bekamen die Schüler zehn Minuten Zeit für das Planen der Geschichte. Die Schüler machten sich Stichpunkte auf einem separaten Blatt. Die Qualität der Planungsnotizen wurde bezüglich der Nennung und der bildlichen Beschreibung der Geschichtenelemente mit jeweils sieben Punkten bewertet. Für jedes richtig genannte Element sowie für jede Ausgestaltung einer 7W-Frage wurde ein Punkt vergeben, so dass sich ebenfalls eine 14-stufige Skala ergab, die bereits zur Bewertung der Bildergeschichten herangezogen wurde (siehe Tabelle 18).

Überdies wurde die Struktur der Planungsnotizen anhand von drei Variablen analysiert. Zunächst wurde der Umfang der geplanten Handlungsabfolge untersucht. Ein Punkt wurde vergeben, wenn mindestens drei Handlungsschritte geplant wurden. Ab vier Handlungsschritten wurden zwei Punkte vergeben. Danach bewerteten die Beurteiler die Reihenfolge der geplanten Handlungsschritte. Wurden die Handlungsschritte bis auf wenige Ausnahmen in der richtigen Reihenfolge aufgelistet, wurde ein Punkt vergeben. Wenn die Reihenfolge komplett richtig war, wurden die Notizen mit zwei Punkten bewertet. Als letztes schätzten die Beurteiler die Gesamtstruktur der Stichpunkte ein. Dabei wurden sämtliche

Stichpunkte zum Anfang, Hauptteil und Ende der Geschichte mit einbezogen. Ähnlich zu den vorherigen Variablen wurde ein Punkt für eine erkennbare, aber nicht vollständige Struktur und zwei Punkte für eine komplett vollständige Struktur verteilt. Ein gesonderter Punkt wurde vergeben, wenn die 7W-Fragen explizit genannt wurden. In der Gesamtheit ergab sich eine Maximalpunktzahl von sieben.

Revisionsprozesse: Nach dem Verfassen der Geschichte hatten die Schüler zehn Minuten Zeit, die Geschichte zu überarbeiten. Die Einfügungen wurden hinsichtlich Verbesserungen des Inhalts und des Ausdrucks ausgewertet, wobei nur den Text verbessernde Einfügungen gezählt wurden. Unter Verbesserungen des Inhalts/Ausdrucks zählten Einfügungen der 7W-Fragen sowie stilistischer Mittel zur Beschreibung der 7W-Fragen.

Schreibbezogenes Wissen: Das schreibbezogene Wissen wurde mittels eines Fragebogens erfasst (siehe CD-Anhang A2), welcher bereits von Glaser und Brunstein (2007b) eingesetzt wurde. Der Fragebogen bezog sich explizit auf die in der USAT- und WW-Bedingung vermittelten Strategien und enthielt folgende vier Fragen:

- (1) Was sollte im ersten Teil einer Geschichte beschrieben werden?
- (2) Was sollte im zweiten Teil einer Geschichte beschrieben werden?
- (3) Was sollte im letzten Teil einer Geschichte beschrieben werden?
- (4) Wie kannst Du eine Geschichte spannend schreiben?

Jede der vier Fragen wurde einzeln quantitativ ausgewertet. Für jede richtig genannte Strategie erhielten die Schüler einen Punkt. Die Gesamtpunktzahl des Fragebogens ergab sich durch Addition der Punkte über die vier Fragen hinweg.

Strategieferne Maße: Im Gegensatz zu strategienahen Maßen erfassen die strategiefernen Maße Kompetenzen, die nicht direkt während des Trainings instruiert und erworben wurden, sondern einen Anforderungstransfer erfordern. Als strategieferne Maße wurden in der Studie I die Narrative Qualität als strategiefernes Schreibmaß und die Erlebnisgeschichten erfasst.

Narrative Qualität/Geschichtenqualität: Abgesehen von den Geschichtelementen wurden die Bildergeschichten anhand eines holistischen Maßes, der Narrativen Qualität, ausgewertet. Die sechsstufige Skala der Narrativen Qualität basiert auf der von Gentile (1992) entwickelten *National Scoring Guide* und wurde bereits in ähnlicher Form in der Pilotstudie eingesetzt. Die Geschichtenqualität und die Narrative Qualität werden im Folgenden äquivalent behandelt. Für jede Stufe wurden Kriterien festgelegt, die erfüllt sein mussten, um die jeweilige Punktstufe zu erreichen (siehe Tabelle 19). Überdies wurde für jede Stufe eine prototypische Bildergeschichte als Vergleichsmaßstab den Beurteilern zur Verfügung gestellt.

Tabelle 19: Bewertungskriterien der Geschichtenqualität (Studie I).

Punktzahl	Ausarbeitung der Geschichte	Bewertungskriterien
1	Auflistung von Ereignissen	<ul style="list-style-type: none"> ▪ Aneinanderreihung mehrerer zusammenhangloser Sätze
2	Wenig ausgearbeitete Geschichte	<ul style="list-style-type: none"> ▪ Auflistung mehrerer Sätze ▪ Ansatzweise inhaltlich vollständig ▪ Wenig sprachliche Ausgestaltung
3	Grundstruktur	<ul style="list-style-type: none"> ▪ Grundstruktur (Anfang, Hauptteil, Abschluss) vorhanden ▪ Teilweise inhaltlich vollständig ▪ Kaum Ausgestaltung vorhanden ▪ Oder: Anfang und Hauptteil vorhanden (aus Zeitgründen wurde das Ende nicht geschafft) jedoch überwiegend inhaltlich vollständig und ansatzweise ausgestaltet ▪ Mängel in Bezug zur Kohärenz (Gedankensprünge, isolierte Elemente)
4	Erweiterte Geschichtenstruktur	<ul style="list-style-type: none"> ▪ Grundstruktur vorhanden und überwiegend inhaltlich vollständig ▪ Teilweise Ausgestaltung vorhanden ▪ Mängel in Bezug auf Kohärenz (Gedankensprünge, isolierte Elemente)
5	Ausgearbeitete Geschichte	<ul style="list-style-type: none"> ▪ Grundstruktur vorhanden und überwiegend inhaltlich vollständig ▪ Ausgestaltung überwiegend gelungen ▪ Kohärenz überwiegend vorhanden
6	Voll entfaltete Geschichte	<ul style="list-style-type: none"> ▪ Grundstruktur vorhanden und überwiegend inhaltlich vollständig ▪ Ausgestaltung überwiegend gelungen ▪ Kohärenz überwiegend vorhanden ▪ Gesamtstruktur bildet eine Einheit ▪ Passung, Übergänge und Sinnzusammenhänge hervorragend realisiert

Erlebnisgeschichten: Als Transfermaß wurde das Schreiben von Erlebnisgeschichten gewählt. Analog dem Randomisierungsverfahren der Bildergeschichten wurden insgesamt drei verschiedene Schreibstimuli zu den drei Messzeitpunkten zufällig den Schülern zugeordnet. Die drei Stimuli waren die Begriffe „Spielen“, „Ferien“ und „Freunde“. Die

Schüler wurden aufgefordert, zehn Minuten die Geschichte zu planen, zwanzig Minuten zu schreiben und zehn Minuten zu überarbeiten. Die Erlebnisgeschichten wurden wie auch die Bildergeschichten hinsichtlich der Geschichtenelemente und der Narrativen Qualität ausgewertet.

Interrater-Reliabilitäten: Alle abhängigen Variablen wurden von der Doktorandin ausgewertet. Eine Psychologiestudentin im Hauptstudium, die mit dem Design und der Fragestellung der Studie nicht vertraut war, analysierte zudem 30% der Daten, so dass für die abhängigen Variablen Interrater-Reliabilitäten zur Einschätzung der Auswertungsobjektivität berechnet wurden. Diese sind in Tabelle 20 aufgelistet. Die Wörter wurden nur einmal gezählt, so dass hierfür keine Interrater-Reliabilitäten ermittelt wurden. Alle Interrater-Reliabilitäten lagen zwischen $r = .71$ und $r = .99$, so dass insgesamt die Übereinstimmungen zwischen den beiden Beurteilern als gut zu bewerten ist.

Tabelle 20: Interrater-Reliabilitäten (Studie I).

	Prätest	Posttest	Follow-Up
Geschichtenelemente	0.85	0.89	0.91
Narrative Qualität	0.81	0.68	0.75
Planen Inhalt/Ausdruck	0.86	0.99	0.93
Planen Struktur	0.83	0.97	0.81
Einfügungen	0.89	0.94	0.85
Wissen	0.88	0.93	0.89
Erlebnisgeschichten Geschichtenelemente	0.74	0.89	0.83
Erlebnisgeschichten Narrative Qualität	0.75	0.71	0.87

6.2.5.2 Kontrollvariablen

Um den Einfluss möglicher Kontrollvariablen auf die Ergebnisse zu erfassen, wurden die Rechtschreibfertigkeiten überprüft und die Nationalität sowie die Muttersprache mittels eines selbst entwickelten Herkunftsfragebogens (siehe CD-Anhang A2) erfragt. Die Erhebung wurde von der Doktorandin im gesamten Klassenkontext durchgeführt. Die Durchführung des Rechtschreibtests orientierte sich an den Angaben der Hersteller.

Rechtschreibung: Zur Erfassung der Rechtschreibung wurde der Diagnostische Rechtschreibtest für vierte Klassen (DRT 4) (Grund et al., 2004) eingesetzt. Der DRT 4 ist ein standardisierter und normierter Schulleistungstest zur Erfassung der Rechtschreibfähigkeiten im vierten Grundschuljahr und kann als Gruppentest in der gesamten Klasse durchgeführt

werden. Der Test ist als Lückentest aufgebaut. Die Schüler müssen 42 Wörter nach Diktat in die Lücken eintragen. Die Reliabilität wird von den Autoren mit $r = .92$ angegeben.

Herkunftsfragebogen: Ein selbst entwickelter Herkunftsfragebogen diente der Erhebung des Geschlechts, der Deutschnote am Ende der dritten Klasse, der Nationalität sowie der Muttersprache. Zur Erfassung der Nationalität wurden die Schüler gefragt, (a) in welchem Land sie geboren sind und (b) in welchem Land die Mutter bzw. der Vater geboren ist. Die Muttersprache wurde durch die Frage nach der Sprache, die sie zu Hause am meisten sprechen, ermittelt. Im CD-Anhang A2 ist der Herkunftsfragebogen abgebildet.

6.3 Ergebnisse

Zur Überprüfung der Hypothesen der Studie I wurden verschiedene statistische Verfahren durchgeführt, die im Folgenden erläutert werden. Die Ergebnisse von Messwiederholungsanalysen, Kovarianzanalysen und Effektstärkenberechnungen werden berichtet.

6.3.1 Statistische Analysen

6.3.1.1 Begründung der Analyseeinheit Klasse

Die Entscheidung, ob die Klasse oder der Schüler als Analyseeinheit gewählt wird, fiel zu Gunsten der Klasse, da die Zuweisung der Klassen zu den Bedingungen sowie sämtliche Instruktionen und Testungen im gesamten Klassenkontext erfolgten. Auf diese Weise kann zudem ausgeschlossen werden, dass die Effekte aufgrund von Unterschieden in den Klassen verzerrt werden. Dieses Vorgehen führt zu einer geringeren statistischen Power, aber erhöht gleichzeitig die interne Validität (Bloom, Bos, & Lee, 1999). Auch Hedges und Hedberg (2007a) halten die Aggregation der Einzelwerte zur Ermittlung von Gruppenwerten als das beste statistische Vorgehen, wenn ein gruppenrandomisiertes Design vorliegt, da dadurch die exaktesten Ergebnisse für den Bedingungsfaktor erzielt werden. Hedges (2007) argumentiert außerdem, dass die Analyse auf Individualebene bei gruppenrandomisierten Verfahren zu einer Überschätzung der Genauigkeit der Ergebnisse und zu antikonservativen Schlussfolgerungen der statistischen Signifikanz der Bedingungseffekte führt. Bei gruppenrandomisierten Verfahren kann auf Individualebene von keiner Unabhängigkeit der Beobachtungen ausgegangen werden.

Infolge der Wahl Klasse als Analyseeinheit sind wegen der reduzierten Stichprobengröße Analysemethoden (z.B. Pfadanalysen), die eine große Stichprobe voraussetzen, nicht durchführbar. Dieser Nachteil ist jedoch für die Fragestellungen des

Dissertationsprojekts weniger bedeutsam, da das Hauptanliegen darin besteht, den Nachweis zu erbringen, dass die Implementierung eines selbstregulatorischen Aufsatztrainings erfolgreich ist und im Vergleich zu einem Training ohne selbstregulatorische Komponenten wirksamer ist. Somit steht die Kovarianzanalyse als statistische Analysemethode im Vordergrund. Auch wenn bei einer geringen Stichprobengröße die Normalverteilungsannahme verletzt sein kann, ist die ANCOVA nicht-parametrischen Verfahren vorzuziehen, da sie gegenüber Verletzungen der Voraussetzungen sehr robust ist (Bortz & Weber, 2005; Glass, Peckham, & Sanders, 1972). Der Nachweis signifikanter Unterschiede ist jedoch durch eine kleine Stichprobe erschwert. Gleichzeitig sprechen Ergebnisse, die eine Signifikanz auf Klassenebene nachweisen, in erheblichem Maße für die Wirksamkeit und Überlegenheit des USAT-Programms. Des Weiteren intendierte die vorliegende Arbeit Generalisierungen auf Klassen- und nicht auf Individualebene, da untersucht wurde, ob Trainer bzw. Lehrer das Klasseniveau steigern können. Die Förderung einzelner Schüler stand dagegen nicht im Fokus.

6.3.1.2 Voranalysen

Trainereffekte: Die Kontrolle möglicher Trainereffekte wurde dadurch realisiert, dass jede Trainerin jeweils eine USAT- sowie eine WW-Bedingung instruierte. Die Berechnung des Effekts des Trainers auf die beiden Bedingungen konnte wegen der geringen Stichprobengröße auf Klassenebene nicht durchgeführt werden, so dass die trainerspezifischen Analysen auf Schülerebene gerechnet wurden. Um den Interaktionseffekt Bedingung x Trainer zu testen, wurden zweifaktorielle Kovarianzanalysen mit den einzelnen Schreibmaßen als abhängige Variablen, den Faktoren Bedingung sowie Trainer als unabhängige Variablen und den jeweiligen Prätest-Werten als Kovariate gerechnet. Für keines der erhobenen abhängigen Schreibmaße ergaben sich signifikante Trainer x Bedingungs-Interaktionseffekte ($p > .11$), so dass die Schlussfolgerung zulässig ist, dass jede der vier Trainerinnen in beiden Bedingungen gleich effektiv und engagiert war. In den folgenden Analysen kann somit der Einfluss von Trainereffekten vernachlässigt werden.

Kontrollvariablen: In einem weiteren Schritt wurden Rechtschreibung, Klassengröße, Geschlecht, Nationalität und Muttersprache Deutsch jeweils als Kovariate in den statistischen Analysen berücksichtigt. Sämtliche statistischen Analysen der Kontrollvariablen wurden auf Klassenebene durchgeführt.

Als erstes wurde überprüft, ob sich die Mittelwerte der Kontrollvariablen für die beiden Bedingungen signifikant unterscheiden. Dazu wurden T-Tests für unabhängige Stichproben mit der zweistufigen Bedingung (USAT, WW) als unabhängige Variable und die

jeweilige Kontrollvariable (Rechtschreibung, Klassengröße, Geschlecht, Nationalität, Muttersprache Deutsch) als abhängige Variable durchgeführt. Wie bereits oben beschrieben, ergaben sich für keine der Variablen signifikante Unterschiede zwischen der USAT- und der WW-Bedingung ($p > .05$).

Schließlich wurde ein kovarianzanalytisches Modell mit dem jeweiligen abhängigen Schreibmaß als abhängige Variable, dem Messzeitpunkt (Posttest, Follow-Up) als Innersubjektfaktor, der Bedingung (USAT, WW) als Zwischensubjektfaktor und der jeweiligen Kontrollvariable (Rechtschreibung, Klassengröße, Geschlecht, Nationalität, Muttersprache Deutsch) sowie der jeweiligen Prätest-Variablen als Kovariate gerechnet. Abgesehen von Erlebniserzählung Narrative Qualität veränderte sich bei keiner abhängigen Variablen die statistische Signifikanz des Haupteffekts Bedingung (siehe Tabelle 21). Bei Revisionen blieben die Bedingungseffekte weiterhin aus (siehe Tabelle 21 und Tabelle 35). Bei Erlebniserzählung Narrativer Qualität war der statistisch bedeutsame Bedingungseffekt bei Berücksichtigung der Kontrollvariablen Rechtschreibung ($p = .05$), Klassengröße ($p = .10$) und Geschlecht ($p = .05$) knapp nicht signifikant. Da die Signifikanzniveaus jedoch nur geringfügig größer als .05 waren, ist der Einfluss der Kontrollvariablen auf die Ergebnisse für Erlebniserzählungen Narrative Qualität als eher unbedeutend einzuschätzen.

Tabelle 21: Signifikanzniveaus für den Bedingungseffekt nach Berücksichtigung der Kontrollvariablen Rechtschreibung, Klassengröße, Geschlecht, Nationalität und Sprache (Studie I).

	Recht- schreibung	Klassen- größe	Geschlecht	Nationa- lität	Sprache
Geschichtenelemente	$p < .01$	$p < .01$	$p < .001$	$p < .001$	$p < .01$
Narrative Qualität	$p < .05$	$p < .01$	$p < .01$	$p < .05$	$p < .01$
Planen Geschichtenelemente	$p < .01$	$p < .01$	$p < .001$	$p < .01$	$p < .01$
Planen Struktur	$p < .05$	$p < .05$	$p < .01$	$p < .05$	$p < .05$
Verbesserungen	$p = .06$	$p = .36$	$p = .11$	$p = .42$	$p = .44$
Wissen	$p < .05$	$p < .05$	$p < .05$	$p < .05$	$p < .05$
Erlebniserzählungen Geschichtenelemente	$p < .05$	$p < .05$	$p < .01$	$p < .05$	$p < .01$
Erlebniserzählungen Narrative Qualität	$p = .05$	$p = .10$	$p = .05$	$p < .05$	$p < .05$

Darüber hinaus wurden Produkt-Moment-Korrelationen zur Überprüfung des Zusammenhangs zwischen den abhängigen Schreibmaßen und den Kontrollvariablen

durchgeführt. Die einzelnen strategienahen und -fernen Schreibmaße zum ersten bis zum dritten Messzeitpunkt wurden jeweils mit den einzelnen Kontrollvariablen (Rechtschreibung, Klassengröße, Geschlecht, Nationalität, Muttersprache) korreliert. Die Kontrollvariablen korrelierten nicht signifikant miteinander ($ps > .05$). Auch für die meisten abhängigen Maße ergaben sich zu keinem der drei Messzeitpunkte signifikante Korrelationen ($ps > .05$). Ausnahmen gab es für folgende Variablen: Die Geschichtenqualität zum ersten Messzeitpunkt korrelierte signifikant mit der Rechtschreibung mit $r = .81$ ($p < .05$). Weiter zeigte sich ein signifikanter Zusammenhang zwischen Planen Geschichtenelemente Prätest und Klassengröße $r = .71$ ($p < .05$). Verbesserungen Posttest ($r = .74$, $p < .05$) und Follow-Up ($r = .76$, $p < .05$) korrelierten ebenfalls signifikant mit der Nationalität.

6.3.1.3 Statistische Hauptanalysen

Zur Überprüfung der Wirksamkeit eines unterrichtsintegrierten selbstregulatorischen Interventionsprogramms im Vergleich zu einem ohne selbstregulatorische Komponenten wurde folgendes statistisches Vorgehen gewählt.

Die fehlenden Werte von Schülern, die zu einem der drei Messzeitpunkte nicht anwesend waren, wurden regressionsanalytisch ersetzt. Keiner der Schüler fehlte an mehr als einer Testung.

Zur Aggregation der Daten wurden für jede der abhängigen Variablen und Kontrollvariablen über alle Schüler hinweg pro Klasse Mittelwerte gebildet. Die so gebildeten Mittelwerte wurden für die statistischen Analysen herangezogen.

Um das Problem der Analyseeinheit bei gruppenrandomisierten Untersuchungen genauer zu untersuchen wurden zu Beginn Intra-Klassen-Korrelationen berechnet. Ferner wurden die Mittelwerte der einzelnen Klassen innerhalb einer Bedingung aufgelistet und grafisch dargestellt, um Klassen mit extremen Ausprägungen in den einzelnen abhängigen Variablen zu identifizieren. Auf diese Weise konnte ausgeschlossen werden, dass einzelne Klassen innerhalb einer Bedingung für die Effekte verantwortlich waren. Schließen wurden Mixed Model-ANCOVA's durchgeführt. Dabei wurden für jede abhängige Variable der Bedingungsfaktor unter Berücksichtigung des Einflusses Schüler innerhalb einer Klasse und Klasse innerhalb einer Bedingung berücksichtigt.

Der Studie lag ein Untersuchungsdesign mit drei Messzeitpunkten zu Grunde. Eine Prätest-Variable, Planen Struktur, variierte signifikant zwischen den beiden Bedingungen ($p < .05$) (siehe Tabelle 34). Für alle anderen abhängigen Variablen ergaben sich keine signifikanten Bedingungseffekte beim Prätest ($ps > .05$). Aus diesem Grund wurde in den als nächstes durchgeführten Messwiederholungsanalysen mit der Schreibleistung

(Geschichtenelemente, Geschichtenqualität, Wortanzahl), den strategienahen Schreibmaßen (Planen, Revisionen, Wissen) und dem Transfermaß Erlebniserzählungen jeweils als abhängige Variable, dem Messzeitpunkt (Posttest, Follow-Up) als Innersubjektfaktor, Bedingung (USAT, WW) als Zwischensubjektfaktor der jeweilige Prätest als Kovariate berücksichtigt.

Anschließend wurden zur Analyse der Unterschiede zwischen den Klassen der beiden Bedingungen alle abhängigen Variablen mit kovarianzanalytischen Modellen mit den jeweiligen Schreibmaßen (getrennt für Posttest und Follow-Up) als abhängige Variable, Bedingung (ISAT, WW) als unabhängige Variable und dem Prätest als Kovariate geprüft. Die Voraussetzungen der Kovarianzanalyse wurden getestet und konnten als gegeben angenommen werden. Insbesondere die Homogenität der Regressionssteigungen wurde bei keiner abhängigen Variablen verletzt, d.h. es trat bei keiner abhängigen Variable eine signifikante Interaktion Prätest x Bedingung auf ($ps > .05$).

Die Wahl der Klasse als Analyseeinheit führte zu einer Reduzierung der statistischen Power, so dass eine einseitige Testung der Ergebnisse legitim wäre. Dennoch wurde eine Adjustierung der Fehlerwahrscheinlichkeiten nicht vorgenommen, da trotz vorher festgelegter Gruppenunterschiede zweiseitig getestet wurde und sich bei einer Verschärfung des α -Niveaus in den zentralen Befunden keine Änderungen ergaben.

Abgesehen von den bereits beschriebenen statistischen Analysen wurden als nächstes zur Beurteilung der praktischen Signifikanz der Ergebnisse um Prätestunterschiede korrigierte Effektstärken basierend auf Cohen's d (Cohen, 1988) berechnet. Die konventionelle Schätzung der Effektstärken führte dabei auf Klassenebene zu inkorrekten Standardfehlern, so dass die Effektstärken sowohl auf Klassen- als auch auf Individualebene berechnet wurden. Neben der besseren Interpretierbarkeit der Maße ermöglichen Effektstärken auf Individualebene zudem einen Vergleich mit anderen Studien, die in der Regel die Effektstärken auf Individualebene angeben.

Zur Prüfung von Moderatoreffekte wurden zunächst Subgruppen der Kontrollvariablen Geschlecht (Mädchen, Jungen), Rechtschreibung (rechtschreibschwach, rechtschreibstark) und Sprache (Deutsch, nicht-Deutsch) gebildet. Anschließend wurden Messwiederholungsanalysen mit den einzelnen strategienahen und -fernen Schreibmaßen als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und die Subgruppen der Kontrollvariablen (Geschlecht, Rechtschreibung, Nationalität, Muttersprache Deutsch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor und den Prätest-Variablen der Subgruppen der Kontrollgruppen sowie der Rechtschreibung als Kovariaten

gerechnet. Um den Effekt der beiden Ausprägungen der jeweiligen Kontrollvariable getrennt auf die abhängigen Variablen zu untersuchen, wurde die Kontrollvariable als Inner- und nicht als Zwischensubjektfaktor berücksichtigt. In jeder Klasse lagen beide Ausprägungen der Kontrollvariablen vor.

6.3.2 Zusätzliche Analysen für gruppenrandomisierte Designs

Zur weiteren Begründung der Analyseeinheit Klasse wurden Intra-Klassen-Korrelationen (ICC's) berechnet, um Aussagen über den Zusammenhang der Klassen innerhalb einer Bedingung zu treffen. Ferner wurden die Mittelwerte der Klassen innerhalb einer Bedingung betrachtet. Schließlich wurden Mixed Model-Analysen durchgeführt, um den Einfluss des Schülers innerhalb einer Klasse und der Klasse innerhalb einer Bedingung zu ermitteln.

6.3.2.1 Intra-Klassen-Korrelationen (ICC's)

Die Intra-Klassen-Korrelation gibt an, wie sehr sich die Einheiten einer Gruppe ähneln und wie groß der Zusammenhang zwischen den Beobachtungen der Mitglieder einer Klasse sind (Murray, Varnell, & Blitstein, 2004). Demnach stellt die Intra-Klassen-Korrelation einen weiteren Kennwert für die Entscheidung, ob die Analyseeinheit ein individueller Wert oder der Mittelwert der Gruppe sein sollte, dar.

Die Berechnung der in Tabelle 22 aufgelisteten Intra-Klassen-Korrelationen erfolgte mit dem Programm *MPlus 5.2.1*. Es wurde für jede Bedingung getrennt Intra-Klassen-Korrelationen (ICC) mit Klasse als Cluster gerechnet, um das Ausmaß zu bestimmen, in welchem sich die Mitglieder einer Klasse ähnlicher sind als Mitglieder einer anderen Klasse. Eine Nullkorrelation spricht dabei für eine Ähnlichkeit der Klassen innerhalb einer Bedingung.

Tabelle 22: Geschätzte Intra-Klassen-Korrelationen (ICC's) getrennt nach Messzeitpunkt und Bedingung (Studie I).

	USAT			WW		
	Prätest	Posttest	Follow-Up	Prätest	Posttest	Follow-Up
Geschichtenelemente	0.012	0.017	0.036	0.097	0.008	0.065
Geschichtenqualität	0.063	0.003	0.088	0.114	0.021	0.004
Wortanzahl	0.186	0.185	0.316	0.158	0.045	0.177
Planen Inhalt	0.037	0.029	0.024	0.042	0.144	0.084
Planen Struktur	0.003	0.004	0.021	0.029	0.046	0.008

Revisionen	0.015	0.082	0.034	0.008	0.002	0.008
Wissen	0.021	0.053	0.034	0.155	0.148	0.079
Erlebniserzählungen (GE)	0.071	0.117	0.119	0.029	0.067	0.085
Erlebniserzählungen (NQ)	0.015	0.046	0.046	0.001	0.002	0.001

Abgesehen von der Wortzahl ergaben sich für die USAT-Bedingung ICC's zwischen .003 und .119 und für die WW-Bedingung Werte zwischen .001 und .155.

Um den Cluster als Analyseeinheit zu legitimieren, geben Hedges und Hedberg (2007a, 2007b) als Optimalwerte ICC's zwischen .05 und .15 an. Aufbauend auf ihren Untersuchungen halten sie jedoch bei größeren Gruppen (Nationen, Schulen) ICC's zwischen .15 und .25 für realistischer. Insgesamt kann geschlussfolgert werden, dass die in Tabelle 22 aufgeführten Intra-Klassen-Korrelationen die Analyseeinheit Klasse für fast alle untersuchten abhängigen Variablen rechtfertigen.

6.3.2.2 Deskriptive Beschreibung der Klassenunterschiede

Für jede abhängige Variable werden die deskriptiven Statistiken getrennt nach Messzeitpunkt, Bedingung und Klasse dargestellt, um potentielle Ausreißer-Klassen innerhalb einer Bedingung zu identifizieren.

6.3.2.2.1 *Geschichtenelemente*

Um auszuschließen, dass Effekte auf einzelne Klassen innerhalb einer Bedingung zurückzuführen sind, sind in Tabelle 23 die deskriptiven Statistiken der einzelnen Klassen für jede Bedingung getrennt nach Messzeitpunkt aufgelistet.

Tabelle 23: Mittelwerte und Standardabweichungen für Geschichtenelemente getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).

	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	3.19	1.47	9.06	2.41	8.31	2.39
Klasse 2	2.45	1.64	7.80	2.89	6.60	2.19
Klasse 3	3.25	1.48	8.45	2.24	6.85	2.68
Klasse 4	3.10	1.70	8.86	2.46	8.00	2.49

WW						
Klasse 5	2.65	1.66	4.90	2.00	3.10	1.77
Klasse 6	2.41	1.18	4.27	2.00	4.14	1.83
Klasse 7	3.64	1.60	4.57	1.09	4.07	1.38
Klasse 8	3.74	1.54	4.61	1.75	4.52	1.88

Tabelle 23 zeigt, dass die Mittelwerte der Klassen innerhalb einer Bedingung nicht stark variierten. Die Mittelwerte für die USAT-Klassen sind in Abbildung 35 und für die WW-Klassen in Abbildung 36 grafisch dargestellt.

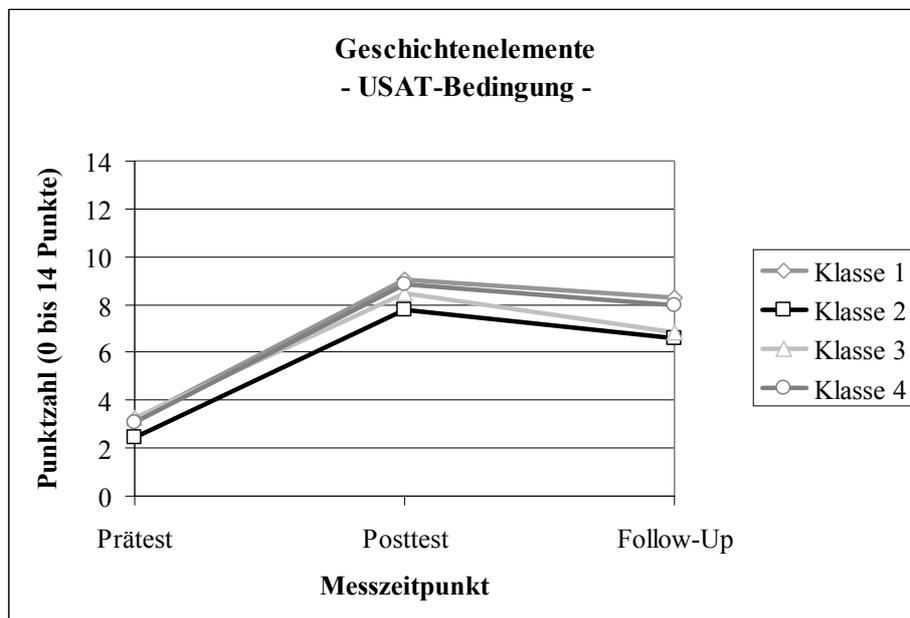


Abbildung 35: Geschichtenelemente der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).

In der WW-Bedingung sind die größten Leistungsunterschiede beim Prätest zwischen den Klassen 6 und 8.

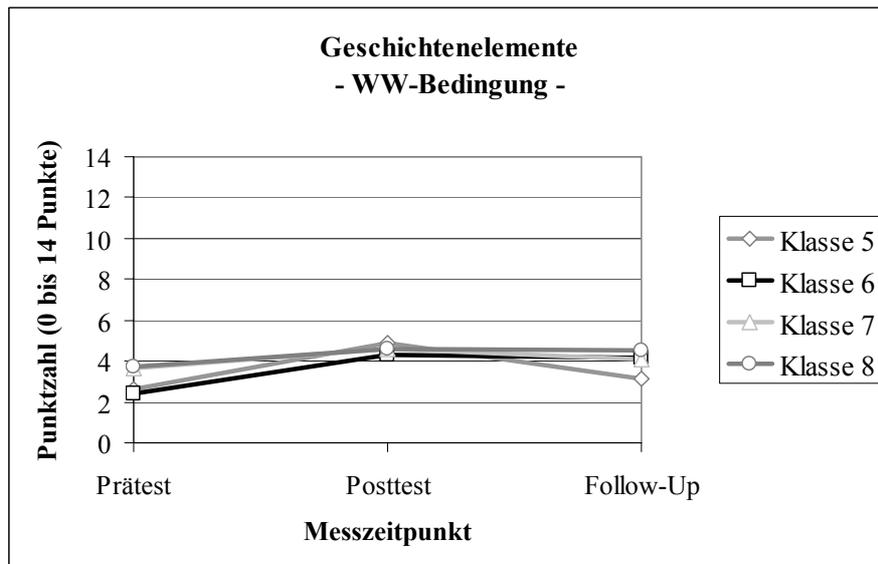


Abbildung 36: Geschichtenelemente der WW-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).

6.3.2.2.2 *Geschichtenqualität*

Um die Unterschiede zwischen den Klassen innerhalb einer Bedingung genauer zu analysieren, sind die Mittelwerte für die Geschichtenqualität für die Klassen nach Bedingung und Messzeitpunkt getrennt dargestellt (siehe Tabelle 24).

Tabelle 24: Mittelwerte und Standardabweichungen für Geschichtenqualität getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).

	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	1.94	0.64	3.63	0.68	3.31	0.61
Klasse 2	1.80	0.77	3.20	0.50	3.00	0.46
Klasse 3	2.35	0.63	3.45	0.65	2.80	0.52
Klasse 4	2.14	0.49	3.38	0.52	3.52	0.49
WW						
Klasse 5	1.75	0.57	2.40	0.81	2.50	0.79
Klasse 6	1.73	0.62	2.59	0.83	2.73	0.79
Klasse 7	2.36	0.49	2.57	0.83	2.50	0.77
Klasse 8	2.17	0.73	2.78	1.02	2.65	0.68

Abbildung 37 und Abbildung 38 bieten eine anschauliche Darstellung der Mittelwerte der Klassen innerhalb der beiden Bedingungen.

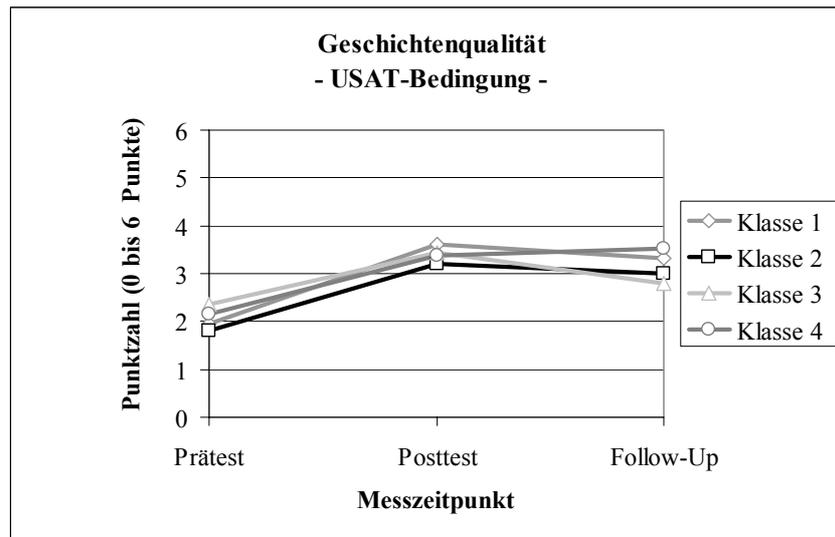


Abbildung 37: Geschichtenqualität der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).

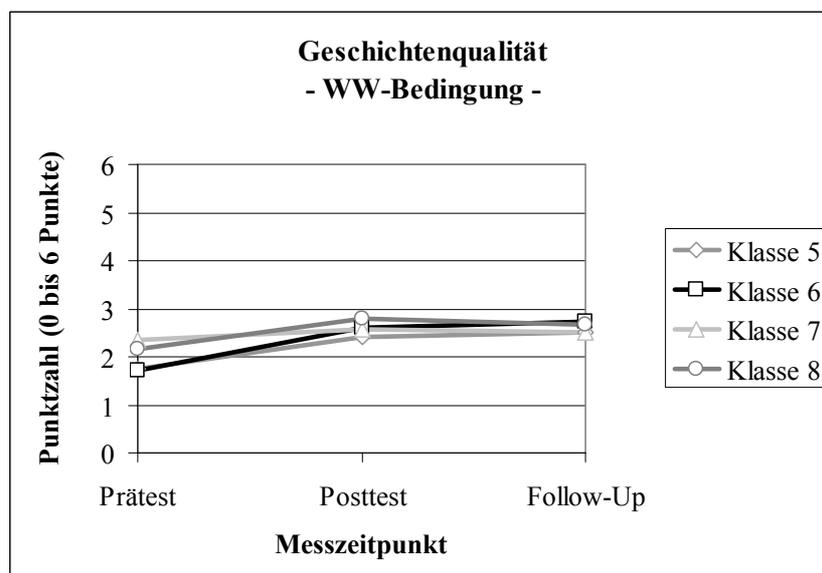


Abbildung 38: Geschichtenqualität der WW-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).

In der USAT-Bedingung ist der größte Unterschied beim Prätest zwischen den Klassen 2 und 3 festzustellen. In der WW-Bedingung ist dieser zwischen den Klassen 6 und 7 zu erkennen. Die Prätestunterschiede sowohl in der USAT- als auch in der WW-Bedingung werden bei den Kovarianzanalysen berücksichtigt. Eine genauere Betrachtung der Leistungsunterschiede zwischen den Klassen der USAT-Bedingung beim dritten Messzeitpunkt zeigt, dass Klasse 3 im Vergleich zu den anderen drei USAT-Klassen am schlechtesten abschnitt, wobei jedoch alle Klassen der USAT-Bedingung einen deutlich höheren Mittelwert aufweisen als die WW-Klassen.

6.3.2.2.3 Wortanzahl

Um potentielle Ausreißer-Klassen zu identifizieren, werden die Mittelwerte und Standardabweichungen für jede Klasse und jeden Messzeitpunkt getrennt betrachtet (siehe Tabelle 25).

Tabelle 25: Mittelwerte und Standardabweichungen für Wortanzahl getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).

	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	109.19	49.64	198.75	49.07	164.31	52.01
Klasse 2	77.35	31.37	153.65	46.73	72.80	45.63
Klasse 3	79.27	19.42	153.41	45.84	123.09	42.53
Klasse 4	69.75	36.42	134.00	77.98	121.70	63.44
WW						
Klasse 5	97.79	32.39	106.50	68.97	88.71	25.67
Klasse 6	105.13	21.16	141.17	85.76	131.87	71.28
Klasse 7	93.00	25.84	140.30	24.58	122.50	19.12
Klasse 8	118.14	24.06	198.24	48.10	211.14	44.16

Bei Betrachtung der USAT-Mittelwertslinien in Abbildung 39 ist kein paralleler Verlauf zu erkennen. Beim Posttest erzielten die Klassen 1 und 4 sowie 3 und 2 sehr ähnliche Leistungen, wobei diese wiederum erheblich von den anderen beiden Klassen abweichen. Beim Follow-Up verschlechterten sich alle Klassen außer Klasse 4.

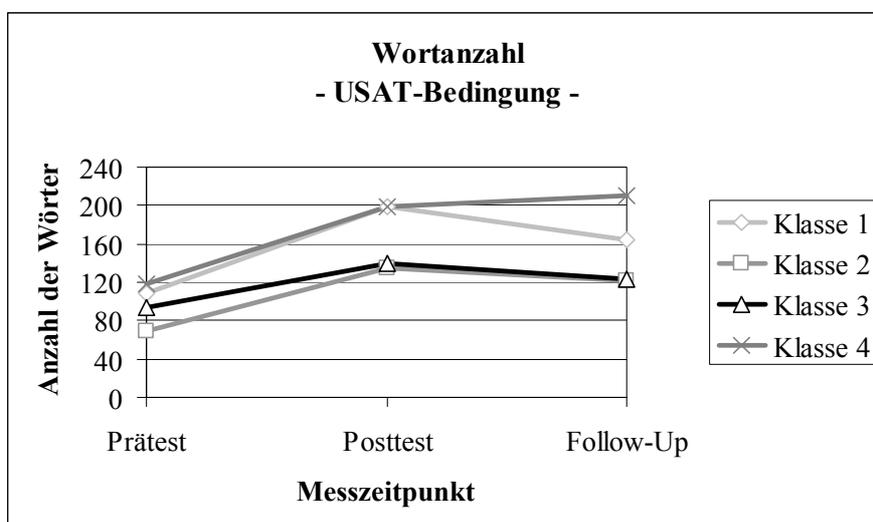


Abbildung 39: Wortanzahl der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).

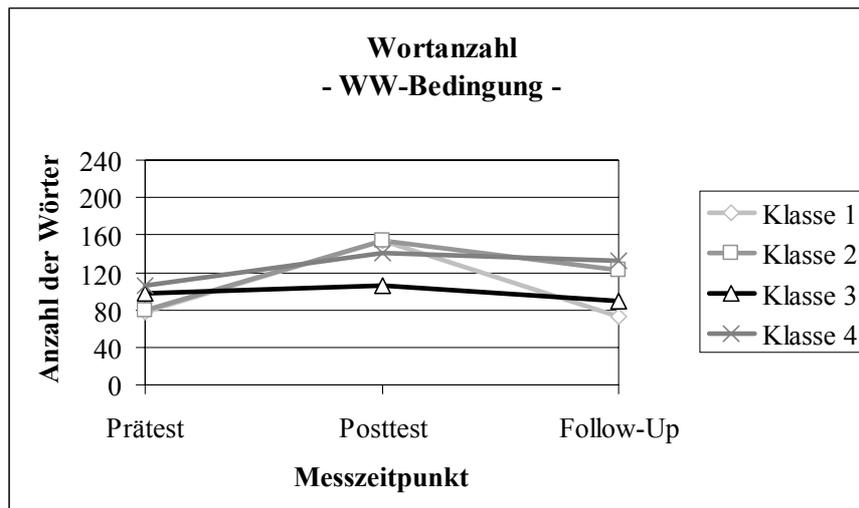


Abbildung 40: Wortanzahl der WW-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).

Ein ähnlich uneinheitliches Bild zeigt sich für die WW-Klassen (siehe Abbildung 40). Auch hier divergieren die Linien zwischen den Klassen. Während die Klasse 3 kaum Veränderungen über die drei Messzeitpunkte aufweist, steigen die Linien der anderen drei Klassen beim Posttest erheblich an. Beim Follow-Up fallen die Leistungen jedoch wieder ab, wobei die Leistungen der Klasse 1 am stärksten einbricht.

6.3.2.2.4 Planen

Als nächstes wurden die Unterschiede zwischen den Klassen für das strategienahe Maß Planen analysiert.

Tabelle 26: Mittelwerte und Standardabweichungen für Planen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).

Planen Geschichtenelemente						
	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	1.19	0.54	7.00	1.93	6.63	1.63
Klasse 2	1.35	0.81	6.00	2.64	6.05	2.80
Klasse 3	1.30	0.86	4.55	3.15	4.75	2.84
Klasse 4	1.81	0.98	5.81	3.67	5.48	3.71

WW						
Klasse 5	1.70	1.17	0.55	0.83	0.90	1.29
Klasse 6	1.59	0.50	1.73	1.32	1.73	0.98
Klasse 7	0.93	0.92	1.14	1.78	1.86	1.10
Klasse 8	1.35	0.98	1.48	0.73	1.22	0.74
Planen Struktur						
USAT						
Klasse 1	1.31	1.78	3.38	1.93	4.44	2.00
Klasse 2	1.40	1.19	3.35	2.03	3.50	2.09
Klasse 3	1.10	1.07	2.90	2.02	2.70	2.18
Klasse 4	1.57	1.21	3.67	2.22	3.76	2.45
WW						
Klasse 5	2.80	1.96	0.75	1.25	0.85	1.42
Klasse 6	2.27	1.63	1.64	1.47	1.05	1.09
Klasse 7	1.57	1.34	1.21	0.80	1.36	1.50
Klasse 8	2.09	1.44	1.65	1.43	1.13	1.14

Zur genaueren Analyse der Unterschiede werden die Mittelwerte der Klassen für Planen Geschichtenelemente und Planen Struktur getrennt für jede Bedingung und für jeden Messzeitpunkt betrachtet (siehe Abbildung 41 und Abbildung 42).

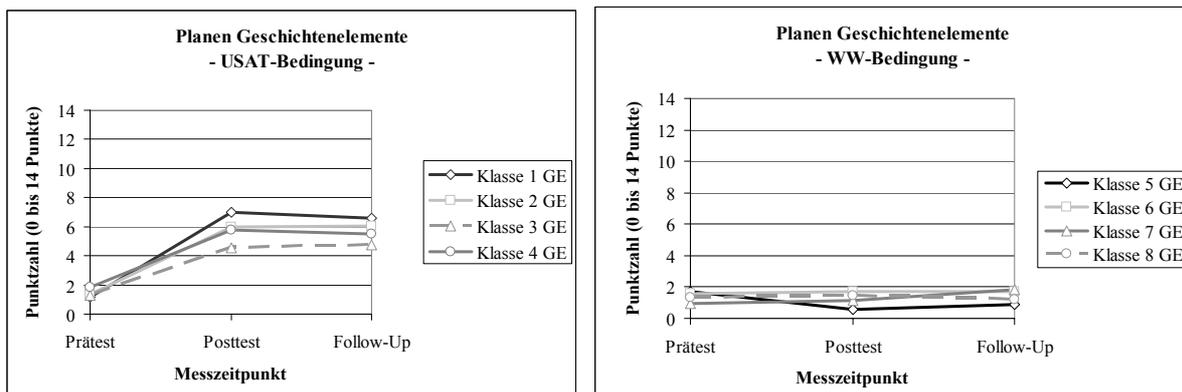


Abbildung 41: Planen Geschichtenelemente als Funktion von Messzeitpunkt und Klasse (Studie I).

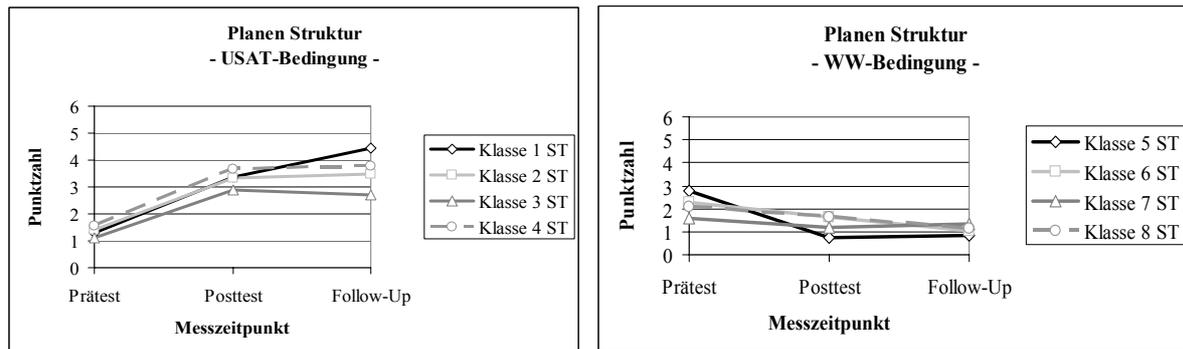


Abbildung 42: Planen Struktur als Funktion von Messzeitpunkt und Klasse (Studie I).

In der WW-Bedingung erzielte die Klasse 5 wesentlich geringere Leistungen bei Planen Geschichtenelemente im Posttest und beim Follow-Up als die anderen WW-Klassen, wobei die Klasse 6 beim Posttest und die Klasse 7 beim Follow-Up am besten abschnitten. Folglich sind in einzelnen Klassen der WW-Bedingung deutliche Verbesserungen vom Prätest zum Posttest bzw. vom Prätest zum Follow-Up zu erkennen. Betrachtet man jedoch die Mittelwerte der Klassen 6 und 7, die zur WW-Bedingung gehören, im Vergleich mit den Mittelwerten der vier Klassen der USAT-Bedingung, wird deutlich, dass sämtliche Mittelwerte der USAT-Bedingung erheblich über denen der Klassen 6 und 7 liegen.

6.3.2.2.5 Revisionen

Die Klassenunterschiede innerhalb einer Bedingung wurden ebenfalls für Revisionen genauer betrachtet. Tabelle 27 enthält die Revisions-Mittelwerte der Klassen der USAT- und WW-Bedingung getrennt nach Messzeitpunkt.

Tabelle 27: Mittelwerte und Standardabweichungen für Revisionen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).

	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	0.13	0.34	1.44	1.36	0.88	1.26
Klasse 2	0.20	0.52	0.75	0.91	0.30	0.57
Klasse 3	0.00	0.00	0.40	0.68	0.40	0.75
Klasse 4	0.24	0.44	0.86	0.96	0.52	0.60
WW						
Klasse 5	0.10	0.31	0.55	0.76	0.70	0.98
Klasse 6	0.14	0.35	0.41	0.96	0.27	0.55
Klasse 7	0.07	0.27	0.07	0.27	0.29	0.61
Klasse 8	0.00	0.00	0.39	0.78	0.48	0.85

Der Tabelle 27 ist zu entnehmen, dass die Leistungen der Klasse 1 beim Posttest deutlich stärker anstiegen als die der drei anderen Klassen der USAT-Bedingung. Gleichzeitig entspricht der Mittelwert der Klasse 3 denen der WW-Klassen. Dies wird bei Betrachtung der Abbildung 43 und 44 deutlich.

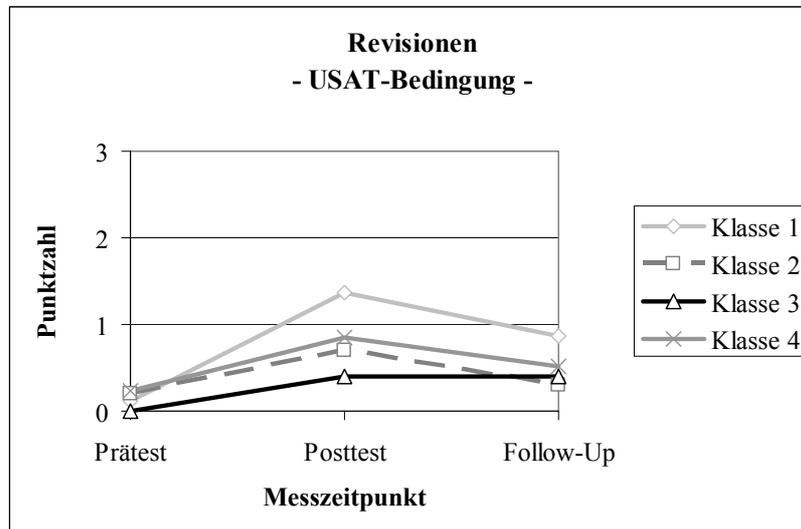


Abbildung 43: Revisionen der USAT-Bedingung in Abhängigkeit von Messzeitpunkt und Klasse (Studie I).

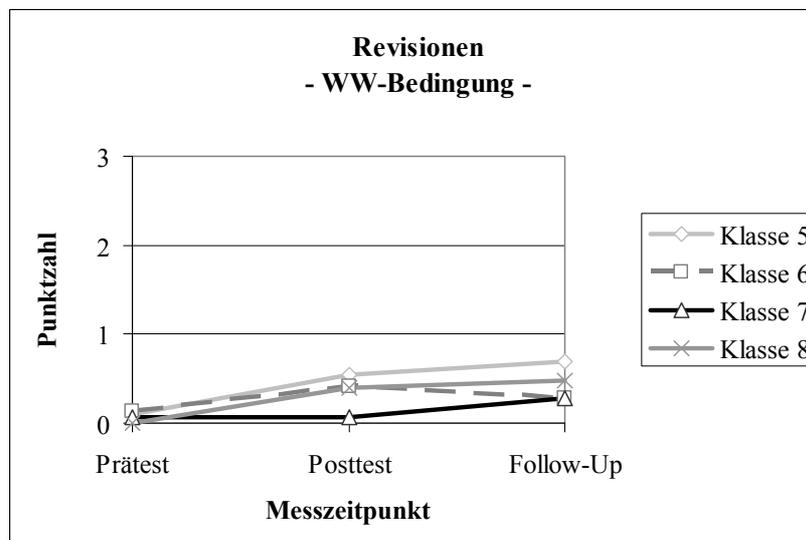


Abbildung 44: Revisionen der WW-Bedingung in Abhängigkeit von Messzeitpunkt und Klasse (Studie I).

6.3.2.2.6 *Schreibbezogenes Wissen*

Als nächstes wurden die Mittelwerte für Wissen der einzelnen Klassen innerhalb einer Bedingung näher untersucht. In Tabelle 28 sind die Mittelwerte der einzelnen Klassen aufgelistet und in den Abbildung 45 (USAT-Bedingung) sowie Abbildung 46 (WW-Bedingung) grafisch dargestellt.

Tabelle 28: Mittelwerte und Standardabweichungen für Wissen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).

	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	1.50	1.26	6.13	2.73	6.00	2.16
Klasse 2	1.60	1.05	4.10	2.73	4.75	2.27
Klasse 3	2.15	1.31	4.95	2.09	4.35	2.39
Klasse 4	1.62	1.02	5.81	2.27	4.90	2.26
WW						
Klasse 5	1.50	1.40	2.40	1.35	2.40	1.14
Klasse 6	2.68	1.52	3.77	1.19	3.09	1.11
Klasse 7	3.21	0.89	3.21	0.89	3.43	1.09
Klasse 8	2.17	0.94	2.83	0.89	2.43	1.44

In Abbildung 45 verlaufen die Mittelwertlinien der einzelnen USAT-Klassen nahezu parallel. Innerhalb der WW-Bedingung erzielten die Klassen 6 und 7 höhere Leistungen als die Klassen 5 und 8, wobei Klasse 5 bereits ein sehr niedriges Ausgangsniveau und Klasse 7 hingegen ein sehr hohes Ausgangsniveau besaß (siehe Abbildung 46).

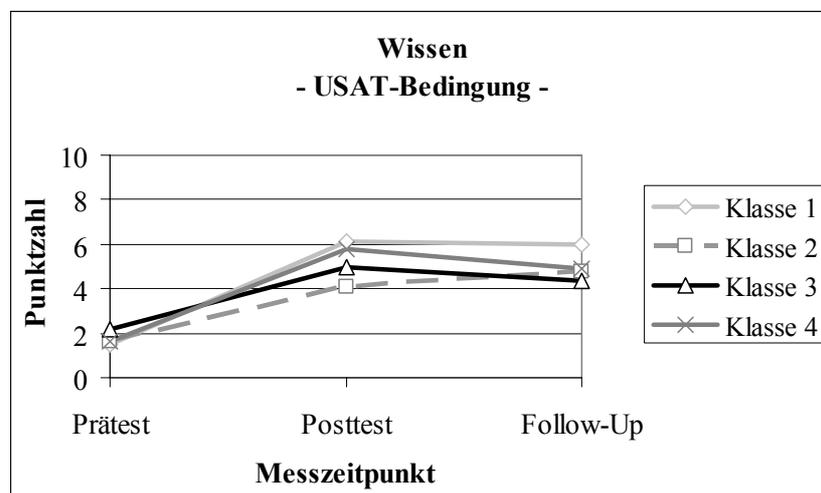


Abbildung 45: Wissen der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).

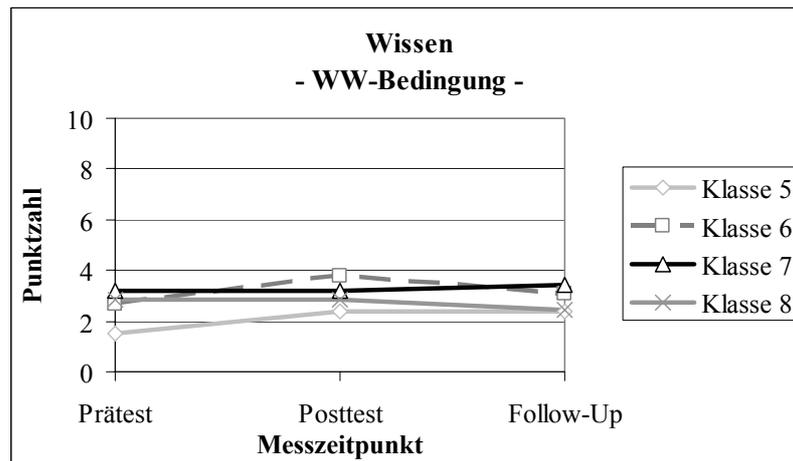


Abbildung 46: Wissen der WW-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie I).

6.3.2.2.7 *Erlebniserzählungen*

Als nächstes wurden die Mittelwerte der einzelnen Klassen innerhalb einer Bedingung für Erlebniserzählungen Geschichtenelemente und Geschichtenqualität genauer betrachtet. Tabelle 29 enthält eine Auflistung der Mittelwerte für Erlebniserzählungen Geschichtenelemente und Geschichtenqualität getrennt nach Bedingung, Klasse und Messzeitpunkt.

Tabelle 29: Mittelwerte und Standardabweichungen für Erlebniserzählungen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie I).

Erlebniserzählung Geschichtenelemente						
	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	3.50	1.51	7.19	3.06	6.13	.34
Klasse 2	3.00	1.03	6.00	2.38	5.60	2.23
Klasse 3	3.65	1.63	6.60	2.26	5.30	2.47
Klasse 4	4.29	1.42	8.71	2.49	7.62	2.48
WW						
Klasse 5	4.10	1.83	3.80	1.96	3.70	1.30
Klasse 6	3.23	1.72	3.18	1.50	3.27	1.35
Klasse 7	3.86	1.17	4.50	1.51	4.14	1.35
Klasse 8	3.10	1.14	4.57	1.85	4.61	1.67

Erlebniserzählung Geschichtenqualität						
USAT						
Klasse 1	1.50	0.63	2.88	0.89	2.44	0.51
Klasse 2	1.60	0.60	2.60	0.82	2.75	0.97
Klasse 3	1.75	0.64	2.45	1.00	2.45	1.10
Klasse 4	1.90	0.62	3.14	0.85	3.10	0.77
WW						
Klasse 5	1.55	0.60	1.80	0.70	1.85	0.75
Klasse 6	1.36	0.58	1.82	0.66	2.18	0.66
Klasse 7	1.36	0.50	2.00	0.88	2.14	0.95
Klasse 8	1.61	0.66	2.00	0.80	2.26	0.81

Zur besseren Veranschaulichung sind die Mittelwerte grafisch in den untenstehenden Abbildungen gezeigt.

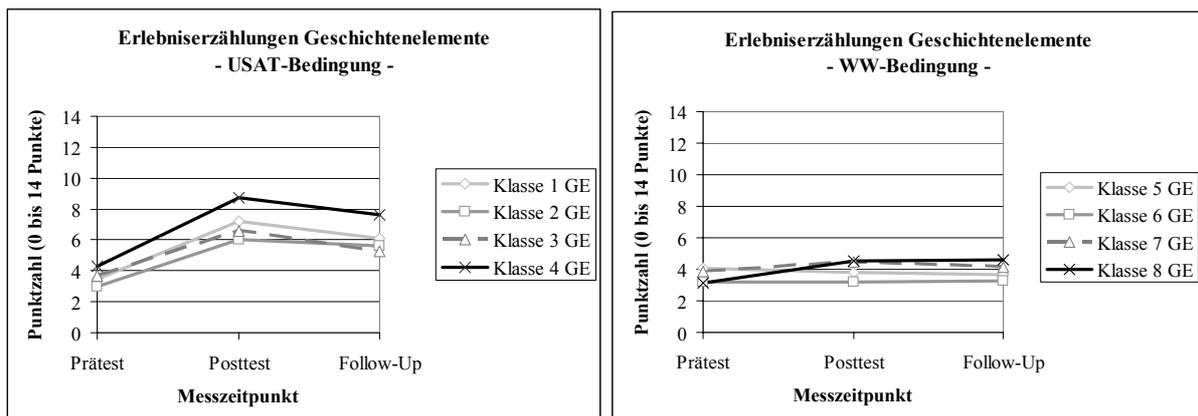


Abbildung 47: Erlebniserzählungen Geschichtenelemente als Funktion von Messzeitpunkt und Klasse (Studie I).

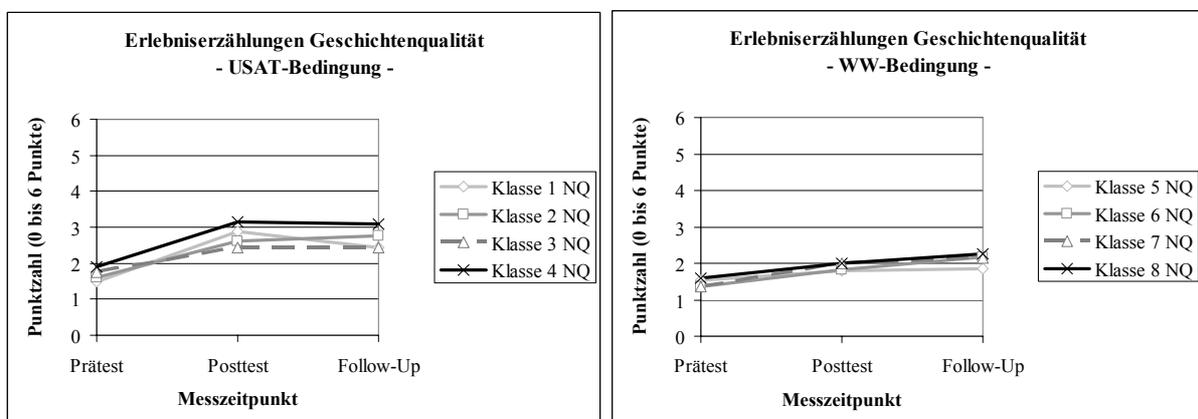


Abbildung 48: Erlebniserzählungen Geschichtenqualität als Funktion von Messzeitpunkt und Klasse (Studie I).

In Abbildung 47 und Abbildung 48 laufen die Mittelwertslinien sowohl für Erlebniserzählungen Geschichtenelemente als auch für Erlebniserzählungen Geschichtenqualität für beide Bedingungen nahezu parallel, wobei die USAT-Linien tendenziell steiler steigen als die WW-Linien.

In der USAT-Bedingung sind die Leistungen der Klasse 4 auffallend und heben sich deutlich von den Leistungen der anderen Klassen ab. In der WW-Bedingung sind die Unterschiede nicht so eindeutig. Während Klasse 6 und 7 kaum Leistungszuwächse von Prätest zum Posttest und vom Prätest zum Follow-Up zeigten, schnitten die Klassen 7 und 8 besser ab. Gleichzeitig waren, wie auch bei den anderen abhängigen Variablen, die Unterschiede zwischen den Bedingungen deutlich größer als innerhalb einer Bedingung.

6.3.2.3 Mixed Model-Analysen

In der Studie I wurde Klasse als Analyseeinheit gewählt. Der wesentliche Grund hierfür ist, dass infolge der Zuweisung von Klassen zu Bedingungen und der Durchführung sämtlicher Instruktionen, Übungen und Testungen auf Klassenebene angenommen werden kann, dass Beobachtungen bzw. Erhebungen auf Schülerebene nicht unabhängig voneinander waren. Vielmehr ist die Wahrscheinlichkeit hoch, dass ein Teil der Schülerleistungen aufgrund der Zugehörigkeit zu einer bestimmten Klasse erklärbar ist. Folglich ist es nicht zulässig, Schüler als Analyseeinheit zu wählen. Mit dem Dilemma der Analyseeinheit beschäftigte sich bereits Hopkins (1982). Zur Berechnung des Einflusses der Klasse innerhalb einer Bedingung bzw. des Schülers innerhalb einer Klasse schlägt er die Durchführung genesteter Varianzanalysen vor.

Aus diesem Grund wurden Mixed Model-ANCOVA's mit Schüler (als zufälliger Faktor) genestet in der Klasse und mit Klasse (als zufälliger Faktor) genestet in der Bedingung (als fester Faktor) durchgeführt. Als Kovariate wurde die jeweilige Prätest-Leistung berücksichtigt. Zur Überprüfung der Hypothese, ob der Einfluss der Klasse innerhalb der Bedingung bzw. des Schülers innerhalb einer Klasse statistisch bedeutsam ist, wurden *Wald Z's* angegeben. Wenn dieser Wert nicht signifikant war, kann der Einfluss unterschiedlicher Klassen innerhalb einer Bedingung als nicht bedeutsam bewertet werden. In Tabelle 30 sind die *Wald Z*- und F-Werte für die einzelnen abhängigen Variablen aufgelistet.

Tabelle 30: Statistische Kennwerte (Wald Z) für den Einfluss Klasse innerhalb Bedingung bzw. Schüler innerhalb Klasse sowie F-Werte für den Haupteffekt Bedingung (Studie I).

	Wald Z				Bedingungseffekt		
	Klasse (Bedingung)		Schüler (Klasse)		F-Wert	FG ¹¹ Nenner	Sig.
	Wald Z	Sig.	Wald Z	Sig.			
Geschichtenelemente	0.28	0.78	4.96	<.001	125.82	5.27	<.001
Geschichtenqualität	0.83	0.41	5.57	<.001	21.78	5.59	<.01
Wortanzahl	1.46	0.14	2.43	<.05	4.45	5.89	0.08
Planen Inhalt	1.05	0.29	6.89	<.001	72.32	5.64	<.001
Planen Struktur	0.55	0.59	3.99	<.001	64.99	6.58	<.001
Revisionen	1.05	0.30	3.09	<.01	1.67	5.60	0.25
Wissen	1.07	0.28	5.00	<.001	30.17	5.98	<.01
Erlebniserzählungen (GE)	1.40	0.16	4.28	<.001	19.43	6.16	<.01
Erlebniserzählungen (NQ)	0.95	0.34	2.43	<.05	19.29	6.36	<.01

Die errechneten Kennwerte für Klasse innerhalb derselben Bedingung lagen zwischen .28 und 1.46, wobei sich bei keiner abhängigen Variable eine statistische Signifikanz ergab. Folglich war der Einfluss der Klasse innerhalb einer Bedingung nicht bedeutsam. Im Gegensatz dazu erwiesen sich sämtliche *Wald Z*'s für Schüler innerhalb einer Klasse als hoch signifikant ($ps < .05$).

Zur Prüfung der Klasse-innerhalb-Bedingung-Effekte schlagen Kromrey und Dickinson (1996) ein Alpha-Niveau von .30 vor, so dass bei einigen abhängigen Variablen der Einfluss der Klasse innerhalb einer Bedingung nicht ausgeschlossen werden kann (siehe Tabelle 30). Weiter argumentieren die Autoren, dass bei Versuchsdesigns, in denen Gruppen anstatt Individuen den Bedingungen zugeordnet werden und in denen die Testungen in Klassen stattfinden, der α -Fehler nur dann adäquat kontrolliert werden kann, wenn die Auswertungen auf Klassenebene erfolgen. Aus diesem Grund wurde in den anschließenden Analysen Klasse als Analyseeinheit gewählt.

Dennoch ergaben die Mixed Model-ANCOVA's abgesehen für Wortanzahl und Revisionen für alle abhängigen Variablen signifikante Bedingungseffekte zu Gunsten der USAT-Bedingung (siehe Tabelle 30). Folglich unterschieden sich die beiden Bedingungen

¹¹ FG = Freiheitsgrade; der Zähler-Freiheitsgrad war bei jeder abhängigen Variablen 1.

signifikant voneinander, wenn der Einfluss der Klasse innerhalb der Bedingung bzw. des Schülers innerhalb der Klasse berücksichtigt wurde.

6.3.3 Schreibleistung

Zur Beurteilung der Schreibleistungen der Schüler wurde die Nennung und Ausgestaltung der 7W-Fragen (Geschichtenelemente) und ein holistisches Maß zur Bewertung der Qualität der Aufsätze (Geschichtenqualität) herangezogen. Überdies wurde der Textumfang durch die Wortanzahl erfasst.

6.3.3.1 Geschichtenelemente

Die Geschichtenelemente wurden auf einer 14-stufigen Skala bewertet. Die Mittelwerte, Standardabweichungen, die um Prätestunterschiede korrigierten Mittelwerte, F-Werte und Effektstärken für beide Bedingungen zu den drei Messzeitpunkten sind in Tabelle 31 aufgelistet.

Tabelle 31: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Geschichtenelemente getrennt nach Bedingung und Messzeitpunkt (Studie I).

	Prätest		Posttest			Follow-Up		
	M	SD	M	SD	M ^{ad}	M	SD	M ^{ad}
USAT-Gruppe	3.00	0.37	8.54	0.56	8.56	7.44	0.84	7.48
WW-Gruppe	3.11	0.56	4.59	0.26	4.57	3.96	0.61	3.92
F-Wert	0.09		168.98***			52.04**		
Effektstärke			9.05			4.74		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=4); WW-Gruppe = Writer's Workshop (n=4); M^{ad} = um Prätest-Werte korrigierte Mittelwerte; *** p < .001, signifikanter Unterschied zwischen den Bedingungen; ** p < .01, signifikanter Unterschied zwischen den Bedingungen; * p < .05, signifikanter Unterschied zwischen den Bedingungen.

Die Messwiederholungsanalyse mit Geschichtenelemente als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) als Innersubjektfaktor, Bedingung (USAT, WW) als Zwischensubjektfaktor und Prätest als Kovariate erzielte einen signifikanten Haupteffekt Bedingung, $F(1, 5) = 124.61$, $p < .001$. Der Haupteffekt Zeit, $F(1, 5) = 1.62$, $p = .26$, die Interaktion Messzeitpunkt x Bedingung, $F(1, 5) = .85$, $p = .40$ sowie die Kovariate Prätest, $F(1, 5) = 2.03$, $p = .21$, erwiesen sich als nicht signifikant.

Wie in der Tabelle 31 ersichtlich, lieferten die Ergebnisse der Kovarianzanalysen mit Geschichtenelemente (jeweils Posttest und Follow-Up) als abhängige Variable, Bedingung (USAT, WW) als unabhängige Variable und dem Prätest als Kovariate ebenfalls

hochsignifikante Ergebnisse sowohl beim Posttest als auch beim Follow-Up zu Gunsten der USAT-Bedingung. Die Prätest-Werte variierten dagegen nicht signifikant in den beiden Bedingungen, $F(1, 5) = .09, p = .78$. Die Prätest-Variable als Kovariate erwies sich sowohl beim zweiten, $F(1, 5) = 1.06, p = .35$, als auch beim dritten Messzeitpunkt, $F(1, 5) = 1.72, p = .25$, als nicht signifikant. Eine grafische Darstellung der Unterschiede zwischen den beiden Bedingungen liefert Abbildung 49.

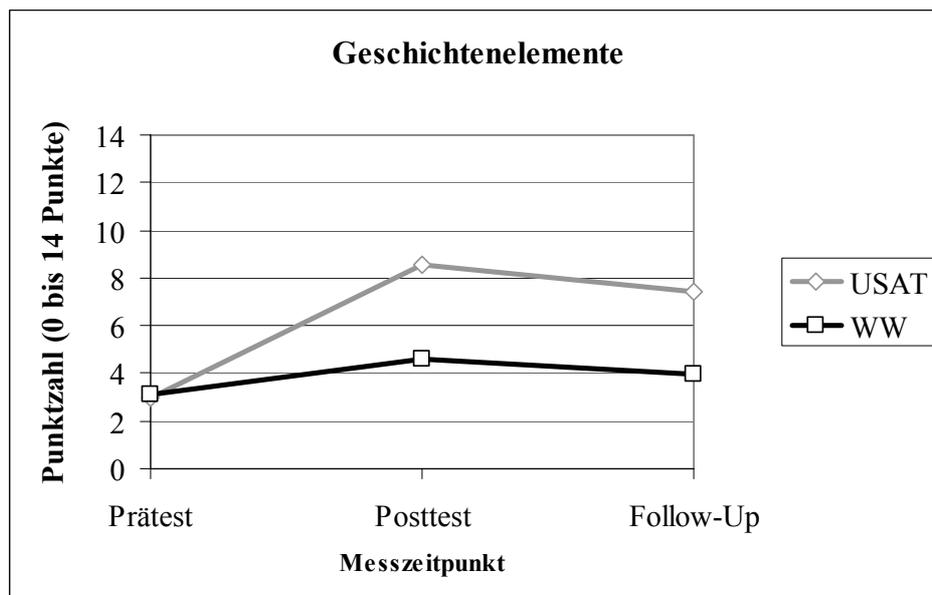


Abbildung 49: Geschichtenelemente in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).

6.3.3.2 Geschichtenqualität

Während es sich bei den Geschichtenelementen um ein strategienahes Maß handelt, erfasst die Geschichtenqualität nicht die im Training vermittelten Schreibstrategien. Dennoch zeigte sich ein statistisch bedeutsamer Zusammenhang zwischen den Geschichtenelementen und der Geschichtenqualität bei allen drei Messzeitpunkten (Prätest: $r = .85, p < .01$; Posttest: $r = .96, p < .01$; Follow-Up: $r = .91, p < .01$).

Nachfolgend wurden dieselben statistischen Analysen wie bei den Geschichtenelementen durchgeführt. In Tabelle 32 sind Mittelwerte, Standardabweichungen, die um die Prätest-Werte korrigierten Mittelwerte, F-Werte und Effektstärken für Geschichtenqualität getrennt nach Bedingung und Messzeitpunkt aufgelistet.

Tabelle 32: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Geschichtenqualität getrennt nach Bedingung und Messzeitpunkt (Studie I).

	Prätest		Posttest			Follow-Up		
	M	SD	M	SD	M ^{ad}	M	SD	M ^{ad}
USAT-Gruppe	2.06	0.24	3.41	0.18	3.41	3.16	0.32	3.16
WW-Gruppe	2.00	0.31	2.59	0.16	2.59	2.59	0.11	2.59
F-Wert	0.08		49.47***			10.93*		
Effektstärke			4.82			2.38		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=4); WW-Gruppe = Writer's Workshop (n=4); M^{ad} = um Prätest-Werte korrigierte Mittelwerte; *** p < .001, signifikanter Unterschied zwischen den Bedingungen; ** p < .01, signifikanter Unterschied zwischen den Bedingungen; * p < .05, signifikanter Unterschied zwischen den Bedingungen.

Eine Messwiederholungsanalyse mit Geschichtenqualität als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) als Innersubjektfaktor, Bedingung (USAT, WW) als Zwischensubjektfaktor und Prätest als Kovariate lieferte abgesehen von dem Haupteffekt Bedingung, $F(1, 5) = 28.89$, $p < .01$, keine signifikanten Effekte für den Messzeitpunkt, $F(1, 5) = 1.07$, $p = .35$, die Interaktion Messzeitpunkt x Bedingung, $F(1, 5) = 1.99$, $p = .22$, und den Prätest, $F(1, 5) = .01$, $p = .93$.

Anschließend wurden mittels einer Kovarianzanalyse mit Geschichtenqualität (jeweils Posttest und Follow-Up) als abhängige Variable, Bedingung (USAT, WW) als unabhängige Variable und dem Prätest als Kovariate die Leistungen in Bezug auf Geschichtenqualität der Klassen der USAT-Bedingung mit denen der WW-Bedingung verglichen. Während sich die Bedingungen beim Prätest nicht unterschieden, $F(1, 5) = .08$, $p = .79$, variierten die Klassen zwischen den beiden Bedingungen beim Posttest und beim Follow-Up signifikant voneinander. Die Klassen der USAT-Bedingung übertrafen dabei deutlich die Klassen der WW-Bedingung (siehe Tabelle 32). Die Kovariate Prätest erwies sich beim Posttest, $F(1, 5) = .95$, $p = .37$, und beim Follow-Up, $F(1, 5) = .25$, $p = .64$, als nicht signifikant. Der in Tabelle 24 festgestellte große Leistungsabfall der USAT-Klasse 3 vom Posttest zum Follow-Up könnte folglich auch für das Absinken des F-Wertes und der Effektstärke beim Follow-Up verantwortlich sein.

In Abbildung 50 ist ersichtlich, dass die Klassen der beiden Bedingungen den gleichen Ausgangspunkt besaßen, jedoch unterschiedlich viel von den Interventionen profitierten, so dass die Leistungen der beiden Bedingungen beim Posttest und beim Follow-Up erheblich divergierten.

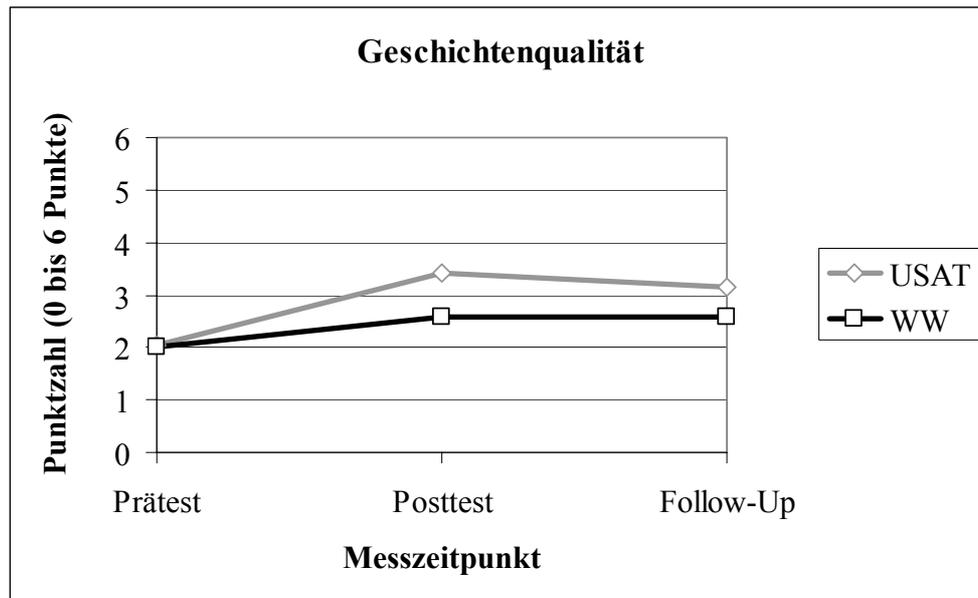


Abbildung 50: Geschichtenqualität in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).

6.3.3.3 Wortanzahl

Als weiteres Maß zur Beurteilung der Schreibleistung wurde die Anzahl der Wörter herangezogen. Die Mittelwerte, Standardabweichungen, die um Prätest-Werte korrigierten Mittelwerte, F-Werte und Effektstärken sind in der nachfolgenden Tabelle 33 dargestellt.

Tabelle 33: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Wortanzahl getrennt nach Bedingung und Messzeitpunkt (Studie I).

	Prätest		Posttest			Follow-Up		
	M	SD	M	SD	M ^{ad}	M	SD	M ^{ad}
USAT-Gruppe	97.52	21.24	167.82	35.51	164.85	154.91	42.44	149.31
WW-Gruppe	89.88	13.72	138.68	22.23	141.65	104.12	27.98	109.75
F-Wert	0.37		1.24			4.12		
Effektstärke			0.98			1.41		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=4); WW-Gruppe = Writer's Workshop (n=4); M^{ad} = um Prätest-Werte korrigierte Mittelwerte; *** p < .001, signifikanter Unterschied zwischen den Bedingungen; ** p < .01, signifikanter Unterschied zwischen den Bedingungen; * p < .05, signifikanter Unterschied zwischen den Bedingungen.

Zunächst wurde eine Messwiederholungsanalyse mit Wortanzahl als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) als Innersubjektfaktor, Bedingung (USAT, WW) als unabhängige Variable und dem Prätest-Wert als Kovariate gerechnet. Sämtliche Effekte waren nicht signifikant (Bedingung: $F(1, 5) = 3.08$, $p = .14$; Messzeitpunkt: $F(1, 5) = 2.56$, $p = .17$; Messzeitpunkt x Bedingung: $F(1, 5) = .76$, $p = .42$; Prätest: $F(1, 5) = 4.04$, $p = .10$).

Zu denselben Ergebnissen gelangten Kovarianzanalysen für Posttest und Follow-Up mit Wortanzahl als abhängige Variable, Bedingung (USAT, WW) als unabhängige Variable sowie Prätest als Kovariate. Die Bedingung erwies sich dabei als nicht signifikant (Prätest: $F(1, 5) = .37, p = .57$, Posttest: $F(1, 5) = 1.24, p = .32$; Follow-Up: $F(1, 5) = 4.12, p = .10$). Auch die Kovariate Prätest war weder beim Posttest, $F(1, 5) = 1.41, p = .29$, noch beim Follow-Up, $F(1, 5) = 5.85, p = .06$, signifikant.

In der nachfolgenden Abbildung 51 sind die Wortanzahl-Mittelwerte für die beiden Bedingungen zu den drei Messzeitpunkten veranschaulicht.

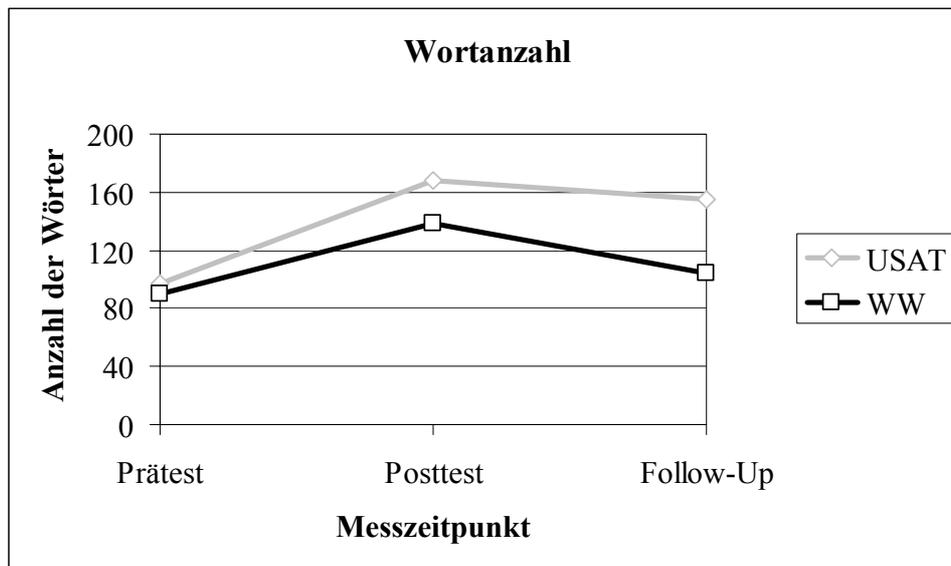


Abbildung 51: Wortanzahl in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).

6.3.4 Strategienaher Schreibmaße

In dem folgenden Abschnitt werden die Ergebnisse strategienaher Maße vorgestellt. Zunächst werden die Planungsaktivitäten untersucht. Anschließend werden die Überarbeitungsnotizen analysiert. Als letztes werden die Ergebnisse für Wissen erläutert.

6.3.4.1 Planen

Die Planungsaktivitäten der Schüler wurden gesammelt und hinsichtlich der Geschichtenelemente (Inhalt und Ausdruck) sowie der Struktur ausgewertet. Die beiden Interventionsprogramme beinhalteten explizite Planungsstrategien, so dass es sich hierbei um ein strategienahes Maß handelt.

Es zeigte sich ein substantieller Zusammenhang zwischen Planen Geschichtenelemente und Planen Struktur sowohl beim zweiten ($r = .97, p < .01$) als auch beim dritten Messzeitpunkt ($r = .98, p < .01$). Beim Prätest bestand jedoch kein signifikanter

Zusammenhang zwischen Planen Geschichtenelemente und Planen Struktur ($r = .51, p = .20$). Dies ist dadurch zu erklären, dass die Schüler in beiden Bedingungen am Anfang wenig bis gar nicht planten. Im Verlauf der beiden Trainings lernten die Schüler verschiedene Strategien zum Planen von Geschichten kennen, wobei die Interventionen sowohl das Planen der Geschichtenelemente und das Planen der Struktur beeinflussten.

In der folgenden Tabelle 34 sind die Mittelwerte, Standardabweichungen und die um Prätest-Werte korrigierten Mittelwerte, F-Werte und Effektstärken für die beiden Bedingungen zu den drei Messzeitpunkten aufgelistet.

Tabelle 34: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Planen getrennt nach Bedingung und Messzeitpunkt (Studie I).

	Prätest		Posttest			Follow-Up		
	M	SD	M	SD	M ^{ad}	M	SD	M ^{ad}
	Planen (Geschichtenelemente)							
USAT-Gruppe	1.41	0.27	5.84	1.01	5.84	5.73	0.80	5.73
WW-Gruppe	1.39	0.34	1.22	0.51	1.22	1.43	0.45	1.42
F-Wert	0.01		57.19**			89.36***		
Effektstärke			5.77			6.63		
	Planen (Struktur)							
USAT-Gruppe	1.35	0.20	3.32	0.32	3.27	3.60	0.72	3.56
WW-Gruppe	2.18	0.51	1.31	0.43	1.37	1.10	0.21	1.14
F-Wert	3.08*		16.92**			13.54*		
Effektstärke			5.30			4.71		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=4); WW-Gruppe = Writer's Workshop (n=4); M^{ad} = um Prätest-Werte korrigierte Mittelwerte; *** $p < .001$, signifikanter Unterschied zwischen den Bedingungen; ** $p < .01$, signifikanter Unterschied zwischen den Bedingungen; * $p < .05$, signifikanter Unterschied zwischen den Bedingungen.

Zu Beginn wurden zwei Messwiederholungsanalysen mit jeweils Planen Geschichtenelemente und Planen Struktur als abhängige Variable, Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, WW) als Zwischensubjektfaktor und der jeweiligen Prätest-Variablen als Kovariate durchgeführt. Für Planen Geschichtenelemente und Planen Struktur ergaben sich dieselben Ergebnisse. Nur der Faktor Bedingung (Geschichtenelemente: $F(1, 5) = 72.63, p < .001$; Struktur: $F(1, 5) = 19.84, p < .01$) erwies sich als signifikant. Der Messzeitpunkt (Geschichtenelemente: $F(1, 5) = 1.20, p = .32$; Struktur: $F(1, 5) = .00, p = 1.00$), die Interaktion Messzeitpunkt x Bedingung (Geschichtenelemente: $F(1, 5) = 1.46, p = .28$; Struktur: $F(1, 5) = .74, p = .43$) sowie die

Kovariate Prätest (Geschichtenelemente: $F(1, 5) = .41, p = .55$; Struktur: $F(1, 5) = .06, p = .81$) waren nicht signifikant.

Kovarianzanalysen mit Planen Geschichtenelemente als abhängige Variable (jeweils getrennt Posttest und Follow-Up), Bedingung (USAT, WW) als unabhängige Variable und dem Prätest-Wert als Kovariate führten beim Posttest und beim Follow-Up zu einem hoch signifikanten Bedingungs-Haupteffekt zu Gunsten der USAT-Bedingung (siehe Tabelle 34) und zu nicht signifikanten Effekten der Kovariaten Prätest (Posttest: $F(1, 5) = .11, p = .75$; Follow-Up: $F(1, 5) = 1.06, p = .35$). Analog durchgeführte Kovarianzanalysen mit Planen Struktur als abhängige Variable (jeweils getrennt Posttest und Follow-Up) lieferten ebenfalls hohe signifikante F-Werte für den Haupteffekt Bedingung (siehe Tabelle 34) und keine Effekte für die Kovariate Prätest (Posttest: $F(1, 5) = .08, p = .79$; Follow-Up: $F(1, 5) = .03, p = .87$). Während sich die Prätest-Werte der Bedingungen bezüglich Planen Geschichtenelemente nicht bedeutsam unterschieden, $F(1, 5) = .01, p = .93$, war der Unterschied bezüglich Planen Struktur signifikant (siehe Tabelle 34).

Abbildung 52 bietet eine grafische Darstellung der Ergebnisse für Planen Geschichtenelemente und Planen Struktur getrennt nach Bedingung und Messzeitpunkt. In der Abbildung ist erkenntlich, dass sich die Klassen der WW-Bedingung hinsichtlich der Planungsaktivitäten nicht vom Ausgangsniveau verbesserten, während die Klassen der USAT-Bedingung deutlich bessere Leistungen sowohl in Bezug zu Planen Geschichtenelemente als auch bezüglich Planen Struktur beim Posttest und beim Follow-Up zeigten.

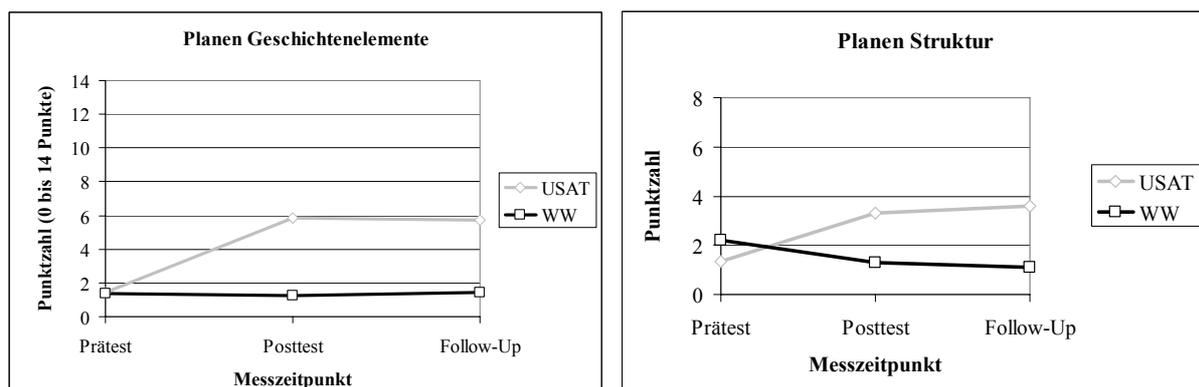


Abbildung 52: Planen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).

6.3.4.2 Revisionen

Als weiteres strategienahes Maß wurden die Einfügungen und Verbesserungen der Schüler gezählt. In den Trainings wurden die Schüler angeleitet und darin unterstützt, den selbst

verfassten Text zu überarbeiten. Der Schwerpunkt der Interventionsprogramme lag jedoch beim Planen.

Tabelle 35 enthält die Mittelwerte, Standardabweichungen und die um Prätest-Werte korrigierten Mittelwerte, F-Werte und Effektstärken für die beiden Bedingungen zu den drei Messzeitpunkten.

Tabelle 35: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Revisionen getrennt nach Bedingung und Messzeitpunkt (Studie I).

	Prätest		Posttest			Follow-Up		
	M	SD	M	SD	M ^{ad}	M	SD	M ^{ad}
USAT-Gruppe	0.14	0.10	0.86	0.43	0.82	0.52	0.25	0.53
WW-Gruppe	0.08	0.06	0.36	0.20	0.40	0.43	0.20	0.43
F-Wert	1.14		2.48			0.28		
Effektstärke			1.49			0.40		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=4); WW-Gruppe = Writer's Workshop (n=4); M^{ad} = um Prätest-Werte korrigierte Mittelwerte; *** p< .001, signifikanter Unterschied zwischen den Bedingungen; ** p< .01, signifikanter Unterschied zwischen den Bedingungen; * p< .05, signifikanter Unterschied zwischen den Bedingungen.

Eine Messwiederholungsanalyse mit Revisionen als abhängige Variable, Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, WW) als Zwischensubjektfaktor und dem Prätest als Kovariate lieferte keine signifikanten Effekte (Bedingung: $F(1, 5) = 1.38, p = .29$; Messzeitpunkt: $F(1, 5) = .07, p = .80$; Messzeitpunkt x Bedingung: $F(1, 5) = 5.61, p = .09$; Prätest: $F(1, 5) = .16, p = .71$).

Kovarianzanalysen mit dem Prätest als Kovariate, Bedingung (USAT, WW) als unabhängige sowie die Post- bzw. Follow-Up-Werte jeweils als abhängige Variable führten ebenfalls zu keinen signifikanten Bedingungseffekten, so dass sich die beiden Bedingungen hinsichtlich der Verbesserungen weder beim Prätest, $F(1, 5) = 1.14, p = .33$, noch beim Posttest, $F(1, 5) = 2.48, p = .18$, und beim Follow-Up, $F(1, 5) = .28, p = .62$, unterschieden. Der Einfluss der Kovariaten Prätest war weder beim Posttest, $F(1, 5) = .58, p = .48$, noch beim Follow-Up, $F(1, 5) = .02, p = .89$, signifikant. Der nicht signifikante Bedingungseffekt beim Posttest könnte durch die Ausreißer-Werte der Klassen 1 und 3 erklärt werden (siehe Tabelle 27). Es ergeben sich jedoch weder beim Posttest noch beim Follow-Up signifikante Bedingungseffekte, wenn die Klasse 3 von den Analysen ausgeschlossen wird. Betrachtet man die Mittelwerte genauer, fällt auf, dass in beiden Bedingungen sehr wenig bzw. größtenteils gar nicht überarbeitet wurde, da sämtliche Mittelwerte kleiner als 1 sind, d.h. im Mittel verbesserten alle Klassen weniger als einmal pro Messzeitpunkt (siehe Tabelle 27).

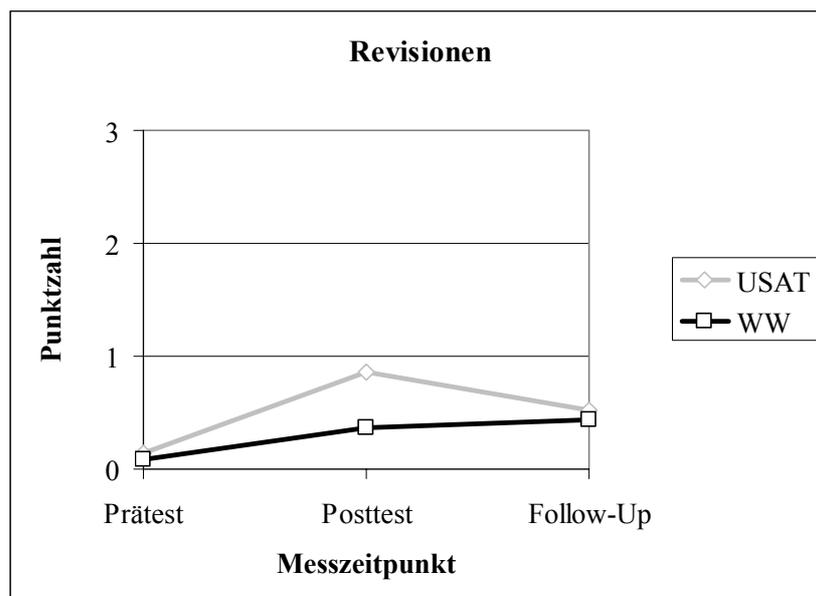


Abbildung 53: Revisionen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).

Abbildung 53 stellt die Mittelwerte für beide Bedingungen zu den drei Messzeitpunkten grafisch dar. Es ist zu erkennen, dass beim Posttest tendenziell die Klassen der USAT-Bedingung besser abschnitten als die Klassen der WW-Bedingung, wobei dieser Unterschied beim Follow-Up wieder verschwand.

6.3.4.3 Schreibbezogenes Wissen

Als nächstes wurde das schreibbezogene Wissen, welches mittels eines Fragebogens erfasst wurde, analysiert. Tabelle 36 enthält die Mittelwerte, Standardabweichungen, die um Prätest-Werte korrigierten Mittelwerte, F-Werte und Effektstärken für beide Bedingungen zu den drei Messzeitpunkten.

Tabelle 36: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Wissen getrennt nach Bedingung und Messzeitpunkt (Studie I).

	Prätest		Posttest			Follow-Up		
	M	SD	M	SD	M ^{ad}	M	SD	M ^{ad}
USAT-Gruppe	1.72	0.29	5.25	0.91	5.38	5.00	0.71	5.11
WW-Gruppe	2.39	0.73	3.05	0.58	2.92	2.84	0.51	2.73
F-Wert	2.94		12.59*			18.14**		
Effektstärke			2.88			3.49		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=4); WW-Gruppe = Writer's Workshop (n=4); M^{ad} = um Prätest-Werte korrigierte Mittelwerte; *** p< .001, signifikanter Unterschied zwischen den Bedingungen; ** p< .01, signifikanter Unterschied zwischen den Bedingungen; * p< .05, signifikanter Unterschied zwischen den Bedingungen.

Eine Messwiederholungsanalyse mit Wissen als abhängige Variable, Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, WW) als Zwischensubjektfaktor und der Prätest-Wert als Kovariate ergab, abgesehen von dem Haupteffekt Bedingung, $F(1, 5) = 17.86$, $p < .01$, keine signifikanten Effekte für den Messzeitpunkt, $F(1, 5) = .00$, $p = .95$, die Interaktion Messzeitpunkt x Bedingung, $F(1, 5) = .03$, $p = .88$, und den Prätest, $F(1, 5) = .53$, $p = .50$.

Kovarianzanalysen mit dem Prätest-Wert als Kovariate, Bedingung (USAT, WW) als unabhängige Variable und jeweils dem Post- bzw. Follow-Up-Wert als abhängige Variable lieferten sowohl beim Posttest als auch beim Follow-Up einen signifikanten Bedingungseffekt zu Gunsten der USAT-Bedingung (siehe Tabelle 36). Die Prätest-Werte unterschieden sich nicht signifikant zwischen den beiden Bedingungen, $F(1, 5) = 2.94$, $p = .14$. Während die Unterschiede zum zweiten Messzeitpunkt bereits auf dem .05-Niveau signifikant waren, vergrößerten sich die Unterschiede beim Follow-Up und wurden auf dem .01-Niveau signifikant. Sowohl beim zweiten als auch beim dritten Messzeitpunkt erwies sich die Kovariate Prätest als nicht signifikant (Posttest: $F(1, 5) = .45$, $p = .53$; Follow-Up: $F(1, 5) = .43$, $p = .54$).

Auch wenn in der WW-Bedingung die Klassen 6 und 7 höhere Leistungen beim Posttest und beim Follow-Up im Vergleich zu den anderen Klassen erzielten, lagen sämtliche USAT-Mittelwerte über den WW-Mittelwerten (siehe Tabelle 28). Folglich sind die Klassenunterschiede für den Bedingungseffekt wenig relevant, da insgesamt die USAT-Klassen den WW-Klassen überlegen waren. Dies bestätigt Abbildung 54, die die Mittelwerte für Wissen getrennt nach Messzeitpunkt und Bedingung darstellt.

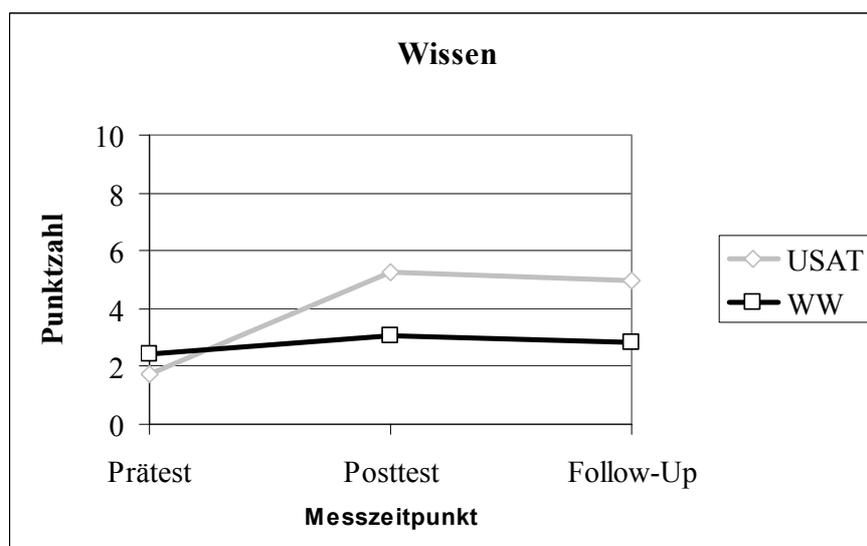


Abbildung 54: Wissen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).

6.3.5 Transferleistungen

6.3.5.1 Erlebniserzählungen

Zur Erfassung der Fähigkeit, die im Training erlernten Strategien auf eine ungelernete Aufgabe zu übertragen, schrieben die Schüler zu drei verschiedenen Stimuli eine Erlebniserzählung. Ebenso wie bei den Bildergeschichten wurden die Anzahl der Geschichtenelemente sowie die Geschichtenqualität der Erlebniserzählungen ausgewertet. Tabelle 37 enthält die Mittelwerte, die um Prätest-Werte korrigierten Mittelwerte, Standardabweichungen, F-Werte und Effektstärken für Erlebniserzählungen Geschichtenelemente und Geschichtenqualität.

Tabelle 37: Mittelwerte, Standardabweichungen, um Prätest-Werte korrigierte Mittelwerte, F-Werte und Effektstärken für Erlebniserzählungen getrennt nach Bedingung und Messzeitpunkt (Studie I).

	Prätest		Posttest			Follow-Up		
	M	SD	M	SD	M ^{ad}	M	SD	M ^{ad}
	Geschichtenelemente (Erlebniserzählung)							
USAT-Gruppe	3.61	0,53	7.13	1.16	7.11	6.16	1.03	6.15
WW-Gruppe	3.58	0.47	4.01	0.65	4.03	3.93	0.57	3.94
F-Wert	0.01		29.48**			14.82*		
Effektstärke			3.32			2.68		
	Geschichtenqualität (Erlebniserzählung)							
USAT-Gruppe	1.69	0.18	2.77	0.31	2.72	2.68	0.31	2.60
WW-Gruppe	1.47	0.13	1.90	0.11	1.95	2.11	0.18	2.19
F-Wert	4.00		12.18*			3.34		
Effektstärke			3.74			2.25		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=4); WW-Gruppe = Writer's Workshop (n=4); M^{ad} = um Prätest-Werte korrigierte Mittelwerte; *** p< .001, signifikanter Unterschied zwischen den Bedingungen; ** p< .01, signifikanter Unterschied zwischen den Bedingungen; * p< .05, signifikanter Unterschied zwischen den Bedingungen.

Als erstes wurde für Erlebniserzählung Geschichtenelemente und Geschichtenqualität jeweils eine Messwiederholungsanalyse gerechnet. Die Messwiederholungsanalyse mit Erlebniserzählung Geschichtenelemente als abhängige Variable, Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, WW) als Zwischensubjektfaktor und Prätest als Kovariate ergab einen signifikanten Haupteffekt Bedingung, $F(1, 5) = 22.01$, $p < .01$, und eine signifikante Interaktion Messzeitpunkt x Bedingung, $F(1, 5) = 24.38$, $p < .01$. Der Messzeitpunkt, $F(1, 5) = 1.83$, $p = .24$, sowie der Prätest, $F(1, 5) = 2.29$, $p = .19$, waren nicht signifikant.

Im Gegensatz dazu erzeugte die Messwiederholungsanalyse mit Erlebniserzählung Geschichtenqualität als abhängige Variable, Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, WW) als Zwischensubjektfaktor und Prätest als Kovariate nur einen signifikanten Haupteffekt Bedingung, $F(1, 5) = 8.66, p < .05$, und keine signifikanten Ergebnisse für die Interaktion Messzeitpunkt x Bedingung, $F(1, 5) = 3.30, p = .13$, den Messzeitpunkt, $F(1, 5) = .23, p = .65$, sowie den Prätest, $F(1, 5) = .95, p = .37$.

Zur Berechnung der Unterschiede zwischen den beiden Bedingungen wurden zusätzlich Kovarianzanalysen für Posttest und Follow-Up getrennt durchgeführt. Zunächst wurde ein kovarianzanalytisches Modell mit Erlebniserzählungen Geschichtenelemente (jeweils getrennt für Posttest und Follow-Up) als abhängige Variable, Bedingung (USAT, WW) als unabhängige Variable und dem Prätest-Wert als Kovariate gerechnet. Während die Unterschiede zwischen den Bedingungen beim Prätest noch nicht signifikant waren, $F(1, 5) = .01, p = .94$, ergab sich ein signifikanter Bedingungshaupteffekt beim Posttest und beim Follow-Up (siehe Tabelle 37). Der Einfluss der Kovariaten Prätest war sowohl beim Posttest, $F(1, 5) = 3.34, p = .13$, als auch beim Follow-Up, $F(1, 5) = 1.36, p = .30$, nicht signifikant.

Dasselbe Vorgehen wurde für die Analyse der Effekte für Erlebniserzählungen Geschichtenqualität gewählt. Die Kovarianzanalysen mit Erlebniserzählungen Geschichtenqualität Posttest bzw. Follow-Up als abhängige Variable, Bedingung (USAT, WW) als unabhängige Variable und dem Prätest als Kovariate lieferten einen signifikanten Haupteffekt für die Bedingung beim Posttest (siehe Tabelle 37), aber nicht beim Follow-Up, $F(1, 5) = 3.34, p = .13$. Die Bedingungsunterschiede beim ersten Messzeitpunkt waren ebenfalls nicht signifikant, $F(1, 5) = 4.00, p = .09$. Die Kovariate Prätest war weder beim Posttest, $F(1, 5) = .42, p = .55$, noch beim Follow-Up, $F(1, 5) = 1.21, p = .32$, signifikant. Abbildung 55 zeigt eine grafische Darstellung der Ergebnisse sowohl für Erlebniserzählungen Geschichtenelemente als auch für die Geschichtenqualität zu den drei Messzeitpunkten.

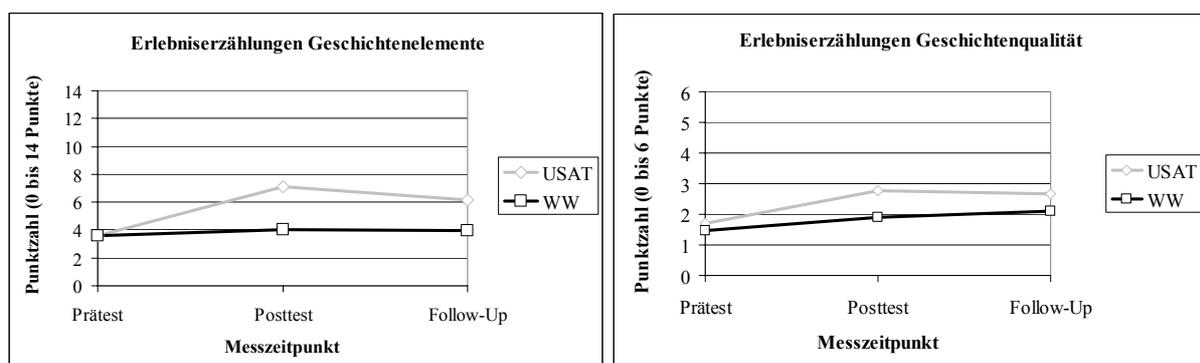


Abbildung 55: Erlebniserzählungen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie I).

6.3.6 Maße der praktischen Signifikanz

In dem vorherigen Abschnitt wurden die Ergebnisse der statistischen Hauptanalysen berichtet, welche auf Klassenebene erfolgten. Die nach Cohen's d (Cohen, 1988) berechneten und um Prätestunterschiede korrigierten Effektstärken zur Beurteilung der praktischen Signifikanz der Unterschiede zwischen den beiden Bedingungen sind in den Tabellen der jeweiligen abhängigen Maße aufgelistet und variierten abgesehen von Revisionen (Posttest: 1.25 und Follow-Up: .40) zwischen 2.25 und 9.05. Nach den allgemeinen Regeln zur Bewertung von Effektstärken, wie sie z.B. Graham und Perin (2007) berichten, gelten Effektstärken größer als .80 als groß, größer als .50 als mittel und größer als .20 als klein. Kleiner als .20 bedeutet kein Effekt. Da die Analyseeinheit jedoch Klasse war, führte die konventionelle Schätzung der Effektstärken zu inkorrekten Standardfehlern, was die sehr hohen Effektstärken erklärt.

Zur besseren Vergleichbarkeit der dargestellten Effekte mit anderen Studien werden daher in Tabelle 38 zusätzlich zu den Effektstärken auf Klassenebene die auf Schülerebene angegeben.

Tabelle 38: Effektstärken nach Cohen's d für Posttest und Follow-Up mit jeweils Klasse und Schüler als Analyseeinheit (Studie I).

	Klasse als Analyseeinheit		Schüler als Analyseeinheit	
	Posttest	Follow-Up	Posttest	Follow-Up
Geschichtenelemente	9.05	4.74	1.81	1.57
Geschichtenqualität	4.82	2.38	1.08	0.82
Wortanzahl	0.98	1.41	0.36	0.83
Planen Geschichtenelemente	5.77	6.63	2.01	1.95
Planen Struktur	5.30	4.71	1.15	1.37
Revisionen	1.49	0.40	0.50	0.09
Wissen	2.88	3.49	1.09	1.17
Erlebniserzählungen Geschichtenelemente	3.32	2.68	1.38	1.16
Erlebniserzählungen Geschichtenqualität	3.74	2.25	1.04	0.70

Wenn Schüler als Analyseeinheit gewählt wird, liegen die Effektstärken abgesehen von Revisionen zwischen .70 und 2.01 (siehe Tabelle 38). Gemäß Graham und Perin (2007) sind somit die Effektstärken als groß bis sehr groß zu bewerten. Für Revisionen ergab sich ein mittlerer Effekt zu Gunsten der USAT-Bedingung beim Posttest, jedoch keinen Effekt beim

Follow-Up. Insgesamt lassen die Effektstärken die Schlussfolgerung zu, dass das USAT-Programm deutlich effektiver war als das WW-Programm.

6.3.7 Differentielle Betrachtung der Bedingungseffekte in Abhängigkeit von den Kontrollvariablen

Im nächsten Abschnitt werden die Effekte der beiden Bedingungen (USAT, WW) in Abhängigkeit von den Kontrollvariablen Geschlecht, Rechtschreibung und Muttersprache Deutsch gerechnet. Es wurden dazu jeweils zwei Untergruppen der Kontrollvariablen gebildet und überprüft, ob die Bedingungseffekte für die Subgruppen weiterhin bestehen blieben. Die Kontrollvariable Nationalität wurde ausgespart, da von der gesamten Stichprobe nur acht Schüler eine andere Nationalität als deutsch angaben. Fünf der Schüler waren in der WW- und drei in der USAT-Bedingung.

6.3.7.1 Geschlecht

Nachdem für die einzelnen abhängigen Maße bereits überprüft wurde, ob sich signifikante Bedingungseffekte ergaben, sollen nun die Effekte für Mädchen und Jungen separat analysiert werden.

Zu diesem Zweck wurden für jede abhängige Variable zwei neue Variablen, eine für Mädchen und eine für Jungen, gebildet. Dazu wurden die Daten auf Schülerebene getrennt für Mädchen und Jungen über die Klassen hinweg aggregiert.

In Tabelle 39 sind die Mittelwerte und Standardabweichungen für alle Klassen der USAT- und der WW-Bedingung getrennt für Mädchen und Jungen und für die drei Messzeitpunkte aufgelistet.

Tabelle 39: Mittelwerte und Standardabweichungen für Mädchen und Jungen getrennt nach Messzeitpunkt und Bedingung (Studie I).

	USAT				WW			
	Mädchen		Jungen		Mädchen		Jungen	
	M	SD	M	SD	M	SD	M	SD
Geschichtenelemente								
Prätest	3.16	0.61	2.73	0.75	3.29	0.72	2.85	0.94
Posttest	9.49	0.70	7.68	0.59	4.91	0.67	4.21	0.37
Follow-Up	8.42	0.98	6.55	0.65	4.65	0.54	3.24	0.78

Geschichtenqualität								
Prätest	2.23	0.25	1.88	0.30	2.11	0.25	1.88	0.39
Posttest	3.88	0.19	2.98	0.16	2.68	0.25	2.49	0.06
Follow-Up	3.50	0.34	2.83	0.28	2.79	0.10	2.41	0.15
Wortanzahl								
Prätest	105.82	22.06	87.97	23.73	93.98	17.65	84.41	15.07
Posttest	191.73	33.86	146.64	35.73	139.31	30.69	136.72	33.31
Follow-Up	177.41	35.09	133.89	48.19	120.76	36.42	85.87	19.31
Planen (Geschichtenelemente)								
Prätest	1.57	0.35	1.24	0.31	1.51	0.10	1.31	0.59
Posttest	6.78	0.75	5.09	1.70	1.35	0.48	1.11	0.54
Follow-Up	6.49	0.31	5.11	1.31	1.35	0.44	1.50	0.46
Planen (Struktur)								
Prätest	1.49	0.50	1.13	0.53	2.25	0.12	2.17	0.94
Posttest	3.74	0.66	2.98	0.50	1.46	0.55	1.16	0.36
Follow-Up	4.12	0.67	3.15	0.83	1.14	0.62	1.09	0.16
Revisionen								
Prätest	0.16	0.10	0.12	0.14	0.11	0.08	0.05	0.06
Posttest	1.21	0.72	0.48	0.19	0.36	0.30	0.34	0.18
Follow-Up	0.75	0.45	0.30	0.24	0.67	0.32	0.19	0.13
Wissen								
Prätest	1.69	0.26	1.72	0.40	2.27	0.50	2.49	0.95
Posttest	5.33	0.76	5.20	1.38	3.06	0.59	3.06	0.59
Follow-Up	5.27	0.72	4.74	0.63	2.86	0.80	2.87	0.40
Geschichtenelemente (Erlebniserzählungen)								
Prätest	3.68	0.56	3.51	0.62	3.72	0.52	3.43	0.50
Posttest	7.73	1.29	6.44	1.45	4.28	0.63	3.72	0.68
Follow-Up	6.87	0.87	5.58	1.29	4.27	0.78	3.60	0.38
Geschichtenqualität (Erlebniserzählungen)								
Prätest	1.70	0.23	1.67	0.17	1.50	0.25	1.43	0.20
Posttest	3.10	0.44	2.44	0.22	2.07	0.34	1.76	0.12
Follow-Up	2.89	0.30	2.52	0.47	2.16	0.24	2.08	0.35

Anschließend wurden (Messwiederholungs-) Kovarianzanalysen mit den prozess- und produktbezogenen Schreibmaßen als abhängige Variable, Messzeitpunkt (Posttest, Follow-

Up) und Geschlecht (Mädchen, Jungen) als Innersubjektfaktoren, den zweistufigen Bedingungsfaktor (USAT, WW) als Zwischensubjektfaktor und den jeweiligen Prätest-Werten (Mädchen, Jungen) als Kovariaten durchgeführt. Im Gegensatz zur Analyse mit Schüler als Analyseeinheit ist die Variable Geschlecht mit der Analyseeinheit Klasse ein Innersubjektfaktor, da jede Klasse sowohl männliche als auch weibliche Schüler umfasst und folglich zwei Ausprägungen des Faktors Geschlecht besitzt. Um den separaten Effekt der Mädchen bzw. Jungen auf die abhängigen Variablen zu bestimmen, muss folglich Geschlecht als Innersubjektfaktor gerechnet werden. Auf Schülerebene konnte das Geschlecht nur als Zwischensubjektfaktor berücksichtigt werden. Die nachfolgende Tabelle 40 fasst die Ergebnisse der Analysen zusammen.

Tabelle 40: Ergebnisse der Analysen zur Prüfung des Faktors Geschlecht als Moderator (Studie I).

	Bed. x Gesch.	Bed.	Gesch.	MZP	Prätest Jungen	Prätest Mäd.
Geschichtenelemente		X				
Geschichtenqualität	X	X				
Wortanzahl						
Planen Inhalt		X				
Planen Struktur		X	X			
Revisionen						
Wissen		X				
Erlebniserzählungen (GE)	X	X	X			
Erlebniserzählungen (NQ)		X				

Anmerkungen: X = signifikanter Effekt ($p < .05$); Bed. x Gesch. = Interaktion Bedingung x Geschlecht; Bed. = Haupteffekt Zwischensubjektfaktor Bedingung; Gesch. = Haupteffekt Innersubjektfaktor Geschlecht; MZP = Haupteffekt Innersubjektfaktor Messzeitpunkt; Prätest Jungen = Effekt der Kovariaten Prätest Jungen; Prätest Mäd. = Effekt der Kovariaten Prätest Mädchen.

Insgesamt bestätigen die Messwiederholungskovarianzanalysen die Ergebnisse der vorangegangenen Analysen. Auch unter separater Berücksichtigung von Mädchen und Jungen ergeben sich dieselben Bedingungseffekte: Abgesehen von Wortanzahl und Revisionen zeigte sich für alle anderen abhängigen Variablen ein signifikanter Bedingungseffekt zu Gunsten der USAT- Bedingung ($p < .05$). Darüber hinaus wurden für einzelne abhängige Variablen signifikante Haupteffekte des Faktors Geschlecht und signifikante Interaktionen Bedingung x Geschlecht festgestellt.

Die Berechnung einer Kovarianzanalysen mit Planen Struktur als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Geschlecht (Mädchen, Jungen) als Innersubjektfaktoren, dem Zwischensubjektfaktor Bedingung (USAT, WW) und den Prätest-Werten für Mädchen und Jungen als Kovariaten führte zu einem signifikanten Effekt des Faktors Geschlecht, $F(1, 4) = 11.54, p < .05$. Daher unterschieden sich die Mädchen und Jungen signifikant voneinander, wobei die Unterschiede zwischen den Bedingungen, $F(1, 4) = 24.10, p < .01$, wesentlich größer sind als die Unterschiede zwischen Mädchen und Jungen innerhalb einer Bedingung.

Auch eine Kovarianzanalyse mit Erlebniserzählungen Geschichtenelemente als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Geschlecht (Mädchen, Jungen) als Innersubjektfaktoren, dem Faktor Bedingung (USAT, WW) als Zwischensubjektfaktor und den jeweiligen Prätest-Werten für Mädchen und Jungen als Kovariaten lieferte einen signifikanten Haupteffekt des Faktors Geschlecht, $F(1, 4) = 12.43, p < .05$. Gleichzeitig zeigte sich eine signifikante Interaktion Bedingung x Geschlecht, $F(1, 4) = 32.93, p < .01$. Insgesamt schrieben die Mädchen in beiden Bedingungen inhaltlich vollständigere Erlebnisgeschichten als die Jungen, wobei der Unterschied in der USAT- Bedingung größer war als in der WW-Bedingung. Die Unterschiede zwischen den Bedingungen, $F(1, 4) = 19.10, p < .05$, sind darüber hinaus größer als innerhalb.

Eine weitere signifikante Interaktion Bedingung x Geschlecht, $F(1, 4) = 102.94, p < .01$, ergab sich nach Berechnung einer Kovarianzanalyse mit Geschichtenqualität als abhängige Variable, dem Messzeitpunkt (Posttest, Follow-Up) und Geschlecht (Mädchen, Jungen) als Innersubjektfaktoren, dem Faktor Bedingung (USAT, WW) als Zwischensubjektfaktor und den Prätest-Werten für Mädchen und Jungen als Kovariaten. Demnach vergrößerte das USAT-Programm im Vergleich zur WW-Bedingung die Unterschiede zwischen Mädchen und Jungen stärker. Mädchen der USAT-Bedingung schnitten deutlich besser als Jungen der USAT-Bedingung ab. Solche großen Unterschiede waren in der WW-Bedingung nicht zu finden.

6.3.7.2 Rechtschreibung

Analog zum Geschlecht werden im nächsten Abschnitt die Effekte für die rechtschreibschwachen und rechtschreibstarken Schüler differenziert betrachtet.

Dazu wurden zunächst die Schüler beider Bedingungen aufgrund ihrer Rechtschreibleistungen einer von zwei Gruppen zugeordnet. Schüler, die im DRT 4 (Grund et al., 2004) einen Wert kleiner als dem 50. Perzentil, welches einem Punktwert von 47 entspricht, erreicht hatten, wurde eine 1 zugeordnet. Schüler, die im DRT 4 (Grund et al.,

2004) einen Wert größer als dem 50. Perzentil erzielt hatten, bekamen die Ziffer 2 zugeordnet. Auf diese Weise wurde eine neue zweistufige Variable Rechtschreibung gebildet. In einem nächsten Schritt wurde der Schülerdatensatz hinsichtlich Rechtschreibung niedrig (RS niedrig) und Rechtschreibung hoch (RS hoch) getrennt und über die Klassen hinweg aggregiert, so dass für jede abhängige Variable ein RS niedrig- und ein RS hoch-Wert gebildet wurde.

Tabelle 41 enthält die Mittelwerte und Standardabweichungen der abhängigen Schreibmaße getrennt für Schüler mit niedrigen und mit hohen Rechtschreibfähigkeiten, für alle drei Messzeitpunkte und für beide Bedingungen.

Tabelle 41: Mittelwerte und Standardabweichungen für rechtschreibschwache und rechtschreibstarke Schüler getrennt nach Messzeitpunkt und Bedingung (Studie I).

	USAT				WW			
	RS niedrig		RS hoch		RS niedrig		RS hoch	
	M	SD	M	SD	M	SD	M	SD
Geschichtenelemente								
Prätest	2.57	0.45	3.36	0.26	2.89	0.43	3.30	0.94
Posttest	7.38	0.83	9.59	0.54	4.25	0.50	5.28	0.98
Follow-Up	6.40	0.96	8.27	1.04	3.89	1.01	4.35	0.62
Geschichtenqualität								
Prätest	1.95	0.33	2.15	0.23	1.92	0.26	2.10	0.34
Posttest	2.93	0.29	3.84	0.24	2.51	0.22	2.73	0.15
Follow-Up	2.90	0.38	3.42	0.40	2.60	0.16	2.68	0.27
Wortanzahl								
Prätest	92.18	27.38	100.96	21.43	85.23	15.43	98.37	13.31
Posttest	151.29	40.67	181.02	35.32	129.53	20.19	154.58	38.42
Follow-Up	140.04	46.85	170.03	39.23	99.36	24.13	110.44	30.22
Planen (Geschichtenelemente)								
Prätest	1.21	0.20	1.59	0.38	1.26	0.35	1.56	0.41
Posttest	4.76	1.25	6.74	1.15	.94	0.46	1.59	0.57
Follow-Up	4.86	1.09	6.45	0.96	1.06	0.37	1.81	0.44
Planen (Struktur)								
Prätest	1.09	0.24	1.58	0.37	2.14	0.55	2.22	0.57
Posttest	3.10	0.38	3.57	0.46	0.99	0.45	1.78	0.52
Follow-Up	3.25	0.59	3.93	1.10	0.94	0.20	1.29	0.28

Revisionen								
Prätest	0.16	0.19	0.13	0.13	0.11	0.09	0.07	0.08
Posttest	0.73	0.34	1.00	0.55	0.25	0.09	0.56	0.42
Follow-Up	0.51	0.18	0.59	0.57	0.37	0.18	0.64	0.54
Wissen								
Prätest	1.65	0.54	1.83	0.14	2.10	0.51	2.68	0.83
Posttest	5.02	1.13	5.42	0.43	2.97	0.70	3.10	0.82
Follow-Up	5.03	0.92	4.93	0.53	2.62	0.46	3.12	0.50
Geschichtenelemente (Erlebniserzählungen)								
Prätest	3.39	0.64	3.88	0.66	3.50	0.61	3.98	1.25
Posttest	6.57	1.43	7.78	1.43	3.85	1.10	4.50	0.58
Follow-Up	5.49	0.65	6.69	1.41	3.71	0.82	4.31	0.28
Geschichtenqualität (Erlebniserzählungen)								
Prätest	1.62	0.28	1.75	0.12	1.46	0.09	1.53	0.39
Posttest	2.40	0.21	3.12	0.44	1.98	0.44	1.96	0.21
Follow-Up	2.54	0.38	2.78	0.27	2.11	0.34	2.21	0.24

Als nächstes wurden Kovarianzanalysen mit den schreibrelevanten Maßen jeweils als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Rechtschreibung (RS niedrig, RS hoch) als Innersubjektfaktoren, Bedingung (USAT, WW) als Zwischensubjektfaktor und den jeweiligen Prätest-Werten für rechtschreibschwache und rechtschreibstarke Schüler als Kovariaten gerechnet. Wie bei der Kontrollvariablen Geschlecht wurde Rechtschreibung als Innersubjektfaktor behandelt, da die rechtschreibschwachen und rechtschreibstarken Schüler innerhalb jeder Klasse im Fokus standen. In Tabelle 42 sind die Ergebnisse zusammengefasst.

Tabelle 42: Ergebnisse der Analysen zur Prüfung des Faktors Rechtschreibung als Moderator (Studie I)

	Bed. x RS	Bed.	RS	MZP	Prä RS niedrig	Prä RS hoch
Geschichtenelemente		X				
Geschichtenqualität	X	X	X			
Wortanzahl						
Planen Inhalt		X				
Planen Struktur		X				
Revisionen						
Wissen		X				

Erlebniserzählungen (GE)		X				
Erlebniserzählungen (NQ)		X				

Anmerkungen: X = signifikanter Effekt ($p < .05$); Bed. x RS. = Interaktion Bedingung x Rechtschreibung; Bed. = Haupteffekt Zwischensubjektfaktor Bedingung; RS = Haupteffekt Innersubjektfaktor Rechtschreibung; MZP = Haupteffekt Innersubjektfaktor Messzeitpunkt; Rechtschreibleistung hoch; Prätest RS niedrig = Effekt der Kovariaten Prätest Rechtschreibleistung niedrig; Prätest RS hoch = Effekt der Kovariaten Prätest.

Wie auch beim Geschlecht bestätigten sich die Ergebnisse der vorangegangenen Analysen. Auch unter separater Betrachtung rechtschreibschwacher und rechtschreibstarker Schüler ergaben sich außer für Wortanzahl und Revisionen signifikante Bedingungseffekte zu Gunsten der USAT-Bedingung ($ps < .05$).

Eine Kovarianzanalyse mit Geschichtenqualität als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Rechtschreibung (RS niedrig, RS hoch) als Innersubjektfaktoren, Bedingung (USAT, WW) als Zwischensubjektfaktor und den Prätest-Werten für rechtschreibschwache und rechtschreibstarke Schüler als Kovariaten lieferte zusätzlich zum Haupteffekt Bedingung, $F(1, 4) = 31.61$, $p < .001$, signifikante Effekte für die Interaktion Bedingung x Rechtschreibung, $F(1, 4) = 108.26$, $p < .001$, und für den Innersubjektfaktor Rechtschreibung, $F(1, 4) = 90.95$, $p < .01$. Folglich profitierten rechtschreibschwache und rechtschreibstarke Schüler unterschiedlich viel von den beiden Bedingungen in Bezug zur Geschichtenqualität. Während rechtschreibstarke Schüler generell bessere Leistungen als rechtschreibschwache Schüler in beiden Bedingungen erbrachten, war der Unterschied in der USAT-Bedingung deutlich höher als in der WW-Bedingung.

6.3.7.3 Sprache

Um zu prüfen, ob die Sprache Deutsch einen Einfluss auf die Effekte hatte, wurden Schüler, die Deutsch zu Hause sprachen, mit Schülern, die eine andere Sprache zu Hause sprachen, verglichen.

Analog zum Geschlecht und zur Rechtschreibung wurde dazu jeder Schüler einer von zwei Gruppen zugeordnet. Die eine Gruppe umfasste Schüler, die nur Deutsch zu Hause sprachen und die andere Gruppe Schüler, die Deutsch und bzw. oder eine andere Sprache zu Hause sprachen. Die Entscheidung, die Einteilung in dieser Weise vorzunehmen, lag darin begründet, dass zwei Klassen keine Schüler hatten, die nicht Deutsch zu Hause sprachen.

Nach Bildung der zweistufigen Variablen Sprache wurde der Schülerdatensatz getrennt für die beiden Gruppen Deutsch und nicht-Deutsch über die Klasse hinweg aggregiert. Tabelle 43 enthält die Mittelwerte und Standardabweichungen für

deutschsprachige und nicht-deutschsprachige Schüler getrennt für die drei Messzeitpunkte und die beiden Bedingungen.

Tabelle 43: Mittelwerte und Standardabweichungen für deutschsprachige und nicht-deutschsprachige Schüler getrennt nach Messzeitpunkt und Bedingung (Studie I).

	USAT				WW			
	Deutsch		nicht-Deutsch		Deutsch		nicht-Deutsch	
	M	SD	M	SD	M	SD	M	SD
Geschichtenelemente								
Prätest	3.17	0.52	2.78	0.63	3.19	0.70	2.58	0.64
Posttest	8.74	0.64	7.47	1.58	4.73	0.59	4.49	0.49
Follow-Up	7.89	1.09	6.01	1.27	4.00	0.72	3.61	0.65
Geschichtenqualität								
Prätest	2.07	0.15	1.89	0.68	2.07	0.33	1.58	0.43
Posttest	3.49	0.23	3.16	0.51	2.63	0.15	2.57	0.36
Follow-Up	3.28	0.30	2.74	0.38	2.53	0.21	2.50	0.45
Wortanzahl								
Prätest	98.94	21.47	91.10	27.33	93.35	10.56	74.54	29.66
Posttest	169.83	34.30	152.35	35.81	142.38	26.64	133.89	23.38
Follow-Up	159.01	40.19	144.83	48.37	102.29	31.90	111.42	21.73
Planen (Geschichtenelemente)								
Prätest	1.38	0.32	1.47	0.41	1.33	0.28	1.48	0.44
Posttest	6.10	0.75	5.22	2.93	1.15	0.48	1.53	0.65
Follow-Up	6.10	0.78	4.16	2.29	1.27	0.39	1.96	0.89
Planen (Struktur)								
Prätest	1.28	0.37	1.14	0.66	2.13	0.44	2.21	0.85
Posttest	3.47	0.29	3.11	1.36	1.21	0.45	1.44	0.52
Follow-Up	3.94	0.72	2.33	1.51	0.91	0.20	2.11	1.34
Revisionen								
Prätest	0.14	0.10	0.16	0.24	0.08	0.06	0.06	0.07
Posttest	0.89	0.38	0.58	0.65	0.44	0.31	0.24	0.22
Follow-Up	0.60	0.30	0.44	0.31	0.47	0.21	0.27	0.26

Wissen								
Prätest	1.72	0.29	1.58	0.17	2.43	0.69	2.61	1.13
Posttest	5.17	0.92	5.25	1.04	3.05	0.52	3.03	0.67
Follow-Up	5.11	0.35	4.30	1.41	2.77	0.58	3.47	1.10
Geschichtenelemente (Erlebniserzählungen)								
Prätest	3.52	0.56	3.47	0.98	3.66	0.51	3.11	0.68
Posttest	7.36	1.14	6.73	1.26	4.09	0.62	4.31	1.44
Follow-Up	6.40	1.05	4.99	0.81	3.97	0.81	3.28	0.92
Geschichtenqualität (Erlebniserzählungen)								
Prätest	1.71	0.27	1.53	0.06	1.54	0.15	1.18	0.30
Posttest	2.94	0.26	2.49	0.49	1.91	0.16	1.72	0.54
Follow-Up	2.69	0.38	2.46	0.42	2.18	0.13	1.71	0.53

Danach wurden Kovarianzanalysen mit den prozess- und produktbezogenen Maßen als abhängige Variablen, Messzeitpunkt (Posttest, Follow-Up) und Sprache (Deutsch, nicht-Deutsch) als Innersubjektfaktoren, Bedingung (USAT, WW) als Zwischensubjektfaktor und den Prätest-Werten für deutschsprechende und nicht-deutschsprechende Schüler als Kovariaten gerechnet. Sprache wurde wie zuvor die Kontrollvariablen Geschlecht und Rechtschreibung auch als Innersubjektfaktor behandelt, weil die deutschsprechenden und nicht-deutschsprechenden Schüler innerhalb einer Klasse untersucht wurden. Eine Zusammenfassung der Ergebnisse bietet Tabelle 44.

Tabelle 44: Ergebnisse der Analysen zur Prüfung des Faktors Sprache als Moderator (Studie I).

	Bed. x Sprache	Bed.	Sprache	MZP	Prä Deutsch	Prä nicht- Deutsch
Geschichtenelemente		X				
Geschichtenqualität		X				
Wortanzahl						
Planen Inhalt		X				
Planen Struktur						
Revisionen						
Wissen		X				

Erlebniserzählungen (GE)	X	X	X			
Erlebniserzählungen (NQ)						

Anmerkungen: X = signifikanter Effekt ($p < .05$); Bed. x Sprache = Interaktion Bedingung x Sprache; Bed. = Haupteffekt Zwischensubjektfaktor Bedingung; Sprache = Haupteffekt Innersubjektfaktor Rechtschreibung; MZP = Haupteffekt Innersubjektfaktor Messzeitpunkt; Prä Deutsch = Effekt der Kovariaten Prätest Deutsch; Prä nicht-Deutsch = Effekt der Kovariaten Prätest nicht-Deutsch.

Im Gegensatz zu den Kontrollvariablen Geschlecht und Rechtschreibung konnten unter separater Betrachtung deutschsprechender und nicht-deutschsprechender Schüler nicht für alle abhängigen Variablen die signifikanten Bedingungseffekte zu Gunsten der USAT-Bedingung bestätigt werden ($ps < .05$).

Eine Kovarianzanalyse mit Planen Struktur als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Sprache (Deutsch, nicht-Deutsch) als Innersubjektfaktoren, Bedingung (USAT, WW) als Zwischensubjektfaktor und den jeweiligen Prätest-Werten für deutschsprechende und nicht-deutschsprechende Schüler als Kovariaten führte zu einem nicht signifikanten Effekt des Zwischensubjektfaktors Bedingung, $F(1, 4) = 2.92, p = .16$.

Auch eine Kovarianzanalyse mit Erlebniserzählungen Geschichtenqualität als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Sprache (Deutsch, nicht-Deutsch) als Innersubjektfaktoren, Bedingung (USAT, WW) als Zwischensubjektfaktor und den jeweiligen Prätest-Werten für deutschsprechende und nicht-deutschsprechende Schüler als Kovariaten führte zu einem knapp nicht signifikanten Effekt des Faktors Bedingung, $F(1, 4) = 6.84, p = .06$.

Folglich führte die Berücksichtigung der Sprache bei Planen Struktur und Erlebniserzählungen Geschichtenqualität dazu, dass der Bedingungseffekt zu Gunsten der USAT-Bedingung verschwand.

Darüber hinaus lieferte eine Kovarianzanalyse mit Erlebniserzählungen Geschichtenelemente als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Sprache (Deutsch, nicht-Deutsch) als Innersubjektfaktoren, Bedingung (USAT, WW) als Zwischensubjektfaktor und den jeweiligen Prätest-Werten für deutschsprechende und nicht-deutschsprechende Schüler als Kovariaten nicht nur einen signifikanten Bedingungseffekt, $F(1, 4) = 19.28, p < .05$, sondern ebenfalls einen signifikanten Bedingung x Sprache-Interaktionseffekt, $F(1, 4) = 16.82, p < .05$, sowie einen signifikanten Effekt des Faktors Sprache, $F(1, 4) = 18.84, p < .05$.

Bezüglich Erlebniserzählungen Geschichtenelemente unterschieden sich somit nicht nur die beiden Bedingungen signifikant voneinander, sondern ebenfalls die

deutschsprechenden und nicht-deutschsprechenden Schüler innerhalb einer Bedingung. Gleichzeitig profitierten die nicht-deutschsprechenden Schüler deutlich mehr von dem USAT-Programm als von dem WW-Programm. Beim Posttest sind die Unterschiede in der USAT-Bedingung zwischen den deutschsprechenden und nicht-deutschsprechenden Schülern deutlich geringer, wobei die nicht-deutschsprechenden Schüler den Leistungszuwachs zum Follow-Up nicht beibehalten konnten.

6.4 Diskussion

Die Ergebnisse der Studie I, die im Folgenden zusammengefasst werden, belegen die Überlegenheit des USAT-Programms im Vergleich zum WW-Programm. Gleichzeitig lassen sich aus der Studie I Schlussfolgerungen für die Konzeption und Durchführung der Studie II ziehen. Am Ende der Diskussion wird Bezug zu den postulierten Hypothesen der Studie I genommen.

6.4.1 Zentrale Befunde der Studie I

Bedingungsunterschiede: Insgesamt schrieben alle Schüler in beiden Bedingungen am Ende des Trainings inhaltlich vollständigere, mehr ausgestaltete und qualitativ bessere Aufsätze.

Darüber hinaus lieferte keine Messwiederholungsanalyse abgesehen von Erlebniserzählungen Geschichtenelemente eine signifikante Bedingung x Messzeitpunkt Interaktion ($ps > .05$), so dass sich weder in der USAT- noch in der WW-Bedingung die Leistungen vom Posttest zum Follow-Up veränderten. Bezüglich Erlebniserzählungen Geschichtenelemente zeigte sich, dass in beiden Bedingungen die Leistungen beim Follow-Up absanken, wobei dies stärker auf die USAT- als auf die WW-Bedingung zutraf.

Hinsichtlich des Vergleichs des USAT- mit dem WW-Programm zeigten sich abgesehen von Wortanzahl, Revisionen und Erlebniserzählungen Geschichtenqualität Follow-Up hoch signifikante Bedingungseffekte zu Gunsten der USAT-Bedingung. Wortanzahl Posttest und Revisionen ausgenommen, ergaben sich auf Schülerebene Effektstärken zwischen .70 und 2.01, was großen Effekten entspricht (Graham & Perin, 2007). Während für Revisionen Follow-Up kein Effekt nachweisbar war, lagen die Effektstärken für Wortanzahl Posttest und Revisionen Posttest in einem kleinen bis mittleren Bereich. Folglich übertrafen die Leistungen der USAT-Bedingung deutlich die der WW-Bedingung. Da sich die beiden Bedingungen nur hinsichtlich selbstregulatorischer Komponenten unterschieden und alle übrigen Trainingsinhalte sowie Rahmenbedingungen konstant gehalten wurden, ist die Überlegenheit des USAT-Programms nur durch die Selbstregulation zu erklären.

Bei den meisten abhängigen Variablen sanken die Effektstärken auf Schülerebene beim Follow-Up ab. Nur bei Wortanzahl, Planen Struktur und Wissen konnte ein leichter Anstieg der Effektstärke festgestellt werden, so dass sich die Unterschiede zwischen den Bedingungen beim Follow-Up verstärkten.

Klassenunterschiede innerhalb einer Bedingung: Bezüglich der Klassenunterschiede lassen sich bei einzelnen abhängigen Variablen Unterschiede zwischen den Klassen einer Bedingung feststellen, wobei diese etwas häufiger in der WW- als in der USAT-Bedingung festgestellt wurden. Gleichzeitig gilt für alle abhängigen Variablen, dass die Unterschiede zwischen den Klassen innerhalb einer Bedingung deutlich geringer waren als die Unterschiede zwischen den beiden Bedingungen, was durch die signifikanten Bedingungseffekte zu Gunsten der USAT-Bedingung der Messwiederholungs- und Kovarianzanalysen bestätigt wurde. Die Klassenunterschiede bezüglich der Erlebniserzählungen Geschichtenelemente lassen sich dadurch erklären, dass diese von den Klassenlehrerinnen selbst und nicht von den Trainerinnen durchgeführt wurden. Die Art der Motivierung und Instruktion bei den Erlebniserzählungen kann daher in diesem Fall größeren Variationen unterliegen als bei den anderen Schreibmaßen.

Insbesondere die teilweisen großen Unterschiede zwischen den Klassen der WW-Bedingung lassen vermuten, dass das WW-Programm weniger erfolgreich war, die unterschiedlichen Ausgangsniveaus der Klassen auszugleichen, so dass signifikante Differenzen auch beim Posttest und beim Follow-Up bestehen blieben. Im Gegensatz dazu glichen sich generell die Klassen der USAT-Bedingung beim Posttest und beim Follow-Up immer mehr an und zeigten größere Ähnlichkeiten als beim Prätest.

Differentialbetrachtung Kontrollvariablen: Um der Frage nachzugehen, ob einzelne Subgruppen der Gesamtstichprobe mehr als andere von den Interventionen profitierten, wurden anhand der Kontrollvariablen Geschlecht (Mädchen/Jungen), Rechtschreibung (rechtschreibschwach/rechtschreibstark) und Sprache (deutsch/nicht-Deutsch) Subgruppen gebildet und anhand von Messwiederholungskovarianzanalysen untersucht.

Bei den meisten abhängigen Variablen konnten keine signifikanten Bedingung x Kontrollvariable-Interaktionen festgestellt werden. Diese sind für die Bewertung der Wirksamkeit der Trainingsprogramme relevant, da bei einer signifikanten Interaktion die Kontrollvariable einen unterschiedlich starken Einfluss auf die Effektivität der Interventionen ausüben würde.

Die festgestellten signifikanten Interaktionen Bedingung x Kontrollvariable betrafen ausschließlich das strategiefere Schreibleistungsmaß Geschichtenqualität (bei Geschlecht

und Rechtschreibung) und das Transfermaß Erlebniserzählungen Geschichtenelemente (bei Geschlecht und Sprache). Obwohl die Unterschiede zwischen den Bedingungen bestehen blieben, produzierte das USAT-Programm bezüglich strategieferner Schreibmaße größere Leistungsdifferenzen zwischen den Subgruppen einer Klasse als das WW-Programm. Dies kann dadurch erklärt werden, dass für den Transfer auf ein untrainiertes Textgenre Schüler mehr kognitive Ressourcen benötigten, die eventuell nicht verfügbar waren, wenn Sprachbarrieren und Rechtschreibschwierigkeiten überwunden werden mussten. Das Schreibmaß Geschichtenqualität erfasst ebenfalls Fertigkeiten (z.B. Kohärenz), die kaum in dem Training vermittelt wurden und eine höhere kognitive Anstrengung erfordern.

Davon abgesehen sind Geschlechtsunterschiede in der Schreibforschung häufig vorzufinden, da Mädchen tendenziell lieber schreiben und auch in der Freizeit eher Schreibaktivitäten nachgehen als Jungen. Überdies waren sämtliche eingesetzte Trainerinnen weiblich und konnten somit stärker als explizite Rollenvorbilder für Mädchen fungieren als für Jungen.

Darüber hinaus führte die Berücksichtigung der Kontrollvariablen Sprache dazu, dass die zuvor signifikanten Bedingungeffekte zu Gunsten der USAT-Bedingung für Planen Struktur und Erlebniserzählungen Geschichtenqualität verschwanden. Für beide waren die Ergebnisse jedoch nur knapp nicht signifikant, so dass der Einfluss der Variablen Sprache als gering zu bewerten ist.

Insgesamt kann jedoch für die meisten abhängigen Variablen geschlussfolgert werden, dass die Leistungsunterschiede zwischen den Bedingungen nicht auf Unterschiede in den Kontrollvariablen, sondern auf Unterschiede in der Wirksamkeit der Interventionen zurückzuführen sind.

Implementationsgenauigkeit: Es kann von einer hohen Qualität der Implementierung ausgegangen werden. Die interne Validität wurde durch die randomisierte Zuteilung der Klassen zu den Bedingungen und der Konzeption einer wirkungsvollen Kontrollgruppe gewährleistet. Überdies erfolgte die Zuteilung gematcht nach den Rechtschreibleistungen und der Klassengröße, so dass zwischen den Bedingungen keine signifikanten Unterschiede diesbezüglich vorlagen. Die externe Validität wurde durch die Analyse von Subgruppen überprüft und kann ebenfalls für das USAT- und das WW-Programm als gesichert gelten. Hinsichtlich der Implementierbarkeit kann zunächst nur geschlussfolgert werden, dass das USAT-Programm mit ganzen Klassen von Trainern durchführbar ist. In einem nächsten Schritt ist es daher notwendig zu prüfen, ob diese Umsetzung auch Lehrern selbst gelingt.

Weiter wurde die soziale Validität beider Programme durch Fragebogenerhebungen am Ende der Trainings gewährleistet. Beide Trainingsprogramme wurden von den Lehrern mit sehr gut bis gut bewertet. Des Weiteren wurden mehrere Maßnahmen zur Gewährleistung der inhaltlichen Vollständigkeit und der Implementationsgenauigkeit in die Durchführung der beiden Trainingsprogramme integriert. Geschulte Tutorinnen beobachteten sämtliche Trainingseinheiten in beiden Bedingungen und konstatierten eine sehr hohe Qualität der Umsetzung. Außerdem wurden alle vorgesehenen Inhalte in beiden Bedingungen von den Trainerinnen realisiert, was durch Checklisten, die die Trainerinnen ausfüllten, erfasst wurde.

6.4.2 Einschränkungen und Schlussfolgerungen für Studie II

Insgesamt erwies sich das USAT-Programm als sehr erfolgreich sowohl in der Förderung der Schreibkompetenz als auch bezüglich der Stabilität und Generalisierbarkeit der Effekte. Im Folgenden werden die zentralen Befunde, Kritikpunkte und Anregungen für die Studie II diskutiert.

Erklärung der zentralen Befunde: Die Ergebnisse der Studie I lassen sich wie folgt zusammenfassen.

(1) *Wirksamkeitsüberprüfung:* Gemäß O'Donnell (2008) muss vor der Prüfung der Effektivität eines Trainingsprogramms der Nachweis der Wirksamkeit erfolgen. Da bislang das SAT-Programm nur in Kleingruppen durchgeführt wurde, konnten die positiven Effekte des Programms nicht auf ganze Klassen verallgemeinert werden. Die Studie I zeigte, dass das USAT-Programm ebenfalls erfolgreich im gesamten Klassenkontext durchführbar ist. Überdies kann von einer hohen Implementationsgenauigkeit ausgegangen werden, da die Trainer in beiden Bedingungen die zentralen Inhalte in hoher Qualität realisierten.

(2) *Überlegenheit selbstregulatorischer Prinzipien:* Des Weiteren intendierte die Studie I, den Nachweis zu erbringen, dass Selbstregulation für den Erfolg eines Schreibtrainings zwingend notwendig ist. Aus diesem Grund unterschieden sich die USAT- und WW-Bedingung inhaltlich lediglich bezüglich der Selbstregulation. Es wurden dieselben Schreibstrategien in beiden Bedingungen vermittelt. Außerdem wurden die gleichen Instruktionsprinzipien verwendet. Zentral war hierbei die sechsstufige SRSD-Instruktion von Harris und Graham (1996), welche ebenfalls in beiden Bedingungen implementiert wurde. Insgesamt lassen sich die Unterschiede zwischen den beiden Bedingungen nur durch die Selbstregulation erklären. Die SRSD-Instruktion ist nur dann erfolgreich, wenn diese mit selbstregulatorischen Komponenten verknüpft wird. Diese Ergebnisse der Studie I erweitern bisherige Erkenntnisse, da abgesehen von der Studie von Glaser (2005) inkrementelle Effekte

der Selbstregulation lediglich für Transferaufgaben nachgewiesen wurden (Sawyer, Graham, & Harris, 1992).

(3) *Kein Bedingungseffekt für Wortanzahl:* Für Wortanzahl konnten weder beim Posttest noch beim Follow-Up Bedingungseffekte nachgewiesen werden. Gleichzeitig waren in der USAT- und in der WW-Bedingung sehr große Unterschiede zwischen den Klassen einer Bedingung feststellbar, was sich ebenfalls in den vergleichsweise hohen ICC's widerspiegelte (siehe Tabelle 22). Der nicht signifikante Bedingungseffekt könnte darauf zurückgeführt werden, dass in der WW-Bedingung zwar extrem lange, jedoch inhaltlich sowie qualitativ schlechte Aufsätze geschrieben wurden. In der USAT-Bedingung dagegen wurden sowohl lange als auch inhaltlich und qualitativ bessere Aufsätze geschrieben.

(4) *Kein Bedingungseffekt für Revisionen:* Auch hinsichtlich Revisionen zeigte sich weder beim Posttest noch beim Follow-Up ein signifikanter Bedingungseffekt. Insgesamt schnitten jedoch die Klassen der USAT-Bedingung besser ab als die Klassen der WW-Bedingung. Die Posttest-Leistungen der USAT-Klassen variierten aber erheblich, so dass eine Klasse mehr von dem Training profitierte als eine andere. Im Vergleich zur Pilotstudie lag der Revisionsmittelwert der Studie I beim Prätest unter und beim Posttest über dem jeweiligen Wert der Pilotstudie, so dass insgesamt angenommen werden kann, dass die Förderung der Revisionsfertigkeiten in der Studie I besser realisiert wurde als in der Pilotstudie. Dies kann u.a. auf Veränderungen der selbstregulatorischen Materialien zurückgeführt werden. Dennoch sind die Leistungen der USAT-Klassen der Studie I als gering zu bewerten, da die Klassen im Mittel weniger als eine inhaltliche Verbesserung beim Posttest machten. Dies ist zum einen dadurch erklärbar, dass sich die Revisionsfertigkeiten erst im späten Jugendalter entwickeln (Feilke, 1993) und zum anderen dadurch, dass die Konzentrationsleistung und die Motivation der Schüler am Ende der Trainingssitzung soweit absanken, dass nur wenige Überarbeitungen vorgenommen wurden.

Kritikpunkte der Studie I: Nichtsdestotrotz lassen sich einige Kritikpunkte identifizieren.

(1) Die Zuteilung der Tutorinnen zu den Trainerinnen erfolgte aus organisatorischen Gründen nicht randomisiert. Die Tutorinnen waren zeitlich sehr eingeschränkt und konnten nicht flexibel zu verschiedenen Zeitpunkten in den Klassen eingesetzt werden. Abgesehen von zwei Klassen einer Schule befanden sich zudem alle weiteren Klassen im Gießener Umland und im Vogelsbergkreis, so dass ein Auto notwendig war, über das die Tutorinnen nicht verfügten. Um Absprachen zwischen Trainerinnen, Tutorinnen und Lehrerinnen zu erleichtern, wurde schließlich eine Tutorin einer Trainerin

zugeordnet. Dies hatte den Vorteil, dass sich Trainerinnen und Tutorinnen besser aufeinander abstimmen konnten und dadurch der Trainingsablauf vereinfacht wurde.

(2) Die Stichprobe umfasste nur acht Klassen. Die geringe Anzahl ist dadurch zu erklären, dass der Aufwand für die Vorbereitung und Durchführung der Trainingsprogramme so groß war, dass eine Trainerin nur eine Klasse pro Bedingung instruieren konnte. Dennoch sollten die Ergebnisse der Studie I mit mehr Klassen und einer konzeptionell anderen Kontrollbedingung repliziert werden. Interessant wäre die Frage, inwieweit das USAT-Programm traditionellen Unterrichtsmethoden überlegen ist.

(3) Da in der Studie I Trainer anstatt Lehrer die beiden Interventionen umsetzten, kann der Versuchsleiter-Effekt nicht ausgeschlossen werden. Dieser besagt, dass Schüler positiv auf den Versuchsleiter reagieren und motivierter sind als wenn die Lehrer selbst das Training umsetzen. Dieser Effekt betrifft jedoch sowohl die USAT- als auch die WW-Bedingung. Insgesamt wurde die Studie I unter kontrollierten Versuchsbedingungen durchgeführt, die nicht direkt identisch sind mit Bedingungen, denen Lehrer in ihren eigenen Klassen ausgesetzt sind. Demzufolge muss in einem nächsten Schritt die Effektivität des USAT-Programms überprüft werden. Dazu sollte das USAT-Programm von Lehrern im natürlichen Klassenkontext umgesetzt und mit dem traditionellen Aufsatzunterricht der vierten Klasse verglichen werden (O'Donnell, 2008). Dieser Ansatz sollte in der Studie II verfolgt werden.

Anregungen für Studie II: Bei der Konzeption und Durchführung der Studie II sollten folgende Punkte berücksichtigt werden.

(1) Die größte Schwierigkeit bei der Durchführung der Studie I bestand darin, die Inhalte in der zeitlich vorgegebenen Zeit zu vermitteln. Auf dieses Problem wiesen die Lehrerinnen beider Bedingungen, die sämtliche Trainingseinheiten beobachteten, in der abschließenden Befragung hin. Obwohl sie sich generell sehr positiv über die beiden Programme äußerten, sahen die meisten Lehrerinnen es aufgrund des Zeitdrucks als schwierig an, die Trainings selbst in den eigenen Klassen umzusetzen. Sie schlugen hierfür einen längeren Instruktionszeitraum von mehreren Monaten vor. Sie empfanden das Instruktionstempo als zu schnell und waren überrascht darüber, wie gut die Schüler trotz der knappen Zeit mitarbeiteten.

(2) Um die Bereitschaft der Lehrer zu erhöhen, die Trainingsprogramme zu implementieren, sollte das USAT-Programm zeitlich entzerrt werden. Da die Zusammenhangsinstruktion erst in der fünften Trainingseinheit erfolgte und somit kaum Wiederholungs- und Übungsmöglichkeiten den Schülern zur Verfügung standen, wendeten

nur sehr wenige Schüler die Zusammenhgangsstrategien in ihren Geschichten an. Somit erscheint das Ersetzen der Zusammenhgangsinstruktion durch eine allgemeine Wiederholungseinheit in der Studie II als sinnvoll. Die Wiederholungseinheit, die in zwei 45minütige Übungseinheiten aufgeteilt sein sollte, könnte als eine Art „Pufferstunde“ dienen, die die Lehrer frei gestalten können. Auch wenn spezifische Übungseinheiten den Lehrern vorgeschlagen werden sollten, müsste den Lehrern die Möglichkeit eingeräumt werden, auf individuelle Probleme der Klasse einzugehen und verpasste Inhalte nachzuholen.

(3) Des Weiteren erscheint die Erhebung eines Transfermaßes für das Hauptanliegen der Studie II als nicht notwendig. Auf der einen Seite zeigten sich positive Effekte des Transfermaßes zu Gunsten der USAT-Bedingung in der Studie I, auf der anderen Seite bedeutet die Erfassung der Erlebniserzählungen eine zusätzliche Arbeitsbelastung für die Lehrer. Da in der Studie II die Lehrer bereits stark mit der Umsetzung der Trainingsinhalte in den eigenen Unterricht gefordert sind, sollte in der Studie II die Erhebung des Transfermaßes Erlebniserzählungen entfallen, um die Belastung der Lehrer zu reduzieren.

6.4.3 Zusammenfassung

Bezug nehmend zu den postulierten Hypothesen können diese als bestätigt angesehen und folgende Aussagen getroffen werden.

(1) **Implementierbarkeit der Trainingsprogramme:** Alle Klassen in beiden Bedingungen profitierten von den jeweiligen Trainings, so dass folglich die Implementierbarkeit von extern entwickelten Schreibtrainings in den regulären Deutschunterricht möglich war.

(2) **Effektivitätsnachweis selbstregulatorischer Prozeduren:** Gleichzeitig konnten bei fast allen abhängigen Maßen, Revisionen und Wortanzahl ausgenommen, signifikante Bedingungseffekte zu Gunsten der USAT-Bedingung festgestellt werden. Da sich die beiden Bedingungen nur hinsichtlich selbstregulatorischer Methoden voneinander unterschieden, kann geschlussfolgert werden, dass Selbstregulation für die Effektivität von Schreibtrainings wichtig ist.

(3) **Stabilität der Effekte:** Abgesehen von Erlebniserzählungen Geschichtenqualität blieben die signifikanten Bedingungseffekte beim Posttest auch nach sechs Wochen stabil.

(4) **Transfereffekte:** Obwohl die USAT- und WW-Klassen dieselben Schreibstrategien erlernten, gelang es den USAT-Klassen signifikant besser, ihr Wissen auf ein untrainiertes Textgenre, Erlebniserzählungen, zu transferieren.

(5) **Generalisierbarkeit der Ergebnisse:** Auch wenn das USAT-Programm bei strategiefernen Maßen größere Differenzen zwischen den Subgruppen der Kontrollvariablen produzierte als das WW-Programm, kann dennoch geschlussfolgert werden, dass die Ergebnisse für Subgruppen der Kontrollvariablen Geschlecht, Rechtschreibung und Sprache größtenteils generalisierbar sind.

(6) **Zusätzliche Analysen:** Intra-Klassen-Korrelationen bzw. Mixed Model-Analysen bestätigten, dass einerseits kein bedeutsamer Zusammenhang zwischen den Klassen einer Bedingung bestand und andererseits, dass der Einfluss der Klasse innerhalb einer Bedingung nicht signifikant war.

7 Studie II: Implementierung durch Lehrer

7.1 Fragestellung und Hypothesen

Die nachfolgende Studie II sollte zeigen, dass die Umsetzung des *Unterrichtsintegrierten Selbstregulatorischen Aufsatztrainings* (USAT) auch von Lehrern erfolgreich realisiert werden kann. Dazu wurde das USAT-Programm mit einer von *Lehrern entwickelten Unterrichtseinheit zum Thema Geschichtschreiben* (LUG) verglichen. Gemäß Gräsel und Parchmann (2004) erfolgte die Implementation des USAT-Programms wie in der Studie I anhand der Top-down-Strategie. Hingegen sollte das LUG-Programm explizit zusammen mit Lehrern konzipiert werden, so dass dieses Vorgehen der symbiotischen Implementationsstrategie entspricht (Gräsel & Parchmann, 2004).

Die Effektivität, Stabilität und Generalisierbarkeit des USAT- und des LUG-Programms wurden anhand von folgenden Variablen untersucht, welche vor und nach den Trainings sowie sechs Wochen nach der Posttestung erhoben wurden:

- Strategienahe Variablen (Geschichtenelemente, Planungsqualität, Anzahl der Revisionen, schreibbezogenes Wissen);
- Strategieferne Variablen (Geschichtenqualität, Wortanzahl).

Unter Anknüpfung an die Generalhypothesen und die Hypothesen der Studie I verfolgt die Studie II folgende Ziele und Hypothesen. Zur genaueren Begründung der Hypothesen wird auf Kapitel 3.1 und Kapitel 6.1 verwiesen.

(1) Sind ein von Experten entwickeltes Trainingsprogramm wie das *Unterrichtsintegrierte Selbstregulatorische Aufsatztraining* (USAT) sowie eine von Lehrern selbst konzipierte *Unterrichtseinheit zum Thema Geschichtschreiben* (LUG) in den Regelunterricht implementierbar?

Analog zur Studie I wurde das USAT-Programm mit einem Kontrollgruppentraining verglichen, welches von Lehrern entwickelt wurde und die traditionellen Unterrichtsmethoden des Aufsatzunterrichts widerspiegelt. Folglich beinhaltet das LUG-Programm Unterrichtsmethoden, mit denen die Lehrer vertraut sind und deren Umsetzung ihnen leicht fallen sollte.

Hypothese 1: Es wurde erwartet, dass Lehrer das USAT- und das LUG-Programm erfolgreich in den Unterricht in den eigenen Klassen integrieren können. Ferner wurde angenommen, dass die beteiligten Lehrer die jeweiligen Trainings gleichermaßen gut in ihren

Klassen umsetzen, so dass sich die Lehrer und Bedingungen nicht hinsichtlich des Inhalts und der Qualität der Implementierung unterscheiden. Zur Sicherung der Implementationsgenauigkeit wurden unterschiedliche Maßnahmen ergriffen.

(2) Ist das *Unterrichtintegrierte Selbstregulatorische Aufsatztraining (USAT)* effektiver in der Förderung schreibbezogener Maße als die von *Lehrern entwickelten Unterrichtseinheit zum Thema Geschichtschreiben (LUG)*?

Die in der Studie I nachgewiesenen signifikanten Effekte zu Gunsten der USAT-Bedingung sollten in der Studie II in einer größeren Stichprobe, mit mehr Klassen und einer konzeptionell anderen Kontrollgruppe repliziert werden. Aus diesem Grund wurde die Effektivität des USAT-Programms mit der Effektivität des LUG-Programms verglichen.

Hypothese 2: Hinsichtlich der strategienahen und –fernen Schreibmaße wurde angenommen, dass das USAT-Programm deutlich bessere Leistungen hervorbringt als das LUG-Programm. Aufgrund der Ergebnisse der Lehrerbefragung (siehe Kapitel 4.3) wurde vermutet, dass Lehrer keine selbstregulatorischen Methoden in das LUG-Programm integrieren, so dass die erwarteten Bedingungeffekte zu Gunsten der USAT-Bedingung den selbstregulatorischen Komponenten zugeschrieben werden können.

(3) Können die Effekte nach sechs Wochen ebenfalls nachgewiesen werden?

Die strategienahen und –fernen Schreibmaße wurden unmittelbar vor und unmittelbar nach der Durchführung der Trainings sowie sechs Wochen nach der Posttestung erhoben. So sollten Aussagen über den zeitlichen Transfer der beiden Trainingsbedingungen getroffen werden.

Hypothese 3: Es wurde erwartet, dass selbstregulatorische Strategien den Erwerb der Schreibstrategien nachhaltiger fördern als der traditionelle Aufsatzunterricht, so dass das USAT-Programm dem LUG-Programm sowohl im unmittelbaren Anschluss an die Trainings überlegen ist und die Effekte auch noch sechs Wochen danach bestehen bleiben.

Analog zur Studie I wurden weitere statistische Analysen berechnet, um zusätzliche Aussagen über die Trainingseffekte zu treffen.

(4) Zusätzliche Analysen

Aufgrund des gruppenrandomisierten Designs der Studie II wurden Mixed Model-Analysen durchgeführt und Intra-Klassen-Korrelationen berechnet.

(5) Prüfung von Moderatoreffekten

Um auszuschließen, dass nur ein Teil der Klasse von den jeweiligen Trainings profitierte, wurden in weiterführenden Analysen anhand der Kontrollvariablen Geschlecht, Rechtschreibung, Sprachverständnis, Nationalität und Muttersprache Deutsch einzelne Subgruppen gebildet und die statistischen Ergebnisse für die Subgruppen separat betrachtet. Es wurde erwartet, dass die Befunde der USAT- und der LUG-Bedingung für die jeweiligen Subgruppen generalisierbar sind.

7.2 Methode

7.2.1 Untersuchungsdesign und Ablauf

Der Studie II lag ein Vortest-Nachtest-Follow-Up-Vergleichsgruppen-Versuchsplan zu Grunde, um die zuvor formulierten Hypothesen zu prüfen. Das USAT-Programm wurde mit dem LUG-Programm kontrastiert. Beide Bedingungen wurden von Klassenlehrern in den eigenen Klassen realisiert. Eine Veranschaulichung des Untersuchungsdesign ist in Abbildung 56 zu finden.

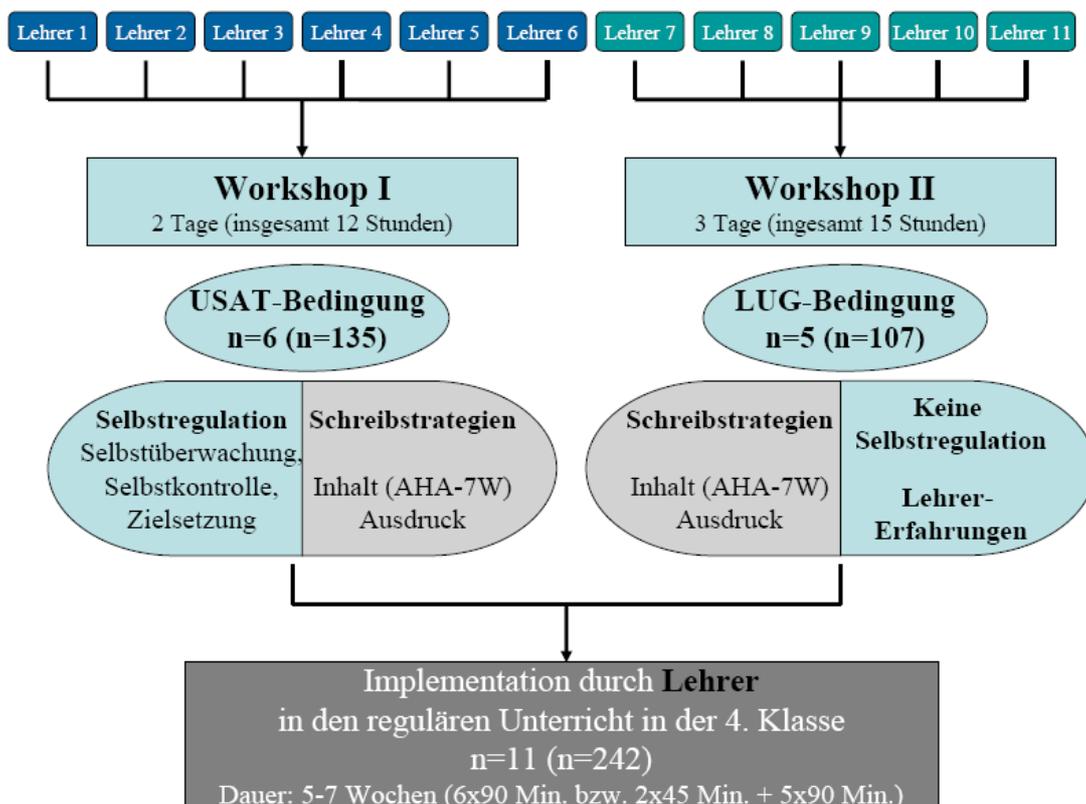


Abbildung 56: Untersuchungsdesign der Studie II.

Die Rekrutierung der Lehrer erfolgte in zwei Phasen. In der ersten Phase wurden Lehrer der vierten Klasse gesucht, die an dem Erlernen und der Umsetzung eines selbstregulatorischen Aufsatztrainings in den eigenen Klassen Interesse hatten. Aufgrund von positiven Erfahrungen der Lehrer der Studie I erklärten sich sechs Lehrer aus zwei Schulen, die bereits in der Studie I involviert waren, bereit, an der Studie II teilzunehmen. Demgegenüber sollten in der zweiten Rekrutierungsphase, die telefonisch erfolgte, Lehrer gefunden werden, die für einen Workshop aufgeschlossen waren, in dem sie basierend auf eigenen Erfahrungen sowie auf Anregungen aus der Pädagogischen Psychologie ein Unterrichtskonzept zum Thema Geschichtschreiben entwickelten. Diese zweite Phase gestaltete sich schwieriger und dauerte dementsprechend länger, da bei einer Vielzahl von Schulen angefragt werden musste. Aus diesem Grund konnte mit der Durchführung der Kontrollbedingung (LUG) erst zwei Wochen nach der USAT-Bedingung gestartet werden. Aufgrund der inhaltlichen Unterschiede zwischen der USAT- und der LUG-Bedingung war eine randomisierte Zuordnung der Lehrer zu den Bedingungen nicht möglich.

Nachdem genügend Lehrer rekrutiert wurden, fanden zunächst die Erhebungen der Kontrollvariablen und anschließend die jeweiligen Workshops statt (siehe Abbildung 14). Direkt im Anschluss an die Workshops wurden die Prätestungen durchgeführt. Danach implementierten die Lehrer das USAT- bzw. das LUG-Programm in die eigenen Klassen. Die absolute Trainingsdauer (12 Unterrichtsstunden, à 45 Minuten) war in beiden Bedingungen identisch, wobei die Lehrer einen Spielraum von fünf bis sieben Wochen zur Verfügung hatten.

Überdies beobachteten in den ersten drei Wochen geschulte Trainer die Durchführung der beiden Trainingsbedingungen. Unmittelbar nach den Trainings wurden die Posttest-Daten erhoben. Eine Follow-Up-Messung fand sechs Wochen nach der Posttestung statt. Die Testungen wurden von der Doktorandin sowie drei Psychologiestudenten im Hauptstudium, die zuvor intensiv in der Durchführung der Testungen geschult worden waren, vorgenommen.

7.2.2 Stichprobe

Die Studie II umfasste insgesamt elf Klassen mit 242 Schülern. Wie in Abbildung 56 ersichtlich ist, waren in der USAT-Bedingung sechs Klassen mit 135 Schülern und in der LUG-Bedingung fünf Klassen mit 107 Schülern beteiligt. Da in der Studie II Lehrer anstatt Trainer die Aufsatzprogramme implementierten, werden im Folgenden zunächst die Lehrer- und anschließend die Schüler-Stichprobe beschrieben.

7.2.2.1 Lehrer

Die USAT-Bedingung umfasste sechs Lehrer aus zwei miteinander kooperierenden Schulen aus Grünberg. Es handelt sich dabei um eine große fünfzügige Schule und dessen einzügige Zweigstelle in einem Nachbarort. In der LUG-Bedingung waren dagegen fünf Lehrer aus insgesamt vier verschiedenen Schulen aus den Landkreisen Gießen und Vogelsberg (Rödgen, Lich, Lauterbauch, Rüdtingshausen) beteiligt.

Von den elf beteiligten Lehrern hatten zehn Lehrer mehr als fünf Jahre Berufserfahrung. Nur ein Lehrer der LUG-Bedingung besaß eine Berufserfahrung von drei bis fünf Jahren. Während in der LUG-Bedingung nur weibliche Lehrkräfte teilnahmen, waren in der USAT-Bedingung fünf Frauen und ein Mann involviert.

In der USAT-Bedingung variierte die Klassengröße zwischen 20 und 27 Schülern und betrug im Mittel 22.50 (SD: 2.81), wohingegen die Größe der LUG-Klassen zwischen 18 und 23 lag, was ein Mittel von 21.40 (SD: 2.07) ergab.

Aufgrund der unterschiedlichen Anforderungen, die das USAT- und das LUG-Programm an die Lehrer stellte, wurden diese nicht zufällig den beiden Bedingungen zugeordnet. Die Lehrer meldeten sich freiwillig entweder für das USAT- oder für das LUG-Programm an.

7.2.2.2 Schüler

Bezogen auf die Gesamtstichprobe von 242 Schülern waren 53% (129) weiblich und 47% (113) männlich. Das Durchschnittsalter betrug insgesamt 9.6 Jahre. 96% der Schüler waren deutscher Herkunft.

Um die Vergleichbarkeit der beiden Bedingungen zu erleichtern, sind in Tabelle 45 die beschreibenden Merkmale der Schülerstichprobe getrennt nach Bedingung dargestellt.

Tabelle 45: Beschreibung der Stichprobe getrennt nach Bedingung (Studie II).

	USAT		LUG	
<i>n</i>	135		107	
Geschlecht	w: 48.1%	m: 51.9%	w: 59.8%	m: 40.2%
Alter (in Jahren)	M: 9.53	SD: 0.58	M: 9.71	SD: 0.61
Nationalität Deutsch	96.3%		93.5%	
Andere Nationalität	3.7%		6.5%	
Deutsch als Muttersprache	74.8%		71.0%	
Andere Muttersprache	7.4%		14.0%	

Deutsch und andere Sprache	17.8%		15.0%	
Klassengröße	M: 22.50	SD: 2.81	M: 21.40	SD:2.07
Rechtschreibung	M: 42.43	SD: 27.70	M: 59.70	SD: 27.33
Sprachverständnis	M: 43.63	SD: 24.78	M: 48.45	SD: 23.93
Fehltermine	M: 0.22	SD: 0.47	M: 0.30	SD: 0.59

Zwischen den beiden Bedingungen (USAT, LUG) bestanden keine signifikanten Unterschiede in Bezug auf Geschlecht, Fehltermine, Nationalität, Muttersprache Deutsch, Sprachverständnis sowie Klassengröße ($ps > .11$). Dagegen unterschieden sich die Bedingungen signifikant hinsichtlich des Alters und der Rechtschreibleistung zu Gunsten der LUG-Bedingung voneinander ($ps < .05$).

7.2.3 Trainingsbedingungen

7.2.3.1 Vergleich der USAT- und der LUG-Bedingung

Bei der Konzeption der LUG-Bedingung wurde beabsichtigt, dem USAT-Programm eine Bedingung gegenüberzustellen, welche die typischen Methoden des Aufsatzunterrichts der vierten Klasse reflektiert. Es sollte der Frage nachgegangen werden, ob die bereits üblichen Methoden effektiv sind und ob das USAT-Programm gegebenenfalls durch die Erfahrungen der Lehrer optimiert werden kann. Konkret ging es darum, ob ein von Schreibexperten konzipiertes Trainingsprogramm effektiver ist als der „normale“ Unterricht.

Um einen kurzen Überblick über den Ablauf einer Trainingseinheit zu gewinnen, sind die prototypischen Einheiten der USAT- und der LUG-Bedingung in Tabelle 46 gegenübergestellt.

Tabelle 46: Prototypische Trainingsinhalte der USAT- und der LUG-Bedingung.

USAT	LUG
Rückmeldung Prozess- (7W-Protokoll) und produktorientierte Rückmeldungen (Leistungsprotokoll) für die Schüler und Lehrer	Rückmeldung Keine prozess- oder produktorientierte Rückmeldungen; allgemeine Rückmeldungen für die Schüler und ausführliche Rückmeldungen für die Lehrer

<p>Motivierung</p> <p>Motivierung der Schüler, die Schreibstrategien anzuwenden (Strategieregulation) und Geschichtenexperte zu werden (Performanzregulation)</p> <p>Loben der Schüler</p> <p>Ernennung des „Besten Bergsteigers“ (siehe Abbildung 30)</p> <p>Berechnung der Punktedifferenz zwischen letzter und vorletzter Geschichte im Leistungsprotokoll</p>	<p>Motivierung</p> <p>Allgemeine Motivierung der Schüler</p>
<p>Wiederholung</p> <p>Wiederholung der Schreibstrategien in Bezug zu den Geschichten der Schülern</p>	<p>Wiederholung</p> <p>Wiederholung der Schreibstrategien anhand verschiedener Übungen</p>
<p>Strategieinstruktion (TE¹² 1 bis TE 5)/ Übungen (TE 6)</p> <p>Kognitives Modellieren der Schreibstrategien in Kombination mit den selbstregulatorischen Prozessen</p>	<p>Strategieinstruktion (TE 1 bis TE 5)/ Übungen (TE 6)</p> <p>Herleitung der Schreibstrategien im Unterrichtsgespräch anhand von Beispielen</p>
<p>Zielsetzung</p>	<p>Keine Zielsetzung</p>
<p>Geschichtschreiben</p> <p>Angeleitetes Planen mit Geschichtenplan</p> <p>Angeleitetes Schreiben mit Geschichtenplan</p> <p>Angeleitetes Überarbeiten mit Checkliste</p>	<p>Geschichtschreiben</p> <p>Angeleitetes Planen mit Hilfsmittel</p> <p>Angeleitetes Schreiben mit Hilfsmittel</p> <p>Angeleitetes Überarbeiten mit Hilfsmittel</p> <p>Hilfsmittel: 7W-Merkkärtchen und Ausdrucksgewürze</p>

Im Gegensatz zur Studie I ist die Gegenüberstellung der prototypischen Sitzungen der beiden Bedingungen dadurch erschwert, dass das USAT-Programm von Experten entwickelt wurde (siehe Kapitel 2.4.1.2, Top-down-Strategie) und das LUG-Programm in Kooperation mit Lehrern entstand und auf deren Erfahrungen basiert (siehe Kapitel 2.4.1.2, symbiotische Strategie). Das LUG-Training spiegelt größtenteils den gängigen Aufsatzunterricht in der vierten Klasse wider. Aus diesem Grund enthält das LUG-Training keine festgelegten Unterrichtsmethoden, da die Lehrer die einzelnen Trainingskomponenten freier gestalten konnten und überdies den Schülern viel Zeit zum Diskutieren einräumten. In der nachfolgenden Abbildung 57 sind die Gemeinsamkeiten und Unterschiede der USAT- und der LUG-Bedingung grafisch dargestellt.

¹² TE = Trainingseinheit; Dauer: 90 Minuten.

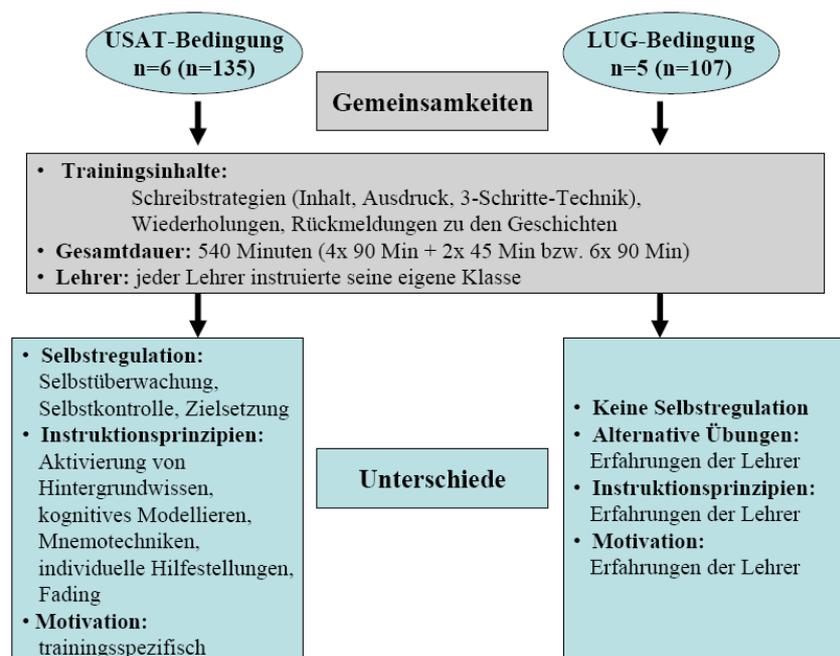


Abbildung 57: Gemeinsamkeiten und Unterschiede der USAT- und der LUG-Bedingung.

Aus Abbildung 57 ist ersichtlich, dass die Trainingsinhalte wie u.a. die Schreibstrategien und die Dauer des Trainings in beiden Bedingungen identisch sind. Darüber hinaus sind selbstregulatorische Prinzipien nur in der USAT- und nicht in der LUG-Bedingung integriert. Die sechs SRSD-Instruktionsprinzipien (siehe Kapitel 2.3.2.2) wurden ebenfalls nur im USAT-Programm durchgeführt. Die Lehrer der LUG-Bedingung verwendeten traditionelle Methoden des Aufsatzunterrichts. Tabelle 47 enthält die wichtigsten Inhalte der Trainings- und Übungseinheiten getrennt nach Bedingung.

Tabelle 47: Vergleich der USAT- und der LUG-Bedingung.

	USAT	LUG
Prätest	Erfassung der Schreibleistung sowie der Planungs- und Überarbeitungsaktivitäten und des schreibbezogenen Wissens	
TE 1	Herleitung der AHA-7W-Strategie Schreiben einer Bildergeschichte	
	Herleitung der 3-Schritte-Technik Modellierung und angeleitetes Üben der Selbstüberwachung beim Planen und Schreiben mit dem Geschichtenplan Modellierung der Zielsetzung mit dem 7W-Protokoll und dem Leistungsprotokoll Planen einer Bildergeschichte	Einführung von 7W-Merkkärtchen Wiederholung der AHA-7W-Strategie Generierung von Ideen für eine Bildergeschichte mit der gesamten Klasse

ÜE 1¹³	Wiederholung der AHA-7W-Strategie und der 3-Schritte Technik, Modellierung und angeleitetes Üben der Selbstbewertung und -kontrolle beim Überarbeiten mit der Checkliste	<i>Wurde nicht durchgeführt</i>
TE 2	Rückmeldungen zu den Geschichten der letzten Trainingseinheit Wiederholung der gelernten Strategien Erarbeitung der Strategien zur Verbesserung des Ausdrucks Angeleitetes Schreiben einer Bildergeschichte	
	Modellierung und angeleitetes Üben der Selbstbewertung und -kontrolle beim Berechnen der Punktedifferenz zwischen der Prätest- und der letzten Geschichte; Einkleben von entsprechenden Belohnungen in das Leistungsprotokoll Rückmeldung und Zielsetzung mit dem Leistungs- und dem 7W-Fragen-Protokoll Modellierung und angeleitetes Üben der Anwendung der Ausdruckskarte Anfang während des Planens, Schreibens und Überarbeitens Bildergeschichte wird von den Schülern nicht nur geschrieben, sondern auch geplant und überarbeitet	Erarbeitung sämtlicher Ausdrucksstrategien anhand der Ausdrucksgewürze Durchführung einer Übung zur Verbesserung der sprachlichen Ausgestaltung Generierung von Ideen für eine Bildergeschichte mit der gesamten Klasse
TE 3	Rückmeldungen zu den Geschichten der letzten Trainingseinheit Wiederholung der erlernten Strategien Angeleitetes Planen, Schreiben und Überarbeiten einer Geschichte	
	Berechnen der Punktedifferenz zwischen der aktuellen Aufsatzbewertung und der Aufsatzbewertung der Prätestgeschichte; Einkleben von Belohnungen in das Leistungsprotokoll Rückmeldung und Zielsetzung mit dem Leistungs- und dem 7W-Fragen-Protokoll Modellierung und angeleitetes Üben der Anwendung der Ausdruckskarte Hauptteil/ Ende während des Planens, Schreibens und Überarbeitens	Herleitung der 3-Schritte-Technik Erarbeiten von Strategien zum Planen/ Generieren von Ideen Planen einer Geschichte mit Hilfe eines Arbeitsblattes Erarbeiten von Bewertungskriterien und Strategien zum Überarbeiten einer Geschichte Überarbeiten einer Geschichte anhand der Ausdrucksgewürze

¹³ ÜE= Übungseinheit; Dauer: 45 Minuten

ÜE 2	<p>Pufferstunde, die von Lehrern frei gestaltet werden konnte</p> <p>vorgegebener Vorschlag für die Stunde: Wiederholung der gelernten Strategien vertiefende Übung zum Überarbeiten mit der Checkliste</p> <p>Als Hausaufgabe sollten Schüler eine Bildergeschichte planen, schreiben und überarbeiten</p>	<i>Wurde nicht durchgeführt</i>
TE 4	Rückmeldungen zu den Geschichten der letzten Trainingseinheit Wiederholung der gelernten Strategien	
	<p>Berechnen der Punktedifferenz zwischen der aktuellen Aufsatzbewertung und der Aufsatzbewertung der Prätestgeschichte; Einkleben von Belohnungen in das Leistungsprotokoll</p> <p>Rückmeldung und Zielsetzung mit dem Leistungs- und dem 7W-Fragen-Protokoll</p> <p>Durchführung einer Übung zur Wiederholung der erlernten Strategien</p> <p>Erarbeitung von Strategien zur Verbesserung des Hauptteils einer Geschichte</p> <p>Begleitetes Planen, Schreiben und Überarbeiten einer Geschichte</p>	<p>Durchführung einer Stationsarbeit zum Thema Planen: verschiedene Methoden zum Generieren von Ideen wurden durchgeführt</p> <p>Einführung des „roten Fadens“ als Hilfsmittel beim Planen einer Geschichte</p> <p>Reflexion der Stationsarbeit in der gesamten Klasse</p> <p>Es wird keine Geschichte geschrieben</p>
TE 5	Wiederholung der gelernten Strategien Planen, Schreiben und Überarbeiten einer Geschichte	
	<p>Berechnen der Punktedifferenz zwischen der aktuellen Aufsatzbewertung und der Aufsatzbewertung der Prätestgeschichte; Einkleben von Belohnungen in das Leistungsprotokoll</p> <p>Rückmeldung und Zielsetzung mit dem Leistungs- und dem 7W-Fragen-Protokoll</p> <p>Wiederholung anhand von drei Schülergeschichten</p> <p>Durchführung eines Schreibquiz zur Selbstkontrolle des schreibbezogenen Wissens</p> <p>Kein Einsatz von Hilfsmitteln beim Verfassen der Bildergeschichte</p>	<p>Einführung der Textlupe (siehe Abbildung 34) zur Bewertung von Geschichten</p> <p>Durchführung einer Übung zur Anwendung der Textlupe</p> <p>Eigenständiges Planen mit dem „roten Faden“, Schreiben und Überarbeiten einer Bildergeschichte unter Verwendung von 7W- Merkkärtchen und Ausdrucksgewürzen</p>

TE 6	<i>Wurde nicht durchgeführt</i>	Rückmeldungen zu den Geschichten der letzten Trainingseinheit Wiederholung der Strategien anhand einer Reizbildgeschichte Kein Einsatz von Hilfsmitteln beim Verfassen der Bildergeschichte
Posttest	Erfassung der Schreibleistung sowie der Planungs- und Überarbeitungsaktivitäten und des schreibbezogenen Wissens	
Follow-Up	Erfassung der Schreibleistung sowie der Planungs- und Überarbeitungsaktivitäten und des schreibbezogenen Wissens	

7.2.3.2 Schreibstrategien

Zentrales Anliegen der Studie II war die Vermittlung derselben Schreibstrategien in der USAT- und in der LUG-Bedingung. Die Instruktion der Schreibstrategien in der USAT-Bedingung wurde wie in der Pilotstudie bzw. in Studie I durchgeführt, so dass für ausführliche Beschreibungen auf das Kapitel 5.3.2.1 und Kapitel 6.2.3.2 verwiesen wird. Es wurden jedoch in der Studie II keine Zusammenhangsstrategien vermittelt.

Trickposter: Die LUG-Klassen erhielten das Trickposter der WW-Bedingung (siehe Abbildung 28). Lediglich die Zusammenhangsstrategien fehlten in den jeweiligen Postern der USAT- und LUG-Bedingung.

3-Schritte-Technik: Die LUG-Lehrerinnen führten die Strategie erst als letztes ein. Sie beschäftigten sich dabei insbesondere mit Übungen zur Generierung von Ideen. Als Merkhilfen wurden ebenfalls wieder die 3-Schritte-Technik-Poster (USAT: siehe Abbildung 15, LUG: siehe Abbildung 29) verwendet.

Inhalt: Die Instruktion der Inhaltsstrategie (AHA-7W) erfolgte in der LUG-Bedingung nicht analog zur USAT-Bedingung anhand einer Beispielgeschichte (siehe CD-Anhang A7.1), sondern wurde im Unterrichtsgespräch und mit Übungen eingeführt.

Ausdruck: Die Lehrer der LUG-Bedingung entschieden sich für eine Ausdrucksinstruktion anhand sogenannter Ausdrucksgewürze (siehe Abbildung 59), die die Schüler beim Planen, Schreiben und Überarbeiten benutzen sollten. Die Gewürze entsprachen dabei den intendierten Schreibstrategien (Adjektive, Verben, Gedanken, Gefühle).

7.2.3.3 Unterrichtsmethoden

Während in der Studie I die Unterrichtsmethoden in den beiden Bedingungen konstant gehalten wurden, variierten diese in der Studie II. Das USAT-Programm basiert auf dem

SRSD-Ansatz von Harris und Graham (1996), wobei das LUG-Programm vornehmlich auf den Erfahrungen der Lehrer aufbaut.

Partnerarbeit: Aus zeitlichen Gründen wurde die Partnerarbeit im USAT-Programm weggelassen. Sowohl in der USAT- als auch in der LUG-Bedingung wurden jedoch vereinzelt Partner- und Gruppenarbeiten eingesetzt. Diese waren jedoch kein expliziter Bestandteil jeder Trainingseinheit.

Instruktionsprinzipien: Im Gegensatz zur LUG-Bedingung wurden von den USAT-Lehrern die sechsstufige Instruktion des SRSD-Ansatzes (Harris & Graham, 1996) eingesetzt. Dazu zählen das kognitive Modellieren, Fading, angeleitetes Üben, selbstständiges Üben und Selbstinstruktionen. Dagegen basierte das LUG-Programm auf den Erfahrungen der Lehrerinnen. Die Instruktionsprinzipien konnten aus zeitlichen Gründen nicht mehr explizit im Workshop festgelegt werden, so dass diese nach Belieben und Ermessen der Lehrerinnen gewählt und eingesetzt wurden. In den meisten Fällen waren dies der Frontalunterricht und das Unterrichtsgespräch.

Wiederholung: Zu Beginn jeder Trainingsstunde wurden die bis dahin erlernten Schreibstrategien wiederholt. Während im USAT-Programm die Wiederholung anhand konkreter Schülergeschichten und Geschichtenbeispielen erfolgte, wiederholten die LUG-Lehrerinnen die Schreibstrategien anhand der eingesetzten Poster und Merkkarten. Die Wiederholungen fanden ebenfalls im gesamten Klassenkontext statt und unterschieden sich in Form und Dauer nicht von den Wiederholungen der USAT-Bedingung.

Motivationsfördernde Komponenten: Wie auch in der Studie I wurde in der USAT-Bedingung der Studie II der Button „Bester Bergsteiger“ (siehe Abbildung 30) eingesetzt und regelmäßig Schülergeschichten vorgelesen, um die Motivation der Schüler zu erhöhen. Im LUG-Workshop wurden die Lehrer darauf hingewiesen, die Schüler für die Mitarbeit an dem Training zu motivieren. Es wurden jedoch keine expliziten Materialien erarbeitet, so dass die Art der Motivierung von den Lehrern frei gestaltet wurde.

Hausaufgaben: In beiden Bedingungen bekamen die Schüler nach jeder Trainingseinheit eine Hausaufgabe. Die USAT-Lehrer konnten dabei aus einem Repertoire an verschiedenen Übungen frei wählen (siehe CD-Anhang A6.1.4). Die LUG-Lehrerinnen überlegten sich gemeinsam im Workshop zu dem LUG-Programm passende Hausaufgaben. Sie griffen dabei auf Übungen zurück, die sie bereits im Unterricht eingesetzt hatten.

7.2.3.4 Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (USAT)

Das im Folgenden skizzierte USAT ist in weiten Teilen mit dem USAT-Programm der Studie I identisch. Zusätzlich wurden die Erfahrungen der Studie I herangezogen, um einzelne

Instruktionen, Materialien und Übungen zu optimieren. Des Weiteren wurden Kürzungen und Veränderungen vorgenommen, um adäquat auf die Bedürfnisse und Voraussetzungen (konstitutioneller und institutioneller Art) der Lehrer einzugehen. Ein wesentliches Ziel bei der Neukonzeptionalisierung des USAT-Programms war die zeitliche Entzerrung der Inhalte. Aus diesem Grund wurden Komponenten identifiziert und gestrichen, die zeitlich aufwendig, aber im Vergleich zu anderen Inhalten weniger gewinnbringend waren. Auf diese Weise entfiel die Instruktion der Zusammenhgangsstrategien. Die Partner- und Gruppenarbeiten fielen ebenfalls weg.

7.2.3.4.1 Workshop

Zur Vorbereitung und Schulung der Lehrer, das USAT-Programm in den regulären Unterricht der vierten Klasse zu implementieren, fand ein zweitägiger Workshop in den Räumen der Justus-Liebig-Universität statt. Die zeitliche Aufteilung des Workshops wurde von den Lehrern bestimmt. Der Workshop wurde geleitet von der Doktorandin und einer Psychologiestudentin im Hauptstudium. Der erste Workshoptag dauerte vier Stunden und der zweite, der zwei Tage später stattfand, acht Stunden. Der Workshop verfolgte folgende Ziele:

- Vermittlung der Inhalte des USAT-Programms;
- Erprobung der Durchführung der Trainingseinheiten;
- Erhöhung der Motivation und Kooperationsbereitschaft der Lehrer;
- Förderung einer positiven Einstellung der Lehrer zum Trainingsprogramm;
- Sensibilisierung der Lehrer dafür, das Training möglichst genau und qualitativ gut zu implementieren.

Tabelle 48 enthält eine Auflistung der im ersten und zweiten Workshop-Tag vermittelten Inhalte.

Tabelle 48: Inhalte des USAT-Workshops.

	Inhalte
Workshop-Tag 1	Ziele des Workshops Theoretische Einführung in die Förderung der Schreibkompetenz bei Viertklässlern Grobe Vorstellung des Trainings (Aufbau, Dauer, Schreibstrategien, selbstregulatorische Methoden) Vorstellung und Erprobung der ersten Trainingseinheit Übung zur prozess- und produktbezogenen Rückmeldung Zusammenfassung

Workshop-Tag 2	Übung zur Instruktion der Zielsetzung Vorstellung und Erprobung der ersten Übungseinheit Vorstellung und Erprobung der Ausdrucksinstruktionen (TE 2/TE 3) Vorstellung und Erprobung der Instruktion der Rückmeldungen Vorstellung und Durchführung der Übungen der vierten und fünften Trainingseinheiten Übung zum Ablauf einer prototypischen Trainingseinheit Übungen zur Förderung selbstregulationsfördernder Einstellungen und der Motivation Abschlussbesprechung und Ausfüllen des Evaluationsbogens Zusammenfassung
-----------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Workshop-Tag 1: Am ersten Workshoptag wurden den Lehrern die theoretischen Grundlagen, auf die das USAT-Programm basiert, vermittelt. Dies geschah in Form eines Vortrags. Anschließend wurden den Lehrern die Struktur und Gliederung des Trainings näher gebracht. Dazu erhielt jeder Lehrer einen Trainingsordner. Dieser umfasste alle für die sechs Trainingseinheiten und zwei Übungseinheiten relevanten Materialien. Für jede Einheit gab es ein kurzes (siehe CD-Anhang A6.1.1) und ein langes Manual (siehe CD-Anhang A6.1.2), Kopiervorlagen für die Schüler sowie Instruktionmaterialien (z.B. Folien) für die Lehrer. Zudem umfasste der Ordner verschiedene Schreibenlässe (siehe CD-Anhang A3) und weitere Arbeitsblätter (siehe CD-Anhang A6.1.4), die die Lehrer als zusätzliche Übungen oder Hausaufgaben einsetzen konnten. Abgesehen von den trainingsrelevanten Materialien und Übungen konnten die Lehrer Handouts und Übungsblätter, die sie während des Workshops bekamen, im Trainingsordner abheften. Weiter befand sich im Ordner eine CD mit Videosequenzen aus der Studie I, die die Instruktionen der wichtigsten Strategien des USAT-Programms zeigten.

Nachdem sich die Lehrer mit dem Ordner vertraut gemacht hatten, wurden die Schreibstrategien und die selbstregulatorischen Methoden des USAT-Programms erklärt. Dazu wurden die wichtigsten Materialien vorgestellt und erläutert.

Als nächstes erfolgte eine genaue Vorstellung der ersten Trainingseinheit. Den Lehrern wurden hierzu Videosequenzen aus der Studie I gezeigt. Anschließend modellierten die Lehrer zu zweit einzelne Aspekte der ersten Trainingseinheit. Sie orientierten sich dabei an dem ausführlichen Manual. Während der Modellierung unterstützten die Doktorandin und die Psychologiestudentin die Lehrer und gaben Hilfestellungen. Die Lehrer notierten sich Anmerkungen oder Schwierigkeiten und besprachen diese anschließend in der gesamten

Gruppe. Auf diese Weise konnten die Lehrer in Absprache mit den Workshopleitern Änderungen vornehmen, die ihnen die Instruktion in ihren eigenen Klassen erleichtern sollten.

Im Anschluss daran wurden die Lehrer an die prozess- und produktbezogene Rückmeldung mit dem 7W- und Leistungsprotokoll herangeführt. Sie bekamen verschiedene Schülertexte aus der Studie I und korrigierten diese anhand der Protokolle.

Am Ende des ersten Workshop-Tages wurden die bisherigen Erkenntnisse zusammengefasst und die bereits erarbeiteten Komponenten des USAT-Programms in einem Schaubild, welches am zweiten Workshop-Tag weiter vervollständigt wurde, dargestellt.

Workshop-Tag 2: Zwei Tage später fand der zweite Workshop-Tag statt, an dem die Trainingseinheiten zwei bis fünf sowie die beiden Übungseinheiten vorgestellt und von den Lehrern erprobt wurden. Der Ablauf ähnelte dem des ersten Workshop-Tages. Jede Trainings- und Übungseinheit wurde zunächst von den Workshopleitern erklärt und die korrekte Instruktion der jeweiligen Strategien wurde den Lehrern gezeigt. Dies geschah entweder durch das Modellieren der Instruktion durch die Workshopleiter oder durch das Zeigen von Videosequenzen. Nach der Klärung von Problemen und Unklarheiten erfolgte eine Erprobungsphase, in der die Lehrer in Kleingruppen die Trainings- bzw. Übungseinheiten durchspielten. Im Anschluss daran traf sich wieder die gesamte Gruppe und besprach die in der Kleingruppe gewonnenen Erkenntnisse, so dass sich die Lehrer in ihren Manualen Notizen machen konnten.

Die Vorstellung und Übung der Einheiten wurde immer wieder durch verschiedene praktische Aufgaben, die den Lehrern wichtige selbstregulatorische Prinzipien näher bringen sollten, unterbrochen. So mussten sie Instruktionsfehler bei der Modellierung der Zielsetzung, die von den Workshop-Leitern gezeigt wurde, identifizieren. In einer weiteren Übung erhielten die Lehrer Schnipsel mit den einzelnen Komponenten einer prototypischen Trainingseinheit, die die Lehrer sortieren sowie ein paar Punkte dazu notieren mussten. Um die Lehrer für die Verwendung von selbstregulationsfördernden Aussagen zu sensibilisieren, formulierten sie Sätze, die sie während der Zielsetzung, der Rückmeldung der Schülergeschichten und der Erläuterung des Geschichtenplans (Selbstüberwachung) und der Checkliste (Selbstkontrolle) verwenden konnten. Diese Formulierungen mündeten in dem selbstregulatorischen Fragebogen (siehe CD-Anhang A6.3.2), den die Lehrer nach jeder Trainingseinheit zur Kontrolle der Implementationsgenauigkeit ausfüllten. Zudem sammelten die Lehrer Ideen zur Motivierung der Schüler.

Am Ende des zweiten Workshop-Tages wurden organisatorische Fragen geklärt und schließlich die wichtigsten Inhalte des Workshops kurz in dem Schaubild des ersten

Workshop-Tages zusammengefasst. Außerdem füllten die Lehrer einen Fragebogen zur Evaluierung des Workshops aus.

7.2.3.4.2 *Selbstregulatorische Methoden*

Hinsichtlich der selbstregulatorischen Methoden wurden keine Veränderungen zur Studie I vorgenommen, da die Ergebnisse der Studie I zeigten, dass die Materialien und Methoden des USAT-Programms bereits sehr gut im Klassenkontext einsetzbar waren.

Zur *Überwachung des Strategieeinsatzes* wurde der Geschichtenplan (siehe Abbildung 23), welcher während des Planens und Schreibens eingesetzt wurde und die Checkliste (siehe Abbildung 31), welche zum Überarbeiten des geschriebenen Textes herangezogen wurde, verwendet.

Die prozessbezogene *Zielsetzung* erfolgte mit dem 7W-Protokoll (siehe Abbildung 32). Mit dem Leistungsprotokoll (siehe Abbildung 33) setzten sich die Schüler globale und auf die Gesamtqualität der Geschichte bezogene Ziele. Die Klebepunkte zur Förderung des schreibbezogenen Fähigkeitsselbstkonzepts (siehe CD-Anhang A5.1.3) wurden ebenfalls wieder verwendet, wobei sie dieses Mal in jeder Trainingseinheit eingesetzt wurden. Im Gegensatz zur Studie I entfiel die Rückseite des Leistungsprotokolls, um die Lehrer zu entlasten und ihnen mehr Zeit für die übrigen Trainingsinhalte zur Verfügung zu stellen.

Die *Selbstbewertung* als weiteres selbstregulatorisches Prinzip wurde ebenfalls durch die Checkliste und die Verwendung der Protokolle erreicht.

7.2.3.4.3 *Ablauf des Trainings*

Wie in Tabelle 47 ersichtlich, umfasst das USAT-Programm fünf Trainingseinheiten mit jeweils 90 Minuten und zwei Übungseinheiten mit jeweils 45 Minuten. Die erste Übungseinheit erfolgte nach der ersten, die zweite nach der dritten Trainingseinheit. Den Lehrern wurde vorgegeben, das gesamte Training in fünf bis sieben Wochen durchzuführen. Genauere Angaben wurden vermieden, um es den Lehrern zu ermöglichen, das USAT-Programm leichter in den Stundenplan einzubauen. Fünf der sechs Lehrer führten das Training in fünf Wochen durch. Jede Woche erfolgte eine Trainingseinheit, wobei in den Wochen eins und drei zusätzlich die Übungseinheiten durchgeführt wurden. Einer der Lehrer benötigte sieben Wochen, da jede Woche nur eine Trainings- oder Übungseinheit umgesetzt wurde.

Die ersten drei Trainingseinheiten und die erste Übungseinheit wurden von der Doktorandin und einer Psychologiestudentin im Hauptstudium, die bereits den Workshop mit geleitet hat, zur Sicherung der Implementationsgenauigkeit beobachtet. In jeder

Trainingseinheit wurde eine Geschichte von den Schülern verfasst, die die Doktorandin, die Psychologiestudentin im Hauptstudium sowie zwei Lehramtsstudenten anhand der 7W- und Leistungsprotokolle bewerteten. Gleichzeitig erhielten die Lehrer wöchentlich eine Übersicht der Geschichtenbewertungen für jeden Schüler einzeln (siehe CD-Anhang A6.1.5). Auf diese Weise waren die Lehrer über das aktuelle Leistungsniveau der Klasse informiert und konnten ihre Instruktionen entsprechend adaptieren. Jede Woche holte die Doktorandin die Trainingsgeschichten ab, so dass die Lehrer dadurch die Möglichkeit hatten, Schwierigkeiten und Fragen zu klären.

Bei der Durchführung der Trainings orientierten sich die Lehrer maßgeblich an den ausführlichen Manualen (siehe CD-Anhang A6.1.2), mit denen sie bereits im Workshop gearbeitet hatten. Jeder der Lehrer berichtete, dass sie sich vor jeder Trainings- und Übungsstunde intensiv mit Hilfe des Materialordners und der Manuale vorbereiteten.

Um Redundanzen zu vermeiden, werden nur die Unterschiede zur Pilotstudie und zur Studie I ausführlich beschrieben. Für genauere Ausführung wird auf Kapitel 5.3.2.4 und Kapitel 6.2.3.4.2 verwiesen.

Trainingseinheit 1 (Schwerpunkt: Inhalt, Planen, Schreiben): In der ersten Trainingseinheit wurden keine Änderungen vorgenommen. Folgende Trainingsinhalte waren mit Studie I identisch:

- Einführung des Trickposters (siehe Abbildung 28) und Überblick über die Inhalte des Schreibtrainings;
- Modellierung der Anwendung des Leistungsprotokolls und Verteilung der Trainingsmappen mit Rückmeldungen zur Prätest-Geschichte;
- Herleitung der AHA-7W-Strategie anhand des Textes „Der schlaue Fuchs“ (siehe CD-Anhang A7.1) und Aufhängen des 7W-Posters (siehe Abbildung 16);
- Einführung des 7W-Protokolls und Verteilung der 7W-Protokolle mit Rückmeldungen zur Prätest-Geschichte;
- Modellierung der Zielsetzung anhand des 7W- und des Leistungsprotokolls;
- Zielsetzung;
- Einführung der 3-Schritte-Technik;
- Modellierung der Anwendung des Geschichtenplans anhand einer Bildergeschichte (siehe CD-Anhang A7.4);
- Planen und Schreiben einer Bildergeschichte mit dem Geschichtenplan.

Übungseinheit 1 (Schwerpunkt: Überarbeiten): Das Einschreiben einer Übungseinheit lag darin begründet, dass aus zeitlichen Gründen und um die Schüler kognitiv

nicht zu überfordern, die Instruktion des dritten Schreibprozesses, des Überarbeitens, nicht mehr in der ersten Trainingseinheit erfolgen konnte. Damit die Schüler dennoch lernten, dass das Verfassen einer guten Geschichte die Prozesse Planen, Schreiben und Überarbeiten umfasst, wurde das Überarbeiten in der ersten Übungseinheit nachgeholt.

Zunächst wiederholte der Lehrer die AHA-7W-Strategie. Anschließend wurde die Anwendung der Checkliste zum Überarbeiten eines selbst geschriebenen Textes von dem Lehrer anhand der Beispielgeschichte „Tim beim Einkaufen“ modelliert (siehe CD-Anhang A7.5). Alle handlungsleitenden Gedanken wurden dabei laut verbalisiert. Nachdem die Schüler die Checkliste verstanden hatten, wurden die in der ersten Trainingseinheit geplanten und geschriebenen Bildergeschichten erneut ausgeteilt. Die Schüler überarbeiteten daraufhin ihre Geschichte mit der Checkliste. Des Weiteren wurde das 3-Schritte-Technik-Poster aufgehängt (siehe Abbildung 15). Am Ende der Stunde füllten die Schüler ein Schreibquiz (siehe CD-Anhang A6.1.3) aus, welches sie hinterher anhand eines Lösungsblattes kontrollierten.

Trainingseinheit 2 (Schwerpunkt: Ausdruck Anfang): Die zweite Trainingseinheit entspricht der dritten Trainingseinheit der Studie I. Es wurde jedoch keine Partnerarbeit durchgeführt. Insgesamt handelt es sich um folgende Punkte, die wie in der Studie I vermittelt wurden:

- Rückgabe der Leistungs- und 7W-Protokolle im Stuhlkreis;
- Einführung und Vergabe des Buttons „Bester Bergsteiger“;
- Modellierung der Berechnung der Punktedifferenz zur Prätestgeschichte im Leistungsprotokoll;
- Berechnung der Punktedifferenz und Einkleben der Klebepunkte;
- Wiederholung der 3-Schritte-Technik;
- Vermittlung der Ausdrucksstrategien Anfang anhand einer vollständigen, nicht ausgestalteten Geschichte und der Ausdruckskarte (siehe CD-Anhang A7.2 und Abbildung 17);
- Zielsetzung;
- Planen, Schreiben und Überarbeiten einer Bildergeschichte unter Verwendung des Geschichtenplans, der Ausdruckskarte und der Checkliste II (siehe Abbildung 31), welche die Inhalts- und Ausdrucksstrategien umfasst.

Trainingseinheit 3 (Schwerpunkt: Ausdruck Hauptteil/Ende): Der Ablauf der dritten Trainingseinheit entsprach dem der zweiten und stimmte mit der vierten Trainingseinheit der Studie I weitestgehend überein:

- Rückgabe der Leistungs- und 7W-Protokolle;
- Ernennung des neuen „Besten Bergsteigers“;
- Berechnung der Punktedifferenz und Einkleben der Klebepunkte;
- Wiederholung der bisher erlernten Strategien, insbesondere der Ausdrucksstrategien Anfang;
- Vermittlung der Ausdrucksstrategien Hauptteil/Ende anhand einer vollständigen, nicht ausgestalteten Geschichte und der Ausdruckskarte (siehe CD-Anhang A7.2 und Abbildung 17);
- Zielsetzung;
- Verfassen (Planen, Schreiben, Überarbeiten) einer neuen Geschichte unter Verwendung des Geschichtenplans, der Ausdruckskarte und der Checkliste II.

Nach der dritten Trainingseinheit war die Instruktion von Schreibstrategien abgeschlossen, so dass es in den darauf folgenden Sitzungen vornehmlich um Wiederholung und Verfestigung des Wissens ging.

Übungseinheit 2 (Schwerpunkt: Wiederholung, Überarbeiten): Die zweite Übungseinheit diente den Schülern dazu, potenzielle Defizite zu beseitigen und Routine in der Anwendung der Schreibstrategien sowie der unterstützenden Materialien zu gewinnen. Darüber hinaus konnten die Lehrer Inhalte nachholen, die sie eventuell in einer der vorherigen Sitzungen nicht geschafft hatten. Da die Schüler in der Regel noch Schwierigkeiten beim Überarbeiten hatten, entschieden sich die Lehrer für Übungen zur Verfestigung der Revisionsfertigkeiten.

Zu Beginn wiederholte der Lehrer die erlernten Schreibstrategien anhand einer Schülergeschichte. Während der Lehrer die Geschichte vorlas, markierten die Schüler mit der Checkliste, welche Strategien sie in der Geschichte wiedererkannt hatten. Der Lehrer diskutierte dabei die korrekte Anwendung der Checkliste und der Schreibstrategien.

Anschließend bekamen die Schüler eine weitere Geschichte („Das Geheimnis“, siehe CD-Anhang A6.1.3), die sie zunächst mit der Checkliste überprüfen und schließlich verbessern sollten. Die Verbesserungsvorschläge wurden hinterher in der Klasse gesammelt und besprochen.

Am Ende der Stunde erhielten die Schüler eine Bildergeschichte, die sie zu Hause unter Bereitstellung aller bekannten Hilfsmaterialien planen, schreiben und überarbeiten sollten. Diese Geschichte wurde von den Lehrern selbst nachgesehen und hinsichtlich der bekannten Kriterien bewertet. Die Lehrer füllten dazu ein Hausaufgabenprotokoll (siehe CD-Anhang A6.1.4) aus.

Trainingseinheit 4 (Schwerpunkt: Wiederholung, Hauptteil): Da der Hauptteil den längsten Teil einer Geschichte ausmacht, wurden in der vierten Trainingseinheit insbesondere Strategien, die im Hauptteil wichtig sind, geübt und diskutiert. Es wurden dabei Inhalte der fünften Trainingseinheit der Studie I integriert:

- Rückgabe der Leistungs- und 7W-Protokolle;
- Ernennung des neuen „Besten Bergsteigers“;
- Berechnung der Punktedifferenz und Einkleben der Klebepunkte.

Anschließend las der Lehrer zunächst den Anfang und das Ende einer Geschichte („Lagerfeuer“, siehe CD-Anhang A6.1.3) vor und wiederholte unter Heranziehung der Checkliste alle relevanten Schreibstrategien. Im gesamten Klassenverband ergänzte er zudem Adjektive und bildliche Beschreibungen. Den Schülern wurde verdeutlicht, dass die ausführliche Beschreibung des Hauptteils zentral für das Verständnis einer Geschichte ist. Weiter erhielten die Schüler Schnipsel mit Sätzen, welche sie in Partnerarbeit so sortieren sollten, dass der Hauptteil der Geschichte entstand. Im Anschluss daran überlegten sich die Schüler Ausgestaltungsmöglichkeiten für den Hauptteil, welche sie am Ende mit dem Lehrer diskutierten. Die anschließenden Inhalte waren mit der Pilotstudie wieder identisch:

- Zielsetzung;
- Verfassen (Planen, Schreiben, Überarbeiten) einer neuen Geschichte unter Verwendung des Geschichtenplans, der Ausdruckskarte und der Checkliste II;
- Hilfestellungen der Lehrer wurden langsam zurückgenommen.

Trainingseinheit 5 (Schwerpunkt: Wiederholung allgemein): Die letzte Trainingseinheit wurde von den Lehrern dazu genutzt, alle Strategien zu wiederholen und die letzten Wissenslücken zu schließen sowie die Schüler zu motivieren, eigenständig das Erlernte anzuwenden und sich weiterhin beim Geschichtschreiben anzustrengen. Sie war zu großen Anteilen mit der sechsten Trainingseinheit der Studie I identisch und bestand aus folgenden Punkten:

- Rückgabe der Leistungs- und 7W-Protokolle;
- Ernennung des neuen „Besten Bergsteigers“;
- Berechnung der Punktedifferenz und Einkleben der Klebepunkte: alle Schüler konnten einen „super Schreiber“- oder „Meister-Schreiber“-Klebepunkt in das Leistungsprotokoll einfügen;
- Wiederholung sämtlicher Schreibstrategien anhand von drei Schülergeschichten (Inhalt, Ausdruck Anfang, Ausdruck Hauptteil/Ende);

- Als weitere Übung zur Festigung des Schreibwissens füllten die Schüler ein weiteres Schreibquiz (siehe CD-Anhang A6.1.3) aus, welches sie selbstständig korrigieren mussten.
- Zielsetzung;
- Planen, Schreiben und Überarbeiten einer Bildergeschichte ohne die Verwendung von Hilfsmitteln (z.B. Geschichtenplan, Ausdruckskarte, Checkliste II).

Analog zur Studie I erhielt zum Schluss jeder Schüler einen persönlichen Brief (siehe CD-Anhang A5.1.3), in dem hervorgehoben wurde, was der Schüler bereits besonders gut kann und womit er noch Schwierigkeiten hat. Es wurde dabei explizit Bezug zu den Schülergeschichten und den erlernten Schreibstrategien genommen. Die Briefe sollten die Schüler motivieren, sich beim Posttest anzustrengen.

7.2.3.5 Lehrerkonzipierte Unterrichtseinheit zum Geschichtschreiben (LUG)

Als Kontrollgruppe wurde in der Studie II die LUG-Bedingung der USAT-Bedingung gegenübergestellt. Inhaltlich sollte die LUG-Bedingung möglichst komplett mit der USAT-Bedingung identisch sein, so dass den Lehrern zunächst im Workshop dieselben Schreibstrategien vermittelt wurden. Gleichzeitig sollte die konkrete Instruktion und Übung der Strategien möglichst optimal den für Lehrer bekannten und üblichen Prinzipien des traditionellen Aufsatzunterrichts entsprechen.

7.2.3.5.1 Workshop

Der Workshop zur Erarbeitung einer von *Lehrern konzipierten Unterrichtseinheit zum Thema Geschichtschreiben* für die vierte Grundschulklasse fand an drei aufeinander folgenden Nachmittagen (à fünf Stunden) in den Räumen der Justus-Liebig-Universität statt. Für den LUG-Workshop wurden drei Stunden länger als für den USAT-Workshop angesetzt, da die Lehrer gemäß der symbiotischen Implementierungsstrategie, welche im Theorieteil ausführlich erklärt wurde (siehe Kapitel 2.4.1.2), ein eigenes Training mit vorgegebenen Inhalten konzipieren mussten. Die Lehrer durften selbst entscheiden, an welchen Tagen und in welcher zeitlichen Aufteilung die Workshops stattfinden sollten.

Der Workshop wurde von zwei männlichen Psychologiestudenten im Hauptstudium geleitet. Das Konzept und der Ablauf des Workshops wurden im Vorhinein mit der Doktorandin besprochen und eingehend geübt. Nach jedem der drei Nachmittage fand zudem eine Besprechung zwischen den Psychologiestudenten und der Doktorandin statt. Auf diese Weise wurde sichergestellt, dass die im Workshop konzipierte Unterrichtseinheit zum

Geschichtschreiben den Anforderungen für eine adäquate Kontrollbedingung des USAT-Programms entsprach.

Zur besseren Vergleichbarkeit der beiden Bedingungen bekamen die LUG-Lehrerinnen bestimmte Vorgaben für die inhaltliche Gestaltung der Unterrichtseinheit. Zum einen sollten die Schreibstrategien 3-Schritte-Technik (Planen, Schreiben, Überarbeiten), Inhalt und Ausdruck Gegenstand des intendierten Trainings sein. Für jede dieser Strategien wurden den Lehrerinnen Übungen zur Verfügung gestellt, die größtenteils aus dem *Writer's Workshop* der Studie I stammten und die die Lehrerinnen anregen sollten, eigene Ideen und Erfahrungen aus der bisherigen Tätigkeit als Lehrer zu generieren und in die Konzeption der Unterrichtseinheit mit einzubringen. Zum anderen wurde den Lehrerinnen vorgegeben, dass die Schüler für jede geschriebene Geschichte Rückmeldungen erhalten und die erlernten Strategien regelmäßig wiederholt werden sollten. Überdies wurden die Lehrerinnen gebeten, in dem Training fünf Geschichten sowie eine weitere Geschichte als Hausaufgabe schreiben zu lassen, wobei die Schüler zehn Minuten für das Planen, zwanzig Minuten für das Schreiben und zehn Minuten für das Überarbeiten zur Verfügung haben sollten. Schließlich wurde ein zeitlicher Rahmen von sechs Trainingseinheiten (à 90 Minuten) festgelegt, wobei jede Woche eine Trainingseinheit stattfinden sollte.

Ausgehend von den beschriebenen inhaltlichen und zeitlichen Vorgaben, die die Lehrerinnen für die Konzeption der Unterrichtseinheit bekamen, lag das Hauptaugenmerk des Workshops auf der Auswahl didaktischer Prinzipien und geeigneter Übungen zur optimalen Vermittlung der vorgegebenen Schreibstrategien. Dabei sollten die Lehrerinnen möglichst viele eigene Erfahrungen integrieren. Im Gegensatz zum USAT-Workshop waren folglich die Inhalte des LUG-Workshops offen und frei gestaltbarer, was der symbiotischen Implementierungsstrategie nach Gräsel und Parchmann (2004) entspricht.

In der nachfolgenden Tabelle 49 sind die Inhalte des LUG-Workshops aufgelistet.

Tabelle 49: Inhalte des LUG-Workshops.

	Inhalte
Workshop-Tag 1	<p>Ziele des Workshops und Erwartungen der Lehrerinnen</p> <p>Theoretische Einführung in die Förderung der Schreibkompetenz bei Viertklässlern</p> <p>Durchführung des ersten Arbeitsblocks (Inhalt I: Erarbeitung der Inhaltsstrategie)</p> <p>Durchführung des zweiten Arbeitsblocks (Inhalt II: Didaktische Umsetzung der Inhaltsstrategie)</p>

Workshop-Tag 2	Durchführung des dritten Arbeitsblocks (Ausdruck) Durchführung des vierten Arbeitsblocks (Planen)
Workshop-Tag 3	Durchführung des fünften Arbeitsblocks (Überarbeiten) Abschlussbesprechung und Ausfüllen des Evaluationsbogens Zusammenfassung

Wie in Tabelle 49 ersichtlich, gliederte sich der Workshop insgesamt in die fünf Arbeitsblöcke (1) Inhalt I (Herleitung der Strategien), (2) Inhalt II (Didaktische Umsetzung), (3) Ausdruck, (4) Planen und (5) Überarbeiten. Der Ablauf der Arbeitsblöcke war dabei immer gleich gestaltet.

- (1) *Fachlicher Impuls* (z.B. Zeitschriftenartikel, Rollenspiel): Zu Beginn eines jeden Arbeitsblocks erhielten die Lehrerinnen einen fachlichen Impuls in Form eines Zeitschriftenartikels oder Buchkapitels bzw. Spiels. Damit setzten sich die Lehrerinnen zunächst alleine auseinander.
- (2) *Brainstorming/Teamarbeit*: In Teams beschäftigten sich die Lehrerinnen anschließend intensiv mit den fachlichen Impulsen. Zum Beispiel erarbeiteten sie Übungen zur Vertiefung der Inhaltsstrategie oder zum Üben des Revidierens und sammelten auf Karten Beispiele zur Verbesserung des Ausdrucks. Darüber hinaus probierten sie mögliche Übungen, die dem *Writer's Workshop* entstammten, aus. So erstellten sie in dem Planen-Arbeitsblock einen Zeitstrahl mit einschneidenden persönlichen Erlebnissen, die als Schreibimpulse dienen könnten.
- (3) *Diskussion/didaktische Umsetzung*: Im Anschluss daran trugen die Lehrerinnen die Ergebnisse der Teamarbeiten vor und diskutierten gemeinsam die konkrete didaktische Umsetzung des Themas des jeweiligen Arbeitsblocks.
- (4) *Ergebnissicherung*: Zum Schluss eines Arbeitsblocks einigten sich die Lehrerinnen auf konkrete Prinzipien, Übungen und Inhalte, die sie in der geplanten Unterrichtseinheit integrieren wollten. Dazu wurden auf der sogenannten „Konzept-Wand“, welche in der Mitte des Seminarraums hing, die Ergebnisse notiert.

Nach Abschluss der fünf Arbeitsblöcke fand am dritten Workshop-Tag die Abschlussbesprechung statt. Dazu wurden sämtliche in den Arbeitsblöcken zusammengetragenen Ergebnisse, die auf der „Konzept-Wand“ abgebildet waren, herangezogen. Als erstes einigten sich die Lehrerinnen auf einen inhaltlichen Ablauf der sechs Trainingsstunden:

- Trainingseinheit 1: Herleitung der AHA-7W-Strategie (Inhalt)
- Trainingseinheit 2: Herleitung der Ausdrucksstrategien
- Trainingseinheit 3: Herleitung der 3-Schritte-Technik
- Trainingseinheit 4: Übungen zum Generieren von Ideen/Planen
- Trainingseinheit 5: Übungen zum Bewerten von Geschichten
- Trainingseinheit 6: Anwendung aller erlernten Strategien

Als nächstes ordneten sie die Ideen auf der „Konzept-Wand“ den einzelnen Trainingseinheiten zu. Schließlich wurden organisatorische Punkte geklärt und abschließend über den Workshop reflektiert. Dazu füllten die Lehrerinnen einen Evaluationsbogen aus.

Im Anschluss an den Workshop fassten die beiden Workshopleiter die Workshop-Ergebnisse in einer Trainingsmappe zusammen. Diese enthielt alle für die Durchführung relevanten Materialien, d.h. für jede Trainingseinheit wurden Manuale (siehe CD-Anhang A6.2.1), Übungen (siehe CD-Anhang A6.2.2 und A6.2.4), Merkhilfen (siehe CD-Anhang A6.2.3), Kopiervorlagen für die Schüler und Instruktionmaterialien für die Lehrerinnen vorbereitet. Vor der Durchführung der LUG-Bedingung wurde den Lehrerinnen die Trainingsmappe zur Verfügung gestellt, so dass sich jede noch einmal individuell auf das Training vorbereiten konnte.

7.2.3.5.2 *Programmspezifische Komponenten*

Die konkrete Umsetzung der im LUG-Workshop festgelegten Trainingsinhalte wurde den Lehrerinnen selbst überlassen, so dass sich die Lehrerinnen nur auf wenige programmspezifische Komponenten einigten.

Rückmeldung: Im Workshop verständigten sich die Lehrerinnen auf ein konkretes Rückmeldesystem, welches in jeder Trainingseinheit eingesetzt wurde. Die von den Schülern im Training verfassten Geschichten wurden, wie auch in der USAT-Bedingung, von universitären Mitarbeitern ausgewertet. Im Workshop einigten sich die Lehrerinnen auf folgende Kategorien für die Bewertung von Aufsätzen: Anfang, Hauptteil, Abschluss, Adjektive, Verben, Satzanfänge, Gefühle/Gedanken, Wiederholungen und Kohärenz. Jede dieser Kategorien wurde von den Studenten auf einer vierstufigen Sterneskala bewertet. Kein Stern bedeutete, dass die Kategorie fehlte, ein Stern entsprach „vorhanden“, zwei Sterne „gut vorhanden“ und drei Sterne „sehr gut vorhanden“. Für jeden Schüler wurde auf diese Weise eine Rückmeldung erstellt, die jedoch nur den Lehrerinnen ausgehändigt wurde (siehe CD-Anhang A6.2.5). Die Schüler bekamen einen Universitätsstempel unter die von ihnen geschriebene Geschichte sowie eine allgemeine Aussage zu der Qualität der Geschichte. Die Aussagen entsprachen denen der WW-Bedingung der Studie I (siehe CD-Anhang A5.2.4).

7.2.3.5.3 *Ablauf des Trainings*

In der Tabelle 47 sind die Inhalte der einzelnen Trainingseinheiten skizziert. Insgesamt sechs Trainingseinheiten mit jeweils einer Dauer von 90 Minuten wurden von den Lehrerinnen durchgeführt. Jede Woche wurde eine Trainingseinheit realisiert, so dass sich das Training insgesamt über einen Zeitrahmen von sechs Wochen erstreckte. Wie auch in der USAT-Bedingung wurden die ersten drei Trainingseinheiten von zwei Psychologiestudenten im Hauptstudium beobachtet. Die beiden Beobachter entsprachen den Leitern des LUG-Workshops. Die Trainingsgeschichten wurden jede Woche abgeholt, so dass während des gesamten Trainingszeitraums ein kontinuierlicher Austausch zwischen den Psychologiestudenten und den Lehrerinnen bestand. Die Schülergeschichten wurden ebenfalls von den beiden Psychologiestudenten im Hauptstudium und einer weiteren Lehramtsstudentin wöchentlich bewertet. Das im Workshop von den Lehrerinnen ausgearbeitete Rückmeldeverfahren wurde dabei angewandt. Die im Workshop erarbeiteten Unterrichtseinheiten und von den zwei Psychologiestudenten verschriftlichten Manuale (siehe CD-Anhang A6.2.1) dienten den Lehrerinnen als Orientierung bei der Durchführung des LUG-Programms.

Im Gegensatz zur Studie I variierte die Kontrollgruppe der Studie II in vielfacher Hinsicht von der USAT-Bedingung, so dass im Folgenden die sechs Trainingseinheiten ausführlich dargestellt werden.

Trainingseinheit 1 (Schwerpunkt: Inhalt): Zu Beginn des LUG-Programms wurden Strategien zur Verbesserung des Inhalts einer Geschichte, die AHA-7W-Strategie, erarbeitet.

Die Stunde begann damit, dass die Lehrerin die Schüler nach den wichtigsten Inhalten einer guten Geschichte fragte. Im gesamten Klassenkontext wurden daraufhin die 7W-Fragen aus dem Gedächtnis der Schüler generiert, so dass am Ende die jeweilige Lehrerin ein AHA-7W-Poster (siehe Abbildung 58), welches fortan im Klassenzimmer hing, präsentierte. Während der Diskussion der Inhaltsstrategien im Workshop entwickelten die Lehrerinnen ein abgewandeltes AHA-7W-Poster, das sich von dem ursprünglichen unter anderem darin unterscheidet, dass anstatt „Ziel“ das „Thema“ als 7W-Frage aufgeführt wird. Weiter kritisierten die Lehrerinnen die Struktur des ursprünglichen 7W-Posters, so dass sie eine andere Darstellung wählten.

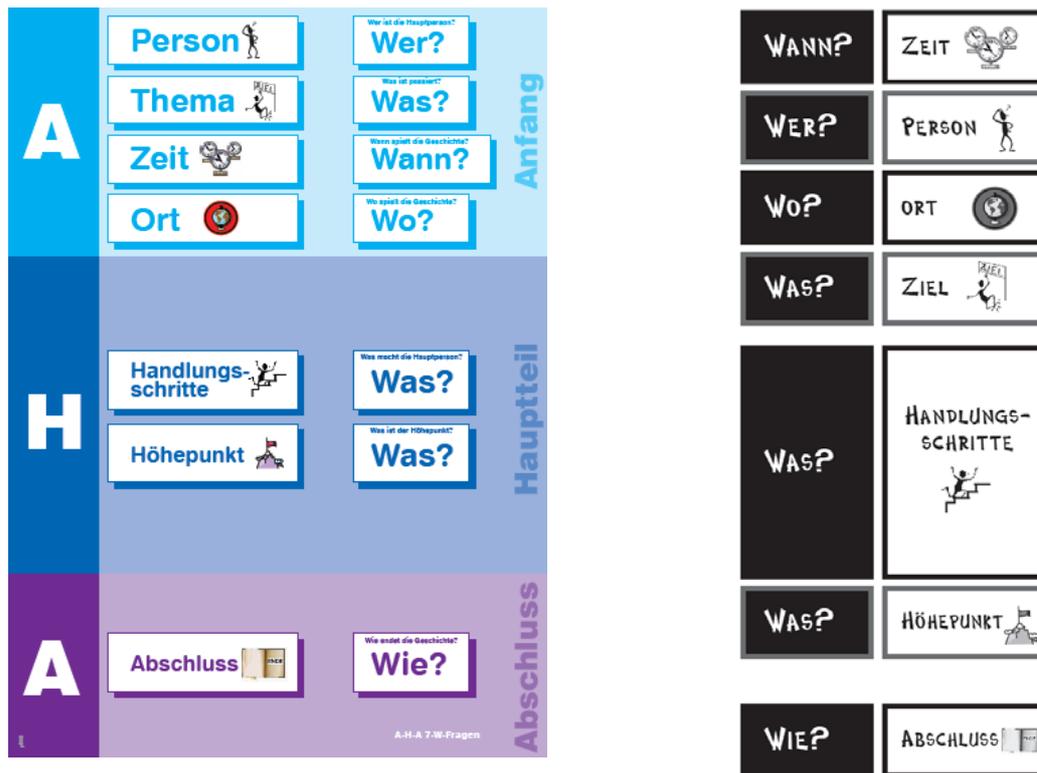


Abbildung 58: AHA-7W-Poster und Merkkärtchen der LUG-Bedingung.

Darüber hinaus erhielten die Schüler AHA-7W-Merkkärtchen (siehe Abbildung 58), die sie beim Verfassen einer Geschichte verwenden sollten. Die Merkkärtchen wurden in der USAT-Bedingung ebenfalls eingesetzt, jedoch zur Herleitung der AHA-7W-Strategie in der ersten Trainingseinheit.

Im Anschluss an die Herleitung der Inhaltsstrategie wurden verschiedene Übungen zur Festigung der erlernten Strategie durchgeführt. Als erstes las die Lehrerin fünf kurze Geschichten vor, in denen jeweils eine der 7W-Fragen fehlte. Die Schüler identifizierten mündlich die fehlenden 7W-Fragen (siehe CD-Anhang A6.2.2). Als nächstes hatten die Schüler die Auswahl zwischen einer von zwei Aufgaben. Entweder sie formulierten analog zu dem Spiel „Jeopardy“ zu zweit oder in einer Gruppe von maximal fünf Personen 7W-Fragen zu vorgegebenen Geschichtenbausteinen (siehe CD-Anhang A6.2.2) oder sie ermittelten in einer inhaltlich vollständigen Geschichte den Anfang, Hauptteil und Abschluss sowie die 7W-Fragen (siehe CD-Anhang A6.2.2).

Abschließend schrieben die Schüler eine Bildergeschichte. Zuvor generierte die Lehrerin im Klassenkontext Ideen, die die Schüler in ihre Geschichten einbauen konnten.

Trainingseinheit 2 (Schwerpunkt: Ausdruck): In der zweiten Trainingseinheit wurden die Strategien zur Verbesserung des Ausdrucks erarbeitet. Dabei handelt es sich um dieselben wie im USAT-Programm, d.h. um Adjektive, Verben, Gedanken und Gefühle.

Zu Beginn wurden die Geschichten der Schüler von der letzten Trainingseinheit zurückgegeben. Als nächstes wurde die AHA-7W-Strategie wiederholt. Dazu las die Lehrerin eine unvollständige Geschichte vor und die Schüler sollten fehlende 7W-Fragen erkennen und in der Geschichte ergänzen.

Die anschließende Herleitung der Ausdrucksstrategien erfolgte in einem Unterrichtsgespräch, an dessen Ende die jeweilige Lehrerin ein Gewürzposter in das Klassenzimmer hing (siehe Abbildung 59). Zusätzlich erhielt jeder Schüler eigene „Ausdrucksgewürze“ als Merkkärtchen (siehe CD-Anhang A6.2.3). Auf jedem Merkkärtchen war eine Gewürzmühle mit einem der folgenden Begriffe abgebildet: Adjektive?, Verben?, Gedanken?, Gefühle?, Satzanfänge?, Wiederholungen?, ausführlich erzählt?. Direkt im Anschluss an die Herleitung der Ausdrucksstrategien erfolgte eine Ausdrucks-Übung (siehe CD-Anhang A6.2.2), bei der die Schüler in einem Lückentext sprachliche Ausgestaltungsmittel ergänzen mussten.



Abbildung 59: Ausdrucksgewürze der LUG-Bedingung.

Am Ende der Stunde sammelte die Lehrerin zusammen mit den Schülern Ideen für eine Geschichte, die die Schüler im Anschluss zu einem vorgegebenen Bild des Malers Paul Klee schreiben sollten. Sie legte dabei besonderen Wert auf die Ausdrucksstrategien und generierte zusammen mit den Schülern vornehmlich zu dem Bild passende Adjektive, Verben, Gedanken. Darauf aufbauend verfassten die Schüler schließlich ihre eigenen Geschichten.

Trainingseinheit 3 (Schwerpunkt: 3-Schritte-Technik): Im Workshop entschieden sich die Lehrerinnen dafür, die 3-Schritte-Technik und die Strategien zum Planen und Überarbeiten erst nach der Einführung der Inhalts- und Ausdrucksstrategien einzuführen. Gemäß dem prototypischen Ablauf fanden in der dritten Trainingseinheit zunächst die Rückgabe der letzten Geschichten und die Wiederholung der zuletzt erlernten Ausdrucksstrategien anhand des Gewürze-Posters statt.

Die Einführung in das Planen einer Geschichte erfolgte anhand eines Arbeitsblatts (siehe Abbildung 60) und einer Bildergeschichte. Zunächst wurden im Unterrichtsgespräch Oberbegriffe generiert, die in die linke Spalte eingetragen wurden. Anschließend überlegte die jeweilige Lehrerin gemeinsam mit den Schülern Verben, Adjektive und ähnliche Begriffe, die zu dem Oberbegriff passten. Am Ende erhielt jeder Schüler ein eigenes Arbeitsblatt, mit dem er zu den vorgegebenen Bildern eine Geschichte plante.

Ideensammeln für deine Geschichte zum Thema _____

Oberbegriff	Finde Verben für den Oberbegriff!	Finde Adjektive für den Oberbegriff!	Welche ähnlichen Wörter fallen dir dazu ein?

Abbildung 60: Arbeitsblatt zum Planen einer Geschichte der LUG-Bedingung.

Nach dem Planen erfolgte eine Schreibphase. Ausgehend von den Notizen formulierten die Schüler Sätze. Die Lehrerin wies sie dabei an, immer eine Reihe freizulassen. Im Anschluss an das Schreiben fand das Überarbeiten statt. Darunter verstanden die LUG-Lehrerinnen das genaue Durchlesen der Geschichten unter besonderer Berücksichtigung der Ausdrucksgewürze und der 7W-Merkkärtchen. Die Schüler wurden aufgefordert, die eigenen Geschichten Schritt für Schritt im Hinblick auf die Ausdrucksgewürze und die 7W-Fragen zu prüfen.

Nachdem in der dritten Trainingseinheit die letzten Schreibstrategien eingeführt wurden, zeigte die Lehrerin am Ende das Trickposter. Dieses entsprach abgesehen von dem Zusammenhang dem Trickposter der WW-Bedingung der Studie I (siehe Abbildung 28). Darüber hinaus wurde das 3-Schritte-Technik-Poster (siehe Abbildung 29), welches ebenfalls in der Studie I in der WW-Bedingung verwendet wurde, in dem Klassenraum aufgehängt.

Trainingseinheit 4 (Schwerpunkt: Planen): In der vierten Trainingsstunde wurde das Planen einer Geschichte vertiefend im Rahmen einer Stationsarbeit behandelt.

Nach der Rückgabe der letzten Geschichte und der Wiederholung aller erlernten Strategien erklärte die Lehrerin die einzelnen Stationen der geplanten Stationsarbeit:

1. Station: Ausgestattet mit einer symbolischen Papierlupe sollten die Schüler als „kleine Detektive“ auf dem Schulhof Ideen für eine Geschichte sammeln und sich diese notieren.
2. Station: Aus einem Kasten mit Karteikarten, auf denen unterschiedliche Begriffen standen, sollten die Schüler eine Karte ziehen, die Augen schließen, sich den Begriff vorstellen und leise Assoziationen generieren.
3. Station: Ein Kartenspiel (siehe CD-Anhang A6.2.2) diente als Anregung für die Ideensammlung. Das Spiel bestand aus drei verschiedenen Kartensorten, die sich in unterschiedlichen Kästen befanden: 1. Ausgangssituation (Symbol: Tor), 2. magische Dinge und Kräfte (Symbol: Schlüssel) und 3. Aufgaben/Prüfungen (Symbol: Schloss). Aus jedem der drei Kästen zog ein Schüler eine Karte und überlegte sich eine Geschichte zu den Bildern.
4. Station: In Anlehnung an die 7W-Fragen war ein Set von Zauberkarten (siehe CD-Anhang A6.2.2) wie folgt sortiert: Gegenstand, Ereignis, Charakter, Ort und Eigenschaft. Die Schüler sollten eine Karte ziehen und diese als Impuls für das Generieren von Ideen verwenden.
5. Station: Im Workshop generierten die Lehrerinnen Geschichten über verschiedene Charaktere (siehe CD-Anhang A6.2.2). Diesen Charakteren mussten die Schüler Eigenschaften zuordnen und notieren.

Folglich lernten die Schüler an den verschiedenen Stationen Methoden zum Generieren von Ideen. Am Ende der Stunde besprach die Lehrerin die Erfahrungen der Schüler. Zudem stellte sie als weiteres Arbeitsblatt den „roten Faden“ (siehe CD-Anhang A6.2.3) als Hilfsmittel beim Planen vor. Der „rote Faden“ sollte es den Schülern erleichtern, die Planungsnotizen in eine chronologische Struktur zu bringen. In den letzten zehn Minuten der Trainingseinheit füllte jeder Schüler einen solchen roten Faden aus. Als Hausaufgabe schrieben die Schüler eine Geschichte zu ihrer geplanten Geschichte.

Trainingseinheit 5 (Schwerpunkt: Überarbeiten): Nach der Rückgabe der letzten Schülergeschichten und der Wiederholung der erlernten Strategien befasste sich die fünfte Trainingseinheit vornehmlich mit dem Thema „Überarbeiten“.

Zu Beginn wurde in der gesamten Klasse die Geschichte „Der doppelte Einkauf“ (siehe CD-Anhang A6.2.2), welche hinsichtlich der erlernten Schreibstrategien mehrere Mängel aufwies, besprochen. Hierzu stellte die Lehrerin die Textlupe (siehe Abbildung 34),

welche bereits in der WW-Bedingung der Studie I eingesetzt wurde, vor. Mit der gesamten Klasse überarbeitete die Lehrerin die Geschichte. Sie zeigte dabei die Anwendung der Textlupe, der 7W-Merkkärtchen und der Ausdrucksgewürze. Die Lehrerin wies die Schüler an, beim Überarbeiten der eigenen Geschichte genauso vorzugehen.

Danach sollte von den Schülern eine neue Geschichte verfasst werden. Zunächst hatten die Schüler zehn Minuten Zeit eine Geschichte zu planen. Sie sollten dabei den in der vierten Trainingseinheit eingeführten „roten Faden“ verwenden. Als Schreibanlass konnten sich die Schüler eine der in der Stationsarbeit der letzten Stunde generierten Ideen wählen. Anschließend hatten die Schüler zwanzig Minuten Zeit die Geschichte zu schreiben. Schließlich überarbeiteten die Schüler ihre Geschichte. Sie fügten Verbesserungen in die leeren Zeilen ein und verwendeten dabei die 7W-Merkkärtchen und die Ausdrucksgewürze. Dafür hatten sie wieder zehn Minuten Zeit.

Trainingseinheit 6 (Schwerpunkt: Wiederholung): Die sechste Trainingseinheit wurde von den LUG-Lehrerinnen explizit als Wiederholungsstunde geplant. Nach der Rückgabe der Schülergeschichten sah diese das Verfassen eines Märchens anhand von Bildern mit Reizwörtern vor. Ähnliche Bilder wurden bereits in der vierten Stationsarbeit der dritten Trainingseinheit verwendet (siehe CD-Anhang A6.2.2). Jedes der Bilder symbolisierte wichtige Inhalte des Märchens und war an die 7W-Fragen angelehnt: Gegenstand, Ereignis, Charakter, Ort und Eigenschaft. Mit der gesamten Klasse diskutierte die Lehrerin die Bilder und wiederholte dabei die Inhalts- und Ausdrucksstrategien. Es wurden gemeinsam Ideen für eine Geschichte generiert. Anschließend planten (zehn Minuten), schrieben (zwanzig Minuten) und überarbeiteten (zehn Minuten) die Schüler ihre eigene Geschichte.

Am Ende fand in der Klasse eine Reflektion der gesamten Unterrichtseinheit zum Geschichtenschreiben statt. Die Schüler konnten ihre Meinungen sagen und Verbesserungsvorschläge bezüglich des Trainings aufschreiben. Die jeweiligen Lehrerinnen händigten die Meinungen der Schüler den beiden Psychologiestudenten aus. Insgesamt äußerten sich die Schüler sehr positiv über das Training.

7.2.4 Genauigkeit der Implementation

Die Gewährung der Implementationsgenauigkeit war ein wichtiger Fokus in der Studie II. Da die Trainingsdurchführung von Lehrern anstatt von Trainern erfolgte, waren für die Sicherstellung der Implementationsgenauigkeit verschiedene Maßnahmen notwendig, die in der Studie I in abgeschwächter Form oder gar nicht angewandt wurden.

Den wichtigsten Beitrag zur Sicherung der Implementationsgenauigkeit leisteten die beiden Workshops, in denen den Lehrern der wissenschaftliche Ansatz der Studie und die

daraus resultierende Notwendigkeit, die Trainings einheitlich nach den vorgegebenen Standards durchzuführen, nahegebracht wurde.

Überdies wurden während der Durchführung der Trainings verschiedene weitere Instrumente zur Sicherung der Implementationsgenauigkeit eingesetzt. Die ersten drei Trainingseinheiten wurden in beiden Bedingungen von geschulten Beobachtern begleitet und hinsichtlich mehrerer für die Implementationsgenauigkeit relevanten Faktoren bewertet. Nach jeder Trainingseinheit füllten die Lehrer außerdem Checklisten zur Überprüfung der inhaltlichen Vollständigkeit aus. Des Weiteren beantworteten die Lehrer sowohl im Anschluss an den Workshop als auch nach Beendigung aller Trainingseinheiten einen Fragebogen zur Evaluation der Workshops und der Trainings.

7.2.4.1 Workshop

Zielsetzung der Workshops: Alle Lehrer in beiden Bedingungen durchliefen einen von zwei Workshops zur optimalen Vorbereitung auf die Trainings, die 12 bzw. 15 Stunden dauerten. In einer Untersuchung von Van Keer und Verhaege (2005) zeigte sich ein 13-stündiger Workshop zur Vorbereitung der Lehrer auf die Implementierung eines Lesetrainings als genauso effektiv wie ein 35-stündiger Workshop, so dass die geplante Workshop-Dauer in beiden Bedingungen als angemessen angesehen werden kann.

Insgesamt verfolgten die Workshops unterschiedliche Schwerpunkte und Zielsetzungen, da das USAT-Programm gemäß der Top-down- und das LUG-Programm gemäß der symbiotischen Implementierungsstrategie entwickelt wurden (Gräsel & Parchmann, 2004). In dem USAT-Workshop ging es vornehmlich darum, die Instruktionen der selbstregulatorischen Methoden und der Schreibstrategien so verständlich wie möglich darzustellen sowie diese von den Lehrern einüben zu lassen. Der Fokus des LUG-Workshops lag dagegen auf der Konzeption einer neuen Unterrichtseinheit zum Thema Geschichtschreiben, welche auf den Schreibstrategien der USAT-Bedingung und den Erfahrungen der Lehrerinnen basieren sollte.

Den beiden Bedingungen gemeinsam waren jedoch die Ziele, die Lehrer von der Wichtigkeit der Studie und der Effektivität der Trainings zu überzeugen und sie zu motivieren, möglichst engagiert im Workshop und im Training mitzuarbeiten. In beiden Workshops wurde insbesondere darauf Wert gelegt, dass sich die Lehrer bereit erklärten, die Trainings einheitlich und so wie in den Manualen beschrieben, umzusetzen. Aus diesem Grund wurden frühzeitig Unklarheiten und Verständnisschwierigkeiten ausgeräumt, so dass sich am Ende der Workshops die Lehrer im Durchschnitt gut vorbereitet fühlten, die Trainings in den eigenen Klassen durchzuführen. In der USAT-Bedingung gaben 50% der

Lehrer an, sich sehr gut vorbereitet zu fühlen und jeweils 17%, was einer Person entsprach, fühlten sich gut vorbereitet, weder gut noch schlecht vorbereitet oder schlecht vorbereitet. In der LUG-Bedingung fanden sich 80% der Lehrer sehr gut und 20% gut vorbereitet.

Evaluation der Workshops: Am Ende des jeweiligen Workshops füllten die Lehrer in beiden Bedingungen einen Evaluationsfragebogen (siehe CD-Anhang A6.3.3.1) aus. Auf einer Likert-Skala von eins („*Stimme überhaupt nicht zu*“) bis fünf („*Stimme sehr zu*“) bewerteten die Lehrer Aussagen zu verschiedenen Workshop relevanten Bereichen. Es handelte sich dabei um Items zur Bewertung des Workshopinhalts (z.B. „*Die Anwendung der Schreibstrategien sind mir deutlich geworden.*“), des Aufbaus bzw. der Organisation des Workshops (z.B. „*Die Zeiteinteilungen für Übungen und Vorträge waren angemessen.*“), der Workshopleiter (z.B. „*Die Leiter waren gut vorbereitet.*“) und der Workshopergebnisse („*Die Schüler werden von dem Trainingsprogramm profitieren.*“). Zudem erhielten die Lehrer auf dem Fragebogen die Gelegenheit, Verbesserungsvorschläge und Kritik aufzuschreiben. Am Ende bewerteten die Lehrer den Workshop sowie das Training auf einer schulüblichen Skala von eins (sehr gut) bis sechs (ungenügend). Die Workshops und Trainings wurden von den USAT-Lehrern (Workshop: 1.17, Training: 1.60) und den LUG-Lehrerinnen (Workshop: 1.33, Training: 1.60) als sehr gut bis gut eingeschätzt, wobei die USAT-Lehrer tendenziell bessere Bewertungen abgaben. Darüber hinaus wurden die Lehrer gefragt, ob sie den Workshop und das Training weiterempfehlen würden, was alle Lehrer in beiden Bedingungen mit „Ja“ beantworteten.

Die Reliabilitäten (Cronbach's α) für die Skalen liegen insgesamt in einem akzeptablen Bereich (Inhalt: $\alpha = .75$, Aufbau: $\alpha = .68$, Leiter: $\alpha = .67$, Ergebnisse: $\alpha = .75$). Tabelle 50 enthält die Mittelwerte und Standardabweichungen der Skalen und der Noten für beide Bedingungen getrennt.

Tabelle 50: Ergebnisse der Workshop-Evaluationen getrennt nach Bedingung (Studie II).

	Inhalt		Aufbau		Leiter		Ergebnisse		Note Workshop		Note Training	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
USAT	4.64	0.35	4.58	0.36	4.88	0.19	4.57	0.44	1.17	0.41	1.33	0.52
LUG	4.48	0.38	4.50	0.47	4.60	0.26	4.14	0.23	1.60	0.55	1.60	0.89
T-Wert	0.61		0.28		4.39		3.81		2.27		0.39	

Grundsätzlich kreuzten alle Lehrer in beiden Bedingungen vornehmlich die höchst möglichen Bewertungen an, was eine hohe Zufriedenheit mit beiden Workshops

widerspiegelt. Zudem lieferten einfaktorielle Varianzanalysen mit Bedingung (USAT, LUG) als unabhängige und die Workshopkomponenten (Inhalt, Aufbau, Leiter, Ergebnisse) sowie die Noten (Workshop, Trainer) jeweils als abhängige Variablen keine signifikanten Ergebnisse ($ps > .07$), so dass sich die Bewertungen der beiden Workshops nicht voneinander unterscheiden.

7.2.4.2 Training

In unmittelbarer zeitlicher Nähe zu den Workshops fand die Durchführung der Trainings in den eigenen Klassen statt. Um zu gewährleisten, dass die Implementation so wie in dem Workshop erarbeitet erfolgte, wurde die Qualität der Implementierung in den ersten drei Unterrichtseinheiten von geschulten Beobachtern anhand von standardisierten Skalen beurteilt. Zudem füllten die Lehrer Checklisten zur Überprüfung der inhaltlichen Vollständigkeit aus. Die Beobachter waren mit den Workshop-Leitern identisch.

7.2.4.2.1 *Unterrichtsbeobachtungen*

Unterrichtsbeobachtungen: Die Beobachter erfassten anhand eines standardisierten Beobachtungsbogens (siehe CD-Anhang A6.3.1) die ersten drei Unterrichtseinheiten der USAT- bzw. der LUG-Bedingung. Basierend auf Helmkes (2009) unterrichtsrelevanten Qualitätsmerkmalen (siehe Kapitel 2.4.2.3) wurden folgende Aspekte wirksamen Unterrichts erfasst:

- *Qualität der Instruktion:* Klarheit/Verständlichkeit, Prägnanz, Struktur/Unterrichtsorganisation, Hervorhebungen zentraler Aspekte, Instruktionstempo;
- *Klassenführung:* Effizienter Umgang bei Störungen, Zeitnutzung, Sicherstellung eines störungsfreien Ablaufs von Stillarbeitsphasen, unterstützendes und überwachendes Lehrerverhalten während der Stillarbeitsphasen;
- *Motivierungsqualität:* Angemessene Reaktion auf Misserfolg, Förderung von Interesse und Neugierde, Bekräftigung und Verstärkung, positives Sozialklima;
- *Individualisierung:* Einsatz von binnendifferenzierenden Maßnahmen, Förderungsorientierung inner- und außerhalb von Stillarbeitsphasen;
- *Aufmerksamkeit der Schüler:* aktives und passives Off-Task-Verhalten;
- *Sonstige Kriterien:* Akzentuierung lernbezogener Strategien, Variabilität von Unterrichtsformen, Verwendung von Stützmaßnahmen, Lehrerengagement, Schülerengagement.

Die Erfassung der Unterrichtsqualität erfolgte sowohl während als auch am Ende des Trainings. Während des Trainings machten sich die Beobachter in einem 20minütigen Abstand Notizen zu den einzelnen Qualitätsmerkmalen und bewerteten den Zeitabschnitt auf einer Likert-Skala von eins (*trifft vollkommen zu*) bis fünf (*trifft gar nicht zu*). Am Ende jeder Trainingseinheit beurteilten die Beobachter schließlich die gesamte Stunde auf einer fünfstufigen Skala. Die Notizen zu den Teilabschnitten wurden dabei für das Gesamturteil herangezogen.

Vor dem Training durchliefen die Doktorandin und die Psychologiestudenten eine Beobachterschulung. Da alle Beobachter sich bereits in der Konzeption der Workshops und der Trainings maßgeblich beteiligt hatten, waren sie bereits mit dem Ablauf der Unterrichtseinheiten vertraut. In einer separaten Sitzung diskutierten die Beobachter die einzelnen zu erfassenden Qualitätsmerkmale. Anschließend wurden gemeinsam zwei verschiedene Videosequenzen aus der Studie I, welche zwei verschiedene Trainerinnen zeigten, angeschaut. Während der Videos füllte jeder Beobachter die Bewertungsbögen aus, die anschließend diskutiert wurden. Unterschiedliche Bewertungstendenzen wurden identifiziert. Zudem wurde sich auf ein gemeinsames Maß an Strenge bzw. Milde bei der Bewertung der Unterrichtsqualität geeinigt.

In der nachfolgenden Tabelle 51 sind die Mittelwerte und Standardabweichungen für die Unterrichtsqualität getrennt nach Bedingung aufgelistet. Da der jeweilige Lehrer und nicht der Schüler beobachtet wurde, war Klasse die Analyseeinheit. Die Bewertungen wurden über alle drei Trainingseinheiten gemittelt. Je niedriger der Mittelwert, desto besser wurde das Qualitätsmerkmal eingeschätzt. Folgende Reliabilitäten (Cronbach's α) ergaben sich für die einzelnen Skalen: Instruktion: $\alpha = .93$, Klassenführung: $\alpha = .89$, Individualisierung: $\alpha = .85$, Motivierung: $\alpha = .94$, Aufmerksamkeit der Schüler: $\alpha = .94$, Sonstiges: $\alpha = .85$. Sämtliche Reliabilitäten sind als hoch einzustufen.

Tabelle 51: Mittelwerte, Standardabweichungen und T-Werte für die Unterrichtsqualität getrennt nach Bedingung (Studie II).

	Qualität		Führung		Individualisierung		Motivierung		Aufmerksamkeit		Sonstiges	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
USAT	2.07	0.69	1.94	0.75	2.15	0.34	1.81	0.64	1.92	0.72	1.70	0.35
LUG	1.95	0.87	2.31	0.99	3.21	0.34	2.17	0.95	2.37	1.31	2.23	0.60
T-Wert	0.60		0.48		26.63		0.57		0.53		3.36	

Anmerkungen: 1 = trifft vollkommen zu; 2 = trifft ziemlich zu; 3 = trifft mittelmäßig zu; 4 = trifft weniger zu; 5 = trifft gar nicht zu.

Für jedes Kriterium der Unterrichtsqualität wurden einfaktorielle Varianzanalysen mit Bedingung (USAT, LUG) als unabhängige und dem jeweiligen Kriterium als abhängige Variable durchgeführt. Insgesamt schnitten die USAT-Lehrer tendenziell besser ab als die LUG-Lehrerinnen. Abgesehen von der Individualisierung ($p < .01$) waren die Unterschiede jedoch nicht signifikant ($ps > .10$), so dass in den nachfolgenden statistischen Berechnungen der Einfluss der Qualität der Implementierung nicht berücksichtigt wurde.

7.2.4.2.2 *Checklisten*

In beiden Bedingungen wurden Checklisten (siehe CD-Anhang A6.3.2) eingesetzt, mit denen überprüft wurde, ob sämtliche geplanten Inhalte von den Lehrern vollständig umgesetzt wurden. Sie dienten somit der Überprüfung der inhaltlichen Vollständigkeit.

In den ersten drei Unterrichtseinheiten füllten die Beobachter die Checklisten aus. In den übrigen Trainingsstunden fungierten die Checklisten als Instrument zur Selbstkontrolle für Lehrer, um die inhaltliche Vollständigkeit zu überprüfen und verpasste Inhalte gegebenenfalls zu identifizieren und nachzuholen. Da die Instruktion sämtlicher Strategien in den ersten drei Einheiten erfolgte, wird im Folgenden die Vollständigkeit nur für diese Stunden berichtet.

Die eingesetzten Checklisten waren jedoch in den beiden Bedingungen unterschiedlich, weil zum einen die Inhalte differierten und zum anderen die inhaltlichen Vorgaben in der USAT-Bedingung strenger waren als in der LUG-Bedingung. Aus diesem Grund werden die Ergebnisse für die beiden Bedingungen getrennt dargestellt.

USAT-Bedingung: In der USAT-Bedingung wurde eine Checkliste zur Überprüfung der inhaltlichen Vollständigkeit und eine zur Überprüfung des Einsatzes selbstregulatorischer Methoden eingesetzt (siehe CD-Anhang A6.3.2).

Inhaltliche Vollständigkeit: Die Checkliste zur Kontrolle der inhaltlichen Vollständigkeit beinhaltete die einzelnen Untereinheiten des USAT-Programms (Einstieg, Rückmeldung, Instruktion, Zielsetzung, Sicherungsphase mit Planen, Schreiben, Überarbeiten), welche von den Lehrern auf einer Skala von null bis drei eingeschätzt wurden. Null bedeutete, dass die Einheit nicht umgesetzt wurde, eins, dass die Einheit teilweise umgesetzt wurde, zwei, dass die Einheit abweichend vom Manual umgesetzt wurde und drei, dass die Einheit äquivalent zum Manual implementiert wurde. Insgesamt kreuzten keine Beobachter und keine Lehrer in keiner Unterrichtseinheit Null an, so dass insgesamt 100% der vorgesehenen Inhalte umgesetzt wurden. In der Tabelle 52 sind die Prozentwerte für die inhaltliche Vollständigkeit für die ersten drei Trainingseinheiten aufgelistet.

Tabelle 52: Prozentwerte für die inhaltliche Vollständigkeit in den ersten drei Trainingseinheiten der USAT-Bedingung (Studie II).

	TE 1	TE 2	TE 3
Teilweise umgesetzt	10%	3.3%	13.3%
Abweichend vom Manual umgesetzt	40%	33.3%	43.3%
Wie im Manual umgesetzt	50%	63.3%	43.3%

Zusammenfassend wurden die ersten drei Unterrichtseinheiten größtenteils so wie im Manual vorgegeben von den Lehrern durchgeführt. Die geringsten Abweichungen vom Manual zeigten die Lehrer bei der Sicherungsphase des Planens, Schreibens und Überarbeitens. Abweichend vom Manual gestalteten die Lehrer die Rückmeldungen und die Instruktionen der Schreibstrategien. Die im Anschluss an die Trainingseinheiten stattgefundenen Gespräche zwischen den Beobachtern und den Lehrern führten ebenfalls dazu, dass Abweichungen vom Manual reduziert wurden.

Selbstregulation: Anhand eines weiteren Fragebogens (siehe CD-Anhang A6.3.2), welcher in den ersten drei Trainingseinheiten von den Beobachtern und anschließend von den Lehrern ausgefüllt wurde, erfolgte die Überprüfung der Verwendung von Ausdrücken zur Förderung der Motivation und von Einstellungen zur Unterstützung selbstregulatorischer Prozesse. Der Fragebogen basiert auf Vorschlägen der Lehrer aus dem Workshop. Zu den einzelnen Komponenten der Selbstregulation wie prozess- und produktbezogene Zielsetzung, Selbstüberwachung, Selbstkontrolle sowie Zusammenhang zwischen Schreibstrategien und Selbstregulation wurden beispielhafte Aussagen, die im Unterricht in ähnlicher Form von den Lehrern verwendet werden sollten, aufgelistet. Am Ende der Trainingseinheit schätzte der Beobachter bzw. der Lehrer auf einer Skala von -2 (*sehr selten*) bis +2 (*sehr oft*) die Häufigkeit des Einsatzes der selbstregulationsfördernden Komponenten ein.

Tabelle 53: Prozentwerte für selbstregulatorische Aussagen in den ersten drei Trainingseinheiten der USAT-Bedingung (Studie II).

	TE 1	TE 2	TE 3
Sehr selten	2.8%	0%	0%
Selten	22.2%	13.9%	25%
Oft	55.6%	50%	36.1%
Sehr oft	19.4%	36.1%	38.9%

In der ersten Trainingseinheit verwendete ein Lehrer keine selbstregulationsfördernde Aussagen im Bereich Selbstüberwachung. Davon abgesehen wurden von allen Lehrern in allen drei Unterrichtseinheiten verbale Ausdrücke verwendet, um den Einsatz und die Wichtigkeit selbstregulatorischer Prinzipien zu unterstützen. Den meisten Lehrern fiel es in der ersten Trainingseinheit noch etwas schwer, selbstregulatorische Äußerungen zu verwenden, da diese für sie wenig geläufig waren bzw. ihnen zunächst ungewohnt vorkamen. Im Verlauf des Trainings änderte sich dies jedoch zunehmend, so dass insgesamt die Förderung der Selbstregulation durch den Einsatz verbaler Aussagen von allen Lehrern umgesetzt wurde.

LUG-Bedingung: In der LUG-Bedingung füllten in den ersten drei Trainingseinheiten die Beobachter und in den weiteren Einheiten die Lehrer eine Checkliste (siehe CD-Anhang A6.3.2) zur Überprüfung der inhaltlichen Vollständigkeit aus. Die Checklisten enthielten eine Auflistung aller für die Unterrichtseinheit vorgesehenen Inhalte und Unterrichtsmaterialien, welche mit null („*wurde nicht eingesetzt*“) oder mit eins („*wurde eingesetzt*“) bewertet wurden. Eine Abstufung nach der Qualität der Umsetzung wie in der USAT-Bedingung wurde nicht vorgenommen, da den Lehrern die konkrete Umsetzung freigestellt wurde.

In der ersten Trainingseinheit wurden alle vorgesehenen Inhalte von allen Lehrern implementiert. In der zweiten und dritten Trainingseinheit wurden über alle Lehrer und Inhalte hinweg jeweils 90% umgesetzt. Insgesamt orientierten sich die Lehrer sehr stark an den vorgegebenen Manualen und setzten das Training größtenteils wie vorgesehen um.

7.2.4.2.3 *Evaluation des Trainings*

Wie auch am Ende des Workshops füllten die Lehrer am Ende des Trainings einen Evaluationsfragebogen aus. Aufgrund der inhaltlichen Unterschiede wichen die Evaluationsbögen der beiden Bedingungen voneinander ab (siehe CD-Anhang A6.3.3.2). Die Themen und die Anzahl der Items waren jedoch identisch. Um etwaige Einstellungsänderungen vor und nach dem Training zu erfassen, waren die Fragebögen wie der Evaluationsfragebogen zum Workshop aufgebaut.

Die Lehrer wurden gebeten, Aussagen zu den Bereichen Trainingsinhalt/Materialien (z.B. „*Die eingesetzten Materialien schätze ich als sehr effektiv ein.*“), Organisation/Aufbau des Trainings (z.B. „*Das Training verlief nach einer klaren Gliederung.*“) und Workshop- bzw. Trainingsergebnisse (z.B. „*Die Schreibleistungen der Schüler sind deutlich angestiegen.*“) auf einer fünfstufigen Likert-Skala von eins („*trifft gar nicht zu*“) bis fünf („*trifft vollständig zu*“) zu bewerten. In der Gesamtheit sind die Reliabilitäten (Cronbach's α)

der Skalen als gut zu bewerten: Inhalt: $\alpha = .93$, Aufbau: $\alpha = .85$, Ergebnisse: $\alpha = .87$. Unter dem Punkt „Sonstiges“ wurden überdies trainingspezifische Aspekte, wie z.B. die durchschnittliche Vorbereitungszeit für das Training, erfragt. Der Evaluationsfragebogen bot zudem ausreichend viel Platz für Verbesserungsvorschläge, Kritik und Anregungen. Die Lehrer wurden gebeten, auf einer fünfstufigen Skala einzuschätzen, inwieweit sie Adaptationen im Bereich Inhalt, Organisation, Lern-/Lehrmethoden und Arbeitsmaterialien vornehmen würden. Am Ende wurden die Lehrer abermals gefragt, welche Note (von eins bis sechs) sie dem Workshop und dem Training geben und ob sie das Training weiterempfehlen würden. Die Noten fielen etwas schlechter aus als nach dem Workshop. Nichtsdestotrotz wurde der USAT-Workshop (1.17) und das USAT-Programm (1.67) als sehr gut bis gut und der LUG-Workshop (2.00) und das LUG-Programm (3.00) als gut bis befriedigend eingeschätzt. Während vor dem Training alle Lehrer das jeweilige Trainingsprogramm weiterempfohlen hätten, würden es nach dem Training alle außer einer Lehrerin der LUG-Bedingung.

In der nachfolgenden Tabelle sind die Mittelwerte und Standardabweichungen für die einzelnen Skalen und die Noten getrennt nach Bedingung aufgelistet.

Tabelle 54: Ergebnisse der Trainings-Evaluationen getrennt nach Bedingung (Studie II).

	Inhalt		Aufbau		Ergebnisse		Note Workshop		Note Training	
	M	SD	M	SD	M	SD	M	SD	M	SD
USAT	4.53	0.36	3.94	0.30	4.47	0.56	1.17	0.41	1.67	0.82
LUG	3.79	0.70	3.67	0.96	3.84	0.80	2.00	0.00	3.00	1.22
T-Wert	5.32		0.46		2.31		20.46		4.68	

Einfaktorielle Varianzanalysen mit den jeweiligen Skalen als abhängige Variable und Bedingung (USAT, LUG) als unabhängige Variable lieferten einen signifikanten Effekt für die Skala Inhalt ($p < .05$) zu Gunsten der USAT-Bedingung. Für die übrigen Skalen ergaben sich keine signifikanten Unterschiede ($ps > .16$). Darüber hinaus bewerteten die USAT-Lehrer im Nachhinein den Workshop signifikant besser als die LUG-Lehrerinnen ($p < .01$). Die Bewertungen der Trainings unterschieden sich jedoch nicht signifikant zwischen den Bedingungen ($p > .06$).

Hinsichtlich der Vorbereitungszeit unterschieden sich die Lehrer der beiden Bedingungen nicht signifikant voneinander ($p > .05$). In der USAT-Bedingung gaben jeweils zwei Lehrer eine Vorbereitungszeit von 0 bis 30 Minuten, 30 bis 60 Minuten und 60 bis 90

Minuten an. In der LUG-Bedingung bereiteten sich jeweils zwei Lehrer durchschnittlich 0 bis 30 bzw. 30 bis 60 Minuten auf die Unterrichtseinheiten vor. Ein Lehrer setzte für die Vorbereitungszeit 60 bis 90 Minuten an.

Bezüglich Adaptationen unterschieden sich die Lehrer der beiden Bedingungen abermals voneinander. Während die USAT-Lehrer im Durchschnitt wenig Adaptationen im Bereich Inhalt ($M= 4.67^{14}$, $SD= .52$) und Lern- bzw. Lehrmethoden ($M= 4.33$, $SD= .52$) vornehmen würden, äußerten die LUG-Lehrerinnen mehrere Verbesserungsvorschläge in diesen Bereichen (Inhalt: $M= 2.20$, $SD= .84$; Lern-/Lehrmethoden: $M= 2.40$, $SD= 1.14$). Die Unterschiede zwischen den beiden Bedingungen waren signifikant (Inhalt: $p < .001$, Lern-/Lehrmethoden: $p < .01$). Hinsichtlich der Organisation (USAT: $M= 2.00$, $SD= 1.67$; LUG: $M= 3.00$, $SD= .71$) und der Arbeitsmaterialien (USAT: $M= 3.50$, $SD= 1.64$; LUG: $M= 2.40$, $SD= 1.14$) formulierten die Lehrer in beiden Bedingungen Kritik sowie Anregungen für Veränderungen. Diese bezogen sich vornehmlich auf die zeitliche Dauer und des Umfangs der Trainings, da sich die Lehrer häufig gestresst fühlten, die vorgesehenen Inhalte in der angegebenen Zeit durchzuführen. Insgesamt würden die Lehrer mehr Unterrichtsstunden für die beiden Trainingsprogramme einplanen.

Ein Vergleich der Evaluationen des Workshops mit denen des Trainings zeigt, dass die Bewertungen der LUG-Lehrerinnen nach dem Training deutlich schlechter ausfallen. In der USAT-Bedingung ist diese Tendenz nicht vorzufinden. Hier wurden größtenteils die Erwartungen des Workshops erfüllt. Die LUG-Lehrerinnen wurden im Anschluss an das Training zu einem Nachtreffen in den Räumen der Justus-Liebig-Universität eingeladen. Die in diesem Zusammenhang vorgebrachte Kritik umfasste vorwiegend die Gestaltung des Materialordners und der Manuale sowie die Auswahl der Schreibanlässe.

7.2.5 Testungsmaterial

Die Testungen sowie das Material entsprachen in großen Teilen denen der ersten Studie, so dass im Folgenden vor allem die vorgenommenen Änderungen beschrieben werden. Für eine ausführliche Darstellung wird auf die Studie I verwiesen (siehe Kapitel 6.2.5).

Die jeweiligen Leiter der Workshops (Doktorandin, drei Psychologiestudenten im Hauptstudium) führten direkt vor und direkt nach dem Training sowie sechs Wochen nach der Posttestung Erhebungen der strategienahen und -fernen Variablen durch. Jeder erhielt ein ausführliches Manual (siehe CD-Anhang A2) zur Durchführung der Testungen und wurde angeleitet, das Manual korrekt anzuwenden, um einen fehlerlosen Ablauf der Testungen zu

¹⁴ Je höher der Mittelwert, desto zufriedener waren die Lehrer mit dem Trainingsprogramm in Bezug zu Inhalt, Organisation, Lern-/Lehrmethoden und Arbeitsmaterialien.

gewährleisten. Zusätzlich begleitete und unterstützte jeweils eine weitere Lehramtsstudentin die Tester. Die Schüler arbeiteten individuell und ohne Hilfestellungen an den jeweiligen Aufgaben. Sie wurden jedoch im gesamten Klassenkontext instruiert.

Schreibanlass: Aufgrund der positiven Erfahrungen mit den *Vater und Sohn*-Geschichten von Plauen (2003) der Studie I wurden diese ebenfalls in der Studie II eingesetzt. Die Randomisierung erfolgte genauso wie in der Studie I.

Ablauf der Testungen: Die Testungen liefen wie in der Studie I ab und waren in beiden Bedingungen völlig identisch. Eine Testung dauerte insgesamt 90 Minuten. Sie umfasste folgende Komponenten:

- Planen (zehn Minuten), Schreiben (zwanzig Minuten) und Überarbeiten (zehn Minuten) einer von drei Bildergeschichten;
- Verteilung von Mandalas für Schüler, die mit dem Planen, Schreiben und Überarbeiten eher fertig waren;
- Durchführung des Wissenstests (zehn Minuten) und des Herkunftsfragebogens.

7.2.5.1 Abhängige Variablen

Sämtliche abhängigen Maße wurden zu allen drei Messzeitpunkten erhoben. Neben der Wortanzahl waren sämtliche strategienahen Maße mit denen der Studie I identisch und können in Kapitel 6.2.5.1 nachgelesen werden:

- Wortanzahl;
- Geschichtenelemente (siehe Tabelle 18);
- Planen Inhalt;
- Revisionen;
- Schreibbezogenes Wissen.

Die Erfassung von Planen Struktur entfiel, weil der zusätzliche Informationsgewinn als überflüssig angesehen wurde.

Narrative Qualität/Geschichtenqualität: In der Studie II wurde als strategiefernes Schreibmaß, welches nicht direkt die in den Trainings vermittelten Schreibstrategien erfasst, die Narrative Qualität erhoben. Die Begriffe Narrative Qualität und Geschichtenqualität werden im Folgenden synonym verwandt. Die Bewertung der Narrativen Qualität wich dabei von der Studie I ab.

In der Studie I wurde die Narrative Qualität auf einer sechsstufigen Skala eingeschätzt. Um die Einschätzung der Narrativen Qualität zu erleichtern, erfolgte die Bewertung der Narrativen Qualität Sprache und der Narrativen Qualität Struktur in Studie II getrennt. Während die Narrative Qualität Sprache sich explizit auf die sprachliche Ausgestaltung, d.h.

die Verwendung eines umfangreichen Wortschatzes und stilistischer Mittel, bezieht, umfasst das Maß Narrative Qualität Struktur die Ausarbeitung der Geschichte in Bezug auf Inhalt und Kohärenz. Die separate Einschätzung der Sprache und der Struktur sollte eine Gleichgewichtung beider Faktoren in dem Gesamtmaß garantieren. Da in der Studie I die Sprache und die Struktur zusammen in der sechsstufigen Skala erfasst wurden, wurde zur besseren Vergleichbarkeit der Ergebnisse der beiden Studien, nur mit dem Gesamtmaß Narrative Qualität gerechnet. Zu allen drei Messzeitpunkten korrelierten auf Klassenebene Narrative Qualität Sprache und Narrative Qualität Struktur signifikant miteinander ($ps < .05$) (Prätest: $r = .64$, Posttest: $r = .63$, Follow-Up: $r = .76$). Sowohl Narrative Qualität Sprache als auch Struktur wurden jeweils auf einer sechsstufigen Skala eingeschätzt, so dass sich insgesamt für die Narrative Qualität eine Maximalpunktzahl von zwölf ergab. Für jede Stufe der sechsstufigen Skalen wurden spezifische Bewertungskriterien festgelegt, welche in der Tabelle 55 und 56 aufgelistet sind. Überdies erhielten die Beurteiler Beispielgeschichten für jede Stufe als Vergleichsmaßstab.

Tabelle 55: Bewertungskriterien der Geschichtenqualität Sprache (Studie II).

Punkt- zahl	Sprachliche Ausarbeitung	Bewertungskriterien
1	Verwendung basaler Sprachelemente	<ul style="list-style-type: none"> ▪ Eintönige Wortwahl, d.h. keine Verwendung von Adjektiven und abwechslungsreichen Verben aus dem Grundwortschatz
2	Anwendung eines narrativen Grundwortschatzes	<ul style="list-style-type: none"> ▪ Ansatzweise Verwendung des trainierten Wortschatzes
3	Narrativer Wortschatz	<ul style="list-style-type: none"> ▪ Überwiegende Verwendung des trainierten Wortschatzes
4	Erweiterter narrativer Wortschatz	<ul style="list-style-type: none"> ▪ Durchgängige Verwendung des trainierten Wortschatzes ▪ Verwendung unterschiedlicher sprachlicher Mittel
5	Ausgearbeiteter narrativer Wortschatz	<ul style="list-style-type: none"> ▪ Sprachliche Ausgestaltung wurde in besonderem Maße umgesetzt ▪ Verwendung von bildlichen Beschreibungen und wörtlicher Rede
6	Voll entfalteter narrativer Wortschatz	<ul style="list-style-type: none"> ▪ Hervorragende und außergewöhnliche Verwendung der sprachlichen Mittel (Funktionalität der Sprache)

Tabelle 56: Bewertungskriterien der Geschichtenqualität Struktur (Studie II).

Punkt-zahl	Strukturelle Ausarbeitung	Bewertungskriterien
1	Auflistung von Ereignissen	<ul style="list-style-type: none"> ▪ Aneinanderreihung mehrerer zusammenhangsloser Sätze
2	Wenig ausgearbeitete Geschichte	<ul style="list-style-type: none"> ▪ Auflistung mehrerer Sätze ▪ Ansatzweise inhaltlich vollständig
3	Grundstruktur	<ul style="list-style-type: none"> ▪ Grundstruktur (Anfang, Hauptteil, Abschluss) vorhanden ▪ Teilweise inhaltlich vollständig ▪ Mängel in Bezug auf Kohärenz (Gedankensprünge, isolierte Elemente)
4	Erweiterte Geschichtenstruktur	<ul style="list-style-type: none"> ▪ Grundstruktur vorhanden und überwiegend inhaltlich vollständig ▪ Mängel in Bezug auf Kohärenz (Gedankensprünge, isolierte Elemente)
5	Ausgearbeitete Geschichte	<ul style="list-style-type: none"> ▪ Grundstruktur vorhanden und überwiegend inhaltlich vollständig ▪ Kohärenz überwiegend vorhanden
6	Voll entfaltete Geschichte	<ul style="list-style-type: none"> ▪ Grundstruktur vorhanden und überwiegend inhaltlich vollständig ▪ Kohärenz überwiegend vorhanden ▪ Gesamtstruktur bildet eine Einheit ▪ Passung, Übergänge und Sinnzusammenhänge hervorragend realisiert

Interrater-Reliabilitäten: Die beschriebenen abhängigen Maße wurden von der Doktorandin ausgewertet. Zudem wertete eine geschulte Psychologiestudentin im Hauptstudium, die mit dem Design und der Fragestellung der Studie nicht vertraut war, eine zufällige Auswahl von 30% der Daten aus. Zur Sicherstellung der Auswertungsobjektivität wurden Interrater-Reliabilitäten berechnet. Die Wörter wurden nur einmal gezählt.

Tabelle 57: Interrater-Reliabilitäten (Studie II).

	Prätest	Posttest	Follow-Up
Geschichtenelemente	0.87	0.91	0.91
Narrative Qualität	0.92	0.82	0.93
Planen	0.87	0.94	0.95
Revisionen	0.92	0.72	0.95
Wissen	0.82	0.93	0.99

Für sämtliche abhängige Maßen ergaben sich akzeptable bis sehr hohe Interrater-Reliabilitäten zwischen $r = .72$ und $r = .99$.

7.2.5.2 Kontrollvariablen

Vor der Prätestung wurden die Kontrollvariablen Rechtschreibung und Sprachverständnis im gesamten Klassenkontext von der Doktorandin und drei Psychologiestudenten im Hauptstudium erhoben (siehe Abbildung 14). Alle Tester orientierten sich an den Durchführungsmanualen der Hersteller. Im Gegensatz zur Studie I wurde das Sprachverständnis zusätzlich erhoben. Es stellt ein Kontrollmaß für den Wortschatz und das Verständnis von Sätzen dar. Überdies wurde der bereits in der Pilotstudie und Studie I eingesetzte Herkunftsfragebogen (siehe CD-Anhang A2, Kapitel 6.2.5.2) zur Erfassung der Nationalität und der Muttersprache durchgeführt.

Rechtschreibung: In der Studie II wurde der Diagnostische Rechtschreibtest für dritte Klassen (DRT 3) (Müller, 2004) durchgeführt. Da die Erhebungen der Kontrollvariablen der Studie II bereits am Anfang des vierten Schuljahrs stattfanden, war nur die Durchführung des DRT 3, welcher für das Ende des dritten und Anfang des vierten Schuljahrs konzipiert und überprüft wurde, zur Erfassung der Rechtschreibfähigkeiten zulässig. Wie auch der DRT 4 ist der DRT 3 ein normierter und standardisierter Schulleistungstest, welcher als Gruppentest durchgeführt werden kann. Der DRT 3 besteht aus 44 in kurze Sätze eingebettete Wörter mit zunehmender Schwierigkeit, die den Schülern diktiert werden und welche diese in die Satzlücken schreiben müssen. Für den DRT 3 (Müller, 2004) wird eine Retest-Reliabilität von $r = .92$ und eine Split- Half- Reliabilität von $r = .95$ angegeben.

Sprachverständnis: Zur Erfassung des Sprachverständnisses wurden die beiden Untertests Wörter und Sätze des Hamburger Schulleistungstests für vierte und fünfte Klassen (Mietzel & Willenberg, 2000) herangezogen. Der HST 4/5 wird ebenfalls als Gruppentest durchgeführt und eignet sich als Test für Schüler beim Übergang von der Grundschule in die Jahrgangsstufe fünf. Die beiden Untertests Wörter und Sätze enthalten Aufgaben nach dem Multiple-Choice-Verfahren, wonach bei vier Antwortalternativen eine richtige Lösung zu finden ist. Beim Untertest Wörter müssen die Schüler zu einem vorgegebenen Begriff ein bedeutungsäquivalentes Wort finden. Der Untertest Sätze enthält mehrere zusammenhängende Sätze, in denen ein grammatikalischer Fehler vorkommen kann, welcher von den Schülern identifiziert werden muss. Die Autoren geben als Reliabilitätsmaß den Alpha-Koeffizienten von Cronbach mit $\alpha = .89$ an.

7.3 Ergebnisse

Um die Ergebnisse der Studie I und II besser vergleichen zu können, entspricht das statistische Vorgehen im Wesentlichen dem der Studie I. Für genauere Ausführungen wird deswegen auf das Kapitel 6.3 verwiesen. Um Aussagen über die Effektivität des USAT-Programms im Vergleich zum LUG-Programm treffen zu können, werden die Ergebnisse von Messwiederholungsanalysen, Kovarianzanalysen und Effektstärkenberechnungen aufgeführt.

7.3.1 Statistische Analysen

7.3.1.1 Begründung der Analyseeinheit Klasse

Analog zur Studie I wurde die Klasse als Analyseeinheit gewählt, da sämtliche Instruktionen und Testungen auf Klassenebene stattfanden und die Zuweisung zu den Bedingungen ebenfalls im gesamten Klassenkontext erfolgte. Die genaue Begründung für dieses Vorgehen wurde ausführlich in der Studie I beschrieben und gilt gleichermaßen für die Studie II (siehe Kapitel 6.3.1.1). Für jede abhängige Variable und Kontrollvariable wurden über alle Schüler hinweg pro Klasse Mittelwerte berechnet, mit denen die folgenden statistischen Analysen durchgeführt wurden.

7.3.1.2 Voranalysen

Kontrollvariablen: Um den Einfluss möglicher Kontrollvariablen auf die statistischen Ergebnisse zu erfassen, wurden die Variablen Klassengröße, Geschlecht, Rechtschreibung, Sprachverständnis, Nationalität und Muttersprache Deutsch jeweils als Kovariate gerechnet. Bei allen statistischen Analysen der Kontrollvariablen wurde die Klasse als Analyseeinheit gewählt.

Zunächst wurden T-Tests für unabhängige Stichproben mit Bedingung (USAT, LUG) als unabhängige Variable und der jeweiligen Kontrollvariable (Klassengröße, Geschlecht, Rechtschreibung, Sprachverständnis, Nationalität, Muttersprache Deutsch) als abhängige Variable gerechnet. Abgesehen von Rechtschreibung unterschieden sich die Bedingungen hinsichtlich der Kontrollvariablen nicht voneinander ($ps > .11$). Für Rechtschreibung ergab sich ein signifikanter Haupteffekt Bedingung zu Gunsten der LUG-Bedingung, $F(1, 9) = 6.63$, $p < .05$.

Weiter wurden Kovarianzanalysen mit den jeweiligen strategienahen und -fernen Schreibmaßen als abhängige Variable, dem Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), der Bedingung (USAT, LUG) als Zwischensubjektfaktor und den einzelnen Kontrollvariablen (Klassengröße, Geschlecht, Rechtschreibung, Sprachverständnis,

Nationalität, Muttersprache Deutsch) sowie den jeweiligen Prätest-Werten als Kovariate durchgeführt. Die statistische Signifikanz des Haupteffekts Bedingung veränderte sich unter Berücksichtigung der Kontrollvariablen Klassengröße, Geschlecht, Nationalität und Muttersprache Deutsch nicht bedeutsam. Dies traf jedoch unter Berücksichtigung der Kontrollvariablen Rechtschreibung zu. Aus diesem Grund und wegen der signifikanten Bedingungseffekte wurde die Rechtschreibung als Kovariate in den nachfolgenden statistischen Hauptanalysen berücksichtigt.

Als letztes wurden Produkt-Moment-Korrelationen gerechnet, um den Zusammenhang zwischen den Kontrollvariablen (Klassengröße, Geschlecht, Rechtschreibung, Sprachverständnis, Nationalität, Muttersprache Deutsch) und den abhängigen Variablen (Prätest, Posttest, Follow-Up) zu ermitteln. Die Kontrollvariable Sprache korrelierte signifikant mit mehreren anderen Kontrollvariablen wie Nationalität ($r = .63, p < .05$), Geschlecht ($r = -.69, p < .05$) und Rechtschreibung ($r = .70, p < .05$). Darüber hinaus ergaben sich signifikante Korrelationen zwischen Klassengröße und Narrativer Qualität Follow-Up ($r = .68, p < .05$), Klassengröße und Revisionen Posttest ($r = .62, p < .05$), Muttersprache Deutsch und Geschichtenelemente Follow-Up ($r = -.65, p < .05$), Sprache und Planen Follow-Up ($r = -.64, p < .05$) und schließlich Sprachverständnis und Wissen Posttest ($r = -.77, p < .05$).

7.3.1.3 Statistische Hauptanalysen

Um die Effektivität und Wirksamkeit des USAT-Programms im Vergleich zum LUG-Programm zu überprüfen, wurde folgendermaßen statistisch vorgegangen:

Zwei Schüler (jeweils einer aus jeder Bedingung) fehlten an zwei Testungen, so dass diese aus der Stichprobe genommen wurden und sich eine Gesamtstichprobe von 242 Schülern ergab. Die fehlenden Werte von Schülern, die nur zu einem der drei Messzeitpunkte abwesend waren, wurden regressionsanalytisch ersetzt.

Zur genaueren Analyse des gruppenrandomisierten Designs der Studie II wurden Intra-Klassen-Korrelationen berechnet, die Mittelwerte der Klassen innerhalb einer Bedingung betrachtet und Mixed Model-ANCOVA's durchgeführt.

Da in den Studien I und II natürlich vorkommende Klassen als Untersuchungseinheiten gewählt wurden, variierte in der Studie II eine Prätest-Variable, Revisionen, signifikant zwischen den beiden Bedingungen ($p < .01$). Für alle anderen abhängigen Maße zeigten sich keine signifikanten Bedingungseffekte beim Prätest. Dies änderte sich auch nicht, wenn Rechtschreibung als Kovariate berücksichtigt wurde ($ps > .05$) (siehe Tabelle 70).

In der Studie II wurden folgende abhängige Variablen untersucht: Schreibleistung (Geschichtenelemente, Geschichtenqualität, Wortanzahl) und strategienahe Schreibmaße (Planen, Revisionen, Wissen). Anknüpfend an Studie I wurden zunächst Messwiederholungsanalysen mit den jeweiligen genannten Maßen als abhängige Variable, dem Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, LUG) als Zwischensubjektfaktor und den Prätest-Variablen sowie Rechtschreibung als Kovariaten gerechnet. Danach wurden kovarianzanalytische Modelle mit den jeweiligen Schreibmaßen einmal für Posttest und einmal für Follow-Up getrennt als abhängige Variable, Bedingung (USAT, LUG) als unabhängige Variable und die Prätest-Variable sowie die Rechtschreibung als Kovariaten herangezogen. Im Vorhinein wurden die Voraussetzungen der Kovarianzanalyse, vor allem die Homogenität der Regressionsgeraden, überprüft. Für keine der abhängigen Variablen ergab sich eine signifikante Prätest x Bedingung-Interaktion ($p > .05$), so dass die Voraussetzungen der Kovarianzanalyse erfüllt wurden.

Obwohl die Klasse als Analyseeinheit die statistische Power reduziert, wurde eine Adjustierung der Fehlerwahrscheinlichkeiten nicht vorgenommen, da sich bei einer Verschärfung des α -Niveaus in den zentralen Befunden keine Änderungen ergaben.

Als nächstes wurde zur Beurteilung der praktischen Signifikanz der Ergebnisse Effektstärken nach Cohen's d (Cohen, 1988) berechnet, welche um Prätestunterschiede korrigiert wurden. Da die konventionelle Schätzung der Effektstärken auf Klassenebene zu inkorrekten Standardabweichungen führte, wurden zur besseren Interpretierbarkeit die Effektstärken sowohl auf Klassen- als auch auf Schülerebene berichtet.

Als letztes wurden in einem weiteren Abschnitt die strategienahen und -fernen Schreibmaße differenzierter für die Subgruppen der Kontrollvariablen Geschlecht (Mädchen, Jungen), Rechtschreibung (rechtschreibschwach, rechtschreibstark) und Sprache (Deutsch, nicht-Deutsch) analysiert, um den Einfluss möglicher Moderatoreffekte zu prüfen.

7.3.2 Zusätzliche Analysen für gruppenrandomisierte Designs

Analog zur Studie I wurden zusätzliche Untersuchungen zur Beschreibung des Problems der Analyseeinheit bei gruppenrandomisierten Designs durchgeführt.

7.3.2.1 Intra-Klassen-Korrelationen (ICC's)

Zunächst wurden Intra-Klassen-Korrelationen (ICC's) berechnet, um den Einfluss unterschiedlicher Klassen innerhalb einer Bedingung zu untersuchen. ICC's geben den Zusammenhang zwischen den Einzelwerten einer Gruppe an. Ein Zusammenhang nahe Null bedeutet, dass die Klassen einer Bedingung unabhängig voneinander sind und Klasse als

Analyseeinheit gewählt werden kann. Tabelle 58 enthält die ICC's für die einzelnen abhängigen Variablen getrennt nach Bedingung.

Tabelle 58: Geschätzte Intra-Klassen-Korrelationen (ICC's) getrennt nach Messzeitpunkt und Bedingung (Studie II).

	USAT			LUG		
	Prätest	Posttest	Follow-Up	Prätest	Posttest	Follow-Up
Geschichtenelemente	0.017	0.040	0.020	0.065	0.028	0.014
Narrative Qualität	0.001	0.005	0.054	0.047	0.006	0.025
Wortanzahl	0.028	0.005	0.046	0.130	0.077	0.050
Planen Inhalt	0.033	0.004	0.058	0.022	0.104	0.095
Revisionen	0.001	0.006	0.104	0.015	0.047	0.095
Wissen	0.099	0.044	0.137	0.203	0.062	0.141

Für die USAT-Bedingung ergaben sich ICC's zwischen .001 und .137. Für die LUG-Bedingung ergaben sich ICC's zwischen .006 und .203. Somit liegen die ICC's der USAT-Bedingung in dem von Hedges und Hedberg (2007a, 2007b) angegebenen optimalen Bereich von maximal .15. Die Werte der LUG-Bedingung sind abgesehen von Prätest Wissen ebenfalls kleiner als .15. Folglich kann für die meisten abhängigen Variablen geschlussfolgert werden, dass in beiden Bedingungen keine Zusammenhänge zwischen den Klassen einer Bedingung bestanden.

7.3.2.2 Deskriptive Beschreibung der Klassenunterschiede

Die Betrachtung der Mittelwerte und Standardabweichungen der einzelnen abhängigen Variablen getrennt nach Messzeitpunkt, Bedingung und Klasse diente dazu, potenzielle Ausreißer-Klassen zu identifizieren.

7.3.2.2.1 *Geschichtenelemente*

Die deskriptiven Statistiken für Geschichtenelemente getrennt nach Messzeitpunkt, Bedingung und Klasse sind in Tabelle 59 aufgelistet.

Tabelle 59: Mittelwerte und Standardabweichungen für Geschichtenelemente getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).

	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	2.08	1.52	9.08	1.75	7.28	1.99
Klasse 2	2.15	1.18	7.65	2.37	7.20	2.35
Klasse 3	2.19	1.66	7.33	3.18	6.52	2.34
Klasse 4	2.38	1.24	8.29	1.98	6.86	2.57
Klasse 5	1.62	1.12	8.71	2.59	8.24	2.23
Klasse 6	2.48	1.63	7.52	2.53	7.56	2.58
LUG						
Klasse 7	1.73	1.32	3.64	1.36	4.45	1.57
Klasse 8	2.39	1.50	4.43	1.41	4.61	1.27
Klasse 9	2.67	1.28	4.10	1.51	4.14	1.46
Klasse 10	2.94	1.21	3.72	1.90	3.67	1.37
Klasse 11	2.87	1.55	4.74	1.51	4.26	2.12

Abbildung 61 ist eine grafische Darstellung der Mittelwerte der USAT-Klassen. Es ist erkennbar, dass Klasse 1 beim Posttest am meisten dazu gewann, jedoch genauso wie Klasse 4 zum Follow-Up einen stärkeren Leistungsabfall im Vergleich zu den anderen USAT-Klassen zeigte.

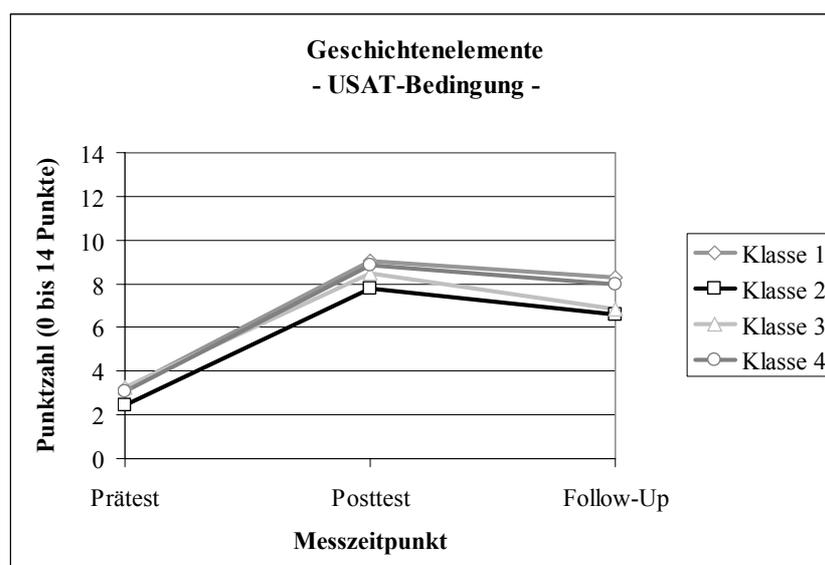


Abbildung 61: Geschichtenelemente der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

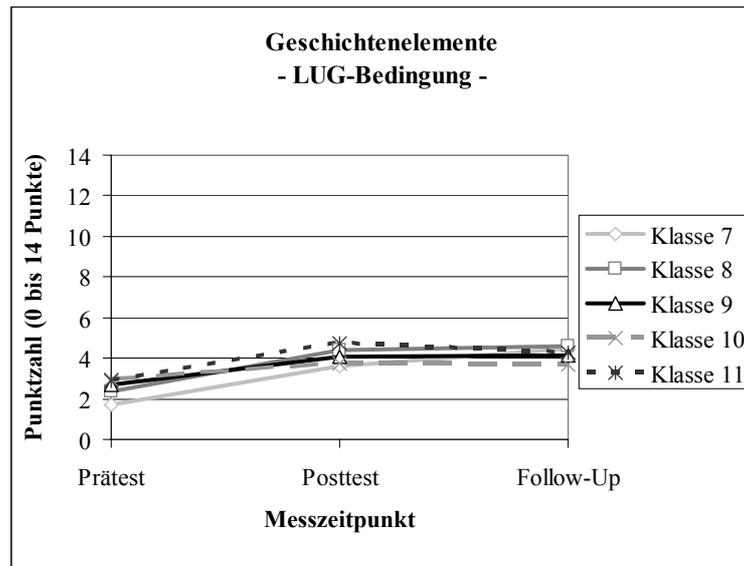


Abbildung 62: Geschichtenelemente der LUG-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

In der Abbildung 62 sind die Geschichtenelemente-Mittelwerte für die LUG-Klassen abgebildet. Klasse 7 schnitt beim Prätest am schlechtesten ab und zeigte überdies den steilsten Verlauf zum Follow-Up. Die Schüler der Klasse 10 zeigten die geringsten Verbesserungen vom Prätest zum Follow-Up.

7.3.2.2.2 *Geschichtenqualität*

Die deskriptiven Statistiken für Geschichtenqualität getrennt nach Messzeitpunkt, Bedingung und Klasse sind in Tabelle 60 zu finden.

Tabelle 60: Mittelwerte und Standardabweichungen für Geschichtenqualität getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).

	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	3.44	0.82	6.04	1.21	5.36	1.11
Klasse 2	2.90	0.85	5.35	0.99	4.90	0.85
Klasse 3	3.29	1.19	5.38	1.47	5.19	1.08
Klasse 4	3.43	0.98	5.33	1.11	5.00	0.95
Klasse 5	3.19	1.08	5.76	1.00	5.71	0.90
Klasse 6	3.33	1.36	5.48	1.55	5.78	1.22

LUG						
Klasse 7	2.86	0.83	5.32	1.13	5.23	0.97
Klasse 8	3.26	1.18	5.57	0.95	5.22	1.31
Klasse 9	3.62	1.07	5.19	1.21	4.81	1.47
Klasse 10	3.61	0.98	5.11	0.83	4.61	0.92
Klasse 11	3.43	1.20	5.39	1.31	4.87	1.36

Abbildung 63 enthält eine grafische Darstellung der Geschichtenqualität-Mittelwerte der USAT-Bedingung. Die Linien verlaufen nahezu parallel, wobei Klasse 1 beim Posttest genauso wie bei den Geschichtenelementen am besten abschnitt. Des Weiteren sind in Tabelle 60 große Leistungsunterschiede beim Follow-Up zwischen der Klasse 6 und Klasse 2 zu erkennen.

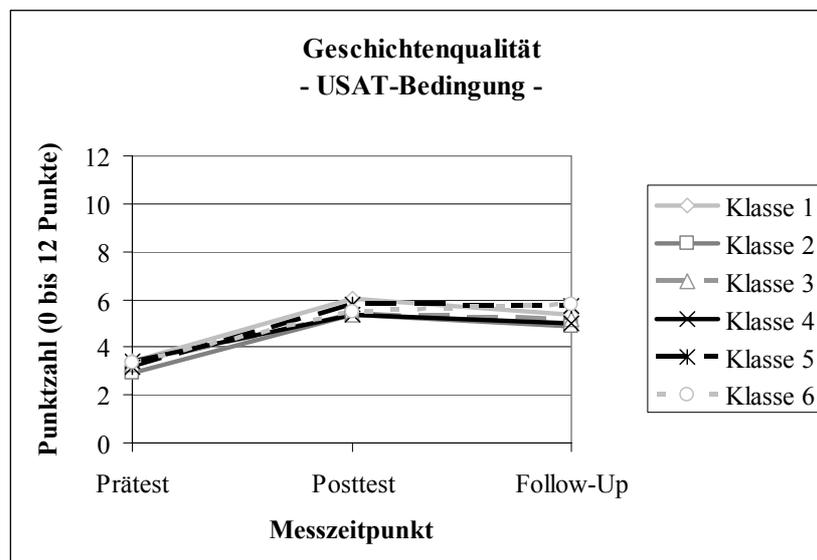


Abbildung 63: Geschichtenqualität der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

Die folgende Abbildung 64 stellt die Geschichtenqualität-Mittelwerte der LUG-Bedingung dar.

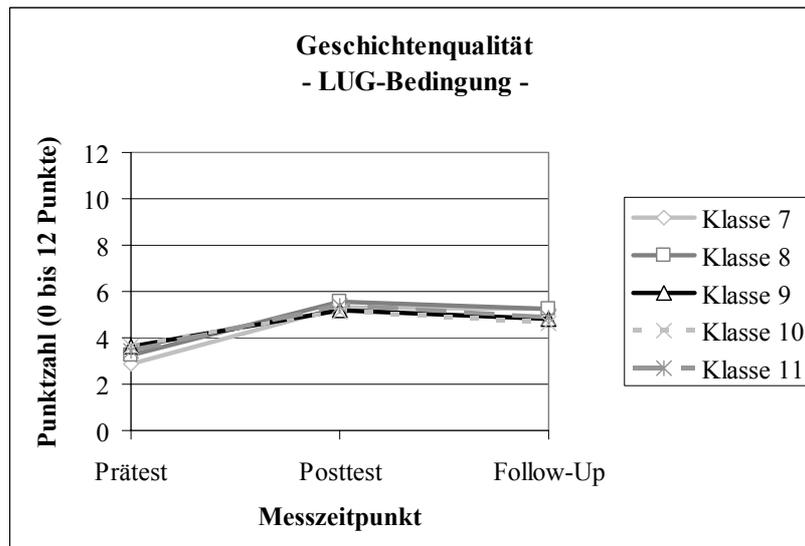


Abbildung 64: Geschichtenelemente der LUG-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

Wie auch bei den Geschichtenelementen schnitt die Klasse 7 beim Prätest deutlich schlechter ab als die übrigen Klassen. Gleichzeitig zeigte die Klasse 7 die höchsten Leistungssteigerungen, da sie beim Follow-Up am besten abschnitt.

7.3.2.2.3 Wortanzahl

Die deskriptiven Statistiken für Wortanzahl getrennt nach Messzeitpunkt, Klasse und Bedingung sind in Tabelle 61 aufgelistet.

Tabelle 61: Mittelwerte und Standardabweichungen für Wortanzahl getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).

	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	74.68	20.81	167.84	51.74	159.72	48.48
Klasse 2	65.60	20.40	152.25	40.57	144.55	38.25
Klasse 3	80.19	50.64	154.62	62.96	154.00	44.38
Klasse 4	74.33	18.00	143.48	41.86	138.81	43.51
Klasse 5	61.57	21.58	156.05	42.02	146.05	42.37
Klasse 6	83.40	43.75	166.00	69.75	181.30	63.81

LUG						
Klasse 7	63.64	19.85	121.68	40.56	121.18	34.45
Klasse 8	69.91	27.27	135.91	43.45	130.30	48.01
Klasse 9	76.81	26.23	103.10	42.40	110.38	40.97
Klasse 10	98.67	29.28	101.61	34.98	103.89	37.67
Klasse 11	74.26	30.13	135.87	52.27	127.39	36.49

In der nachfolgenden Abbildung 65 sind die Klassen-Mittelwerte für die USAT-Bedingung und in Abbildung 66 für die LUG-Bedingung dargestellt. Die Linien der USAT-Klassen verlaufen nahezu parallel. In der LUG-Bedingung dagegen stechen die Leistungen der Klasse 10 hervor. Diese Klasse wies zu allen drei Messzeitpunkten eine sehr ähnliche Anzahl von Wörtern auf und zeigte somit kaum Leistungsveränderungen.

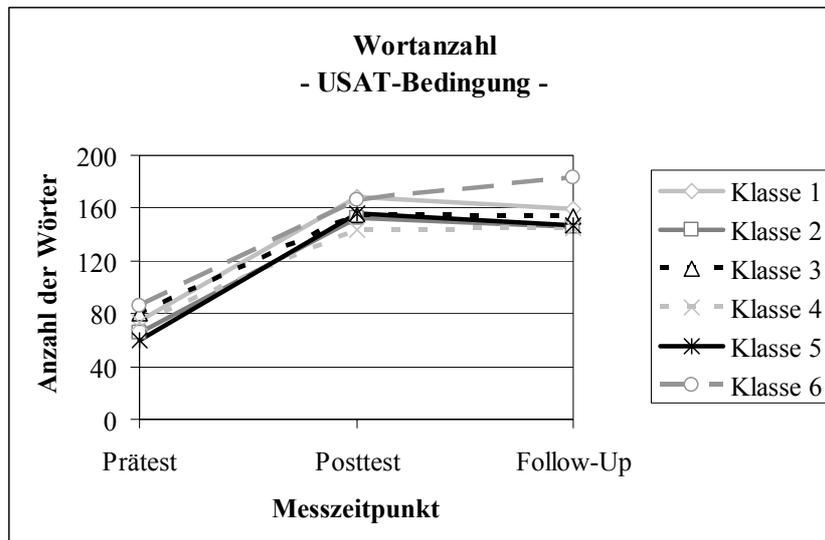


Abbildung 65: Wortanzahl der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

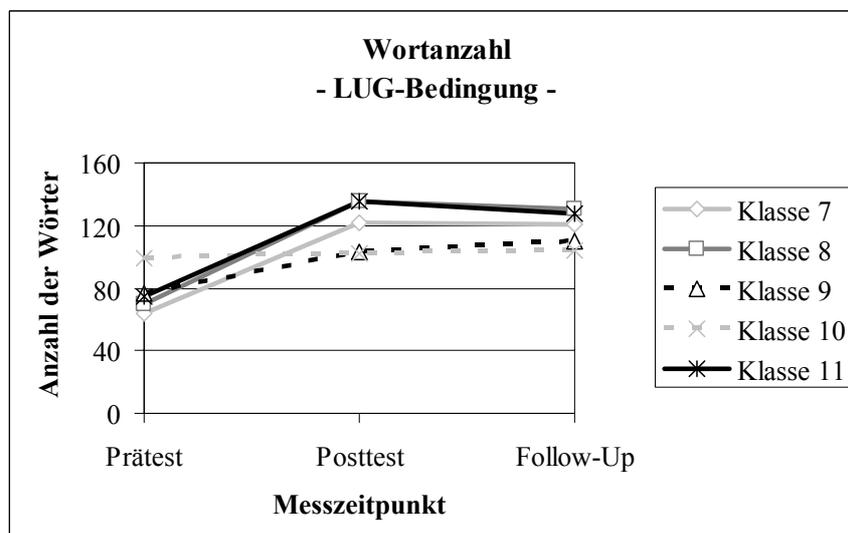


Abbildung 66: Wortanzahl der LUG-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

7.3.2.2.4 *Planen*

Die für Planen ermittelten deskriptiven Statistiken getrennt nach Messzeitpunkt, Bedingung und Klasse sind in Tabelle 62 dargestellt. Die Abbildungen 67 (USAT) und 68 (LUG) stellen die jeweiligen Mittelwerte der Bedingungen grafisch dar.

Tabelle 62: Mittelwerte und Standardabweichungen für Planen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).

	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	2.08	1.00	6.80	2.45	6.40	2.50
Klasse 2	1.45	1.05	6.55	1.82	6.45	2.78
Klasse 3	2.00	0.89	6.10	2.30	5.24	1.51
Klasse 4	1.95	1.16	6.48	2.14	5.76	2.02
Klasse 5	2.24	1.09	6.43	2.29	7.57	2.13
Klasse 6	2.48	1.45	7.07	2.62	6.41	2.12
LUG						
Klasse 7	2.18	1.18	2.91	1.38	3.05	1.09
Klasse 8	2.35	0.88	4.04	1.30	3.57	1.20
Klasse 9	2.14	0.65	3.57	1.66	2.57	1.21
Klasse 10	1.94	0.80	2.33	1.08	2.22	1.11
Klasse 11	2.70	1.26	3.43	1.73	2.70	1.33

Die Klassen 1, 5 und 6 schnitten insgesamt am besten ab. In der Abbildung 67, die die Planen-Mittelwerte der einzelnen USAT-Klassen darstellt, wird deutlich, dass die Klasse 5 sich beim Follow-Up erheblich steigerte und dass gleichzeitig die Klasse 3 sich verschlechterte.

Im Gegensatz zur Abbildung 67 verlaufen die Mittelwertslinien der LUG-Klassen in Abbildung 68 nicht steil nach oben. Während sich die Klassen 7 und 8 vom ersten bis zum dritten Messzeitpunkt deutlich verbesserten, verläuft die Linie der Klasse 10 nahezu waagrecht. Dieses Ergebnis verstärkt den bereits bei den Variablen Geschichtenelemente und Geschichtenqualität gewonnenen Eindruck über die Leistungsunterschiede zwischen den LUG-Klassen.

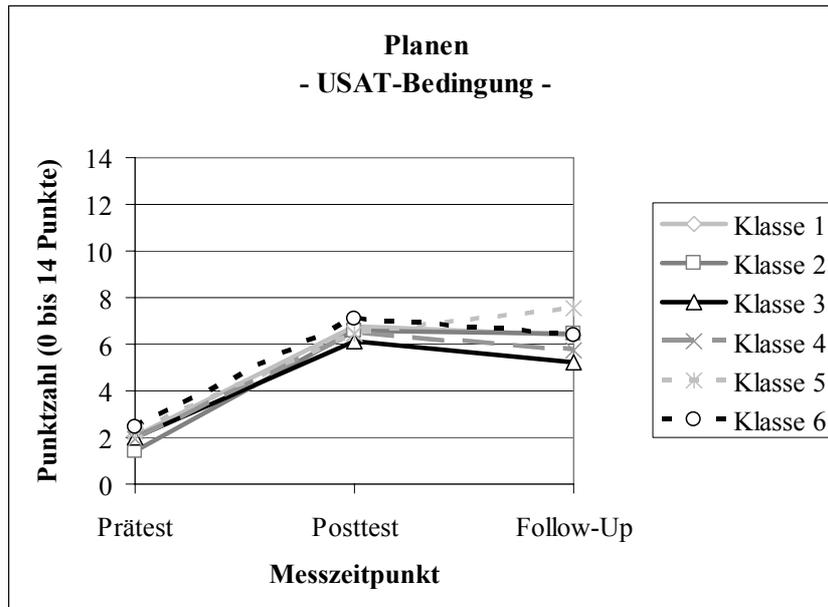


Abbildung 67: Planen der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

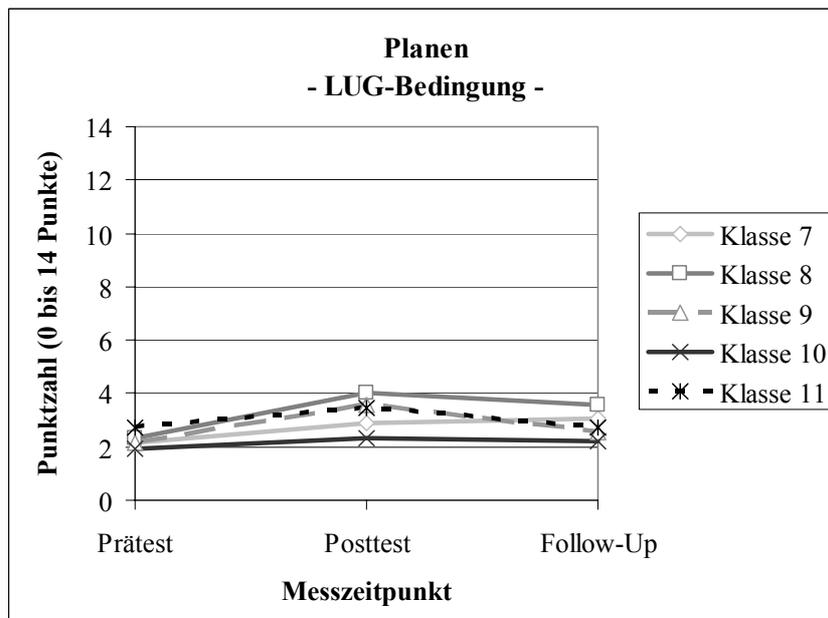


Abbildung 68: Planen der LUG-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

7.3.2.2.5 Revisionen

Die nachfolgende Tabelle 63 enthält die Mittelwerte und Standardabweichungen für Revisionen getrennt nach Messzeitpunkt, Bedingung und Klasse. Grafische Darstellungen der Klassenmittelwerte getrennt für jede Bedingung sind in Abbildung 69 (USAT) und 70 (LUG) zu finden.

Tabelle 63: Mittelwerte und Standardabweichungen für Revisionen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).

	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	0.24	0.52	1.28	1.43	3.04	2.30
Klasse 2	0.25	0.55	0.90	1.07	0.95	1.19
Klasse 3	0.24	0.62	1.29	1.85	1.33	1.53
Klasse 4	0.29	0.64	0.71	0.90	1.90	1.73
Klasse 5	0.14	0.48	1.57	1.36	1.57	1.54
Klasse 6	0.26	0.59	1.56	1.63	1.48	1.40
LUG						
Klasse 7	0.86	1.28	0.68	1.09	1.59	2.02
Klasse 8	1.13	1.89	1.48	1.56	0.57	1.08
Klasse 9	0.38	0.92	0.86	1.06	0.71	0.78
Klasse 10	0.78	1.00	0.39	0.61	0.22	0.55
Klasse 11	0.43	0.51	0.83	1.03	1.35	1.30

Die in Abbildung 69 veranschaulichten Revisionsmittelwerte der USAT-Klassen weisen abgesehen von Klasse 1 und 4 einen ähnlichen Verlauf auf. Beim Posttest schnitten die Klassen 2 und 4 wesentlich schlechter ab als die Klassen 5 und 6, welche im Durchschnitt fast eine inhaltliche Verbesserung mehr machten. Beim Follow-Up holte die Klasse 4 den Leistungsrückstand zu den anderen Klassen auf, während die Klasse 2 weiterhin schlecht abschnitt. Klasse 1 hebt sich durch einen extremen Leistungsanstieg beim Follow-Up von den anderen Klassen ab.

Für die Revisionsmittelwerte der LUG-Bedingung ergab sich ein anderes Bild. In Abbildung 70 ist kein ähnlicher Verlauf der Klassen zu erkennen. Während Klasse 10 beim zweiten und dritten Messzeitpunkt schlechter abschnitt als beim Prätest, steigerte sich die Klasse 11 kontinuierlich. Klasse 8 erzielte die höchsten Leistungen beim Posttest, verschlechterte sich jedoch stark beim Follow-Up. Bei Klasse 7 konnte dagegen ein großer Leistungszugewinn beim Follow-Up festgestellt werden.

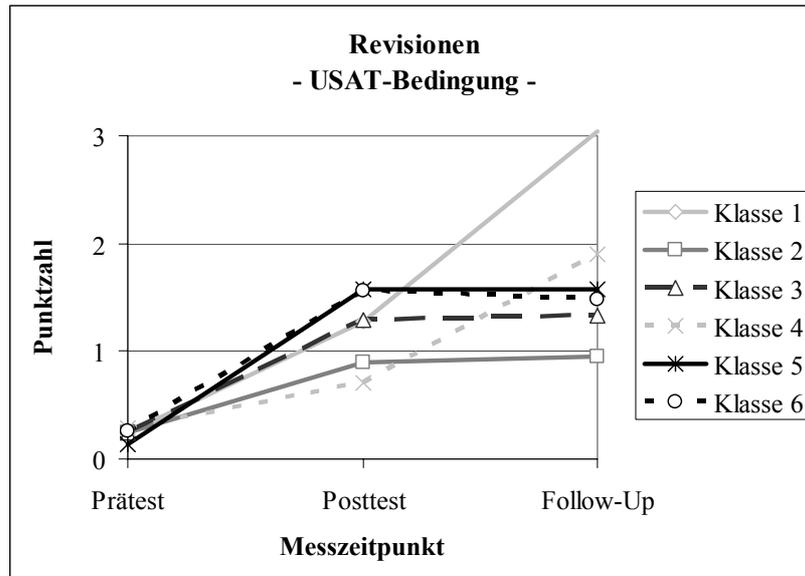


Abbildung 69: Revisionen der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

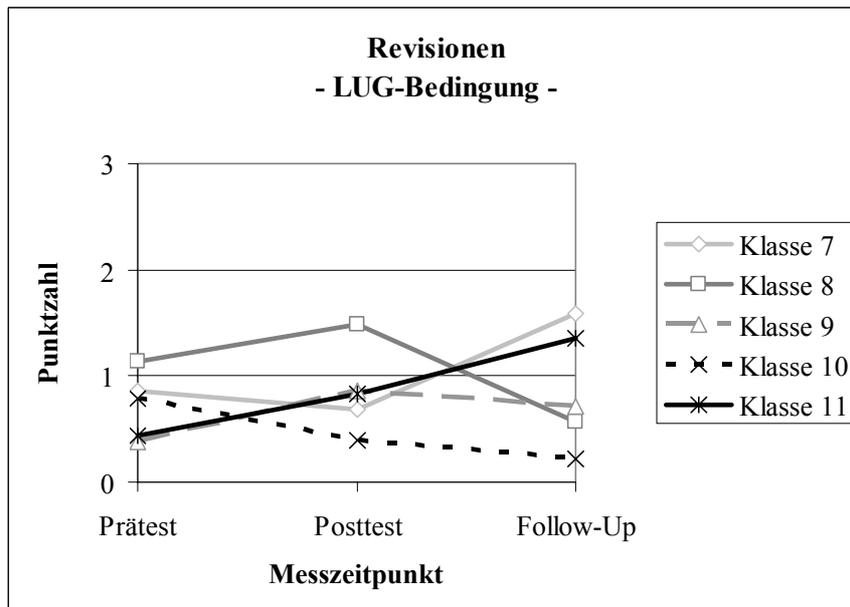


Abbildung 70: Revisionen der LUG-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

7.3.2.2.6 Schreibbezogenes Wissen

Die deskriptiven Statistiken für die einzelnen Klassen, getrennt nach Messzeitpunkt und Bedingung, sind in Tabelle 64 aufgeführt.

Tabelle 64: Mittelwerte und Standardabweichungen für Wissen getrennt nach Messzeitpunkt, Bedingung und Klasse (Studie II).

	Prätest		Posttest		Follow-Up	
	M	SD	M	SD	M	SD
USAT						
Klasse 1	1.00	.817	5.92	2.18	5.96	1.90
Klasse 2	1.85	0.99	5.85	2.18	6.45	1.90
Klasse 3	1.81	0.98	4.90	2.17	4.19	2.46
Klasse 4	2.05	1.47	5.43	2.82	6.33	2.27
Klasse 5	1.43	0.87	6.81	1.33	7.14	1.15
Klasse 6	1.11	0.97	6.07	2.48	5.19	2.54
LUG						
Klasse 7	0.95	0.84	4.32	2.30	4.82	2.06
Klasse 8	2.04	1.33	6.39	2.41	6.74	1.89
Klasse 9	0.67	0.73	5.10	1.92	4.86	2.33
Klasse 10	0.89	0.76	5.11	2.00	3.78	2.13
Klasse 11	1.78	1.09	5.48	2.04	4.57	2.25

Bei den in Abbildung 71 dargestellten Wissens-Mittelwerten für die USAT-Bedingung liegen die Leistungen der Klasse 5 am höchsten und die der Klasse 3 am niedrigsten. Insbesondere beim dritten Messzeitpunkte weichen die Mittelwerte der USAT-Klassen ziemlich weit voneinander ab. Dennoch zeigten alle Klassen vom Prätest zum Posttest und vom Prätest zum Follow-Up einen Wissensanstieg, wenn auch in unterschiedlicher Höhe.

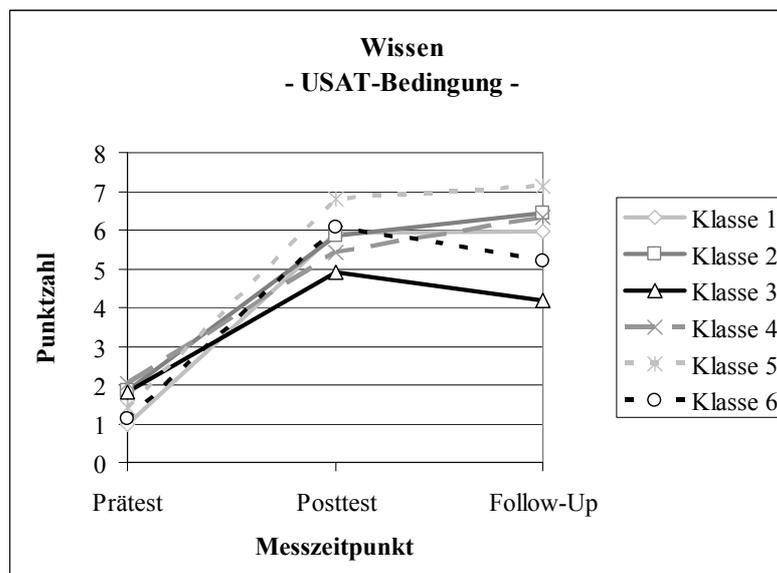


Abbildung 71: Wissen der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

In Abbildung 72 zeigen alle Mittelwertslinien einen erheblichen Leistungsanstieg beim Posttest. Beim Follow-Up weichen jedoch die Linien sehr voneinander ab. Insbesondere die Klasse 8 zeigte zu allen drei Messzeitpunkten deutlich höhere Leistungen als die anderen Klassen. Während die Wissensergebnisse in den Klassen 9, 10 und 11 beim Follow-Up abfielen, steigerten sich die Klassen 7 und 8 weiter. Insgesamt liegen die Leistungen aller LUG-Klassen beim Posttest und Follow-Up deutlich über dem Ausgangsniveau. Insgesamt verlaufen die Mittelwertslinien der USAT- und der LUG-Bedingung ähnlich steil.

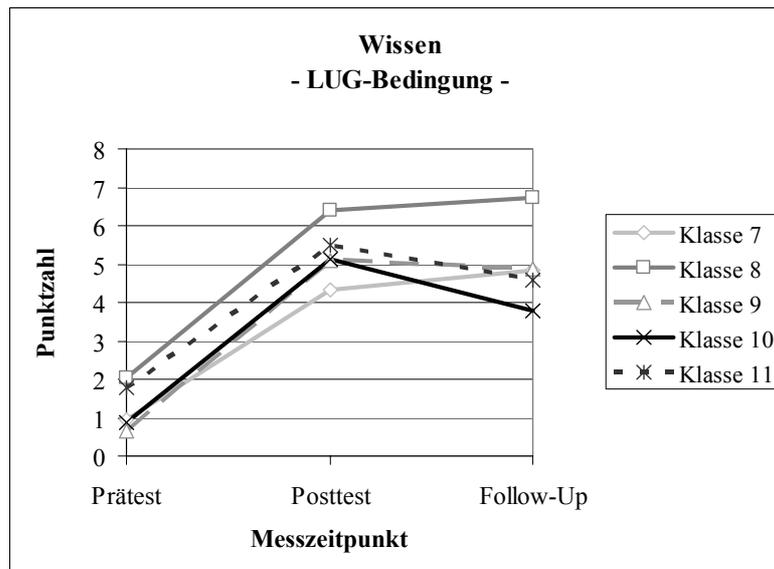


Abbildung 72: Wissen der USAT-Bedingung als Funktion von Messzeitpunkt und Klasse (Studie II).

7.3.2.3 Mixed Model-Analysen

Wie auch in der Studie I wurde aufgrund der Zuteilung von Klassen zu den Bedingungen und der Durchführung des Trainings auf Klassenebene die Klasse als Analyseeinheit gewählt. Auf diese Weise wurden Klassenunterschiede berücksichtigt. Gleichzeitig reduzierten sich die Stichprobenanzahl und die Standardabweichungen drastisch. Um den Effekt unterschiedlicher Klassen innerhalb einer Bedingung zu berechnen, wurde überprüft, ob der verschachtelte Faktor „Klasse in Bedingung“ bzw. „Schüler in Klasse“ signifikant ist. Deswegen wurden Mixed Model-ANCOVA's mit Schüler (als zufälliger Faktor) genestet in der Klasse und mit Klasse (als zufälliger Faktor) genestet in der Bedingung (als fester Faktor) nach Hopkins (1982) gerechnet. Der jeweilige Prätest-Wert sowie die Rechtschreibleistung wurden als Kovariaten mit einbezogen. In Tabelle 65 sind die *Wald Z*-Werte für die einzelnen abhängigen Variablen abzulesen. Ein nicht signifikanter *Wald Z*-Wert bedeutet, dass der Einfluss unterschiedlicher Klassen innerhalb einer Bedingung nicht bedeutsam war.

Tabelle 65: Statistische Kennwerte (Wald Z) für den Einfluss Klasse innerhalb Bedingung bzw. Schüler innerhalb Klasse sowie F-Werte für den Haupteffekt Bedingung (Studie II).

	Wald Z				Bedingungseffekt		
	Klasse (Bedingung)		Schüler (Klasse)		F-Wert	FG ¹⁵ Nenner	Sig.
	Wald Z	Sig.	Wald Z	Sig.			
Geschichtenelemente	0.92	0.36	5.52	<.001	141.14	8.65	<.001
Narrative Qualität	1.05	0.29	5.59	<.001	5.25	8.57	<.05
Wortanzahl	1.31	0.19	7.18	<.001	20.77	8.56	<.01
Planen Inhalt	1.14	0.26	4.42	<.001	126.30	8.83	<.001
Revisionen	1.14	0.25	2.91	<.01	7.49	9.31	<.05
Wissen	1.53	0.13	6.70	<.001	2.10	8.92	<.05

Wie auch in Studie I waren sämtliche Kennwerte für Klasse innerhalb einer Bedingung nicht signifikant und für Schüler innerhalb einer Klasse signifikant. Folglich variierten die Klassen innerhalb einer Bedingung nicht bedeutsam, während sich die Schülerleistungen stark innerhalb einer Klasse unterschieden. Eine Analyse auf Schülerebene wäre demnach durchführbar. Den Empfehlungen von Kromrey und Dickinson (1996) folgend, wurde jedoch aufgrund des vorliegenden Versuchsdesigns im Folgenden auf Klassenebene gerechnet. Nichtsdestotrotz führten sämtliche Mixed Model-ANCOVA's zu signifikanten Bedingungseffekten zu Gunsten der USAT-Bedingung (siehe Tabelle 65).

7.3.3 Schreibleistung

Analog zur Studie I wurde die Schreibleistung anhand eines strategienahen und eines strategiefernen Schreibmaßes erfasst. Zusätzlich wurden die Wörter gezählt. Das strategienahe Maß der Geschichtenelemente wird als Erstes beschrieben. Anschließend folgen die Ergebnisse für das strategieferne Maß Geschichtenqualität. Zum Schluss werden die Ergebnisse für die Wortanzahl präsentiert.

7.3.3.1 Geschichtenelemente

Die Geschichtenelemente wurden anhand einer 14-stufigen Skala eingeschätzt. Tabelle 66 enthält die Mittelwerte, Standardabweichungen, die um die Kovariaten korrigierten Mittelwerte, F-Werte und Effektstärken für beide Bedingungen zu den drei Messzeitpunkten.

¹⁵ FG = Freiheitsgrade; der Zähler-Freiheitsgrad war bei jeder abhängigen Variablen 1.

Tabelle 66: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Geschichtenelemente getrennt nach Bedingung und Messzeitpunkt (Studie II).

	Prätest			Posttest			Follow-Up		
	M	SD	M ^{adR}	M	SD	M ^{ad}	M	SD	M ^{ad}
USAT-Gruppe	2.15	0.30	2.02	8.10	0.71	8.32	7.28	0.59	7.11
LUG-Gruppe	2.52	0.49	2.68	4.13	0.47	3.86	4.23	0.36	4.43
F-Wert	4.85			47.75***			32.11**		
Effektstärke				6.63			6.09		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=6), LUG-Gruppe = Lehrerkonzipierte Unterrichtseinheit zum Geschichtens Schreiben (n=5); M^{adR} = um Rechtschreibung adjustierte Mittelwerte; M^{ad} = um Prätest-Werte und Rechtschreibung korrigierte Mittelwerte; *** p< .001, signifikanter Unterschied zwischen den Bedingungen; ** p< .01, signifikanter Unterschied zwischen den Bedingungen; * p< .05, signifikanter Unterschied zwischen den Bedingungen.

Die Messwiederholungsanalyse mit Geschichtenelemente als abhängige Variable, Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, LUG) als Zwischensubjektfaktor und Prätest sowie Rechtschreibung als Kovariaten führte zu einem signifikanten Bedingungseffekt, $F(1, 7) = 51.59$, $p < .001$, und zu nicht signifikanten Ergebnissen für die Kovariaten (Prätest: $F(1, 7) = .41$, $p = .54$, Rechtschreibung: $F(1, 7) = .34$, $p = .58$). Der Messzeitpunkt erwies sich als nicht signifikant, $F(1, 7) = 3.42$, $p = .11$. Darüber hinaus ergab sich eine signifikante Interaktion Messzeitpunkt x Bedingung, $F(1, 7) = 10.82$, $p < .05$. Dies ist dadurch zu erklären, dass beim Follow-Up die Leistungen der USAT-Klassen absanken und die der LUG-Klassen anstiegen, so dass sich der Unterschied zwischen den Bedingungen vom Posttest zum Follow-Up deutlich reduzierte. Dies ist ebenfalls in der Abbildung 73 erkennbar, in der die Geschichtenelemente getrennt für Bedingung und Messzeitpunkt dargestellt sind.

Die Kovarianzanalysen mit Geschichtenelemente (jeweils Posttest und Follow-Up) als abhängige Variable, Bedingung (USAT, LUG) als unabhängige Variable, Prätest und Rechtschreibung als Kovariaten bestätigten den signifikanten Bedingungshaupteffekt zu Gunsten der USAT-Bedingung beim Posttest und beim Follow-Up (siehe Tabelle 66). Zu keinem der beiden Zeitpunkte lieferten die Kovariaten Prätest (Posttest: $F(1, 7) = .09$, $p = .77$, Follow-Up: $F(1, 7) = 3.07$, $p = .12$) und Rechtschreibung (Prätest: $F(1, 7) = 2.18$, $p = .18$, Posttest: $F(1, 7) = 1.25$, $p = .30$, Follow-Up: $F(1, 7) = .09$, $p = .77$) signifikante Ergebnisse. Zudem zeigte sich die Kovariate Rechtschreibung auch bei der Kovarianzanalyse mit Prätest als abhängige Variable als nicht signifikant, $F(1, 7) = 2.18$, $p = .18$. Die Bedingungsunterschiede beim Prätest waren ebenfalls nicht signifikant, $F(1, 7) = 4.85$, $p = .06$.

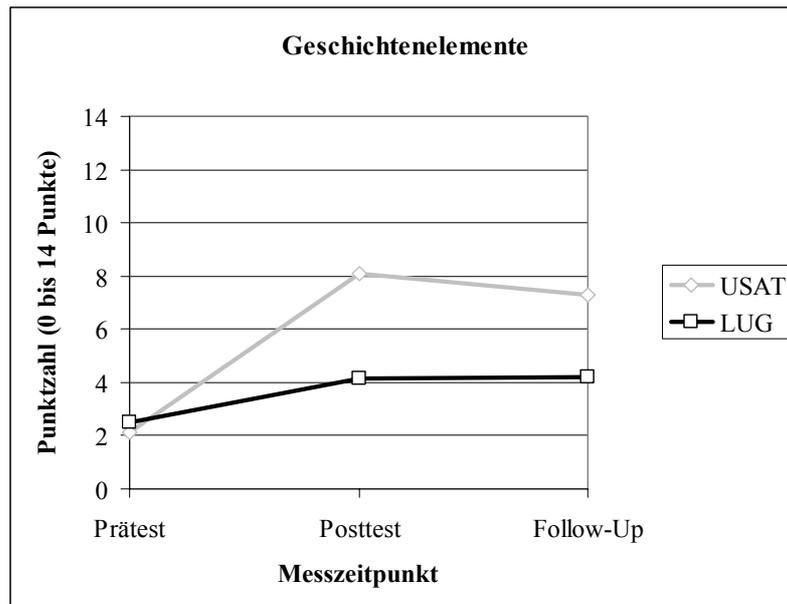


Abbildung 73: Geschichtenelemente in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II).

7.3.3.2 Geschichtenqualität

Das strategiefremde Maß zur Beurteilung der Schreibleistung, die Geschichtenqualität, wurde auf einer zwölfstufigen Skala eingeschätzt. Wie auch in der Studie I korrelierten die Geschichtenelemente und Geschichtenqualität zu allen drei Messzeitpunkten signifikant miteinander (Prätest: $r = .72, p < .05$; Posttest: $r = .65, p < .05$; Follow-Up: $r = .70, p < .01$). Im Vergleich zur Studie I fielen die Korrelationen jedoch deutlich geringer aus (siehe Kapitel 6.3.3.2). In Tabelle 67 sind die Mittelwerte, Standardabweichungen, die um die Kovariaten korrigierten Mittelwerte, F-Werte und Effektstärken aufgelistet.

Tabelle 67: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Geschichtenqualität getrennt nach Bedingung und Messzeitpunkt (Studie II).

	Prätest			Posttest			Follow-Up		
	M	SD	M ^{adR}	M	SD	M ^{ad}	M	SD	M ^{ad}
USAT-Gruppe	3.26	0.20	3.25	5.56	0.28	5.65	5.32	0.36	5.35
LUG-Gruppe	3.36	0.31	3.37	5.32	0.18	5.21	4.95	0.27	4.91
F-Wert	0.31			5.32			2.41		
Effektstärke				1.00			1.15		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=6), LUG-Gruppe = Lehrerkonzipierte Unterrichtseinheit zum Geschichtens Schreiben (n=5); M^{adR} = um Rechtschreibung adjustierte Mittelwerte; M^{ad} = um Prätest-Werte und Rechtschreibung korrigierte Mittelwerte; *** $p < .001$, signifikanter Unterschied zwischen den Bedingungen; ** $p < .01$, signifikanter Unterschied zwischen den Bedingungen; * $p < .05$, signifikanter Unterschied zwischen den Bedingungen.

Die in Abbildung 74 dargestellten Mittelwertlinien für die beiden Bedingungen und die drei Messzeitpunkte verlaufen sehr nah, so dass auf dem ersten Blick kein großer Unterschied zwischen den Bedingungen erkennbar ist.

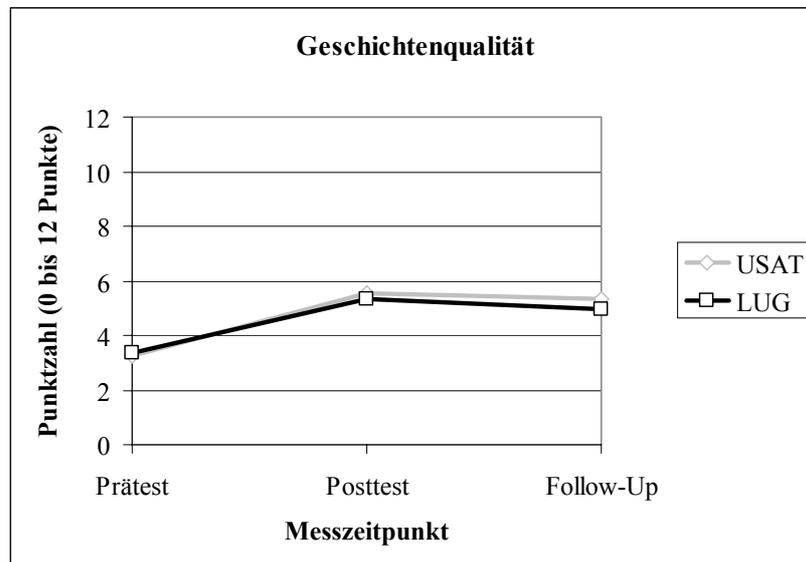


Abbildung 74: Geschichtenqualität in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II).

Dies wurde durch die Messwiederholungsanalyse mit Geschichtenqualität als abhängige Variable, Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, LUG) als Zwischensubjektfaktor und Prätest sowie Rechtschreibung als Kovariaten bestätigt. Der Bedingungseffekt war knapp nicht signifikant, $F(1, 7) = 4.35, p = .08$. Überdies zeigte sich keine signifikante Messzeitpunkt x Bedingung- Interaktion, $F(1, 7) = .00, p = 1.00$, und kein signifikanter Messzeitpunkteffekt, $F(1, 7) = .60, p = .46$. Auch die Kovariaten erwiesen sich als nicht signifikant (Prätest: $F(1, 7) = .16, p = .70$, Rechtschreibung: $F(1, 7) = 1.15, p = .32$).

Ähnliche Ergebnisse lieferten die sowohl für Posttest als auch für Follow-Up durchgeführten Kovarianzanalysen mit Geschichtenqualität als abhängige Variable, Bedingung (USAT, LUG) als unabhängige Variable und den beiden Variablen Prätest und Rechtschreibung als Kovariaten. Zu beiden Messzeitpunkten ergaben sich keine signifikanten Bedingungshaupteffekte (Posttest: $F(1, 7) = 5.32, p = .06$; Follow-Up: $F(1, 7) = 2.41, p = .17$). Auch die Kovariaten Prätest und Rechtschreibung übten weder beim Posttest (Prätest: $F(1, 7) = .01, p = .94$; Rechtschreibung: $F(1, 7) = 2.57, p = .15$) noch beim Follow-Up (Prätest: $F(1, 7) = .42, p = .54$; Rechtschreibung: $F(1, 7) = .27, p = .62$) einen signifikanten Einfluss aus. Die nicht bedeutsamen Prätest-Unterschiede zwischen den Bedingungen, $F(1, 7) = .31, p = .59$, wurden ebenfalls nicht signifikant von der Kovariaten Rechtschreibung, $F(1, 7) = .03, p = .86$, beeinflusst.

7.3.3.3 Wortanzahl

Als nächstes wurde die Anzahl der Wörter untersucht. Die Mittelwerte, Standardabweichungen, die um die Kovariaten korrigierten Mittelwerte, F-Werte und Effektstärken sind getrennt nach Bedingung und Messzeitpunkt in Tabelle 68 illustriert.

Tabelle 68: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Wortzahl getrennt nach Bedingung und Messzeitpunkt (Studie II).

	Prätest			Posttest			Follow-Up		
	M	SD	M ^{ad}	M	SD	M ^{ad}	M	SD	M ^{ad}
USAT-Gruppe	73.30	8.36	69.43	156.71	9.06	156.06	154.07	15.24	154.74
LUG-Gruppe	76.66	13.28	81.30	119.63	16.82	120.41	118.63	11.24	117.83
F-Wert	2.24			8.08*			7.04*		
Effektstärke				2.83			2.60		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=6), LUG-Gruppe = Lehrerkonzipierte Unterrichtseinheit zum Geschichtschreiben (n=5); M^{adR} = um Rechtschreibung adjustierte Mittelwerte; M^{ad} = um Prätest-Werte und Rechtschreibung korrigierte Mittelwerte; *** p < .001, signifikanter Unterschied zwischen den Bedingungen; ** p < .01, signifikanter Unterschied zwischen den Bedingungen; * p < .05, signifikanter Unterschied zwischen den Bedingungen.

Eine Messwiederholungsanalyse mit Wortanzahl als abhängige Variable, Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, LUG) als unabhängige Variable sowie Prätest und Rechtschreibung als Kovariaten lieferten einen signifikanten Bedingungseffekt, $F(1, 7) = 19.45, p < .01$. Alle anderen Variablen erwiesen sich als nicht signifikant (Messzeitpunkt: $F(1, 7) = .91, p = .37$; Messzeitpunkt x Bedingung: $F(1, 7) = .01, p = .92$; Prätest: $F(1, 7) = .14, p = .72$; Rechtschreibung: $F(1, 7) = .96, p = .36$).

Die als nächstes durchgeführten Kovarianzanalysen für Posttest und Follow-Up mit Wortanzahl als abhängige Variable, Bedingung (USAT, LUG) als unabhängige Variable sowie Prätest und Rechtschreibung als Kovariaten bestätigten den signifikanten Bedingungseffekt beim Posttest und beim Follow-Up (siehe Tabelle 68), wobei die Unterschiede beim Prätest nicht bedeutsam variierten, $F(1, 7) = 2.24, p = .17$. Die Kovariaten Rechtschreibung (Prätest: $F(1, 7) = 2.71, p = .14$; Posttest: $F(1, 7) = .00, p = .99$; Follow-Up: $F(1, 7) = .02, p = .88$) und Prätest (Posttest: $F(1, 7) = .75, p = .42$; Follow-Up: $F(1, 7) = .00, p = .97$) erwiesen sich als nicht signifikant.

Die nachfolgende Abbildung 75 ist eine grafische Darstellung der Unterschiede zwischen der USAT- und LUG-Bedingung zu den drei Messzeitpunkten.

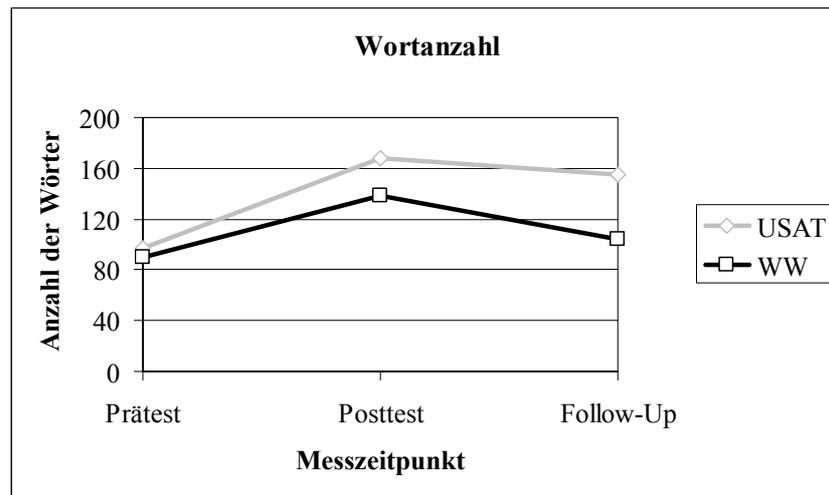


Abbildung 75: Wortanzahl in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II).

7.3.4 Strategienahe Schreibmaße

Als nächstes wurden die Ergebnisse für die schreibbezogenen Maße Planungsaktivitäten, Revisionen und Wissen erläutert. Diese stellen genauso wie das Schreibleistungsmaß Geschichtenelemente ein strategienahes Maß dar.

7.3.4.1 Planen

Die Planungsaktivitäten der Schüler wurden hinsichtlich der 7W-Fragen ausgewertet. Es wurde einerseits das Nennen der 7W-Fragen und andererseits die sprachliche Ausgestaltung der 7W-Fragen in den Planungsnotizen auf jeweils einer siebenstufigen Skala bewertet, so dass sich eine Maximalpunktzahl von 14 ergab. Die Mittelwerte, Standardabweichungen, die um die Kovariaten korrigierten Mittelwerte, F-Werte und Effektstärken getrennt nach Bedingung und Messzeitpunkt sind in Tabelle 69 aufgelistet.

Tabelle 69: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Planen getrennt nach Bedingung und Messzeitpunkt (Studie II).

	Prätest			Posttest			Follow-Up		
	M	SD	M ^{adR}	M	SD	M ^{ad}	M	SD	M ^{ad}
USAT-Gruppe	2.03	0.34	2.03	6.57	0.34	6.82	6.30	0.78	6.46
LUG-Gruppe	2.26	0.28	2.27	3.26	0.66	2.96	2.82	0.51	2.63
F-Wert	0.76			117.00***			40.18***		
Effektstärke				6.52			5.17		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=6), LUG-Gruppe = Lehrerkonzipierte Unterrichtseinheit zum Geschichtens Schreiben (n=5); M^{adR} = um Rechtschreibung adjustierte Mittelwerte; M^{ad} = um Prätest-Werte und Rechtschreibung korrigierte Mittelwerte; *** p < .001, signifikanter Unterschied zwischen den Bedingungen; ** p < .01, signifikanter Unterschied zwischen den Bedingungen; * p < .05, signifikanter Unterschied zwischen den Bedingungen.

Es wurde eine Messwiederholungsanalyse mit Planen als abhängige Variable, Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, LUG) als Zwischensubjektfaktor und den beiden Variablen Prätest und Rechtschreibung als Kovariaten durchgeführt, die einen hoch signifikanten Bedingungseffekt, $F(1, 7) = 92.55, p < .001$, und keine Effekte der beiden Kovariaten Prätest, $F(1, 7) = 1.73, p = .24$, und Rechtschreibung, $F(1, 7) = 1.40, p = .28$, lieferte. Zudem ergab sich keine signifikante Messzeitpunkt x Bedingung- Interaktion, $F(1, 7) = .00, p = .96$, und keinen signifikanten Effekt des Innersubjektfaktors Messzeitpunkt, $F(1, 7) = .07, p = .81$. Die Unterschiede zwischen den beiden Bedingungen sind in Abbildung 65 grafisch dargestellt.

Kovarianzanalysen für jeweils Posttest und Follow-Up getrennt mit Planen als abhängige Variable, Bedingung (USAT, LUG) als unabhängige Variable und Prätest sowie Rechtschreibung als Kovariaten bestätigten die Ergebnisse der Messwiederholungsanalyse. Es ergaben sich signifikante Bedingungshaupteffekte zu Gunsten der USAT-Bedingung beim Posttest und beim Follow-Up (siehe Tabelle 69). Dagegen übten die Kovariaten weder beim Posttest (Prätest: $F(1, 7) = 2.79, p = .14$; Rechtschreibung: $F(1, 7) = 2.83, p = .14$) noch beim Follow-Up (Prätest: $F(1, 7) = .57, p = .48$; Rechtschreibung: $F(1, 7) = .32, p = .59$) einen signifikanten Einfluss aus. Bei den nicht signifikanten Prätest-Unterschieden zwischen den Bedingungen, $F(1, 7) = .76, p = .41$, zeigte sich die Kovariate Rechtschreibung ebenfalls als nicht signifikant, $F(1, 7) = .00, p = .97$.

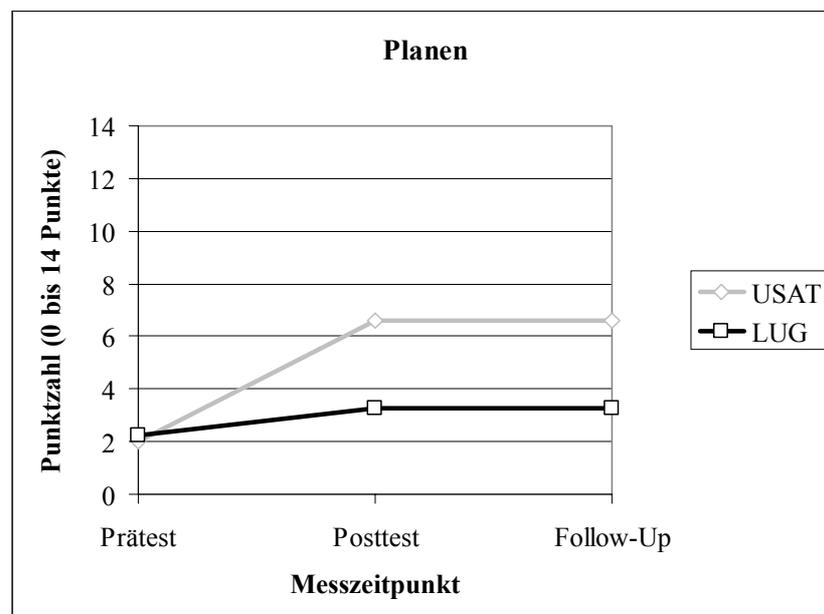


Abbildung 76: Planen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II).

7.3.4.2 Revisionen

Als weiteres strategienahes Maß wurden die Verbesserungen und Einfügungen (Nennung und Ausgestaltung der 7W-Fragen) der Schüler ausgewertet. In Tabelle 70 sind die Mittelwerte, Standardabweichungen, die um die Kovariaten korrigierten Mittelwerte, F-Werte und Effektstärken getrennt nach Bedingung und Messzeitpunkt aufgelistet.

Tabelle 70: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Revisionen getrennt nach Bedingung und Messzeitpunkt (Studie II).

	Prätest			Posttest			Follow-Up		
	M	SD	M ^{adR}	M	SD	M ^{ad}	M	SD	M ^{ad}
USAT-Gruppe	0.24	0.05	0.25	1.22	0.35	1.38	1.71	0.72	1.86
LUG-Gruppe	0.72	0.31	0.70	0.85	0.40	0.66	0.89	0.57	0.71
F-Wert	6.51*			2.97			2.91		
Effektstärke				0.99			1.25		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=6), LUG-Gruppe = Lehrerkonzipierte Unterrichtseinheit zum Geschichtschreiben (n=5); M^{adR} = um Rechtschreibung adjustierte Mittelwerte; M^{ad} = um Prätest-Werte und Rechtschreibung korrigierte Mittelwerte; *** p < .001, signifikanter Unterschied zwischen den Bedingungen; ** p < .01, signifikanter Unterschied zwischen den Bedingungen; * p < .05, signifikanter Unterschied zwischen den Bedingungen.

Eine Messwiederholungsanalyse mit Revisionen als abhängige Variable, Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, LUG) als Zwischensubjektfaktor und die beiden Variablen Prätest und Rechtschreibung als Kovariaten führte zu einem signifikanten Bedingungshaupteffekt, $F(1, 7) = 6.46$, $p < .05$. Die Kovariaten Prätest, $F(1, 7) = .01$, $p = .92$, und Rechtschreibung, $F(1, 7) = 4.16$, $p = .08$, waren nicht signifikant. Auch der Effekt des Innersubjektfaktors Messzeitpunkt, $F(1, 7) = .10$, $p = .76$, und die Interaktion Bedingung x Messzeitpunkt, $F(1, 7) = .26$, $p = .63$, erwiesen sich als nicht signifikant.

Im Gegensatz dazu lieferten die Kovarianzanalysen für Posttest sowie Follow-Up mit Revisionen als abhängige Variable, Bedingung (USAT, LUG) als unabhängige Variable und den Variablen Prätest und Rechtschreibung als Kovariaten weder beim Posttest, $F(1, 7) = 2.97$, $p = .13$, noch beim Follow-Up, $F(1, 7) = 2.91$, $p = .13$, einen signifikanten Bedingungshaupteffekt. Die beiden Kovariaten Prätest (Posttest: $F(1, 7) = .32$, $p = .59$, Follow-Up: $F(1, 7) = .22$, $p = .66$) und Rechtschreibung (Prätest: $F(1, 7) = .05$, $p = .82$, Posttest: $F(1, 7) = .80$, $p = .40$, Follow-Up: $F(1, 7) = 2.79$, $p = .14$) übten ebenfalls keinen signifikanten Einfluss aus. In Tabelle 70 ist erkennbar, dass sich die Prätest-Werte der beiden Bedingungen

signifikant voneinander unterschieden, wobei die Kovariate Rechtschreibung einen unbedeutenden Einfluss ausübte, $F(1, 7) = .05, p = .82$.

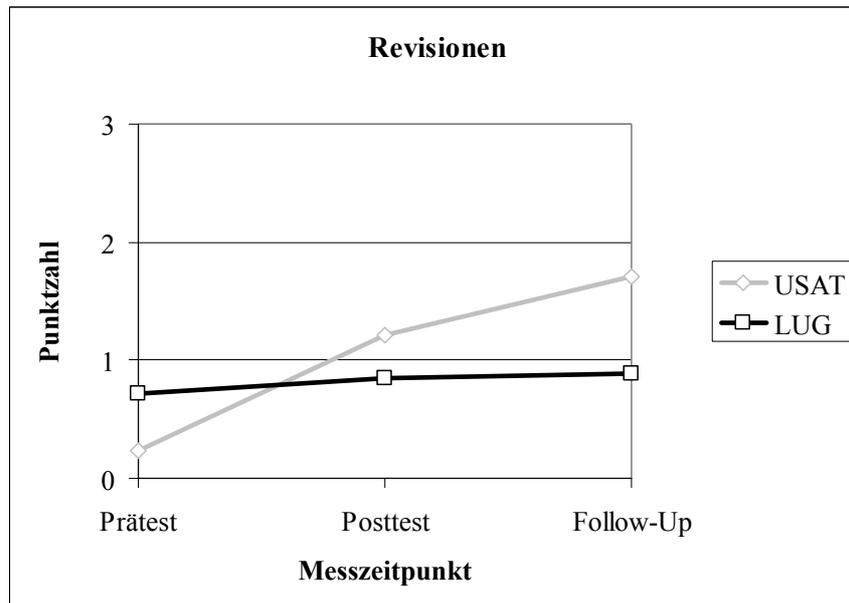


Abbildung 77: Revisionen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II).

In Abbildung 77 sind die Mittelwerte für Revisionen getrennt nach Messzeitpunkt und Bedingung illustriert. Während die Leistungen in den USAT-Klassen steil anstiegen, verläuft die LUG-Linie nahezu waagrecht.

7.3.4.3 Schreibbezogenes Wissen

Die Mittelwerte, Standardabweichungen, die um die Kovariaten korrigierten Mittelwerte, F-Werte und Effektstärken für Wissen sind der Tabelle 71 zu entnehmen.

Tabelle 71: Mittelwerte, Standardabweichungen, um die Kovariaten (Prätest, Rechtschreibung) korrigierte Mittelwerte, F-Werte und Effektstärken für Wissen getrennt nach Bedingung und Messzeitpunkt (Studie II).

	Prätest			Posttest			Follow-Up		
	M	SD	M ^{adR}	M	SD	M ^{ad}	M	SD	M ^{ad}
USAT-Gruppe	1.54	0.43	1.51	5.83	0.64	5.80	5.88	1.05	6.10
LUG-Gruppe	1.27	0.61	1.31	5.28	0.75	5.32	4.95	1.09	4.68
F-Wert	0.21			0.62			3.10		
Effektstärke				0.80			0.87		

Anmerkungen: USAT-Gruppe = Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining (n=6), LUG-Gruppe = Lehrerkonzipierte Unterrichtseinheit zum Geschichtschreiben (n=5); M^{adR} = um Rechtschreibung adjustierte Mittelwerte; M^{ad} = um Prätest-Werte und Rechtschreibung korrigierte Mittelwerte; *** p < .001, signifikanter Unterschied zwischen den Bedingungen; ** p < .01, signifikanter Unterschied zwischen den Bedingungen; * p < .05, signifikanter Unterschied zwischen den Bedingungen.

Es wurde eine Messwiederholungsanalyse mit Wissen als abhängige Variable, Messzeitpunkt als Innersubjektfaktor (Posttest, Follow-Up), Bedingung (USAT, LUG) als Zwischensubjektfaktor und Prätest sowie Rechtschreibung als Kovariaten gerechnet. Die Ergebnisse zeigten keinen signifikanten Bedingungseffekt, $F(1, 7) = 2.01$, $p = .20$, und keine signifikanten Effekte der Kovariaten Prätest, $F(1, 7) = 1.12$, $p = .32$, und Rechtschreibung, $F(1, 7) = .73$, $p = .42$. Die Interaktion Bedingung x Messzeitpunkt erwies sich ebenfalls als nicht signifikant, $F(1, 7) = 3.33$, $p = .11$, während der Messzeitpunkt selbst einen signifikanten Einfluss ausübte, $F(1, 7) = 6.15$, $p < .05$.

Sowohl für Posttest als auch für Follow-Up wurden jeweils Kovarianzanalysen mit Wissen als abhängige Variable, Bedingung (USAT, LUG) als unabhängige Variable und den beiden Variablen Prätest und Rechtschreibung als Kovariaten gerechnet. Wie auch bei der Messwiederholungsanalyse ergab sich weder beim Posttest, $F(1, 7) = .62$, $p = .46$, noch beim Follow-Up, $F(1, 7) = 3.10$, $p = .12$, ein signifikanter Bedingungshaupteffekt. Gleichzeitig übten die Kovariaten Prätest (Posttest: $F(1, 7) = .39$, $p = .55$, Follow-Up: $F(1, 7) = 1.66$, $p = .24$) und Rechtschreibung (Prätest: $F(1, 7) = .07$, $p = .80$, Posttest: $F(1, 7) = .00$, $p = .97$, Follow-Up: $F(1, 7) = 1.92$, $p = .21$) zu keinem Zeitpunkt einen signifikanten Einfluss aus. Vor dem Training unterschieden sich die USAT- und die LUG-Bedingung in Bezug zum schreibbezogenen Wissen nicht signifikant voneinander, $F(1, 7) = .21$, $p = .66$.

Abbildung 78 veranschaulicht die Wissens-Mittelwerte für die USAT- und die LUG-Bedingung getrennt für die drei Messzeitpunkte. Die Linien verlaufen nahezu parallel, wobei die Klassen der USAT-Bedingung höhere Leistungen erzielten als die LUG-Klassen.

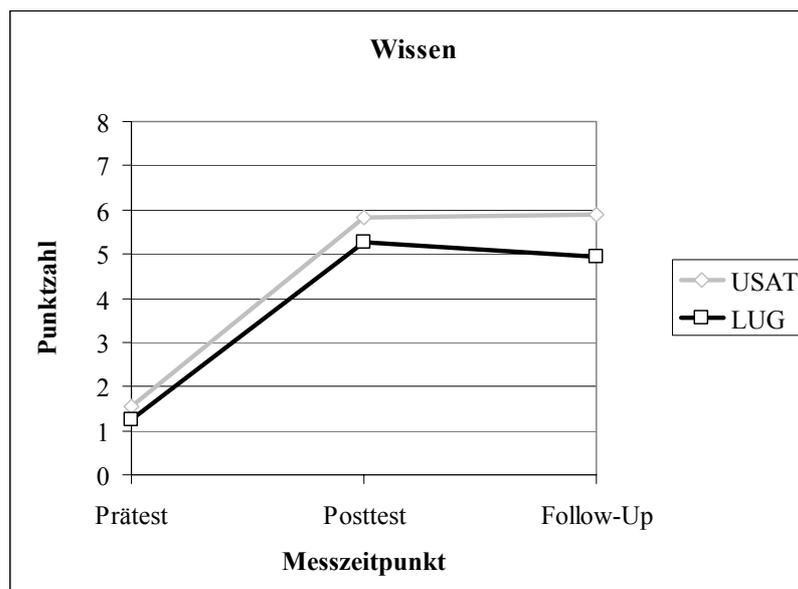


Abbildung 78: Wissen in Abhängigkeit von Messzeitpunkt und Bedingung (Studie II).

7.3.5 Maße der praktischen Signifikanz

Die auf Klassenebene berechneten und um Prätestunterschiede korrigierten Effektstärken nach Cohen's d (Cohen, 1988) wurden bereits in den jeweiligen Abschnitten der einzelnen abhängigen Maße berichtet. Aufgrund der gewählten Analyseeinheit Klasse führte die Schätzung der Effektstärken zu inkorrekten Standardfehlern, so dass sehr hohe Effektstärken (zwischen .80 und 6.63) berichtet wurden, die jedoch schwer zu interpretieren sind. Allgemein werden Effektstärken größer als .80 als groß, größer als .50 als mittel und größer als .20 als klein beurteilt (Graham & Perin, 2007a). Wenn die Effektstärken kleiner als .20 ausfallen, liegt kein Effekt vor. Demnach handelt es sich bei sämtlichen Effektstärken der Studie I und II um sehr große Effekte, was jedoch teilweise in Widerspruch dazu steht, dass nicht signifikante Bedingungeffekte berichtet wurden. Aus diesem Grund wurden zusätzlich Effektstärken auf Schülerebene berechnet und diesen vergleichend den Effektstärken auf Klassenebene gegenübergestellt (siehe Tabelle 72).

Tabelle 72: Effektstärken nach Cohen's d für Posttest und Follow-Up mit jeweils Klasse und Schüler als Analyseeinheit (Studie II).

	Klasse als Analyseeinheit		Schüler als Analyseeinheit	
	Posttest	Follow-Up	Posttest	Follow-Up
Geschichtenelemente	6.63	6.09	1.87	1.47
Geschichtenqualität	1.00	1.15	0.20	0.34
Wortanzahl	2.83	2.60	0.68	0.79
Planen	6.52	5.17	1.66	1.82
Revisionen	0.99	1.25	0.28	0.52
Wissen	0.80	0.87	0.24	0.36

Die Effektstärken auf Schülerebene fielen deutlich geringer aus als die auf Klassenebene. Dennoch ergaben sich sehr große Effektstärken für Geschichtenelemente und Planen, einen mittleren Effekt für Wortanzahl sowie kleine Effektstärken für Geschichtenqualität, Revisionen und Wissen. Insgesamt kann geschlussfolgert werden, dass hinsichtlich aller untersuchten abhängigen Variablen das USAT-Programm besser abschnitt als das LUG-Programm.

7.3.6 Differentielle Betrachtung der Bedingungsseffekte in Abhängigkeit von den Kontrollvariablen

Im folgenden Abschnitt werden die Effekte der USAT- und der LUG-Bedingung für Untergruppen der Kontrollvariablen Geschlecht, Rechtschreibung, Sprachverständnis und Muttersprache Deutsch separat gerechnet, um einen Moderatoreffekt der Kontrollvariablen ausschließen zu können. Da von der gesamten Stichprobe nur elf Schüler eine andere Nationalität als deutsch angaben, ist die Vergleichsgruppe zu gering, so dass die Kontrollvariable Nationalität weggelassen wurde. In der USAT-Bedingung gaben vier Schüler und in der LUG-Bedingung sieben Schüler eine andere Nationalität als Deutsch an.

7.3.6.1 Geschlecht

Zur differentiellen Beurteilung der Leistungen von Mädchen und Jungen wurden für jede abhängige Variable eine Mädchen- sowie eine Jungen-Variable gebildet, indem auf Schülerebene die Daten getrennt für Mädchen und Jungen über die Klassen hinweg aggregiert wurden. Tabelle 73 enthält die Mittelwerte und Standardabweichungen für die drei Messzeitpunkte und die beiden Bedingungen getrennt für Mädchen und Jungen.

Tabelle 73: Mittelwerte und Standardabweichungen für Mädchen und Jungen getrennt nach Messzeitpunkt und Bedingung (Studie II).

	USAT				LUG			
	Mädchen		Jungen		Mädchen		Jungen	
	M	SD	M	SD	M	SD	M	SD
Geschichtenelemente								
Prätest	2.70	0.54	1.71	0.41	2.69	0.55	2.21	0.62
Posttest	8.85	0.73	7.32	1.28	4.42	0.36	3.72	0.72
Follow-Up	7.98	0.71	6.53	1.44	4.44	0.34	3.80	0.62
Geschichtenqualität								
Prätest	3.72	0.28	2.85	0.31	3.59	0.43	2.97	0.23
Posttest	6.07	0.37	5.08	0.44	5.62	0.24	4.86	0.26
Follow-Up	5.55	0.57	5.12	0.55	5.18	0.34	4.62	0.39
Wortanzahl								
Prätest	86.97	21.92	61.95	8.13	81.89	17.43	67.60	12.31
Posttest	167.60	27.68	137.94	20.64	131.77	13.69	92.74	17.03
Follow-Up	175.64	21.62	141.10	16.30	136.65	18.09	98.74	10.03

Planen								
Prätest	2.26	0.41	1.80	0.38	2.35	0.48	2.07	0.25
Posttest	6.88	0.87	6.26	0.48	3.45	0.61	3.02	0.80
Follow-Up	6.60	0.68	6.09	1.15	2.85	0.74	2.69	0.46
Revisionen								
Prätest	0.32	0.15	0.16	0.09	0.84	0.46	0.50	0.35
Posttest	1.50	0.68	1.06	0.35	0.91	0.56	0.73	0.21
Follow-Up	1.79	0.70	1.61	0.79	1.07	0.80	0.63	0.44
Wissen								
Prätest	1.68	0.57	1.35	0.47	1.39	0.78	1.04	0.58
Posttest	6.29	1.00	5.30	1.04	5.25	0.86	5.46	0.91
Follow-Up	6.34	0.92	5.40	1.33	5.23	1.08	4.61	1.19

Anschließend wurde ein Set von Messwiederholungsanalysen mit den einzelnen strategienahen und -fernen Schreibmaßen als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Geschlecht (Mädchen, Jungen) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor und den Variablen Prätest (Mädchen, Jungen) und Rechtschreibung als Kovariaten gerechnet. Geschlecht wurde als Innersubjektfaktor behandelt, da in einer Klasse sowohl Mädchen als auch Jungen waren und die Effekte separat für diese beiden Gruppen analysiert wurden. Analog wurden auch die anderen Kontrollvariablen (Rechtschreibung, Sprachverständnis, Sprache) als Innersubjektfaktoren in den jeweiligen Analysen berücksichtigt. Die nachfolgende Tabelle 74 enthält die Zusammenfassung der Ergebnisse der Analysen.

Tabelle 74: Ergebnisse der Analysen zur Prüfung des Faktors Geschlecht als Moderator (Studie II).

	Bed. x Gesch.	Bed.	Gesch.	MZP	Prätest Jungen	Prätest Mäd.	Prätest RS
Geschichtenelemente		X					
Geschichtenqualität							
Wortanzahl		X					
Planen Inhalt		X					
Revisionen		X					
Wissen							

Anmerkungen: X = signifikanter Effekt ($p < .05$); Bed. x Gesch. = Interaktion Bedingung x Geschlecht; Bed. = Haupteffekt Zwischensubjektfaktor Bedingung; Gesch. = Haupteffekt Innersubjektfaktor Geschlecht; MZP = Haupteffekt Innersubjektfaktor Messzeitpunkt; Prätest Jungen = Effekt der Kovariaten Prätest Jungen; Prätest Mäd. = Effekt der Kovariaten Prätest Mädchen; Prätest RS = Effekt der Kovariaten Rechtschreibung.

Der Tabelle 74 ist zu entnehmen, dass die Ergebnisse der vorangegangenen Kovarianzanalysen auch nach separater Betrachtung von Jungen und Mädchen bestehen blieben. Abgesehen von Geschichtenqualität und Wissen ergaben sich signifikante Bedingungseffekte zu Gunsten der USAT-Bedingung ($ps < .05$). Sämtliche anderen Effekte waren nicht signifikant ($ps > .05$). Folglich profitierten Mädchen und Jungen in gleicher Weise von den beiden Bedingungen, wobei das USAT-Programm, außer für Geschichtenqualität und Wissen, weiterhin deutlich effektiver war als das LUG-Training.

7.3.6.2 Rechtschreibung

Analog zum Geschlecht wurde der Schülerdatensatz hinsichtlich rechtschreibschwacher und rechtschreibstarker Schüler getrennt über die Klassen hinweg aggregiert, so dass für jede abhängige Variable ein Wert für rechtschreibschwache und ein Wert für rechtschreibstarke Schüler gebildet wurde. Als rechtschreibschwach galten Schüler, die im DRT 3 (Müller, 2004) einen Wert kleiner dem 50. Perzentil (54 Punkte) erreicht hatten. Dementsprechend erzielten rechtschreibstarke Schüler im DRT 3 einen Wert gleich oder größer dem 50. Perzentil (54 Punkte).

Die folgende Tabelle 75 enthält die Mittelwerte und Standardabweichungen der rechtschreibschwachen (RS niedrig) und der rechtschreibstarken (RS hoch) Schüler getrennt nach Messzeitpunkt und Bedingung.

Tabelle 75: Mittelwerte und Standardabweichungen für rechtschreibschwache und rechtschreibstarke Schüler getrennt nach Messzeitpunkt und Bedingung (Studie II).

	USAT				LUG			
	RS niedrig		RS hoch		RS niedrig		RS hoch	
	M	SD	M	SD	M	SD	M	SD
Geschichtenelemente								
Prätest	1.92	0.37	2.56	0.56	1.83	1.07	2.55	0.47
Posttest	7.56	0.95	9.34	0.86	3.95	0.55	4.10	0.76
Follow-Up	6.66	0.74	8.48	1.00	3.89	1.22	4.26	0.51
Geschichtenqualität								
Prätest	3.06	0.32	3.68	0.51	2.84	0.62	3.50	0.39
Posttest	5.30	0.33	6.13	0.56	4.65	1.23	0.50	0.27
Follow-Up	5.07	0.32	5.89	0.69	4.36	0.88	5.07	0.55

Wortanzahl								
Prätest	69.71	9.64	79.81	17.25	67.37	19.98	77.64	10.77
Posttest	146.10	16.73	184.51	35.03	109.52	26.28	120.20	18.28
Follow-Up	145.29	13.55	174.91	34.10	101.41	25.47	115.45	24.47
Planen								
Prätest	1.90	0.15	2.18	0.61	2.05	0.35	2.26	0.47
Posttest	6.17	0.49	7.41	1.68	2.80	0.57	3.33	0.84
Follow-Up	5.73	0.84	7.50	0.99	2.66	0.76	3.08	0.34
Revisionen								
Prätest	0.25	0.10	0.30	0.15	0.40	0.41	0.76	0.36
Posttest	0.90	0.43	1.67	1.02	0.47	0.57	0.94	0.40
Follow-Up	1.58	0.76	2.15	0.66	0.69	0.53	0.95	0.75
Wissen								
Prätest	1.49	0.52	1.76	0.52	1.09	0.75	1.31	0.66
Posttest	5.28	0.79	6.83	0.81	5.62	1.45	5.18	0.75
Follow-Up	5.35	1.16	7.02	0.93	4.40	1.49	5.23	0.98

Anschließend wurden mehrere Kovarianzanalysen mit den einzelnen Schreibmaßen als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Rechtschreibung (RS niedrig, RS hoch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor und den Prätest-Werten jeweils für rechtschreibschwache und rechtschreibstarke Schüler als Kovariaten gerechnet. Die Ergebnisse sind in der nachfolgenden Tabelle 76 zusammengefasst.

Tabelle 76: Ergebnisse der Analysen zur Prüfung des Faktors Rechtschreibung als Moderator (Studie II).

	Bed. x RS	Bed.	RS	MZP	Prä RS niedrig	Prä RS hoch
Geschichtenelemente	X	X				
Geschichtenqualität						
Wortanzahl		X		X		
Planen Inhalt		X				
Revisionen		X				
Wissen	X					

Anmerkungen: X = signifikanter Effekt ($p < .05$); Bed. x RS. = Interaktion Bedingung x Rechtschreibung; Bed. = Haupteffekt Zwischensubjektfaktor Bedingung; RS = Haupteffekt Innersubjektfaktor Rechtschreibung; MZP = Haupteffekt Innersubjektfaktor Messzeitpunkt; Prätest RS niedrig = Effekt der Kovariaten Prätest Rechtschreibleistung niedrig; Prätest RS hoch = Effekt der Kovariaten Prätest Rechtschreibleistung hoch.

Die signifikanten Bedingungseffekte zu Gunsten der USAT-Bedingung sowie die nicht signifikanten Bedingungseffekte aus den vorangegangenen Analysen wurden bestätigt. Außer für Geschichtenqualität und Wissen waren die Bedingungseffekte für alle abhängigen Variablen signifikant ($p < .05$).

Zusätzlich wurden bei folgenden abhängigen Variablen signifikante Interaktionen Bedingung x Rechtschreibung festgestellt:

Eine Kovarianzanalyse mit Geschichtenelemente als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Rechtschreibung (RS niedrig, RS hoch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor und der Variablen Prätest (RS niedrig, RS hoch) als Kovariate ergab eine signifikante Interaktion Bedingung x Rechtschreibung, $F(1, 7) = 13.10$, $p < .01$. Eine genaue Betrachtung der Tabelle 75 ergibt, dass sowohl die rechtschreibschwachen als auch die rechtschreibstarken Schüler sich in Bezug auf Geschichtenelemente in beiden Bedingungen verbesserten, wobei die USAT-Bedingung deutlich besser abschnitt als die LUG-Bedingung. Die Unterschiede zwischen rechtschreibschwachen und rechtschreibstarken Schüler sind jedoch in der USAT-Bedingung größer als in der LUG-Bedingung, so dass eine signifikante Bedingung x Rechtschreibung-Interaktion zustande kam.

Des Weiteren lieferte eine Kovarianzanalyse mit Wissen als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Rechtschreibung (RS niedrig, RS hoch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor und der Variablen Prätest (RS niedrig, RS hoch) als Kovariate ebenfalls eine signifikante Bedingung x Rechtschreibung-Interaktion, $F(1, 7) = 6.63$, $p < .05$. Anhand der in Tabelle 75 aufgelisteten Mittelwerte und Standardabweichungen ist die signifikante Interaktion dadurch zu begründen, dass beim Posttest die rechtschreibschwachen Schüler der USAT-Bedingung geringere Leistungen zeigten als die rechtschreibschwachen Schüler der LUG-Bedingung. Es kann somit geschlossen werden, dass die rechtschreibschwachen Schüler besonders wenig in der USAT-Bedingung und besonders viel in der LUG-Bedingung hinsichtlich schreibbezogenen Wissens dazulernten. Gleichzeitig überragten die rechtschreibstarken USAT-Schüler sowohl beim Posttest als auch beim Follow-Up beide Subgruppen der LUG-Bedingung

Überdies führte eine Kovarianzanalyse mit Wortanzahl als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Rechtschreibung (RS niedrig, RS hoch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor und der Variablen Prätest (RS niedrig, RS hoch) als Kovariate zu einem signifikanten Effekt des

Innersubjektfaktors Messzeitpunkt, $F(1, 7) = 6.68, p < .05$. In Tabelle 75 wird deutlich, dass die Wortanzahl in beiden Bedingungen sowie für beide Subgruppen vom Posttest zum Follow-Up absank.

7.3.6.3 Sprachverständnis

Als nächstes wurden die schreibbezogenen Leistungen für Schüler mit niedrigem und Schüler mit hohem Sprachverständnis (SV) untersucht.

Dazu wurde der Schülerdatensatz aufgeteilt in eine Gruppe mit niedrigen Leistungen im Sprachverständnistest (HST 4/5; Mietzel & Willenberg, 2000) und in eine Gruppe mit hohen Leistungen im selbigen Test. Wie auch bei der Rechtschreibung wurde das 50. Perzentil, welches einem Punktwert von 41 entsprach, als Cut-Off-Point gewählt. Nachdem die beiden Gruppen gebildet worden waren, wurde der Datensatz über die Klassen hinweg aggregiert, so dass für jedes abhängige Schreibmaß zwei Werte (SV niedrig, SV hoch) gebildet wurden. In der nachfolgenden Tabelle 77 sind die Mittelwerte und Standardabweichungen der Leistungen sprachverständnisschwacher und –starker Schüler getrennt nach Messzeitpunkt und Bedingung aufgelistet.

Tabelle 77: Mittelwerte und Standardabweichungen für sprachverständnisschwache und sprachverständnisstarke Schüler getrennt nach Messzeitpunkt und Bedingung (Studie II).

	USAT				LUG			
	SV niedrig		SV hoch		SV niedrig		SV hoch	
	M	SD	M	SD	M	SD	M	SD
Geschichtenelemente								
Prätest	2.12	0.39	2.23	0.30	2.38	0.47	2.74	0.95
Posttest	7.18	1.26	9.31	1.01	4.18	0.47	4.34	0.69
Follow-Up	6.48	0.98	8.15	0.78	4.31	0.27	4.28	0.55
Geschichtenqualität								
Prätest	2.98	0.19	3.62	0.46	3.11	0.44	3.66	0.48
Posttest	5.25	0.54	5.92	0.25	5.28	0.19	5.53	0.35
Follow-Up	4.93	0.53	5.77	0.49	4.63	0.16	5.30	0.55
Wortanzahl								
Prätest	68.48	6.66	76.82	13.34	74.33	11.37	78.36	17.41
Posttest	149.17	17.92	163.35	7.15	122.65	18.04	122.19	12.40
Follow-Up	141.23	21.16	164.94	15.63	115.01	7.98	122.41	15.75

Planen								
Prätest	1.92	0.28	2.09	0.35	2.17	0.24	2.40	0.39
Posttest	5.76	0.58	7.59	0.68	3.17	0.51	3.57	1.02
Follow-Up	5.58	1.00	7.19	0.91	2.92	0.71	2.87	0.39
Revisionen								
Prätest	0.22	0.16	0.31	0.19	0.63	0.47	0.71	0.23
Posttest	0.94	0.49	1.38	0.67	0.66	0.33	1.02	0.68
Follow-Up	1.34	0.66	2.19	0.96	0.59	0.30	1.12	0.82
Wissen								
Prätest	1.60	0.45	1.62	0.59	1.29	0.52	1.33	0.70
Posttest	5.23	0.92	6.45	0.85	5.62	1.06	5.08	0.91
Follow-Up	5.17	1.15	6.66	1.11	4.98	1.31	5.14	0.94

Im Anschluss daran wurde ein Set von (Messwiederholungs-) Kovarianzanalysen mit den jeweiligen Schreibmaßen als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Sprachverständnis (SV niedrig, SV hoch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor und den beiden Variablen Prätest (jeweils für SV niedrig und SV hoch) und Rechtschreibung als Kovariaten gerechnet. Eine Zusammenfassung der Ergebnisse der Analysen ist in Tabelle 78 vorzufinden.

Tabelle 78: Ergebnisse der Analysen zur Prüfung des Faktors Sprachverständnis als Moderator (Studie II).

	Bed. x SV	Bed.	SV	MZP	Prä SV niedrig	Prä SV hoch	Prätest RS
Geschichtenelemente		X					
Geschichtenqualität							
Wortanzahl		X					
Planen Inhalt	X	X					
Revisionen							
Wissen				X			

Anmerkungen: X = signifikanter Effekt ($p < .05$); Bed. x SV = Interaktion Bedingung x Sprachverständnis; Bed. = Haupteffekt Zwischensubjektfaktor Bedingung; SV = Haupteffekt Innersubjektfaktor Sprachverständnis; MZP = Haupteffekt Innersubjektfaktor Messzeitpunkt; Prä SV niedrig = Effekt der Kovariaten Prätest Sprachverständnis niedrig; Prätest SV hoch = Effekt der Kovariaten Prätest Sprachverständnis hoch; Prätest RS = Effekt der Kovariaten Rechtschreibung.

Die Ergebnisse der Messwiederholungskovarianzanalysen bestätigten größtenteils die vorangegangenen Analysen. Für Geschichtenelemente, Wortanzahl und Planen zeigten sich

signifikante Bedingungseffekte zu Gunsten der USAT-Bedingung ($p < .01$). Zusätzlich zu Geschichtenqualität und Wissen konnten keine Bedingungseffekte für Revisionen festgestellt werden ($p > .05$).

Demnach bewirkte eine Kovarianzanalyse mit Revisionen als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Sprachverständnis (SV niedrig, SV hoch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor sowie den Prätest-Werten (SV niedrig, SV hoch) und Rechtschreibung als Kovariaten, dass der ursprüngliche Bedingungshaupteffekt, $F(1, 6) = 2.96$, $p = .14$, verschwand. Insofern ist im Zuge einer separaten Analyse sprachverständnisschwacher und –starker Schüler der Unterschied zwischen den USAT- und den LUG-Klassen nicht mehr statistisch bedeutsam.

Des Weiteren lieferte eine Kovarianzanalyse mit Planen als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Sprachverständnis (SV niedrig, SV hoch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor und den Prätest-Werten (SV niedrig, SV hoch) jeweils für sprachverständnisschwache und –starke Schüler sowie Rechtschreibung als Kovariaten eine signifikante Sprachverständnis x Bedingungs-Interaktion ($F(1, 6) = 8.48$, $p < .05$). Eine genaue Betrachtung der Mittelwerte in Tabelle 77 ergibt, dass sowohl sprachverständnisstarke als auch sprachverständnisschwache USAT-Schüler die jeweiligen Subgruppen der LUG-Bedingung übertrafen, aber gleichzeitig deutliche Leistungsunterschiede zwischen den sprachverständnisschwachen und sprachverständnisstarken Schülern der USAT-Bedingung vorlagen. Demgegenüber variierten die Mittelwerte der sprachverständnisschwachen und –starken Schüler der LUG-Bedingung kaum.

Außerdem führte eine Kovarianzanalyse mit Wissen als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Sprachverständnis (SV niedrig, SV hoch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor sowie den Prätest-Werten (SV niedrig, SV hoch) und Rechtschreibung als Kovariaten zu einem signifikanten Effekt des Innersubjektfaktors Messzeitpunkt, $F(1, 6) = 6.54$, $p < .05$. Bei Betrachtung der Mittelwerte in Tabelle 77 ist auffällig, dass beim Posttest die sprachverständnisschwachen Schüler der LUG-Bedingung besser waren als die der USAT-Bedingung und dass diese wiederum stark zum Follow-Up hin abfielen, so dass beim dritten Messzeitpunkt die sprachverständnisschwachen Schüler der USAT-Bedingung denen der LUG-Bedingung überlegen waren.

7.3.6.4 Sprache

Im nächsten Abschnitt wurden die Schreibmaße der USAT- und der LUG-Klassen in Abhängigkeit von der Muttersprache Deutsch analysiert.

Zunächst wurde der Schülerdatensatz in zwei Gruppen eingeteilt. Während die eine Gruppe aus Schülern bestand, die nur Deutsch zu Hause sprachen, bestand die zweite Gruppe aus Schülern, die Deutsch und/oder eine andere Sprache zu Hause sprachen. Die Daten wurden für jede der beiden Gruppen getrennt über die Klassen hinweg aggregiert, so dass für jedes strategienahen und -ferne Schreibmaß ein Wert für nur Deutsch und für Deutsch/andere Sprachen gebildet wurde. Die nachfolgende Tabelle 79 enthält die Mittelwerte und Standardabweichungen der einzelnen abhängigen Schreibmaße getrennt nach Muttersprache Deutsch, Bedingung und Messzeitpunkt.

Tabelle 79: Mittelwerte und Standardabweichungen für deutschsprechende und nicht-deutschsprechende Schüler getrennt nach Messzeitpunkt und Bedingung (Studie II).

	USAT				LUG			
	Deutsch		nicht-Deutsch		Deutsch		nicht-Deutsch	
	M	SD	M	SD	M	SD	M	SD
Geschichtenelemente								
Prätest	2.15	0.35	2.13	0.40	2.54	0.53	2.39	0.52
Posttest	8.06	0.69	8.22	1.45	4.24	0.58	3.76	1.00
Follow-Up	7.25	0.43	7.40	1.48	4.30	0.35	4.04	0.75
Geschichtenqualität								
Prätest	3.25	0.28	3.28	0.15	3.44	0.31	3.22	0.49
Posttest	5.49	0.17	5.70	0.65	5.45	0.26	5.00	0.27
Follow-Up	5.25	0.28	5.61	0.84	5.09	0.34	4.66	0.29
Wortanzahl								
Prätest	72.93	10.49	74.58	9.70	77.49	12.85	77.27	21.34
Posttest	153.57	5.48	167.49	32.53	121.49	17.78	115.28	15.97
Follow-Up	151.02	13.34	120.04	12.24	161.80	33.38	115.05	12.16
Planen								
Prätest	2.03	0.39	1.99	0.45	2.22	0.33	2.37	0.44
Posttest	6.46	0.32	6.82	1.30	3.33	0.77	3.13	0.43
Follow-Up	6.21	0.79	6.65	1.32	2.82	0.50	2.90	0.78

Revisionen								
Prätest	0.25	0.11	0.15	0.16	0.72	0.29	0.70	0.43
Posttest	1.17	0.41	1.41	1.07	0.95	0.54	0.67	0.21
Follow-Up	1.60	0.68	2.10	0.78	1.01	0.70	0.56	0.36
Wissen								
Prätest	1.59	0.55	1.33	0.47	1.42	0.72	0.96	0.56
Posttest	5.84	0.65	5.70	1.63	5.69	1.06	4.41	0.54
Follow-Up	5.84	1.13	5.92	1.32	5.37	1.30	4.12	1.31

Schließlich wurden eine Reihe von (Messwiederholungs-) Kovarianzanalysen mit den jeweiligen Schreibmaßen als abhängige Variablen, Messzeitpunkt (Posttest, Follow-Up) und Sprache (Deutsch, nicht-Deutsch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor und den jeweiligen Prätest-Werten für deutschsprechende und nicht-deutschsprechende Schüler sowie die Rechtschreibung als Kovariaten durchgeführt.

Tabelle 80: Ergebnisse der Analysen zur Prüfung des Faktors Sprache als Moderator (Studie II).

	Bed. x Sprache	Bed.	Sprache	MZP	Prä Deutsch	Prä nicht-Deutsch	Prätest RS
Geschichten-elemente		X			X	X	
Geschichten-qualität							
Wortanzahl		X					
Planen Inhalt		X					
Revisionen							
Wissen							

Anmerkungen: X = signifikanter Effekt ($p < .05$); Bed. x SV = Interaktion Bedingung x Sprache; Bed. = Haupteffekt Zwischensubjektfaktor Bedingung; Sprache = Haupteffekt Innersubjektfaktor Sprache; MZP = Haupteffekt Innersubjektfaktor Messzeitpunkt; Prä Deutsch = Effekt der Kovariaten Prätest Deutsch; Prätest nicht-Deutsch = Effekt der Kovariaten Prätest nicht-Deutsch; Prätest RS = Effekt der Kovariaten Rechtschreibung.

Analog zur Kontrollvariable Sprachverständnis bestätigten die Kovarianzmesswiederholungsanalysen die signifikanten Bedingungseffekte zu Gunsten der USAT-Bedingung für Geschichtenelemente, Wortanzahl und Planen ($ps < .05$). Für die abhängigen Variablen Geschichtenqualität, Wissen und Revisionen konnten keine Bedingungseffekte festgestellt werden ($ps > .05$).

Folglich führte eine Kovarianzanalyse mit Revisionen als abhängige Variable Messzeitpunkt (Posttest, Follow-Up) und Sprache (Deutsch, nicht-Deutsch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor und den jeweiligen Prätest-Werten (Deutsch, nicht-Deutsch) sowie Rechtschreibung als Kovariaten zu einem knapp nicht signifikantem Bedingungseffekt, $F(1, 6) = 5.51, p = .06$. Demnach erwies sich der Unterschied zwischen der USAT- und der LUG-Bedingung als nicht mehr signifikant, sobald deutsch- und nicht-deutschsprechende Schüler getrennt analysiert wurden.

Des Weiteren erbrachte eine Kovarianzanalyse mit Geschichtenelemente als abhängige Variable, Messzeitpunkt (Posttest, Follow-Up) und Sprache (Deutsch, nicht-Deutsch) als Innersubjektfaktoren, Bedingung (USAT, LUG) als Zwischensubjektfaktor und den jeweiligen Prätest-Werten (Deutsch, nicht-Deutsch) sowie Rechtschreibung als Kovariaten neben dem signifikanten Bedingungseffekt, $F(1, 6) = 40.90, p < .01$, auch signifikante Effekte der Kovariaten Prätest Deutsch, $F(1, 6) = 12.35, p < .05$, und Prätest nicht-Deutsch, $F(1, 6) = 10.42, p < .05$.

7.4 Diskussion

Die Studie II intendierte zum einen die Replizierbarkeit der Befunde der Studie I und stellte zum anderen einen weiteren Schritt zur Umsetzung des USAT-Programms unter realen Alltagsbedingungen des Deutschunterrichts dar. Nach Absolvierung eines Workshops setzten Lehrer das USAT-Programm in den eigenen Klassen um. Des Weiteren wurde in der Studie II dem USAT-Programm ein alternatives Trainingsprogramm (LUG) gegenübergestellt, welches die traditionellen Unterrichtsmethoden der vierten Klasse widerspiegelt und ebenfalls von den Klassenlehrern implementiert wurde. Als nächstes werden die Ergebnisse der Studie II zusammengefasst, Einschränkungen diskutiert und schließlich Bezug genommen zu den postulierten Hypothesen.

7.4.1 Zentrale Befunde der Studie II

Bedingungseffekte: Insgesamt profitierten sowohl die USAT- als auch die LUG-Klassen von den jeweiligen Trainings und konnten ihre Leistungen bezüglich der meisten abhängigen Variablen direkt im Anschluss an das Training verbessern.

Gleichzeitig waren die USAT-Klassen bezüglich der Maße Geschichtenelemente, Wortanzahl, Planen und Revisionen signifikant den LUG-Klassen überlegen. Für Geschichtenqualität und Wissen konnte kein signifikanter Bedingungseffekt nachgewiesen werden. Auch wenn im Vergleich zur Studie I kein Bedingungseffekt für Geschichtenqualität und Wissen erzielt wurde, ist dennoch die Schlussfolgerung zulässig, dass das USAT-

Programm sich als effektiver erwies als das LUG-Programm. Dies bestätigten Effektstärken, die auf Schülerebene einen großen Effekt (1.47 bis 1.87) für Geschichtenelemente und Planen, einen mittleren Effekt für Wortanzahl (.68 bis .79) sowie einen kleinen Effekt (.20 bis .52) für Geschichtenqualität, Revisionen und Wissen anzeigten.

Geschichtenelemente ausgenommen, zeigte sich außerdem bei keiner abhängigen Variablen eine signifikante Messzeitpunkt x Bedingung-Interaktion ($p > .05$). Bezüglich Geschichtenelemente war ein Leistungsabfall beim Follow-Up in der USAT-Bedingung zu verzeichnen. Dies wurde durch die Effektstärkenberechnung auf Schülerebene bestätigt. Während die Effektstärke für Geschichtenelemente vom Posttest zum Follow-Up abfiel, stiegen bei allen anderen abhängigen Variablen die Effektstärken an.

Insbesondere die Überlegenheit des USAT-Programms bei prozessbezogenen Variablen wie Planen und Revisionen belegt die Effektivität von selbstregulatorischen Methoden. Im Gegensatz zur USAT-Bedingung wurden in der LUG-Bedingung keine selbstregulatorischen Prinzipien beim Planen, Schreiben und Überarbeiten eingesetzt. Folglich kann die Überlegenheit des USAT-Programms auf die Selbstregulation zurückgeführt werden.

Klassenunterschiede innerhalb einer Bedingung: Das Training wurde von den einzelnen Klassenlehrern durchgeführt. Im Gegensatz zu geschulten Trainern integrierten die Lehrer ihre individuellen Unterrichtsstile, mit denen die Schüler bereits vertraut waren, in die jeweiligen Schreibtrainings. Darüber hinaus verfügten die Lehrer über ein unterschiedliches Repertoire an didaktischem wie auch fachlichem Wissen und waren unterschiedlich stark motiviert. Zudem war die Schulung der Lehrer nicht mit der Schulung der Trainerinnen der Studie I vergleichbar, so dass bestehende Lehrer-Unterschiede nicht vollkommen aufgehoben werden konnten, ohne den natürlichen Klassenkontext und den gewohnten Umgang der Lehrer mit ihren Schülern zu stören. Aus diesem Grund zeigten sich in der Studie II insgesamt mehr Klassenunterschiede als in der Studie I.

Bei der Betrachtung der Klassen der USAT-Bedingung ist auffällig, dass bei allen abhängigen Variablen die Unterschiede beim Follow-Up größer wurden. Da die Lehrer in dem Zeitraum zwischen Posttest und Follow-Up die Schüler weiterhin unterrichteten, liegt die Vermutung nahe, dass einige Lehrer, unbewusst oder bewusst, die trainingsspezifischen Schreibstrategien in ihrem regulären Deutschunterricht übernommen und dadurch wiederholt hatten. Alle Lehrer in beiden Bedingungen wurden jedoch explizit darauf hingewiesen, bis zum Follow-Up keine der Schreibstrategien im Unterricht zu behandeln. Auch wenn dies nicht kontrolliert wurde, ist davon auszugehen, dass die meisten Lehrer sich an diese Vorgabe

hielten, da sie den wissenschaftlichen Ansatz der Studie unterstützten und angaben, sich nach der Durchführung des USAT-Programms vornehmlich mit anderen Bereichen des Deutschunterrichts beschäftigt zu haben.

Eine weitere Ursache könnte jedoch darin begründet liegen, dass nur ein Teil der Lehrer es schaffte, die selbstregulatorischen Prinzipien und Schreibstrategien so bei den Schülern zu verfestigen, dass diese auch noch sechs Wochen später bei diesen abrufbar waren. Folglich scheint die Effektivität des USAT-Programms bei der Umsetzung durch Lehrer in erheblichem Ausmaß von den Eigenschaften und Fertigkeiten der Lehrer selbst abzuhängen. Ein Beleg hierfür ist, dass hinsichtlich der Maße Geschichtenelemente, Planen und Wissen jeweils die Klasse 5 am besten und die Klasse 3 am schlechtesten abschnitt. Besonders auffällig ist ebenfalls der starke Leistungsanstieg der Klasse 1 bei Revisionen. Demzufolge wäre es sinnvoll, die Unterschiede im Verhalten und in der Instruktion der Lehrer 3 und 5 zu analysieren und zu vergleichen.

Im Gegensatz dazu unterschieden sich die Klassen der LUG-Bedingung weniger stark beim Follow-Up und zeigten dagegen in mehreren abhängigen Variablen Unterschiede in ihren Ausgangsniveaus. Diese wurden jedoch in den Kovarianzanalysen berücksichtigt.

Betrachtet man die Abbildungen mit den Mittelwerten der USAT- und der LUG-Bedingung getrennt nach Klasse, fällt auf, dass insgesamt die Klassenmittelwerte der USAT-Bedingung einen parallelen Verlauf zeigten. Demgegenüber variierten die Verläufe der Linien der LUG-Bedingung sehr stark. Dies war vor allem bei Planen und Revisionen auffällig, wo manche LUG-Klassen gar nichts und andere wiederum sehr viel im Verlauf des Trainings dazulernten.

Differentialbetrachtung Kontrollvariablen: Um Aussagen darüber zu treffen, ob die beiden Bedingungen in Abhängigkeit von den Kontrollvariablen unterschiedlich effektiv waren, können die Interaktionen Bedingung x Kontrollvariable herangezogen werden.

Für die Kontrollvariablen Geschlecht und Muttersprache Deutsch konnten für keine abhängige Variable signifikante Interaktionen festgestellt werden.

Bezüglich der Rechtschreibung ergaben sich signifikante Bedingung x Rechtschreibung-Interaktionen für die abhängigen Variablen Geschichtenelemente und Wissen. Für die Kontrollvariable Sprachverständnis erwies sich die Interaktion Bedingung x Sprachverständnis für die abhängige Variable Planen als signifikant.

Folglich ist festzustellen, dass bezüglich der meisten abhängigen Variablen die Kontrollvariablen Geschlecht, Rechtschreibung, Sprachverständnis und Muttersprache Deutsch den gleichen Einfluss auf die USAT- und LUG-Bedingung ausübten. Bei einigen

wenigen Maßen produzierte jedoch das USAT-Programm größere Differenzen zwischen den Subgruppen der Kontrollvariablen als das LUG-Programm, wobei die USAT-Klassen meistens die LUG-Klassen übertrafen (Geschichtenelemente, Planen).

Überdies können die nicht signifikanten Bedingungseffekte beim Wissen teilweise auf die Kontrollvariable Rechtschreibung zurückgeführt werden, da beim Posttest rechtschreibschwache USAT- und LUG-Schüler dieselben Leistungen erbrachten. Die rechtschreibstarken USAT-Schüler waren dagegen den LUG-Schülern sowohl beim Posttest als auch beim Follow-Up überlegen

Des Weiteren verschwand der signifikante Bedingungseffekt für Revisionen, wenn jeweils das Sprachverständnis oder die Muttersprache berücksichtigt wurde. Der in den vorherigen Analysen festgestellte Bedingungseffekt für Revisionen war jedoch nur knapp signifikant und nicht separat für den zweiten und dritten Messzeitpunkt feststellbar. Nichtsdestotrotz scheinen das Sprachverständnis und die Muttersprache Deutsch die Revisionsfertigkeiten zu beeinflussen.

Implementationsgenauigkeit: Die Qualität der Implementierung wurde durch verschiedene Sicherungsmaßnahmen gewährleistet.

Da die LUG-Lehrerinnen gewisse Vorgaben bekamen und dadurch potenzielle Störfaktoren konstant gehalten wurden, kann generell von einer hohen internen Validität ausgegangen werden. Überdies kann die Studie I, die unter stärker kontrollierten Bedingungen durchgeführt wurde, als weiterer Beleg dafür herangezogen werden, dass die ermittelten Bedingungsunterschiede in den abhängigen Variablen auf Differenzen in den Interventionen und nicht auf etwaige konfundierende Variablen zurückzuführen sind.

Des Weiteren kann eine hohe externe Validität der Studie II angenommen werden. Die berichteten Ergebnisse sind auf Subgruppen der Stichprobe (Mädchen/Jungen, rechtschreibstark/rechtschreibstark, sprachverständnisstark/sprachverständnischwach, deutsch/nicht-Deutsch) generalisierbar. Außerdem erfolgte der Nachweis, dass das USAT-Programm auch mit Lehrern und nicht nur mit geschulten Trainern durchführbar ist. Folglich kann davon ausgegangen werden, dass andere Lehrer ebenfalls nach der Absolvierung eines Workshops das USAT-Programm umsetzen könnten. Die untersuchte Lehrerstichprobe ist jedoch nicht vollkommen repräsentativ für die gesamte Population, da abgesehen von einem Mann nur Frauen teilnahmen und alle Lehrer über erhebliche Berufserfahrungen verfügten. Folglich wäre die Implementierung durch Männer und unerfahrene Lehrer wie Referendare für eine weitere Generalisierung der Befunde notwendig.

Die soziale Validität wurde schließlich durch Beobachtungsbögen am Ende der Workshops und der Trainings überprüft. Hier zeigten sowohl die Lehrer der USAT- als auch der LUG-Bedingung eine große Zufriedenheit mit den Workshops und den Trainings, wobei die USAT-Lehrer den Workshop und das Training am Ende besser einstufen als die LUG-Lehrerinnen. Bis auf eine LUG-Lehrerin würden zudem alle Lehrer das jeweilige Trainingsprogramm weiterempfehlen. Die beim Nachtreffen der LUG-Lehrerinnen hervorgebrachte Kritik umfasste vor allem die Materialien und die Verständlichkeit der Manuale. Da das LUG-Programm neu konzipiert wurde, müsste es in nachfolgenden Studien überarbeitet werden. Insgesamt ist von einer hohen sozialen Akzeptanz auf Seiten der Lehrer für beide Bedingungen auszugehen, wobei das USAT-Programm, welches vielfach revidiert wurde, besser abschnitt als das LUG-Programm. Nichtsdestotrotz sollten in zukünftigen Studien die soziale Validität auf Seiten der Schüler erhoben werden. In der LUG-Bedingung wurde dies bereits ansatzweise dadurch realisiert, dass die Lehrer die Schüler in der letzten Trainingseinheit nach ihrer Meinung fragten. Die Schüler äußerten sich dabei positiv über das LUG-Programm.

Die Sicherstellung der Implementationsgenauigkeit ist ein weiterer wesentlicher Aspekt bei der Beurteilung der Implementierungsqualität. In der Studie II wurden verschiedene Maßnahmen zur Sicherstellung der Implementationsgenauigkeit integriert. Abgesehen von den beiden Workshops wurden die ersten Trainingseinheiten, in denen die wichtigsten Strategien vermittelt wurden, in beiden Bedingungen beobachtet. Anhand der Kriterien Klarheit der Instruktion, Klassenführung, Individualisierung, Motivierung, Aufmerksamkeit der Schüler und Sonstiges wurde die Qualität der Umsetzung bewertet. Die USAT-Lehrer erzielten etwas bessere Beurteilungen als die LUG-Lehrerinnen. Es ergab sich jedoch nur ein signifikanter Unterschied zu Gunsten der USAT-Bedingung bezüglich Individualisierung. Die inhaltliche Vollständigkeit wurde mit Checklisten, die zunächst von den Beobachtern und anschließend von den Lehrern ausgefüllt wurden, festgestellt. In beiden Bedingungen wurden die wesentlichen Trainingsinhalte vermittelt. Die Beobachtungen wurden von den USAT-Lehrern als hilfreich und unterstützend wahrgenommen, wobei dies in der LUG-Bedingung weniger zutraf. Es ist daher unklar, ob das Vorhandensein von Beobachtern das Lehrer- und Schülerverhalten beeinflusste. Eine Alternative zu Beobachtungen wären Videoaufnahmen, die eventuell weniger stark auf das Unterrichtsgeschehen einwirken.

7.4.2 Einschränkungen

Im Folgenden werden die wichtigsten Befunde diskutiert und Kritikpunkte identifiziert, die in zukünftigen Studien berücksichtigt werden sollten.

Effektivität des USAT-Programms: Während in Studie I der Nachweis erbracht wurde, dass das USAT-Programm erfolgreich durch Trainer in kontrollierten Versuchsbedingungen in den Regelunterricht implementierbar ist, ermöglicht Studie II die Schlussfolgerung, dass das USAT-Programm auch von Lehrern selbst durchführbar ist. Deswegen kann davon ausgegangen werden, dass das USAT-Programm auch von anderen Personen und in anderen Klassen umsetzbar ist. Aufgrund der positiven Effekte des Programms stellt das USAT-Programm eine Bereicherung des aktuellen Aufsatzunterrichts dar.

Überlegenheit selbstregulatorischer Prinzipien: Zunächst sind die ausbleibenden Bedingungseffekte für die abhängigen Variablen Geschichtenqualität und Wissen erwartungswidrig und nicht mit den Ergebnissen der Studie I konform. Da die USAT-Mittelwerte der Studie I tendenziell eher niedriger ausfallen als die USAT-Mittelwerte der Studie II, sind die fehlenden Bedingungseffekte nicht auf eine schlechtere Umsetzung des USAT-Programms durch die Lehrer, sondern vielmehr durch eine sehr starke LUG-Kontrollbedingung zu erklären. Die besonders gute Qualität der Implementierung des USAT-Programms wurde überdies durch die Unterrichtsbeobachtungen bestätigt. Insgesamt setzte sich die LUG-Bedingung aus leistungsstarken Klassen zusammen, die bezüglich Geschichtenqualität und Wissen nahezu die gleichen Leistungen hervorbrachten wie die Klassen der USAT-Bedingung. Wie bereits oben diskutiert, ist der fehlende Bedingungseffekt für Wissen teilweise auf die Kontrollvariable Rechtschreibung zurückzuführen.

Des Weiteren könnte eine weitere mögliche Erklärung für das besonders gute Abschneiden der LUG-Klassen in den Variablen Geschichtenqualität und Wissen darin liegen, dass die LUG-Lehrerinnen sehr von dem selbst konzipierten Schreibprogramm überzeugt und dementsprechend motiviert waren, die Inhalte möglichst optimal in dem eigenen Unterricht umzusetzen.

Replizierbarkeit der Befunde der Studie I: Die Befunde der Studie I konnten nur teilweise repliziert werden. Die signifikanten Bedingungseffekte für Geschichtenqualität und Wissen wurden nicht in der Studie II repliziert. Für Geschichtenelemente und Planen ergaben sich jedoch wie in der Studie I signifikante Unterschiede zu Gunsten der USAT-Bedingung. Die berichteten Effektstärken waren dabei etwas geringer als in der Studie I. Darüber hinaus

erzielte die Studie II im Gegensatz zur Studie I signifikante Bedingungseffekte zu Gunsten der USAT-Bedingung für Wortanzahl und Revisionen.

Aktuelle Situation im Aufsatzunterricht: Das LUG-Programm spiegelt den aktuellen Aufsatzunterricht in der vierten Klasse wider, so dass aus der Studie II folgende Schlussfolgerungen gezogen werden können.

(1) *Fehlende selbstregulatorische Prinzipien:* In dem LUG-Workshop wurden keine Ideen für selbstregulatorische Prinzipien von den LUG-Lehrerinnen hervorgebracht, was mit den Ergebnissen von vorangegangenen Studien einhergeht (Cutler & Graham, 2008) und die Ergebnisse der Lehrerbefragung bestätigt (siehe Kapitel 4.3).

Dass das Konzept der Selbstregulation bislang unter den beteiligten Lehrern und in deren Unterricht wenig verbreitet ist, lässt sich damit erklären, dass die Lehrer keine klaren Konzepte für die Instruktion von Planungs- und Überarbeitungsfertigkeiten besaßen. Am Anfang des Workshops wurden die USAT-Lehrer nach Ideen und bisherigen Erfahrungen gefragt. Sie gaben an, lediglich den Schülern Zeit für das Planen und Überarbeiten einzuräumen und die Geschichtenideen bzw. Entwürfe in der Klasse oder in Kleingruppen diskutieren zu lassen. In derselben Weise äußerten sich die LUG-Lehrerinnen. Die Konzeption von Trainingseinheiten zum Planen und Überarbeiten fiel ihnen dementsprechend schwer. Am Ende entwickelten die LUG-Lehrerinnen eine Trainingseinheit Planen, in der es nur um das Generieren von Ideen ging und eine Trainingseinheit Überarbeiten, die lediglich die Einführung von Bewertungskriterien, auf die die Schüler beim Revidieren achten sollten, umfasste. Lediglich eine Lehrerin machte den Vorschlag für die Einführung einer Checkliste. Diese Idee wurde jedoch nicht weiter von den anderen Lehrerinnen verfolgt. Gleichzeitig ist aus den Erfahrungen mit den USAT-Lehrern zu erwarten, dass der Einsatz von Checklisten wenig effizient gewesen wäre, da die Lehrer diese nicht als ein selbstregulatorisches Hilfsmittel wahrgenommen und somit entsprechende Einstellungen bei den Schülern nicht gefördert hätten. Darüber hinaus war den Lehrern die Notwendigkeit, Schreibstrategien mit selbstregulatorischen Strategien explizit zu verknüpfen, neu und unbekannt.

Für die Erarbeitung einer Rückmeldeprozedur bestätigte sich ebenfalls der Eindruck fehlender selbstregulatorischer Kenntnisse auf Seiten der Lehrer. Die LUG-Lehrerinnen entwarfen eine sehr lehrerzentrierte Rückmeldung. Während die Schüler nur allgemeine Aussagen über die Qualität des geschriebenen Textes und einen Stempel der Universität bekamen, erhielten die Lehrerinnen ein ausführliches Feedback zu jedem Schüler zu den einzelnen Schreibstrategien. Die LUG-Lehrerinnen gingen davon aus, dass sich die Schreibleistungen und Überarbeitungsfertigkeiten nur dann verbessern ließen, wenn sie und

nicht die Schüler die Leistungen kontrollierten. Sie trauten es den Schülern anscheinend nicht zu, selbst ihr schreibbezogenes Verhalten in Abhängigkeit von den Rückmeldungen zu regulieren. Für die USAT-Lehrer war es deshalb ebenfalls eine ungewohnte Erfahrung, die Schüler sich selbst überwachen, kontrollieren und eigene Ziele setzen zu lassen. Die Ergebnisse der Studie I und II belegen jedoch, dass Schüler der vierten Klasse sehr gut in der Lage sind, ihr Schreibverhalten selbst zu regulieren, so dass selbstregulatorische Strategien zu größeren Lernzuwächsen führten als Schreibtrainings ohne Selbstregulation. In der Studie II sind die größten Leistungsunterschiede zwischen der USAT- und der LUG-Bedingung hinsichtlich Geschichtelemente, Planen und Revisionen zu verzeichnen. Diese Differenzen sind ebenfalls allein darauf zurückzuführen, dass USAT-Schüler sich beim Planen, Schreiben und Überarbeiten kontrollierten und im Gegensatz dazu die LUG-Schüler beim Planen, Schreiben und Überarbeiten von dem Lehrer reguliert wurden.

(2) *Gute Kenntnisse über die Instruktion von Schreibstrategien:* Insgesamt entwickelten die LUG-Lehrerinnen, den Vorgaben entsprechend, gute Ideen für die Umsetzung der vorgegebenen Schreibstrategien in ihren Unterricht, so dass die Schüler qualitativ bessere Aufsätze schrieben und über mehr Strategiewissen verfügten. Zum Beispiel die Übungen zur Wiederholung der 7W-Fragen (z.B. Identifikation von 7W-Fragen in verschiedenen Texten) entsprachen denen der USAT-Bedingung, so dass die inhaltliche Abgrenzung zwischen den beiden Bedingungen weniger deutlich war als in der Studie I. In gleicher Weise waren die im USAT-Workshop vermittelten Schreibstrategien ebenfalls nicht komplett neu für die USAT-Lehrer.

Folglich besaßen die LUG-Lehrerinnen bereits über ausreichend Wissen und Fertigkeiten, die Schreibstrategien erfolgreich zu realisieren und die Schreibkompetenzen von Viertklässlern zu steigern. Gleichzeitig herrscht im aktuellen Deutschunterricht eine defizitäre Vermittlung von selbstregulatorischen Prozeduren, was zur Folge hat, dass die Schüler wenig oder kaum ihre Texte planen und überarbeiten und dadurch geringere Schreibleistungen erbringen als sie potenziell könnten.

Bezüglich der Bereicherung des USAT-Programms durch Ideen des LUG-Programms konnten keine Komponenten oder Übungen identifiziert werden, die zukünftig in das USAT-Programm aufgenommen werden sollten.

Kritikpunkte der Studie II: Folgende Einschränkungen der Studie II lassen sich aufführen.

(1) *Keine zufällige Zuordnung der Lehrer zu den Bedingungen:* Im Gegensatz zur Studie I wurde in Studie II die Zuteilung der Klassen zu den Bedingungen nicht anhand der

Klassengröße und der Rechtschreibung gematcht. Um die Bereitschaft zur Mitarbeit und zur korrekten Durchführung der Trainings zu erhöhen, meldeten sich die Lehrer ihren Interessen und Erfahrungen entsprechend konkret für ein bestimmtes Programm bzw. einen Workshop an. So stammten alle USAT-Lehrer aus einer Schule, während die LUG-Lehrerinnen aus vier unterschiedlichen Schulen kamen. Des Weiteren meldeten sich sechs Lehrer für die USAT- und nur fünf Klassen für die LUG-Bedingung, so dass den Ergebnissen eine unterschiedliche Stichprobenanzahl in den Bedingungen zu Grunde lag.

Die nicht randomisierte Zuteilung hatte weiterhin zur Folge, dass signifikante Differenzen zwischen den Bedingungen bezüglich des Alters und der Rechtschreibung vorlagen. Bei einer Randomisierung wären diese gematcht worden. Hinsichtlich des Sprachverständnisses ergaben sich jedoch keine signifikanten Unterschiede zwischen den Bedingungen. Da sich in der vierten Klasse die Leistungen der Schüler sehr schnell ändern, kann das bessere Abschneiden der LUG-Klassen im Rechtschreibtest auf das höhere Alter der LUG-Schüler zurückgeführt werden. Dennoch ist davon auszugehen, dass Schüler der USAT-Bedingung während der Textproduktion stärker von Rechtschreibschwierigkeiten behindert wurden als Schüler der LUG-Bedingung. In der Studie I dagegen lagen aufgrund des Matchings keine signifikanten Rechtschreibunterschiede zwischen den Bedingungen vor.

(2) *Korrektur der Schüleraufsätze durch universitäre Mitarbeiter:* Als nächste Limitierung kann die Korrektur der Schüleraufsätze durch universitäre Mitarbeiter angeführt werden. Die Studie II belegte die generelle Umsetzbarkeit des USAT- und des LUG-Programms durch Lehrer. Dennoch müssen weitere Entwicklungen erfolgen, um eine vollständige Implementierung durch Lehrer zu ermöglichen. Während die LUG-Lehrerinnen das Rückmeldeverfahren selbst konzipierten, lernten die USAT-Lehrer im Workshop das Ausfüllen der 7W- und Leistungsprotokolle. Eine Schülergeschichte, die die Schüler als Hausaufgabe schrieben, bewerteten die USAT-Lehrer ebenfalls eigenständig. Folglich wäre es prinzipiell für die Lehrer beider Bedingungen möglich gewesen, die Schülergeschichten selbst zu lesen und Rückmeldungen zu geben. Wie auch die LUG-Lehrerinnen erhielten die USAT-Lehrer stichpunktartig zu jeder Geschichte und zu jedem Schüler eine Zusammenfassung der Leistungen, so dass sie etwaige Schwierigkeiten der Klasse in ihrem Unterricht berücksichtigen konnten. Aufgrund des wissenschaftlichen Anspruchs der Studie II und zur Sicherstellung der internen Validität, wurden möglichst viele Störfaktoren kontrolliert. Eine Korrektur der Aufsätze durch die Lehrer hätte zu Schwierigkeiten bezüglich der Vergleichbarkeit der Rückmeldungen innerhalb einer Bedingung geführt, da die Lehrer eventuell zusätzlich andere Standards herangezogen und die Form der Rückmeldung an die

reguläre Rückmeldekultur der Klasse angepasst hätten. Weiter wäre die Arbeitsbelastung durch die wöchentliche Korrektur von Aufsätzen für die Lehrer enorm gewesen, so dass eine schlechtere Vorbereitung auf das Training die Folge gewesen wäre und die Lehrer die Trainingseinheiten nicht in dem vorgegebenen Zeitrahmen von fünf bis sieben Wochen realisiert hätten. Aus zeitlichen Gründen und zur Reduktion der Belastung der Lehrer wurden somit die Schüleraufsätze extern korrigiert. In weiterführenden Studien müsste untersucht werden, inwieweit durch Veränderung des USAT-Programms, z.B. durch eine Verlängerung des Durchführungszeitraums, den Lehrern ermöglicht werden kann, das USAT-Programm vollständig, inklusive der Korrektur der Aufsätze, in den eigenen Klassen zu implementieren.

(3) *Unterrichtsbeobachtungen nur in den ersten drei Trainingseinheiten:* Ein weiterer Kritikpunkt an der Studie II liegt darin begründet, dass nur die ersten drei Trainingseinheiten beobachtet wurden und dass die Beobachter nicht zufällig den Klassen zugeordnet wurden. Ein Beobachter untersuchte immer dieselbe Klasse, so dass die Beobachter der USAT- und der LUG-Bedingung variierten. Die Beobachter waren mit den Workshop-Leitern identisch. Dieses Vorgehen hatte den Vorteil, dass die Beobachter sich sehr gut mit dem jeweiligen Trainingsprogramm auskannten und dadurch den Lehrern Tipps geben konnten. Außerdem konnte auf diese Weise von einem guten Verhältnis zwischen Beobachtern und Lehrern ausgegangen werden, so dass diese die Beobachtungen nicht als störend, sondern als Hilfestellung im Unterricht ansahen. Des Weiteren nahmen alle Beobachter an einer gemeinsamen Schulung teil. Nichtsdestotrotz sollte in zukünftigen Untersuchungen jeder Beobachter eine gleiche Anzahl an Klassen in beiden Bedingungen untersuchen.

Wie bereits beschrieben, wurden in den drei beobachteten Einheiten die wesentlichen Strategien instruiert. In den übrigen Trainingseinheiten fanden hauptsächlich Wiederholungen und Übungen statt. Aus diesem Grund ist davon auszugehen, dass die Qualität der Implementierung sehr gut durch die beobachteten Trainingseinheiten beurteilt werden konnte. Dennoch ist nicht völlig auszuschließen, dass die Lehrer von den vorgegebenen Manualen abwichen. Dies trifft vor allem auf die LUG-Bedingung zu. Die Analyse der inhaltlichen Vollständigkeit anhand der Checklisten ergab, dass die USAT-Lehrer sehr gewissenhaft das Training umsetzten und dass die LUG-Lehrerinnen ebenfalls größtenteils die vorgesehenen Inhalte in den ersten drei Unterrichtseinheiten realisierten. Die Checklisten der USAT-Lehrer waren jedoch wesentlich detaillierter als die der LUG-Lehrerinnen. Diese enthielten nur die Auflistung der zu verwendenden Materialien, wobei jedoch die Art und Weise des Gebrauchs nicht explizit vorgeschrieben wurde. Insgesamt ist das LUG-Konzept als offen zu

charakterisieren, da die konkrete Umsetzung der Inhalte und Übungen den Lehrern selbst überlassen wurde. Die teilweise großen Klassenunterschiede innerhalb der LUG-Bedingung sind ebenfalls dadurch zu begründen. In einer nachfolgenden Studie müsste das LUG-Programm überarbeitet werden. Die Einfügung von konkreten Instruktionsprinzipien und motivationsfördernden Komponenten würde eine bessere Vergleichbarkeit der Umsetzung innerhalb der Bedingung gewährleisten. Dies ist jedoch nur in einem deutlich umfangreicheren Workshop möglich. Insbesondere beim Planen und Revidieren zeigten die LUG-Klassen stark divergierende Leistungen.

(4) *Kontrolle der außerhalb des Trainings ablaufenden Instruktionen:* Ferner wurde in der Studie II nicht kontrolliert, ob zwischen den einzelnen Trainingseinheiten sowie nach dem Posttest weitere Instruktionen, Wiederholungen und Übungen erfolgten. Auch wenn alle Lehrer darauf hingewiesen wurden, weitere Schreibaktivitäten zu vermeiden, kann eine konstante Instruktions- und Übungszeit nicht zweifelsfrei angenommen werden.

7.4.3 Zusammenfassung

Zusammenfassend lassen sich folgende Aussagen bezüglich der in der Fragestellung postulierten Hypothesen treffen:

(1) **Implementierbarkeit des USAT- und LUG-Programms durch Lehrer:** In beiden Bedingungen wurde eine hohe Implementationsgenauigkeit erreicht, so dass geschlussfolgert werden kann, dass Aufsatztrainings sowohl nach der Top-down- als auch nach der symbiotischen Strategie implementierbar sind.

(2) **Effektivitätsnachweis selbstregulatorischer Prozeduren:** Abgesehen von Geschichtenqualität und Wissen übertrafen die USAT-Klassen die LUG-Klassen signifikant in den erfassten Schreibmaßen. Folglich kann geschlussfolgert werden, dass das USAT-Programm mit selbstregulatorischen Komponenten bezüglich produktbezogener (Wortanzahl, Geschichtenelemente) und prozessbezogener Schreibmaße (Planen, Revidieren) besser abschnitt als ein traditionelles Aufsatzprogramm wie LUG, welches keine selbstregulatorischen Komponenten umfasste. Nichtsdestotrotz konnte der Bedingungseffekt nicht für alle abhängigen Variablen nachgewiesen werden.

(3) **Stabilität der Effekte:** Während für Geschichtenelemente eine Verschlechterung der Leistungen der USAT-Bedingung vom Posttest zum Follow-Up festgestellt wurde, veränderten sich bei allen anderen Variablen die Leistungen beim Follow-Up nicht, so dass die Effekte als stabil zu bewerten sind.

(4) **Zusätzliche Analysen:** Die Intra-Klassen-Korrelationen wiesen für keine der abhängigen Variablen einen bedeutsamen Zusammenhang der Klassen innerhalb einer

Bedingung nach. Auch Mixed-Model Analysen zeigten für keine der abhängigen Variablen eine Signifikanz des Effekts Klasse innerhalb Bedingung.

(5) **Generalisierbarkeit der Ergebnisse:** Insgesamt kann geschlussfolgert werden, dass die Ergebnisse für Subgruppen der Kontrollvariablen Geschlecht, Rechtschreibung, Sprachverständnis und Sprache generalisierbar sind. Lediglich bezüglich der Rechtschreibung und des Sprachverständnisses traten signifikante Interaktionen Bedingung x Kontrollvariable auf. Leistungsschwache Schüler profitierten dabei weniger vom USAT-Programm als leistungsstarke. Nichtsdestotrotz blieben die signifikanten Bedingungseffekte zu Gunsten der USAT-Bedingung erhalten.

8 Gesamtdiskussion

In der vorliegenden Arbeit wurde die Wirksamkeit des *Unterrichtsintegrierten Selbstregulatorischen Aufsatztrainings* (USAT) im Vergleich zu Schreibprogrammen ohne selbstregulatorische Komponenten überprüft. Die Implementationsforschung im Bereich Schreiben steckt erst in den Anfängen, so dass die Arbeit wichtige Hinweise für weitere Perspektiven bietet und Aussagen über die ökologische Validität von Interventionsstudien mit Kleingruppen zulässt. Im Folgenden werden zunächst die am Anfang postulierten Hypothesen und die Qualität der Implementationsstudien diskutiert. Anschließend werden Einschränkungen und zukünftige Forschungsanliegen erörtert. Am Ende werden praktische Implikationen herausgearbeitet.

8.1 Diskussion der Hypothesen

Insgesamt lassen sich die Hauptanliegen des Dissertationsprojekts auf zwei Punkte reduzieren. Erstens sollte ein unterrichtsintegriertes selbstregulatorisches Aufsatztraining entwickelt und hinsichtlich der Wirksamkeit anhand von prozess- und produktbezogenen Schreibmaßen überprüft werden. Zweitens intendierte die vorliegende Arbeit, den Nachweis zu erbringen, dass selbstregulatorische Komponenten für die Effektivität von Schreibprogrammen zwingend notwendig sind. Aus diesem Grund wurden dem USAT-Programm Kontrollgruppen gegenübergestellt, die keine Selbstregulation umfassten. Um die Ziele optimal zu erreichen und valide Aussagen zu treffen, wurden vier Untersuchungen durchgeführt:

- (1) **Lehrerbefragung:** Zu Beginn des Dissertationsprojekts wurden zwölf Lehrer hinsichtlich der Inhalte und des Umfangs des Aufsatzunterrichts in der vierten Klasse befragt.
- (2) **Pilotstudie:** Nach der Konzeption des USAT-Programms wurde dieses in einer vierten Klasse erprobt.
- (3) **Studie I:** Insgesamt vier Trainerinnen setzten das USAT-Programm und ein vergleichbares Schreibtraining ohne selbstregulatorische Komponenten (*Writer's Workshop*, WW) in jeweils vier Klassen um.
- (4) **Studie II:** Sechs Lehrer absolvierten einen mehrtägigen Workshop, in dem sie die Durchführung des USAT-Programms erlernten. Anschließend setzten die sechs Lehrer das USAT-Programm in ihrem eigenen Unterricht um. Gleichzeitig durchliefen fünf andere Lehrer einen Workshop, in dem sie basierend auf den

Schreibstrategien des USAT-Programms ein eigenes Aufsatztraining (*Lehrerkonzipierte Unterrichtseinheit zum Thema Geschichtschreiben*, LUG) entwickelten und anschließend in den eigenen Klassen umsetzten.

8.1.1 Implementierung in den Regelunterricht

Das in dem vorliegenden Dissertationsprojekt untersuchte USAT-Programm basiert auf dem *Selbstregulatorischen Aufsatztraining* (SAT) von Glaser (2005), welches wiederum auf dem SRSD-Ansatz von Harris und Graham (1996) aufbaut. Sowohl das SRSD- als auch das SAT-Programm erwiesen sich als sehr erfolgreich in der Förderung der Schreibkompetenz. Das SAT-Programm wurde jedoch bislang nur von geschulten Lerntrainern in Kleingruppen durchgeführt. Auch das SRSD-Programm wurde hauptsächlich in Kleingruppen mit lernschwachen Schülern umgesetzt. Die einzigen Implementationsstudien im Bereich Schreiben wurden von De La Paz und Graham (2002) sowie Torrance und Mitarbeitern (2007) vorgelegt. Das vorliegende Dissertationsprojekt adressierte somit bisher vernachlässigte Forschungsfragen, die jedoch einen erheblichen praktischen Wert besitzen, da erfolgreich implementierte, effektive Schreibtrainings einen Gewinn für Lehrer, die täglich mit den Schwierigkeiten des Unterrichtens konfrontiert werden, darstellen.

Der Eindruck fehlender Aufsatzprogramme in Deutschland bestätigt sich im Rahmen der durchgeführten Lehrerbefragung, bei der sich die Mehrzahl der befragten Lehrer für Fortbildungen in dem Bereich Aufsatzunterricht und Selbstregulation interessierten. Im Anschluss an die Interviewstudie wurde das USAT-Programm zunächst entwickelt und anschließend von Trainerinnen in einer vierten Klasse umgesetzt. Die Ergebnisse der Pilotstudie zeigten, dass die Schüler am Ende des Trainings vollständigere, stilistisch und qualitativ bessere sowie längere Aufsätze schrieben und bessere Planungsnotizen anfertigten, so dass die Implementierung durch Trainerinnen in der Pilotklasse gut realisiert wurde. In der Studie I und II konnten die Ergebnisse der Pilotstudie mit einer größeren Stichprobe von acht bzw. elf Klassen repliziert werden. Sowohl die Implementierung durch geschulte Lerntrainerinnen als auch durch Lehrer selbst konnte erfolgreich umgesetzt werden.

Bei der Implementierung der USAT- und WW-Programme in Studie I handelte es sich gemäß Gräsel und Parchmann (2004) um eine Top-down-Strategie, während das LUG-Programm in Studie II einer symbiotischen Implementierungsstrategie entstammte. Die Top-down-Strategie impliziert, dass die Ziele, Methoden und Inhalte von Externen erarbeitet und von Lehrern umgesetzt werden. Die Studien der vorliegenden Arbeit zeigen, dass die Top-down-Strategie zur Implementierung der USAT- und WW-Programme sehr erfolgreich war, da beide Programme zu Leistungssteigerungen führten und darüber hinaus sehr positiv von

den Lehrern aufgenommen und in den normalen Deutschunterricht integriert wurden. Dies spiegelt sich in der hohen sozialen Validität, die durch Evaluationsbögen am Ende der Workshops und der Trainings erfasst wurde, wider. Gleichzeitig erwies sich auch die symbiotische Strategie, die der LUG-Bedingung zu Grunde lag, als geeignet, Schreibtrainingskonzepte zu entwickeln und erfolgreich in den Deutschunterricht zu integrieren.

Überdies wurden in der Pilotstudie sowie in den Studien I und II vielfältige Maßnahmen zur Sicherung der Implementationsgenauigkeit eingesetzt. Dazu zählen Methoden zur Überprüfung der inhaltlichen Vollständigkeit und der Qualität der Implementierung. Die Ergebnisse zeigten, dass die Trainerinnen und Lehrer sämtliche wesentlichen Inhalte qualitativ gut umsetzten. Auch ergaben sich keine signifikanten Unterschiede zwischen den Trainerinnen und Bedingungen. Folglich können unterschiedliche Leistungen auf Unterschiede in den Bedingungen und nicht auf eine ungenaue oder schlechte Implementierung zurückgeführt werden.

8.1.2 Effektivitätsnachweis selbstregulatorischer Strategien

Der *Good Information Processor* (Pressley, Borkwski, & Schneider, 1989) verfügt nicht nur über ein großes Repertoire an Strategien und Wissen, sondern besitzt zudem metakognitives Wissen, um die Strategien motiviert einzusetzen und an seine eigenen Leistungen zu glauben. Aus diesem Grund wurden in dem vorliegenden USAT-Programm nicht nur Schreibstrategien vermittelt, sondern diese explizit mit selbstregulatorischen Strategien kombiniert. Es handelt sich dabei um Maßnahmen der Selbstüberwachung (Geschichtenplan), der Zielsetzung (7W- und Leistungsprotokoll) und der Selbstbewertung (Checkliste).

Auch wenn theoretisch von inkrementellen Effekten der Selbstregulation ausgegangen wird, sind diese noch nicht hinreichend wissenschaftlich nachgewiesen worden (De La Paz, 2007). Die Studien von Glaser und Kollegen (Glaser, 2005; Glaser & Brunstein, 2007a; Glaser & Brunstein, 2007b; Glaser et al., 2009), welche jedoch nur in Kleingruppen mit Lerntrainern durchgeführt wurden, sind bis jetzt die einzigen methodisch sorgfältig durchgeführten Untersuchungen, die belegten, dass ein selbstregulatorisches Schreibtraining generell bessere Leistungen hervorbringt als ein reines Schreibstrategietraining.

Deswegen wurden in der vorliegenden Arbeit dem USAT-Programm Kontrollbedingungen ohne selbstregulatorische Komponenten gegenübergestellt. Im Folgenden werden die Leistungsunterschiede zwischen der USAT-Bedingung und der jeweiligen Kontrollbedingung diskutiert.

Schreibbezogenes Wissen: Die Aneignung von deklarativem Wissen in einem bestimmten Bereich stellt eine notwendige Voraussetzung für die Entwicklung von bereichsspezifischer Kompetenz dar (Alexander, Graham, & Harris, 1998). In der Studie I verfügten nur die USAT-Klassen beim Posttest und beim Follow-Up über mehr schreibbezogenes Wissen als beim Prätest, so dass die Unterschiede zwischen den beiden Bedingungen signifikant waren. Dieses Ergebnis ist konsistent mit anderen SRSD-Studien, die belegen, dass selbstregulatorische Strategien die Aneignung von Strategiewissen erleichtern (Graham et al., 2005).

Dieser Bedingungseffekt konnte in der Studie II jedoch nicht repliziert werden. Die USAT- als auch die LUG-Klassen verfügten beim Posttest und beim Follow-Up über ähnlich viel schreibbezogenes Wissen. Gleichzeitig bestanden signifikante Unterschiede in der Rechtschreibleistung zwischen den beiden Bedingungen. Auch wenn der Einfluss der Rechtschreibung als Kovariate berücksichtigt wurde, besteht die Vermutung, dass den LUG-Klassen der Wissenserwerb generell leichter viel als den USAT-Klassen. Nichtsdestotrotz ist der fehlende Leistungsunterschied zwischen der USAT- und LUG-Bedingung nicht überraschend, da in beiden Bedingungen dieselben Schreibstrategien vermittelt wurden. In einem Wissenstest, der nur selbstregulatorische Fertigkeiten erfasst, würden vermutlich die USAT-Klassen die LUG-Klassen übertreffen.

Die Ergebnisse gehen jedoch einher mit denen von Graham, Harris und Mason (2005) sowie Glaser und Brunstein (2007b), welche ebenfalls bezüglich des deklarativen Wissens keine signifikanten Unterschiede zwischen den Treatment-Gruppen feststellten. Da jedoch die USAT-Klassen die LUG-Klassen bezüglich der prozessbezogenen Schreibmaße überragten, ist die Schlussfolgerung möglich, dass nur Schüler der USAT-Klassen ihr Wissen erfolgreich beim Planen und Schreiben einsetzen konnten. Folglich ermöglichten selbstregulatorische Prozeduren den Transfer von Wissen in die Produktion eines Textes (Glaser & Brunstein, 2007b). Darüber hinaus zeigten Untersuchungen im Bereich Sport von Cleary und Zimmerman (2001) und Kitsantas und Zimmerman (2002), dass Experten und Nicht-Experten sich nicht hinsichtlich des bereichsspezifischen Wissens unterscheiden, sondern nur bezüglich des unterschiedlichen Gebrauchs von selbstregulatorischen Methoden.

Ferner wurden im LUG-Workshop einige Schreibstrategien von den Lehrerinnen überarbeitet. Zum Beispiel wurde das *Ziel* durch das *Thema* ersetzt. Folglich ist es möglich, dass die von den LUG-Lehrerinnen überarbeiteten Schreibstrategien nicht adäquat mit dem Wissenstest erfasst wurden. In nachfolgenden Studien sollte deswegen ein neuer Wissenstest

konzipiert und evaluiert werden, der aus mehreren Items besteht und eine genauere Messung des schreibbezogenen Wissens sowie des Selbstregulationswissens ermöglicht.

Planen und Revisionen: Schreibanfänger zeigen im Gegensatz zu Schreibexperten keine Planungs- und Überarbeitungsaktivitäten und weisen erhebliche Defizite bei der Überwachung und Steuerung der am Schreiben beteiligten Prozesse auf (Graham & Harris, 2000; Graham & Harris, 2003; Scardamalia et al., 1984). Dies konnte in der Pilotstudie sowie in den Studien I und II bestätigt werden, da beim Prätest alle Schüler kaum planten und überarbeiteten. Im Gegensatz dazu entwickelten sich die Planungs- und Überarbeitungsfertigkeiten im Verlauf der Trainings sehr unterschiedlich in den beiden Bedingungen, wobei sich für die Studie I und II ähnliche Ergebnisse ergaben.

Nur Klassen, die das USAT-Programm absolvierten, fertigten im Anschluss an das Training und sechs Wochen später qualitativ bessere Planungsnotizen an. Weder dem WW- noch dem LUG-Programm gelang es, eine Veränderung der Planungsaktivitäten bei den Schülern hervorzurufen, so dass sich beim Posttest und beim Follow-Up erhebliche Unterschiede zwischen den Bedingungen ergaben. Folglich bewirkte die Vermittlung von selbstregulatorischen Strategien einen direkten Anstieg der Planungsfertigkeiten, welche für hohe Schreibleistungen notwendig sind, da sie das Arbeitsgedächtnis entlasten. Die berichteten Ergebnisse bestätigen damit die Befunde von vorangegangenen Studien (Englert et al., 2009; Kellogg, 2008).

Im Gegensatz zum Planen ergaben sich in den beiden Hauptstudien etwas unterschiedliche Ergebnisse für Revisionen. Das Überarbeiten ist eine kognitiv sehr anspruchsvolle Tätigkeit, da der produzierte Text hinsichtlich der eigenen Ziele und der Erwartungen der Leserschaft evaluiert und korrigiert werden muss (Fitzgerald, 1987). Schreibnovizen zeigen in der Regel nur Revisionen, die die Oberflächenstruktur (Schönschrift, Rechtschreibung, Grammatik) betreffen (Graham & Harris, 2003). Im Gegensatz zur Pilotstudie und zur Studie I konnte in der Studie II nach Durchführung einer Messwiederholungsanalyse ein Unterschied zwischen der USAT- und der LUG-Bedingung festgestellt werden.

Insgesamt lässt sich feststellen, dass alle Schüler sehr wenig überarbeiteten, was darauf zurückzuführen ist, dass die Überarbeitungsfertigkeiten erst im Jugendalter voll ausgebildet werden (Feilke, 1993). Ein explizites Revisionstraining mit älteren Schülern (sechste Klasse) erbrachte dementsprechend größere Effekte (Glaser & Brunstein, 2007a). Eine weitere Möglichkeit zur Steigerung der Revisionsfertigkeiten wäre die Setzung konkreter Revisionsziele (z.B. *„Heute möchte ich mindestens drei Überarbeitungen*

machen.“). Nichtsdestotrotz belegen die Ergebnisse der Studie II, dass selbstregulatorische Strategien die Anwendung von Revisionsstrategien erleichterten.

Schreibleistung: Die Schreibleistung wurde anhand der Wortanzahl, eines strategienahen Maßes (Geschichtenelemente) und eines strategiefernen Maßes (Geschichtenqualität) bewertet.

Hinsichtlich der Geschichtenelemente waren die Ergebnisse der Studie I und II identisch: Sowohl beim Posttest als auch beim Follow-Up zeigte sich ein hoch signifikanter Bedingungseffekt zu Gunsten der USAT-Bedingung.

In der Studie I schrieben die USAT-Klassen im Anschluss an das Training und sechs Wochen später ebenfalls qualitativ bessere Aufsätze als die WW-Klassen. Diese Effekte konnten in der Studie II weder beim Posttest noch beim Follow-Up repliziert werden. Ein umgekehrtes Bild ergab sich für die Wortanzahl. Nur in der Studie II verfassten die USAT-Klassen signifikant längere Texte als die LUG-Klassen. In der Studie I konnte kein signifikanter Unterschied zwischen der USAT- und der WW-Bedingung nachgewiesen werden.

Auch wenn Geschichtenelemente und Geschichtenqualität sowohl in der Studie I als auch in der Studie II signifikant miteinander korrelierten, fielen die Korrelationsquotienten in der Studie II deutlich geringer aus. In der Studie II wurde die Sprache und die Struktur getrennt für die Geschichtenqualität bewertet, während dies in der Studie I zusammen erfolgte. Das veränderte Vorgehen bei der Auswertung könnte folglich ein Grund für die ausbleibenden Effekte in der Studie II sein. Wahrscheinlicher ist es jedoch, dass den LUG-Klassen die Aneignung und Umsetzung von Schreibstrategien leichter fielen als den USAT-Klassen. Aus diesem Grund sollten in nachfolgenden Studien weitere Kontrollvariablen wie z.B. die Intelligenz erhoben werden. Intelligenz korreliert mit der Verarbeitungsgeschwindigkeit (Klauer, 2006), so dass die Möglichkeit besteht, dass die LUG-Klassen aufgrund von höheren Intelligenzwerten die erlernten kognitiven Strategien besser erlernen und ausführen konnten.

Transferleistung: Der Transfer der erlernten Schreibstrategien auf ein untrainiertes Textgenre wie Erlebniserzählungen wurde nur in der Studie I untersucht. Wie auch die Bildergeschichten wurden die Erlebniserzählungen ebenfalls hinsichtlich der Geschichtenelemente und der Geschichtenqualität ausgewertet. Es ergaben sich für beide Schreibleistungsmaße signifikante Unterschiede zwischen der USAT- und der WW-Bedingung beim Posttest und beim Follow-Up. Folglich gelang es den USAT-Klassen wesentlich besser, die Schreibstrategien auf das ungeübte Textgenre Erlebniserzählungen zu

transferieren. Damit bestätigt die Studie I bereits vorangegangene Studien, die den Nachweis erbrachten, dass selbstregulatorische Strategien den Transfer von Schreibstrategien auf ungeübte Aufgaben vereinfachen (Glaser, 2005; Glaser et al., 2009; Sawyer et al., 1992).

Effektstärken: Nach Cohen's d (Cohen, 1988) wurden Effektstärken sowohl auf Klassen- als auch auf Schülerebene berechnet. Da die Analyseeinheit Klasse die Höhe der Standardabweichungen bei der Effektstärkenberechnung unterschätzte, fielen die Effektstärken sehr groß aus, was die Interpretierbarkeit erschwerte. Um die Vergleichbarkeit mit anderen Studien zu ermöglichen, wurden die Effektstärken auch auf Schülerebene berichtet.

Abgesehen von Wortanzahl (Posttest) und Revisionen (Posttest, Follow-Up) zeigten sich für alle abhängigen Variablen der Studie I beim Posttest und beim Follow-Up substantielle Effektstärken zwischen .70 und 2.01. Für Revisionen ergab sich beim Posttest ein mittlerer Effekt (.50), wobei dieser beim Follow-Up (.09) verschwand. Auch hinsichtlich der Wortanzahl zeigte sich beim Posttest ein kleiner Effekt (.36).

In Studie II führten dagegen die starken Verbesserungen der LUG-Klassen hinsichtlich Wissen und Geschichtenqualität nur zu kleinen Effekten (zwischen .20 und .36) beim Posttest und beim Follow-Up. Auch bezüglich Wortanzahl und Revisionen sind die Effektstärken beim Posttest (.68 und .28) und beim Follow-Up (.79 und .52) als klein bis mittel groß zu bezeichnen. Für Geschichtenelemente und Planen konnten jedoch die Effektstärken der Studie I repliziert werden, da hier große Effekte zwischen 1.47 und 1.87 errechnet wurden.

Abschließende Bewertung der Effektivität selbstregulatorischer Strategien: In der Gesamtheit lassen die Ergebnisse der Studie I und II den Schluss zu, dass das USAT-Programm deutlich den Schreibtrainings ohne Selbstregulation überlegen war. Dies traf insbesondere auf Geschichtenelemente und auf die prozessbezogenen Variablen wie Planen und Revisionen zu. Gleichzeitig bleibt jedoch die Frage offen, welche der eingesetzten selbstregulatorischen Methoden speziell für die Effekte verantwortlich waren.

Schließlich belegen die Ergebnisse der Pilotstudie, der Studie I und II zudem die ökologische Validität von Kleingruppentrainings, da die in Kleingruppentrainings erzielten Effekte im gesamten Klassenkontext ebenfalls nachgewiesen wurden. Insgesamt ähneln die von Glaser und Kollegen (Glaser, 2005; Glaser & Brunstein, 2007a; Glaser & Brunstein, 2007b; Glaser et al., 2009) auf Schülerebene ermittelten Effektstärken denen der Studie I und II, welche ebenfalls auf Schülerebene gerechnet wurden.

8.2 Anmerkungen zur Qualität der Implementierungen

Zur Beurteilung der Qualität der Implementierung der eingesetzten Trainingsprogramme müssen die interne Validität, externe Validität, soziale Validität sowie die Implementationsgenauigkeit sichergestellt werden.

Interne Validität: Gemäß Campbell (1969) ist zur Sicherung der internen Validität eine randomisierte Zuweisung der Versuchspersonen zu den Bedingungen und eine geeignete Kontrollgruppe notwendig. Bei den Studien der vorliegenden Arbeit handelt es sich jedoch um Versuchsdesigns mit natürlich vorkommenden Stichproben, so dass eine randomisierte Zuteilung der Schüler zu den Bedingungen nicht möglich war.

In der Studie I wurden jedoch die Klassen zufällig den beiden Bedingungen zugeordnet, wobei die Zuteilung gematcht nach Rechtschreibleistungen und Klassengröße erfolgte. Infolgedessen waren die Rechtschreibleistungen und die Klassengrößen in der USAT- und der WW-Bedingung gleich verteilt. Auch hinsichtlich der übrigen Kontrollvariablen (Nationalität, Muttersprache Deutsch) zeigten sich keine signifikanten Unterschiede zwischen den Bedingungen. Da jedoch eine Prätest-Variable (Planen Struktur) bedeutsam zwischen der USAT- und der WW-Bedingung variierte, wurden die Daten kovarianzanalytisch mit den jeweiligen Prätest-Werten als Kovariate ausgewertet.

Im Gegensatz dazu war in der Studie II die randomisierte Zuteilung der Klassen zu den Bedingungen nicht möglich, da die Lehrerrekutierung speziell auf die USAT- und die LUG-Bedingung ausgerichtet war. Dies hatte zur Folge, dass signifikante Unterschiede bezüglich des Alters, in den Rechtschreibleistungen und in dem Prätest Revisionen zwischen den beiden Bedingungen vorlagen. Bezüglich der anderen Kontrollvariablen (Sprachverständnis, Klassengröße, Geschlecht, Nationalität, Muttersprache Deutsch) und den übrigen Prätest-Werten ergaben sich jedoch keine Bedingungseffekte. In sämtlichen statistischen Analysen wurden die Rechtschreibleistung und der Prätest als Kovariaten berücksichtigt. Da Lehrer der USAT-Bedingung ein komplett fertiges Trainingsprogramm erlernten und Lehrer der LUG-Bedingung ein eigenes Trainingsprogramm erarbeiten mussten, war es notwendig, die Interessen der Lehrer bei der Zuweisung zu den Bedingungen zu berücksichtigen, damit diese motiviert die jeweiligen Trainingsprogramme umsetzten. Des Weiteren unterschieden sich die Lehrer der Studie II nicht hinsichtlich der Berufserfahrungen. Insgesamt kann sowohl in der Studie I als auch in der Studie II von einer Vergleichbarkeit der jeweiligen Bedingungen hinsichtlich der untersuchten Klassen und Lehrer ausgegangen werden. Nichtsdestotrotz sollten in nachfolgenden Studien mehr Kontrollvariablen wie z.B.

die Intelligenz erfasst werden, wenn eine randomisierte Zuteilung der Klassen zu den Bedingungen nicht möglich ist.

Hinsichtlich des zweiten Kriteriums für interne Validität, der adäquaten Kontrollgruppe, sind die WW-Bedingung der Studie I sowie die LUG-Bedingung der Studie II als starke und für die Fragestellung geeignete Kontrollgruppen zu bewerten, so dass die von Pressley und Harris (1994) geforderten glaubwürdigen Kontrollgruppen gewährleistet wurden.

In der Studie I lag der Fokus auf dem Nachweis der Wirksamkeit des USAT-Programms im Vergleich zu einem Aufsatztraining ohne selbstregulatorische Komponenten. Abgesehen von der Selbstregulation waren alle Schreibstrategien und Instruktionsprinzipien in der USAT- und in der WW-Bedingung identisch. Dagegen fokussierte die Studie II einen Vergleich des USAT-Programms mit dem regulären Aufsatzunterricht. Zur besseren Vergleichbarkeit wurden den LUG-Lehrerinnen ebenfalls einige Vorgaben (Schreibstrategien, Dauer und Umfang des Trainings, Rückmeldung, Motivation) gegeben, wobei jedoch keine selbstregulatorischen Komponenten in das LUG-Trainingskonzept mit einfließen.

Insgesamt ist die Kontrolle von Störvariablen bei Implementationsstudien weitaus schwieriger als in anderen experimentellen Studien. Vor allem die Umsetzung des USAT-Programms durch Lehrer kann eine Vielzahl von Störvariablen produzieren. Dennoch wurde die interne Validität in der Studie I und II durch vielfältige Maßnahmen (siehe unten, Implementationsgenauigkeit) gewährleistet, so dass davon auszugehen ist, dass die Unterschiede in den abhängigen Variablen zwischen den Bedingungen auf Unterschiede in der Intervention, speziell der Selbstregulation, zurückzuführen sind.

Externe Validität: Damit langfristig eine flächendeckende Verbreitung des USAT-Programms in den deutschen Grundschulen möglich ist, sollte sich das USAT-Programm in mehreren Schulen, bei unterschiedlichen Schülergruppen und in unterschiedlichen Orten als effektiv erweisen (O'Donnell, 2008). Demnach muss eine Generalisierbarkeit der Ergebnisse durch eine hohe externe Validität der Untersuchung gewährleistet werden.

Aufgrund der Durchführung einer Pilotstudie und zweier Hauptstudien nahmen in der vorliegenden Arbeit insgesamt 20 Klassen aus zwölf verschiedenen Schulen aus den Landkreisen Gießen und Vogelsberg teil. Darüber hinaus wurden die Trainingsprogramme sowohl von Trainern als auch von Lehrern umgesetzt. In allen Studien wurden die Trainings in die natürliche Klassenumgebung integriert, so dass die Rahmenbedingungen weitestgehend mit den realen Bedingungen des Schulalltags identisch waren. In der zweiten Studie setzte eine repräsentative Stichprobe von elf Lehrern in ihren eigenen Klassen das jeweilige

Trainingsprogramm erfolgreich um, so dass davon auszugehen ist, dass andere Lehrer aus anderen Schulen nach Absolvierung eines Workshops ebenfalls erfolgreich die Trainingsprogramme implementieren können.

Nichtsdestotrotz sollte in nachfolgenden Studien die Implementierbarkeit des USAT-Programms anhand einer größeren Lehrer- und Schülerstichprobe überprüft werden, die hinsichtlich des Alters, Kompetenzniveaus, kulturellen Hintergrunds und des Geschlechts heterogen ist. Die Lehrer, die an der Studie II teilnahmen, verfügten bereits über mehrjährige Erfahrungen im Unterrichten. Zudem waren, abgesehen von einer Person, alle Lehrer weiblich. Auch die Untersuchung von Schülern anderer Jahrgangsstufen ist wünschenswert, da die Lehrer in der Lehrerbefragung angaben, vornehmlich in der dritten Klasse das Aufsatzschreiben zu behandeln. Bei den untersuchten Schülern handelte es sich in der Regel um durchschnittlich begabte Viertklässler. Auch wenn in den beteiligten Klassen auch lernbehinderte Schüler und Schüler mit Lese- und Rechtschreibschwäche vorkamen, wurden diese als solche nicht explizit identifiziert. Für einzelne Subgruppen (Mädchen/Jungen, rechtschreibschwach/rechtschreibstark, sprachverständnisschwach/sprachverständnisstark, Muttersprache Deutsch/nicht-Deutsch) blieben jedoch die Effekte weiterhin bestehen.

Soziale Validität: Abgesehen von der internen und externen Validität sollte die soziale Validität einer Intervention ebenfalls berücksichtigt werden (Deshler & Schumaker, 1994). Deshalb wurden im Anschluss an die Studie I sowie II die beteiligten Lehrer nach der Meinung zu den jeweiligen Trainings gefragt. Insgesamt bewerteten sämtliche Lehrer in beiden Studien die jeweiligen Programme mit sehr gut bis befriedend. Folglich kann von einer hohen Akzeptanz der Trainings ausgegangen werden. Auf diese Weise kann zudem ausgeschlossen werden, dass die Lehrer zwischen den Trainingseinheiten oder nach dem Posttest, als keine Trainer oder Beobachter anwesend waren, negative Einstellungen bezüglich der Trainings auf die Schüler projizierten.

Gleichzeitig ist einschränkend hervorzuheben, dass die soziale Validität nur auf Lehrerebene und nicht auf Schülerebene erhoben wurde. Diese ist jedoch ebenfalls als wesentlich anzusehen, da selbstregulatorische Prozesse die Selbstwirksamkeitserwartung der Schüler positiv beeinflussen (Zimmerman & Risemberg, 1997). In der LUG-Bedingung der Studie II befragten die Lehrerinnen die Schüler am Ende des Trainings nach ihren Meinungen. Generell fanden die Schüler das Training interessant und abwechslungsreich. Die USAT-Lehrer berichteten ebenfalls, dass die Schüler sich sehr positiv über das Training geäußert hatten.

Implementationsgenauigkeit: Unter Implementationsgenauigkeit wird das Ausmaß, in dem die Interventionen wie geplant umgesetzt wurden, verstanden. Nur bei einer Sicherstellung der Implementationsgenauigkeit können die Unterschiede in den abhängigen Variablen durch Unterschiede in den unabhängigen Variablen erklärt werden. In diesem Sinne stellt die Implementationsgenauigkeit eine weitere Maßnahme zur Gewährleistung der internen Validität dar.

Sowohl in der Pilotstudie als auch in den Studien I und II wurden vielfältige Instrumente eingesetzt. Die inhaltliche Vollständigkeit wurde anhand von Checklisten mit den wichtigsten Inhalten überprüft, die entweder von den Trainerinnen bzw. den Lehrern oder von den Beobachtern ausgefüllt wurden. In allen Studien wurden die wesentlichen Inhalte und Strategien vermittelt.

Des Weiteren wurden zur Überprüfung der Qualität der Implementierung in allen drei Untersuchungen die Durchführungen der Trainings von geschulten universitären Mitarbeitern unter Verwendung standardisierter Bögen beobachtet. Als Kriterien wurden die Instruktionsqualität, die Motivierung, die Individualisierung und die Klassenführung herangezogen. Insgesamt kann von einer sehr hohen Qualität der Implementierung durch Trainerinnen und Lehrer ausgegangen werden. Kritisch muss jedoch angemerkt werden, dass die Anwesenheit von Beobachtern in den Klassen das Lehrer- und Schülerverhalten beeinflusst haben könnte. Alternativ dazu wäre die Durchführung von Videoanalysen möglich.

8.3 Grenzen und Perspektiven

Mit dem in der vorliegenden Arbeit untersuchten USAT-Programm wurde erstmalig ein Schreibtraining in den regulären Deutschunterricht implementiert. Die berichteten Ergebnisse belegen die Wirksamkeit von implementierten Schreibprogrammen und stellen zudem Ansatzpunkte für weitere Forschungsfragen dar. Gleichzeitig lassen sich einige Grenzen der dargestellten Studien identifizieren.

Langzeiteffekte: Graham und Harris (2000) postulierten, dass mit zunehmendem Alter und Beschulung selbstregulatorische Fähigkeiten zunehmen, so dass bereits erlernte Methoden ineffizient werden und erweitert werden müssen. In der vorliegenden Studie erwiesen sich die erlernten selbstregulatorischen Strategien als wirkungsvoll. Es sind jedoch noch weitere Untersuchungen notwendig, um zu überprüfen, ob die vermittelten Strategien der Selbstüberwachung, Selbstbewertung und Zielsetzung in späteren Schuljahren ebenfalls hilfreich sind und von den Schülern angewandt werden. Zudem wäre es lohnenswert zu

testen, ob es den Schülern gelingt, die selbstregulatorischen Prinzipien in neuen Situationen anzuwenden. Die Autoren gehen davon aus, dass Schüler mehr selbstregulatorische Methoden benutzen, je älter sie werden. Interessant wäre die Untersuchung der Frage, ob die Schüler der USAT-Bedingung einen Entwicklungsvorsprung gegenüber Schülern, die ein Trainingsprogramm ohne Selbstregulation absolviert haben, besitzen. Insgesamt kann erst durch umfangreiche Langzeitstudien mit mehreren Messzeitpunkten ein Entwicklungsmodell der Schreibkompetenz erstellt werden, welches als Grundlage für die Konzeption von effektiven selbstregulatorischen Aufsatztrainings fungieren kann.

Erhebung von Transfereffekten: Mit dem vorangegangenen Punkt geht die Frage nach dem Transfer der erlernten Strategien auf ungeübte Aufgaben einher. In der Studie I konnte bereits der Nachweis erbracht werden, dass selbstregulatorische Komponenten den Transfer der Schreibstrategien auf ein untrainiertes Textgenre (Erlebniserzählung) erleichtern. Neben einer Replizierbarkeit des Befundes sollte auch der Transfer auf andere Bereiche wie Lese- und Hörverständnis untersucht werden. In den Untersuchungen von Olive (2002) wurde festgestellt, dass das Lesen als Vorläuferfertigkeit für das Schreiben angesehen werden kann. Des Weiteren scheinen Hören, Lesen und Schreiben zumindest teilweise auf ähnliche mentale Prozesse und Wissenssysteme zu basieren (Fitzgerald & Shanahan, 2000). So können Probleme im Hörverständnis, welches sich bereits früh in der kindlichen Entwicklung herausbildet, mit späteren Schreibschwierigkeiten einhergehen (Shanahan, 2006). Ein verbindendes Glied könnte der Gebrauch und das Verständnis des Wortschatzes sein. Der Nachweis bedeutsamer Zusammenhänge zwischen dem Hör- und Leseverständnis sowie der Schreibkompetenz würde die Grundlage für die Entwicklung eines grundlegenden Modells des Schriftspracherwerbs schaffen.

Einzelfalluntersuchungen: Neben (quasi-)experimentellen Versuchsdesigns stellen Einzelfalluntersuchungen eine weitere Option zur Überprüfung der Wirksamkeit von Interventionsmaßnahmen dar (Rogers & Graham, 2008; Schumaker & Deshler, 2003). In Einzelfalluntersuchungen durchläuft nur eine geringe Anzahl von Personen zeitversetzt das Training. Zusätzlich zu einer Leistungserhebung vor und nach dem Training finden während des gesamten Trainings kontinuierliche Messungen statt, so dass die Leistung eines Schülers, der bereits das Training begonnen hat, mit der Leistung eines Schülers, der noch wartet, verglichen wird (*Multiple Baseline Design*). Einzelfalluntersuchungen können somit dann durchgeführt werden, wenn alle Schüler einer Klasse aus schulpolitischen oder ethischen Gründen einer Intervention ausgesetzt werden müssen. Durch eine genaue Beschreibung der Stichprobe und der Rahmenbedingungen können zudem Aussagen über die

Generalisierbarkeit getroffen werden. Des Weiteren erfassen Einzelfalluntersuchungen detailliert die Veränderungen bei individuellen Schülern während der Intervention. Insgesamt ermöglichen Einzelfalluntersuchungen wissenschaftlich fundierte und methodisch sorgfältig erarbeitete Aussagen über den Einfluss einer Maßnahme auf beobachtete Leistungsveränderungen, wenn nur kleine Stichproben verfügbar sind. Dies ist vor allem bei speziellen und seltenen Schülergruppen der Fall. Auch bei lerngestörten Schülern oder Schülern mit Teilleistungsschwächen ist es schwierig, eine große Stichprobe zu gewinnen, die aus Schülern mit denselben Symptomen besteht. Um somit Aussagen über die Effektivität des USAT-Programms für leistungsschwache Schüler treffen zu können, sollten demnach Einzelfalluntersuchungen durchgeführt werden.

Prüfung des Einflusses von Mediatoren: Darüber hinaus besteht die Möglichkeit, dass die Effekte des USAT-Programms auf die Schreibleistung durch andere abhängige Variablen wie z.B. die Steigerung der Planungsfertigkeiten vermittelt wurden. Solche gerichteten Hypothesen lassen sich jedoch nur durch Pfadanalysen überprüfen, wozu größere Stichproben mit mehreren Klassen notwendig sind. Des Weiteren sollte in weiterführenden Studien der Einfluss von motivationalen Faktoren auf die Schreibleistung fokussiert werden (Hidi & Boscolo, 2006). Insbesondere die Selbstwirksamkeitserwartung sollte erfasst und als potenzielle Mediatoren überprüft werden, da sich in Pfadanalysen von Zimmerman und Kollegen direkte sowie indirekte Effekte von Selbstwirksamkeit über Selbstregulation auf die Schreibleistung zeigten (Zimmerman et al., 1992; Zimmerman & Bandura, 1994). Dazu muss jedoch zunächst ein adäquates Instrument zur Messung der Selbstwirksamkeitserwartung bei Schreibanfängern konstruiert werden. Dies gestaltet sich deswegen schwierig, weil Schreibnovizen generell ihre Schreibleistungen überschätzen und gleichzeitig wenig Feedback über ihre tatsächlichen Schreibkompetenzen besitzen, so dass die Ausbildung einer adäquaten Selbstwirksamkeitserwartung erst in den Anfängen steckt (Graham & Harris, 2003; Pajares, 2003). Um jedoch ein Wirkungsmodell zu testen, welches den in den Studien berichteten Effekten zugrunde liegt, ist die Selbstwirksamkeitserwartung eine wesentliche Einflussgröße. Somit sind weitere Studien mit größeren Stichproben und weiteren abhängigen Variablen erforderlich.

Erfassung selbstregulatorischer Prozesse: In der vorliegenden Arbeit wurden selbstregulatorische Strategien durch die Erhebung der Planungs- und Überarbeitungsfertigkeiten, die wiederum in Verbindung mit der Schreibleistung stehen, erfasst. Gleichzeitig sind weitere qualitative (z.B. Denkprotokolle, Interviews, Beobachtungen) und quantitative (z.B. Fragebögen, Verhaltensspuren) Methoden denkbar.

Vor allem am Anfang der Schreibforschung wurden häufig Protokolle des „Lauten Denkens“ herangezogen, die jedoch heute bezüglich vielfältiger Aspekte als problematisch anzusehen sind. Zum einen kann die Textproduktion durch das laute Formulieren der Gedanken gestört werden und zum anderen sind viele Prozesse dem Schreiber nicht bewusst oder kommen ihm zu einfach vor. Auch Interviews und Beobachtungen sind für die Erfassung ungünstig, da kognitive und metakognitive Prozesse häufig unbewusst und verdeckt ablaufen. Ein weiteres Problem mit Interviews sowie mit Fragebögen liegt zudem darin, dass die verbale Ausdrucksfähigkeit die Ergebnisse erheblich verzerrt. Demgegenüber stellen Verhaltensspuren eine wichtige Alternative zur Analyse der beim Schreiben ablaufenden selbstregulatorischen Prozesse dar (Winne & Perry, 2000). Anhand einer computerbasierten Aufzeichnung der Planungs-, Übersetzungs- und Überarbeitungshandlungen kann im Nachhinein die Entwicklung des Textprodukts nachvollzogen werden. Auf diese Weise könnte detaillierter überprüft werden, ob und in welcher Form das USAT-Programm Auswirkungen auf die Schreibprozesse ausübte. Des Weiteren können die Prozessdaten als weitere Mediatoren in Pfadanalysen berücksichtigt werden (Glaser & Brunstein, 2007a).

Kombination des USAT-Programms mit kooperativen Lernformen: In der Pilotstudie und in der Studie I wurden Partner- und Gruppenarbeiten zur Wiederholung und Übung der Schreibstrategien und der selbstregulatorischen Elemente eingesetzt. Es wurde erwartet, dass dadurch die Motivation der Schüler gesteigert und schwache Schüler durch starke Schüler unterstützt werden, was eine Entlastung der Lehrer ermöglichen sollte. In der Pilotstudie zeigte sich, dass den Schülern das Prinzip des kooperativen Lernens zunächst näher gebracht und eingeübt werden musste, bevor diese effektiv voneinander profitierten. Der dadurch entstandene Zeitaufwand konnte nicht durch den Output der Gruppenarbeiten gerechtfertigt werden, so dass in der Studie I die Gruppenarbeit durch Partnerarbeit ersetzt wurde. In der Studie II entfiel diese komplett, da den Lehrern ausreichend Zeit für die Instruktion und den Schülern für das Verfassen der Geschichten eingeräumt werden sollte.

Nichtsdestotrotz äußerten sich sowohl die Schüler und insbesondere die Lehrerin der Pilotstudie äußerst positiv über die Gruppenarbeiten. Auch Graham und Kollegen belegten bezüglich des Transfers auf untrainierte Textgenres und planungsspezifischen Wissens die Überlegenheit eines SRSD-Programms mit Partnerarbeit gegenüber einem reinen SRSD-Programm und einem *Writer's Workshop* (Graham et al., 2005). Die Konstanthaltung der Übungs- und Instruktionszeiten muss dabei in allen Bedingungen gewährleistet werden, so dass bei der Konzeption der Kontrollgruppe ebenfalls zusätzliche Elemente integriert werden müssen. Insgesamt ist davon auszugehen, dass ein USAT- plus Gruppenarbeit-Programm

mehr als sechs Wochen in Anspruch nehmen würde. In den 13 wissenschaftlich fundierten Empfehlungen für den Schreibunterricht wird eine durchschnittliche Effektstärke von .75 für kollaborative Lernumgebungen angegeben (Graham & Harris, 2009; Graham & Perin, 2007b; Graham et al., 2009). Somit sind positive Effekte des USAT-Programms durch die Hinzufügung der Partnerarbeit zu erwarten.

Erfassung der Schreibstrategien und der selbstregulatorischen Methoden im Deutschunterricht: Die Kooperation und Miteinbeziehung von Lehrern ist in der Implementationsforschung ein entscheidender Faktor für die Effektivität und Akzeptanz von neuen Unterrichtsmethoden. Aus diesem Grund wurde bereits von Beginn der vorliegenden Arbeit an intensiv mit Lehrern zusammengearbeitet. Die Lehrerbefragung zielte darauf ab, die aktuellen Methoden im Aufsatzunterricht zu erörtern. Auch die Studie II intendierte die Entwicklung eines Unterrichtskonzepts, welches die traditionellen Methoden zur Instruktion von Schreibstrategien widerspiegelt. In beiden Untersuchungen zeigte sich, dass die Lehrer über wenig selbstregulatorisches Wissen und Fertigkeiten verfügten. Dennoch kristallisierten sich nach genauerem Nachfragen in der Lehrerbefragung Methoden (z.B. Wochenplan) heraus, die denen der Selbstregulation ähnlich sind. Demnach gaben 33% der Lehrer an, solche Methoden in ihrem Unterricht zu benutzen. Es ist somit zu vermuten, dass einige selbstregulatorische Elemente bereits im Deutschunterricht integriert sind, dass diese jedoch nicht als solche von den Lehrern wahrgenommen werden. Um folglich ein genaueres Bild über die aktuelle Situation im Deutschunterricht zu gewinnen, sind weitere Studien notwendig. In strukturierten Interviews sollten die Lehrer über ihr Wissen bezüglich der Selbstregulation befragt werden. Ferner sollten die Einstellungen und Interessen in diesem Bereich beleuchtet werden. Zusätzlich sind jedoch Unterrichtsbeobachtungen erforderlich, um präzise Aussagen über die im Deutschunterricht tatsächlich verwendeten Schreibstrategien und selbstregulatorischen Methoden zu treffen. Es sollte dabei auf eine große und repräsentative Lehrer- und Schülerstichprobe zurückgegriffen werden. Die national repräsentative Befragung von Grundschullehrern von Cutler und Graham (2008) könnte dabei als Orientierung dienen.

Transfer der erlernten Unterrichtsmethoden auf andere Fächer: In der Studie II erlernten Lehrer die Instruktion selbstregulatorischer Methoden in Kombination mit Schreibstrategien für das Verfassen von Bildergeschichten. Nichtsdestotrotz können die selbstregulatorischen Methoden auch mit anderen Textgenres und in anderen schulischen Bereichen wie Lesen und Rechnen eingesetzt werden. In zukünftigen Studien sollte deswegen untersucht werden, inwieweit der USAT-Workshop die Lehrer befähigt, die erlernten

selbstregulatorischen Strategien auch auf andere Themen und Fächer zu transferieren. Erst wenn das Prinzip der Selbstregulation von Lehrern verinnerlicht wurde und unter Lehrern weitläufig akzeptiert und geschätzt wird, kann Selbstregulation ein integraler Bestandteil des Unterrichts werden.

8.4 Praktische Implikationen

Aufgrund der positiven und beeindruckenden Effekte, die das USAT-Programm auf die Schreibleistungen und schreibbezogenen Leistungen (Planen, Überarbeiten, Wissen) ausübte, ist eine Verbreitung der Intervention in den Grundschulen erstrebenswert, um langfristige und stabile Verbesserungen der Schreibkompetenz bei den Schülern zu erzeugen. Die sich aus dem vorliegenden Dissertationsprojekt ergebenden Implikationen werden nachfolgend kurz skizziert.

Betreuung der Lehrer während der Durchführung des Trainings: In der Studie II führten Lehrer das USAT-Programm in ihren eigenen Klassen durch. Gleichzeitig wurden sie von universitären Mitarbeitern unterstützt. Zum einen korrigierten diese jede Woche die Schüleraufsätze und zum anderen standen sie den Lehrern als Ansprechpersonen zur Verfügung. Die Erfahrungen der Studie II zeigten, dass die USAT-Lehrer stark auf diese Hilfen in den ersten drei Trainingseinheiten zurückgriffen. Sobald keine Beobachter mehr im Unterricht anwesend waren, fühlten sich die Lehrer selbstsicherer und gewannen sehr schnell Routine in der Durchführung der Trainings. Nichtsdestotrotz ergibt sich die Frage, inwieweit die Lehrer ohne diese Hilfestellungen befähigt gewesen wären, die Trainings durchzuführen. Der umfangreiche Materialordner sowie die ausführlichen Manuale ermöglichen es den Lehrern bereits sehr gut, sich auf die Durchführung der Trainingseinheiten vorzubereiten. Dennoch sollten zukünftig während der Durchführung regelmäßige Treffen stattfinden, in denen die Lehrer ihre Erfahrungen mit anderen Kollegen austauschen und Fragen klären können. Gleichzeitig würde auf diese Weise die Genauigkeit der Implementierung erörtert und überprüft werden.

Die Korrektur der Schüleraufsätze könnte ebenfalls von den Lehrern selbst realisiert werden, weil diese das Ausfüllen der 7W- und Leistungsprotokolle in dem Workshop erlernten. Es müsste den Lehrern jedoch die Möglichkeit eingeräumt werden, das Training über einen längeren Zeitraum, z.B. ein halbes Jahr, durchzuführen, da ansonsten die wöchentliche Arbeitsbelastung zu groß wäre.

Vermittlung von theoretischen Grundlagen: Auf die Teilnahme an einem Workshop kann ebenfalls nicht verzichtet werden. Ein zentraler Befund der Lehrerbefragung

und der Studie II liegt darin, dass im aktuellen Deutschunterricht keine selbstregulatorischen Strategien vermittelt werden und dass die Lehrer wenig über Selbstregulation wissen. Auch Cutler und Graham (2008) kamen bei einer Nationalerhebung in Grundschulen zu denselben Schlussfolgerungen. Folglich müssen den Lehrern im Workshop zunächst das Prinzip der Selbstregulation und die Wichtigkeit selbstregulatorischer Prozesse für die Kompetenzentwicklung vermittelt werden. Ansonsten würden die Materialien nicht wie vorgesehen und empirisch überprüft von den Lehrern umgesetzt werden.

Veränderte Lehrerrolle: Das USAT-Programm erfordert ein verändertes Lehrer- sowie Schülerverhalten. Das kognitive Modellieren war den USAT-Lehrern zunächst sehr fremd. Mit zunehmender Übung verschwand jedoch dieser Eindruck, so dass die Lehrer dem Modellieren sehr positiv gegenüberstanden. Darüber hinaus zeigte der LUG-Workshop, dass die Lehrer eine vom ihm gesteuerte Kontrolle und Regulation des Lernprozesses bevorzugen und den Schülern es zunächst nicht zutrauen, ihr Lernverhalten selbst zu regulieren. Auch die USAT-Lehrer waren erstaunt darüber, wie gut die Selbstüberwachung, Zielsetzung und Selbstkontrolle bei den Schülern funktionierte. Demnach müssen weitverbreitete Einstellungen von Lehrern revidiert werden, so dass die Lehrerzentrierung zu Gunsten einer Schülerzentrierung in dem USAT-Programm durch Lehrer realisiert wird.

Fortbildung von Lehrern: Des Weiteren wäre die Integration des USAT-Programms in die Lehrerausbildung wünschenswert. Momentan wird der mangelnde Praxisbezug im Lehramtsstudium und im Referendariat beklagt (Helmke, 2009). Die Erlernung eines konkreten Trainingskonzepts, welches zudem wissenschaftlich evaluiert wurde, würde den Lehramtsstudierenden und Referendaren die Umsetzung der Theorie in die Praxis erheblich erleichtern. Des Weiteren wäre zu erwarten, dass die Studenten und Referendare die selbstregulatorischen Methoden des USAT-Programms auch auf andere Fächer und Bereiche ausweiten würden. Auf diese Weise könnten die Erkenntnisse der Wissenschaft denjenigen zugänglich gemacht werden, die täglich mit den Schwierigkeiten des Unterrichtsalltags konfrontiert werden. Der sich daraus ergebende praktische Nutzen wäre erheblich, da Lehrer ihr Wissen auf die Schüler übertragen könnten. Im Vergleich zur Förderung von Schülern stellt die Fortbildung von Lehrern einen effizienteren Forschungsansatz zur Verbesserung der Schulleistungen dar, weil ein Lehrer eine Gruppe von 20 bis 30 Schülern unterrichtet. In diesem Sinne agieren Lehrer als Multiplikatoren des Lernens, so dass eine größere Anzahl von Schülern von der Wissenschaft profitieren würde.

9 Zusammenfassung

Schreiben wird als eine Problemlöseaufgabe angesehen, deren Beherrschung im Gegensatz zum Sprechen mühsam ist und wesentlich mehr Zeit und Übung erfordert (Kellogg, 2008). Die Grundlagenforschung im Bereich Schreiben zeigte, dass Schreibanfänger ihre Texte kaum planen und überarbeiten, über wenig Strategie- und metakognitives Schreibwissen verfügen und ihre Schreibleistungen nicht überwachen (Graham & Harris, 2003).

Zur Überwindung der Probleme von Schreibnovizen entwickelten Harris und Graham (1996) den *Self-Regulated Strategy Development* (SRSD)- Ansatz. SRSD beinhaltet eine kombinierte Vermittlung von Schreibstrategien mit selbstregulatorischen Methoden. In vier Metaanalysen wurden große Effekte des SRSD-Programms im Vergleich zu anderen Schreibtrainings nachgewiesen (Graham, 2006; Graham & Harris, 2003; Graham & Perin, 2007a; Rogers & Graham, 2008). Darauf aufbauend entwickelten Glaser und Kollegen ein *Selbstregulatorisches Aufsatztraining* (SAT), welches sich in Kleingruppen mit Lerntrainern als erfolgreicher erwies als ein Schreibtrainingsprogramm ohne selbstregulatorische Komponenten (Glaser, 2005; Glaser & Brunstein, 2007a; Glaser & Brunstein, 2007b; Glaser et al., 2009). Bislang fehlten Ansätze, das SAT-Programm in den regulären Deutschunterricht zu implementieren. Ferner konnte der inkrementelle Effekt der Selbstregulation, insbesondere für ganze Klassen, noch nicht ausreichend wissenschaftlich belegt werden (De La Paz, 2007).

In der vorliegenden Arbeit wurde ein *Unterrichtsintegriertes Selbstregulatorisches Aufsatztraining* (USAT) konzipiert, welches die Vermittlung von Schreibstrategien zur Verbesserung des Inhalts, des Ausdrucks und des Zusammenhangs einer Bildergeschichte mit den selbstregulatorischen Strategien der Selbstüberwachung, Selbstkontrolle und der Zielsetzung kombiniert. Um valide Aussagen über die Wirksamkeit eines Trainingsprogramms im Regelunterricht treffen und Unterschiede in den abhängigen Variablen (z.B. Schreibleistung) eindeutig auf Unterschiede in den unabhängigen Variablen (z.B. Selbstregulation) zurückführen zu können, müssen Maßnahmen zur Sicherstellung der Implementationsgenauigkeit getroffen werden (Griffith et al., 2009). Dazu zählen die Überprüfung der inhaltlichen Vollständigkeit und der Qualität der Implementierung (O'Donnell, 2008).

Aus den bisherigen theoretischen Erkenntnissen resultierten folgende Fragen, die in der vorliegenden Arbeit in vier Untersuchungen beantwortet wurden:

- (1) Sind Aufsatztrainings, die von Experten entwickelt wurden und sich in Studien mit Kleingruppen bewährten, erfolgreich in den Regelunterricht integrierbar?

- (2) Sind selbstregulatorische Komponenten für die Effektivität eines Schreibtrainings zwingend erforderlich?

Als erstes wurde eine Lehrerbefragung mit zwölf Lehrerinnen durchgeführt, um bereits von Beginn der Arbeit an Lehrer in die Entwicklung des USAT-Programms zu integrieren. Ein Ergebnis der Lehrerbefragung war, dass sich ein Großteil für Fortbildungen in dem Bereich Schreibunterricht und Selbstregulation interessierte. Anhand der Aussagen der befragten Lehrer konnte zudem geschlussfolgert werden, dass selbstregulatorische Prinzipien bislang im Deutschunterricht nicht eingesetzt werden und dass keine strukturierten Manuale zur Instruktion von Schreibstrategien vorliegen.

In einem zweiten Schritt erprobten vier geschulte Trainerinnen die Durchführung des USAT-Programms in einer vierten Klasse ($n=21$). Die Trainingseinheiten (sechs x 90 Minuten) wurden von einer Trainerin abwechselnd realisiert, wobei die übrigen die Vollständigkeit und Qualität der Umsetzung kontrollierten. Im Anschluss an das Training schrieben die Schüler Geschichten, die inhaltlich vollständiger, sprachlich besser ausgestaltet, länger und qualitativ hochwertiger als die Geschichten vor dem Training waren.

Als Drittes wurde die Wirksamkeit des USAT-Programms in acht vierten Klassen ($n=156$) aus fünf verschiedenen Grundschulen untersucht. Die Durchführung erfolgte durch vier geschulte Trainerinnen, wobei jede Trainerin eine Experimental- und eine Kontrollgruppe unterrichtete. Anhand der Rechtschreibleistungen und der Klassengröße wurden die Klassen zufällig den beiden Bedingungen zugeteilt. Zur Überprüfung der Implementationsgenauigkeit wurde außerdem jede Trainerin von einer Tutorin begleitet, die anhand von standardisierten Bögen eine sehr gute Implementationsgenauigkeit in beiden Bedingungen attestierte. Die Experimentalgruppe durchlief das USAT-Programm. Die Kontrollgruppe absolvierte ebenfalls ein Schreibtrainingsprogramm (*Writer's Workshop*, WW), welches dieselben Schreibstrategien und Instruktionsprinzipien wie das USAT-Programm umfasste. Die selbstregulatorischen Komponenten wurden jedoch durch Übungen zur Förderung der Schreibfreude und der Phantasie ersetzt. Beide Trainings umfassten wie in der Pilotstudie sechs Einheiten (à 90 Minuten), die einmal wöchentlich stattfanden. Bezüglich Geschichtenelemente, Geschichtenqualität, Planen, Wissen und eines Transfermaßes (Erlebniserzählungen) übertrafen die USAT-Klassen signifikant die WW-Klassen beim Posttest und beim Follow-Up (sechs Wochen später). Abgesehen von Revisionen (Posttest, Follow-Up) und Wortanzahl (Posttest) ergaben sich auf Schülerebene berechnete Effektstärken zwischen .70 und 2.01.

In der vierten und letzten Phase der Arbeit wurde die Implementierung einen Schritt weitergeführt, indem elf Lehrer ($n=242$) aus sechs verschiedenen Schulen Schreibtrainings in ihren eigenen Klassen selbst umsetzten. Dazu nahmen sechs Lehrer an einem Workshop teil, in dem sie das USAT-Programm kennenlernten und erprobten. Eine weitere Gruppe von fünf Lehrerinnen entwickelte im Rahmen eines anderen Workshops eine *Lehrerzentrierte Unterrichtseinheit zum Thema Geschichtschreiben* (LUG), welche auf denselben Schreibstrategien wie das USAT-Programm und den Erfahrungen der Lehrer basierte. Sie integrierten keine selbstregulatorischen Strategien in das LUG-Programm. Die Trainingsdauer war in beiden Bedingungen identisch (sechs x 90 Minuten bzw. fünf x 90 Minuten + zwei x 45 Minuten). Um eine hohe Implementationsgenauigkeit zu garantieren, wurden Checklisten zur Überprüfung der inhaltlichen Vollständigkeit ausgefüllt und Unterrichtsbeobachtungen in den ersten drei Einheiten durchgeführt. Alle Lehrer realisierten qualitativ gut die wesentlichen Inhalte der beiden Programme. Die Schreibleistungen und schreibbezogenen Maße wurden unmittelbar vor, unmittelbar nach sowie sechs Wochen nach der Posttestung erhoben. Bezüglich der Ergebnisse der Studie I konnten die Befunde nur teilweise repliziert werden, da keine signifikanten Bedingungseffekte für die Geschichtenqualität und für das schreibbezogene Wissen festgestellt wurden. Im Gegensatz zur Studie I schrieben die USAT-Klassen jedoch signifikant längere Aufsätze und überarbeiteten mehr als die LUG-Klassen. Die auf Schülerebene errechneten Effektstärken (zwischen .68 und 1.87) lassen ausgenommen für Geschichtenqualität, Revisionen und Wissen einen großen Effekt des USAT-Programms erkennen.

Über alle vier Untersuchungen hinweg können folgende Schlussfolgerungen aus der vorliegenden Arbeit gezogen werden.

- Im aktuellen Aufsatzunterricht der vierten Klasse werden bislang nur wenige selbstregulatorische Strategien angewendet.
- Die Implementierung von extern entwickelten Interventionsmaßnahmen erwies sich sowohl mit geschulten Trainerinnen als auch mit Lehrern als sehr erfolgreich.
- Es konnte der Nachweis erbracht werden, dass Aufsatztrainings mit selbstregulatorischen Komponenten Aufsatztrainings ohne Selbstregulation überlegen sind.

Die vorliegende Arbeit zeigt, wie der herkömmliche Aufsatzunterricht substanziell sowie nachhaltig verbessert werden kann. Es wäre daher lohnenswert, die gewonnenen wissenschaftlichen Erkenntnisse langfristig in die Lehrerausbildung zu integrieren.

Literatur

- Abbott, R. D., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary and intermediate grade writers. *Journal of Educational Psychology, 85*, 478–508.
- Alamargot, D., & Chanquoy, L. (2001). *Studies in writing: Through the models of writing* (Vol. 9). Dordrecht: Kluwer Academic Publishers.
- Alexander, P. A., Graham, S., & Harris, K. R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review, 10*, 129–154.
- Altenburg, E. (2006). Stimmige Texte erkennen- Verknüpfungen realisieren. *Praxis Grundschule, 29*, 16–22.
- Baddeley, A. D. (1986). *Working memory*. Oxford, U.K.: Clarendon.
- Barab, S., & Squire, K. (2004). Introduction: Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*, 1–14.
- Bereiter, C. (1980). Development in writing. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 73–93). Hillsdale, NJ: Erlbaum.
- Bereiter, C., Burtis, P. J., & Scardamalia, M. (1988). Cognitive operations in constructing main points in written composition. *Journal of Memory and Language, 27*, 261–278.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum.
- Berninger, V. W., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 345–363). New York: Guilford Press.
- Berninger, V. W., Garcia, N. P., & Abbott, R. D. (2009). Multiple processes that matter in writing instruction and assessment. In G. A. Troia (Ed.), *Instruction and assessment for struggling writers. Evidence-based practices* (pp. 15–50). New York: Guilford Press.
- Berninger, V. W., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., et al. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology, 94*, 291–304.

- Berninger, V. W., Vaughan, K. B., Abbott, R. D., Abbott, S. P., Woodruff Rogan, L., Brooks, A., et al. (1997). Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. *Journal of Educational Psychology, 89*, 652–666.
- Berninger, V. W., & Winn, W. D. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 96–114). New York: Guilford Press.
- Bessell, A. G., Burke, M. C., Plaza, M. P., Lee, O., & Schumm, J. S. (2008). The educational reform rating rubric: Example of a new tool for evaluating complex school reform initiatives. *Field Methods, 20*, 283–295.
- Bloom, H. S., Bos, J. M., & Lee, S.-W. (1999). Using cluster random assignment to measure program impacts: Statistical implications for the evaluation of education programs. *Evaluation Review, 23*, 445–469.
- Blumenfeld, P., Fishman, B. J., Krajcik, J., & Marx, R. W. (2000). Creating usable innovations in systemic reform: Scaling Up technology-embedded project-based science in urban schools. *Educational Psychologist, 35*, 149–164.
- Boekaerts, M. (1999). Self-regulated learning: where we are today. *International Journal of Educational Research, 31*, 445–457.
- Bortz, J., & Weber, R. (2005). *Statistik für Human- und Sozialwissenschaftler: Mit 242 Tabellen* (6th ed.). Springer-Lehrbuch. Heidelberg: Springer Medizin.
- Bruder, S., Perels, F., & Schmitz, B. (2004). Selbstregulation und elterliche Hausaufgabenunterstützung: Die Evaluation eines Elterntrainings für Kinder der Sekundarstufe I. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 36*, 139–146.
- Brunstein, J. C., & Spörer, N. (2006). Selbstgesteuertes Lernen. In H. D. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (3rd ed., pp. 677–685). Weinheim: Beltz.
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin, 54*, 297–312.
- Campbell, D. T. (1969). Reforms as experiments. *American Psychologist, 24*, 409–429.
- Cleary, T. J., & Zimmerman, B. J. (2001). Self-regulation differences during athletic practice by experts, non-experts, and novices. *Journal of Applied Sport Psychology, 13*, 185–206.

- Cleary, T. J., & Zimmerman, B. J. (2004). Self-regulation empowerment program: A school-based program to enhance self-regulated and self-motivated cycles of students learning. *Psychology in the Schools, 41*, 537–550.
- Cognition and Technology Group at Vanderbilt (1990). Anchored instruction and its relationship to situated cognition, *19*, 2–10.
- Cognition and Technology Group at Vanderbilt (1992). The jasper deries as an example of anchored instruction: Theory, program description, and assessment data. *Educational Psychologist, 27*, 291–315.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. B. Resnick & R. Glaser (Eds.), *Knowing, learning, and instruction. Essays in honor of Robert Glaser* (pp. 453–494). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology, 100*, 907-919.
- Danoff, B., Harris, K. R., & Graham, S. (1993). Incorporating strategy instruction within the writing process in the regular classroom: Effects on the writing of students with and without learning disabilities. *Journal of Reading Behavior, 25*, 295–322.
- Day, J. D. (1986). Teaching summarization skills: Influences of student ability level and strategy difficulty. *Cognition and Instruction, 3*, 193–210.
- De La Paz, S. (2007). Managing cognitive demands for writing: Comparing the effects of instructional components in strategy instruction. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 23*, 249–266.
- De La Paz, S., & Graham, S. (1997). Effects of dictation and advanced planning instruction on the composing of students with writing and learning problems. *Journal of Educational Psychology, 89*, 203–222.
- De La Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology, 94*, 687–698.
- De La Paz, S., Swanson, P. N., & Graham, S. (1998). The contribution of executive control to the revising by students with writing and learning difficulties. *Journal of Educational Psychology, 90*, 448-460.

- Delp, P., Müller, I., & Stöveken, H. (2004). *Gute Noten in Deutsch: Rechtschreiben/Aufsatz: Klasse 4. Schülerhilfe*. Königswinter: Tandem Verlag.
- Deshler, D. D., & Schumaker, J. B. (1994). Grounding intervention research in the larger context of schooling: A response to Pressley and Harris. *Educational Psychology Review, 6*, 215–222.
- Design-Based Research Collective (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*, 5–8.
- Dießner, H. (2003). *Reisen ins Abenteuerland: Phantasiereisen für Erwachsene, Jugendliche und Kinder*. Paderborn: Junfermann Verlag.
- Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively?: A meta-analysis on self-regulation training programmes. *Educational Research Review, 3*, 101–129.
- Drieschner, E. (2009). *Bildungsstandards praktisch: Perspektiven kompetenzorientierten Lehrens und Lernens*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Duffy, G. G. (1993). Rethinking strategy instruction: Four teachers' development and their low achievers' understandings. *The Elementary School Journal, 93*, 231–247.
- Englert, C. S., Okolo, C. M., & Mariage, T. V. (2009). Informational writing across the curriculum. In G. A. Troia (Ed.), *Instruction and assessment for struggling writers. Evidence-based practices* (pp. 132–161). New York: Guilford Press.
- Englert, C. S., Raphael, T. M., Anderson, L. M., Anthony, H. M., & Stevens, D. D. (1991). Making strategies and self-talk visible: Writing instruction in regular and special education classrooms. *American Educational Research Journal, 28*, 337–372.
- Englert, C. S., Raphael, T. E., & Anderson, L. M. (1992). Socially mediated instruction: Improving students' knowledge and talk about writing. *The Elementary School Journal, 92*, 411–449.
- Euler, D., & Sloane, P. F. E. (1998). Implementation als Problem der Modellversuchsforschung. *Unterrichtswissenschaft, 26*, 312–326.
- Feilke, H. (1993). Schreibentwicklungsforschung: Ein kurzer Überblick unter besonderer Berücksichtigung der Entwicklung prozeßorientierter Schreibfähigkeiten. *Diskussion Deutsch, 129*, 17–34.
- Ferretti, R. P., MacArthur, C. A., & Dowdy, N. S. (2000). The effects of an elaborated goal on the persuasive writing of students with learning disabilities and their normally achieving peers. *Journal of Educational Psychology, 92*, 694–702.

- Fitzgerald, J. (1987). Research on revision in writing. *Review of Educational Research*, 57, 481–506.
- Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist*, 35, 39–50.
- Fix, M. (1999). „Deine Geschichte find ich irgendwie komisch!“ Schreibkonferenzen als Ausgangspunkt für Sprachreflexion. *Praxis Schule 5-10*, 10, 24–29.
- Fletcher, R. J., & Portalupi, J. (2008). *Writing workshop: The essential guide*. Portsmouth, NH: Heinemann.
- Fullan, M., & Pomfret, A. (1977). Research on curriculum and instruction implementation. *Review of Educational Research*, 47, 335–397.
- Geijsel, F., Slegers, P., Leithwood, K., & Jantzi, D. (2002). Transformational leadership effects on teachers' commitment and effort toward school reform. *Journal of Educational Administration*, 41, 228–256.
- Geijsel, F., Slegers, P., van den Berg, R., & Kelchtermans, G. (2001). Conditions fostering the implementation of large-scale innovation programs in schools: Teachers' perspectives. *Educational Administration Quarterly*, 37, 130–166.
- Gentile, C. (1992). *Exploring new methods for collecting students' school-based writing: NAEP's 1990 portfolio study*. Washington D.C.: National Center for Educational Statistics.
- Gersten, R., & Baker, S. (2001). Teaching expressive writing to students with learning disabilities: A meta-analysis. *The Elementary School Journal*, 101, 251–272.
- Gillet, J. W., & Beverly, L. (2001). *Directing the writing workshop: An elementary teacher's handbook*. New York: Guilford Press.
- Glaser, C. (2005). *Förderung der Schreibkompetenz bei Grundschulern: Effekte einer integrierten Vermittlung kognitiver Schreibstrategien und selbstregulatorischer Fertigkeiten*. Dissertationsschrift, Universität Potsdam, Potsdam.
- Glaser, C., & Brunstein, J. C. (2007a). Förderung von Fertigkeiten zur Überarbeitung narrativer Texte bei Schülern der 6. Klasse: Effekte von Revisionsstrategien und selbstregulatorischen Prozeduren. *Zeitschrift für Pädagogische Psychologie*, 21, 51–63.
- Glaser, C., & Brunstein, J. C. (2007b). Improving fourth-grade students' composition skills: Effects of strategy instruction and self-regulation procedures. *Journal of Educational Psychology*, 99, 297–310.

- Glaser, C., Keßler, C., & Brunstein, J. C. (2009). Förderung selbstregulierten Schreibens bei Viertklässlern: Effekte auf strategiebezogene, holistische und subjektive Maße der Schreibkompetenz. *Zeitschrift für Pädagogische Psychologie*, 23, 5–18.
- Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, 42, 237–288.
- Graham, S. (1990). The role of production factors in learning disabled students' compositions. *Journal of Educational Psychology*, 82, 781–791.
- Graham, S. (1997). Executive control in the revising of students with learning and writing difficulties. *Journal of Educational Psychology*, 89, 223–234.
- Graham, S. (2006). Strategy instruction and the teaching of writing: A meta-analysis. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 187–207). New York: Guilford Press.
- Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). The role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology*, 89, 170–182.
- Graham, S., & Harris, K. R. (1989). Components analysis of cognitive strategy instruction: Effects on learning disabled students' compositions and self-efficacy. *Journal of Educational Psychology*, 81, 353–361.
- Graham, S., & Harris, K. R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist*, 35, 3–12.
- Graham, S., & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 323–344). New York: Guilford Press.
- Graham, S., & Harris, K. R. (2009). Evidence-based writing practices: Drawing recommendations from multiple sources. *British Journal of Educational Psychology Monograph Series II*, 6, 97–114.
- Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write? Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology*, 92, 620–633.

- Graham, S., Harris, K. R., & Fink Chorzempa, B. (2002). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology, 94*, 669–686.
- Graham, S., Harris, K. R., MacArthur, C. A., & Schwartz, S. (1991). Writing and writing instruction for students with learning disabilities: Review of a research program. *Learning Disability Quarterly, 14*, 89–114.
- Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology, 30*, 207–241.
- Graham, S., Harris, K. R., & Troia, G. A. (1998). Writing and self-regulation: Cases from the self-regulated strategy development model. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning. From teaching to self-reflective practice* (pp. 20–41). New York: Guilford Press.
- Graham, S., MacArthur, C. A., & Schwartz, S. (1995). Effects of goal setting and procedural facilitation on the revising behavior and writing performance of students with writing and learning problems. *Journal of Educational Psychology, 87*(2), 230–240.
- Graham, S., Olinghouse, N. G., & Harris, K. R. (2009). Teaching composing to students with learning disabilities: Scientifically supported recommendations: Instruction and assessment for struggling writers: Evidence-based practices. In G. A. Troia (Ed.), *Instruction and assessment for struggling writers. Evidence-based practices* (pp. 165–186). New York: Guilford Press.
- Graham, S., & Perin, D. (2007a). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*, 445–476.
- Graham, S., & Perin, D. (2007b). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools- A report to Carnegie Corporation of New York*. Washington D.C.: Alliance for Excellent Education.
- Gräsel, C., Fussangel, K., & Parchmann, I. (2006). Lerngemeinschaften in der Lehrerfortbildung: Kooperationserfahrungen und -überzeugungen von Lehrkräften. *Zeitschrift für Erziehungswissenschaft, 9*, 545–561.
- Gräsel, C., & Parchmann, I. (2004). Implementationsforschung - oder: der steinige Weg, Unterricht zu verändern. *Unterrichtswissenschaft, 32*, 194–214.

- Griffith, A. K., Duppong Hurley, K., & Hageman, J. L. (2009). Treatment integrity of literacy interventions for students with emotional and/or behavioral disorders: A review of literature. *Remedial and Special Education, 30*, 245–255.
- Grund, M., Haug, G., & Naumann, C. L. (2004). *DRT 4: Diagnostischer Rechtschreibtest für 4. Klassen* (2nd ed.). Göttingen: Beltz.
- Hacker, D. J., & Tenet, A. (2002). Implementing reciprocal teaching in the classroom: Overcoming obstacles and making modifications. *Journal of Educational Psychology, 94*, 699–718.
- Häcker, H. O., & Stapf, K.-H. (Eds.) (2004). *Dorsch: Psychologisches Wörterbuch* (14th ed.). Bern: Huber.
- Hager, W. (2000a). Wirksamkeits- und Wirksamkeitsunterschiedshypothesen, Evaluationsparadigmen, Vergleichsgruppen und Kontrolle. In W. Hager, J.-L. Patry, & H. Brezing (Eds.), *Aus dem Programm Huber. Evaluation psychologischer Interventionsmaßnahmen. Standards und Kriterien: ein Handbuch* (pp. 180–201). Bern: Huber.
- Hager, W. (2000b). Zur Wirksamkeit von Interventionsprogrammen: Allgemeine Kriterien der Wirksamkeit von Programmen in einzelnen Untersuchungen. In W. Hager, J.-L. Patry, & H. Brezing (Eds.), *Aus dem Programm Huber. Evaluation psychologischer Interventionsmaßnahmen. Standards und Kriterien: ein Handbuch* (pp. 153–168). Bern: Huber.
- Hager, W., Patry, J.-L., & Brezing, H. (Eds.) (2000). *Aus dem Programm Huber. Evaluation psychologischer Interventionsmaßnahmen: Standards und Kriterien: ein Handbuch*. Bern: Huber.
- Harris, K. R., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation* (2nd ed.). Cambridge, Mass: Brookline Books.
- Harris, K. R., & Graham, S. (2009). Self-regulated strategy development in writing: Premises, evolution, and the future. *British Journal of Educational Psychology Monograph Series II, 6*, 113–135.
- Harris, K. R., Graham, S., & Mason, L. H. (2003). Self-regulated strategy development in the classroom: Part of a balanced approach to writing instruction for students with disabilities. *Focus on Exceptional Children, 35*, 1–17.

- Harris, K. R., Graham, S., & Mason, L. H. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. *American Educational Research Journal*, 43, 295–340.
- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, 66, 99–136.
- Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Lawrence Erlbaum.
- Hayes, J. R. (2006). New directions in writing theory. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 28–40). New York: Guilford Press.
- Hayes, J. R., & Flower, L. S. (1980a). Writing as problem solving. *Visible Language*, 14, 288–299.
- Hayes, J. R., & Flower, L. S. (1980b). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale, NJ: Erlbaum.
- Hedges, L. V. (2007). Correcting a significance test for clustering. *Journal of Educational and Behavioral Statistics*, 32, 151–179.
- Hedges, L. V., & Hedberg, E. C. (2007a). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, 29, 60–87.
- Hedges, L. V., & Hedberg, E. C. (2007b). Intraclass correlations for planning group randomized experiments in rural education. *Journal of Research in Rural Education*, 22, 1–15.
- Helmke, A. (2006). Unterrichtsqualität. In H. D. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (3rd ed., pp. 812–820). Weinheim: Beltz.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts; Franz Emanuel Weinert gewidmet*. Seelze-Velber: Klett Kallmeyer.

- Herzmann, P., Sparka, A., & Gräsel, C. (2006). Implementationsforschung zur Lesekompetenz: Wie Wissenschaftler und Lehrkräfte gemeinsam an der Leseförderung arbeiten. In S. Rahm, I. Mammes, & M. Schratz (Eds.), *Schulpädagogische Forschung: Perspektiven innovativer Ansätze ; Bd. 2. Organisations- und Bildungsprozessforschung* (pp. 97–108). Innsbruck: StudienVerlag.
- Hickey, D. T., Moore, A. L., & Pellegrino, J. W. (2001). The motivational and academic consequences of elementary mathematics environments: Do constructivist innovations and reforms make a difference? *American Educational Research Journal*, 38, 611–652.
- Hidi, S., & Boscolo, P. (2006). Motivation and writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 144–157). New York: Guilford Press.
- Hillocks, G. (1984). What works in teaching composition: A meta-analysis of experimental treatment studies. *American Journal of Education*, 93, 133–170.
- Hopkins, K. D. (1982). The unit of analysis: Group means versus individual observations. *American Educational Research Journal*, 19, 5–18.
- Hsieh, P., Acee, T., Chung, W.-H., Hsieh, Y.-P., Kim, H., Thomas, G. D., et al. (2005). Is educational intervention research on the decline? *Journal of Educational Psychology*, 97, 523–529.
- Hull, G., & Rose, M. (1989). Rethinking remediation: towards a social-cognitive understanding of problematic reading and writing. *Written Communication*, 6, 139–154.
- Jaycox, L. H., McCaffrey, D. F., Ocampo, B. W., Shelley, G. A., Blake, S. M., Peterson, D. J., et al. (2006). Challenges in the evaluation and implementation of school-based prevention and intervention programs on sensitive topics. *American Journal of Evaluation*, 27, 320–336.
- Johnson, K. (2004). *Writing like writers: Guiding elementary children through a writer's workshop*. Waco, Tx: Prufrock Press Inc.
- Käferle, V. (2007). Schreibkonferenzen fördern Schreibkompetenzen. Kinder überarbeiten Texte und präsentieren ihre Überlegungen. *Grundschulmagazin*, 75, 29–34.

- Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–71). Mahwah, NJ: Lawrence Erlbaum.
- Kellogg, R. T. (1999). Components of working memory in text production. In M. Torrance & G. C. Jeffery (Eds.), *The cognitive demands of writing. Processing capacity and working memory in text production* (pp. 42–61). Amsterdam: Amsterdam University Press.
- Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research, 1*, 1–26.
- Kernaghan, K., & Woloshyn, V. E. (1995). Providing grade one students with multiple spelling strategies: Comparison between strategy instruction, strategy instruction with metacognitive information, and traditional language arts. *Applied Cognitive Psychology, 9*, 157–166.
- Kitsantas, A., & Zimmerman, B. J. (2002). Comparing self-regulatory processes among novice, non-expert, and expert volleyball players: A microanalytic study. *Journal of Applied Sport Psychology, 14*, 91–105.
- Kitsantas, A., Zimmerman, B. J., & Cleary, T. (2000). The role of observation and emulation in the development of athletic self-regulation. *Journal of Educational Psychology, 92*, 811–817.
- Klauer, K. J. (2006). Intelligenz und Begabung. In H. D. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (3rd ed., pp. 275–280). Weinheim: Beltz.
- Kromrey, J. D., & Dickinson, W. B. (1996). Detecting unit of analysis problems in nested designs: Statistical power and Type I error rates of the F-Test for groups-within-treatments effects. *Educational and Psychological Measurement, 56*, 215–231.
- Kruse, N. (2006). Rückmeldungen beim Textschreiben: Oder: Wie kann eine gute Rückmeldekultur zu Texten in der Klasse entstehen? *Grundschulunterricht, 53*, 14–17.
- Kultusministerkonferenz (2005). *Bildungsstandards im Fach Deutsch für den Primarbereich: (Jahrgangsstufe 4) ; [Beschluss vom 15.10.2004]. Beschlüsse der Kultusministerkonferenz*. Neuwied: Luchterhand.
- Labuhn, A. S., Bögeholz, S., & Hasselhorn, M. (2008). Lernförderung durch Anregung der Selbstregulation im naturwissenschaftlichen Unterricht. *Zeitschrift für Pädagogische Psychologie, 22*, 13–24.

- Levy, C. M., & Marek, P. (1999). Testing components of Kellogg's multicomponent model of working memory in writing: The role of the phonological loop. In M. Torrance & G. C. Jeffery (Eds.), *The cognitive demands of writing. Processing capacity and working memory in text production* (pp. 25–41). Amsterdam: Amsterdam University Press.
- Lipson, M. Y., Mosenthal, J., Daniels, P., & Woodside-Jiron, H. (2000). Process writing in the classrooms of eleven fifth-grade teachers with different orientations to teaching and learning. *The Elementary School Journal, 101*, 209–231.
- Mayer, R. E. (2001). What good is educational psychology? The case of cognition and instruction. *Educational Psychologist, 36*, 83–88.
- Mayring, P. (2001). Kombination und Integration qualitativer und quantitativer Analyse. *Forum Qualitative Sozialforschung, 2*, Art. 6.
- Mayring, P. (2008). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (10th ed.). Weinheim: Beltz.
- McCutchen, D. (1988). "Functional Automacity" in children's writing: A problem of metacognitive control. *Written Communication, 5*, 306–324.
- McCutchen, D. (2006). Cognitive factors in the development of children's writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 115–130). New York: Guilford Press.
- McKenney, S., Nieveen, N., & van den Akker, J. (2006). Design research from a curriculum perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational Design Research* (pp. 67–90). London: Routledge.
- Menauer, S. (2006). Hilfen für den Schulalltag: Kinder mit Textproduktionsstörungen. *Praxis Grundschule, 29*, 6–11.
- Mietzel, G., & Willenberg, H. (2000). *HST 4/5: Hamburger Schulleistungstest für 4. und 5. Klassen*. Göttingen: Hogrefe.
- Mokhlesgerami, J., Souvignier, E., Rühl, K., & Gold, A. (2007). Naher und weiter Transfer eines Unterrichtsprogramms zur Förderung der Lesekompetenz in der Sekundarstufe I. *Zeitschrift für Pädagogische Psychologie, 21*, 169–180.
- Molfese, V. J., Beswick, J., Molnar, A., & Jacobi-Vessels, J. (2006). Alphabetic skills in preschool: A preliminary study of letter naming and letter writing. *Developmental Neuropsychology, 29*, 5–19.

- Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *The American Journal of Evaluation, 24*, 315–340.
- Müller, R. (2004). *DRT 3: Diagnostischer Rechtschreibtest für 3. Klassen* (4th ed.). Göttingen: Beltz.
- Murray, D. M., Varnell, S. P., & Blitstein, J. L. (2004). Design and analysis of group-randomized trials: A review of recent methodological developments. *American Journal of Public Health, 94*, 423–432.
- Nystrand, M. (2006). The social and historical context for writing research. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 11–27). New York: Guilford Press.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research, 78*, 33–84.
- Olive, T., & Kellogg, R. T. (2002). Concurrent activation of high- and low-level writing processes. *Memory & Cognition, 30*, 594–600.
- Page-Voth, V., & Graham, S. (1999). Effects of goal setting and strategy use on the writing performance and self-efficacy of students with writing and learning problems. *Journal of Educational Psychology, 91*, 230–240.
- Pajares, F. (2003). Self-efficacy beliefs, motivation and achievement in writing: A review of the literature. *Reading & Writing Quarterly, 19*, 139–158.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*, 117–175.
- Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist, 26*, 89–101.
- Perels, F., Gürtler, T., & Schmitz, B. (2005). Training of self-regulatory and problem-solving competence. *Learning and Instruction, 15*, 123–139.
- Persky, H. R., Daane, M. C., & Jin, Y. (2003). *The nation's report card: Writing 2002*. Washington D.C.: U.S. Department of Education, National Center for Educational Statistics.
- Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research, 31*, 459–470.

- Piolat, A., Olive, T., & Kellogg, R. T. (2004). Cognitive effort during note taking. *Applied Cognitive Psychology, 19*, 291–312.
- Plauen, E. O. (2003). *Vater und Sohn: Sämtliche Streiche und Abenteuer*. Konstanz: Südverlag.
- Pollington, M. F., Wilcox, B., & Morrison, T. G. (2001). Self-perception in writing: The effects of writing workshop and traditional instruction on intermediate grade students. *Reading Psychology, 22*, 249–265.
- Pressley, M., Borkowski, J. G., & Schneider, W. (1989). Good information processing: What it is and how education can promote it. *International Journal of Educational Research, 13*, 857–867.
- Pressley, M., & Harris, K. R. (1994). Increasing the quality of educational intervention research. *Educational Psychology Review, 6*, 191–208.
- Pritchard, R. J., & Honeycutt, R. L. (2006). The process approach to writing instruction. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 275–290). New York: Guilford Press.
- Reid, R. C., & Lienemann, T. (2006). *Strategy instruction for students with learning disabilities. What works for special needs learners*. New York: Guilford Press.
- Reinmann-Rothmeier, G., & Mandl, H. (1998). Wenn kreative Ansätze versanden: Implementation als verkannte Aufgabe. *Unterrichtswissenschaft, 26*, 292–311.
- Rogers, L. A., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology, 100*, 879–906.
- Rozendall, J. S., Minnaert, A., & Boekaerts, M. (2005). The influence of teacher perceived administration of self-regulated learning on students' motivation and information-processing. *Learning and Instruction, 15*, 141–160.
- Saddler, B., & Graham, S. (2005). The effects of peer-assisted sentence-combining instruction on the writing performance of more and less skilled young writers. *Journal of Educational Psychology, 97*, 43–54.
- Sawyer, R. R., Graham, S., & Harris, K. R. (1992). Direct teaching, strategy instruction, and strategy instruction with explicit self regulation: Effects on the composition skills and self-efficacy of students with learning disabilities. *Journal of Educational Psychology, 84*, 340–352.
- Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflective processes in written composition. *Cognitive Science, 8*, 173–190.

- Schiffler, A., & Hübner, S. (2000). Allgemeine Standards für die Evaluationspraxis: Die Standards des "Joint Committee on Standards for Educational Evaluation" und ihre Anwendung auf praktische Aspekte bei der Evaluation von psychologischen Interventionsmaßnahmen. In W. Hager, J.-L. Patry, & H. Brezing (Eds.), *Aus dem Programm Huber. Evaluation psychologischer Interventionsmaßnahmen. Standards und Kriterien: ein Handbuch* (pp. 141–151). Bern: Huber.
- Schumacher, S. (2001). *Heut' reisen wir ins Paradies: Phantasie Reisen und Erzählgeschichten für Grundschul Kinder*. Stuttgart: Calver Verlag.
- Schumaker, J. B., & Deshler, D. D. (2003). Designs for applied educational research. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 483–500). New York: Guilford Press.
- Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal*, 33, 359–382.
- Schunk, D. H., & Schwartz, C. W. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology*, 18, 337–354.
- Schwanitz, D. (2002). *Bildung: Alles, was man wissen muß*. München: Goldmann.
- Shadish, W. R. (2002). Revisiting field experimentation: Field notes for the future. *Psychological Methods*, 7, 3–18.
- Shanahan, T. (2006). Relations among oral languages, reading and writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 171–183). New York: Guilford Press.
- Simon, E. (1997). Die Schreibkonferenz. Eine Alternative zum traditionellen Unterricht. *Grundschule*, 29, 51–54.
- Snyder, J., Bolin, F., & Zumwalt, K. (1992). Curriculum implementation. In P. W. Jackson & P. W. Jackson (Eds.), *Handbook of research on curriculum. A project of the American Educational Research Association* (pp. 402–435). New York: Macmillan.
- Souvignier, E., & Mokhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction*, 16, 57–71.
- Souvignier, E., & Trenk-Hinterberger, I. (2007). Ein Drei-Stufen-Modell zur Implementation neuer Unterrichtskonzepte in den Schulalltag. In M. Krämer & S. B. K. Preiser (Eds.), *Psychologiedidaktik und Evaluation VI* (pp. 197–204). Göttingen: V & R unipress.

- Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, *124*, 240–261.
- Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *Advances in discourse processes. New directions in discourse processing* (2nd ed., pp. 53–120). Norwood, NJ: Ablex.
- Timperley, H. S., & Phillips, G. (2003). Changing and sustaining teachers' expectations through professional development in literacy. *Teaching and Teacher Education*, *19*, 627–641.
- Tolchinsky, L. (2006). The emergence of writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 83–95). New York: Guilford Press.
- Torrance, M., Fidalgo, R., & García, J. (2007). The teachability and effectiveness of cognitive self-regulation in sixth-grade writers. *Learning and Instruction*, *17*, 265–285.
- Troia, G. A. (Ed.) (2009a). *Instruction and assessment for struggling writers: Evidence-based practices*. New York: Guilford Press.
- Troia, G. A. (2009b). Introduction. In G. A. Troia (Ed.), *Instruction and assessment for struggling writers. Evidence-based practices* (pp. 1–11). New York: Guilford Press.
- Troia, G. A., Lin, S.-J. C., Monroe, B. W., & Cohen, S. (2009). The effects of writing workshop instruction on the performance and motivation of good and poor writers. In G. A. Troia (Ed.), *Instruction and assessment for struggling writers. Evidence-based practices* (pp. 77–107). New York: Guilford Press.
- Uerdingen, M. (2006). Rückmeldungen zu schriftsprachlichen Lernleistungen im Anfangsunterricht. *Grundschulunterricht*, *53*, 18–23.
- van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). Introducing educational design research. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational Design Research* (pp. 3–7). London: Routledge.
- van den Berg, R., & Ros, A. (1999). The permanent importance of the subjective reality of teachers during educational innovation: A concerns-based approach. *American Educational Research Journal*, *36*, 879–906.
- van Keer, H., & Verhaeghe, J. P. (2005). Comparing two teacher development programs for innovating reading comprehension instruction with regard to teachers' experiences and student outcomes. *Teaching and Teacher Education*, *21*, 543–562.

- Walker, D. (2006). Toward productive design studies. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational Design Research* (pp. 8–13). London: Routledge.
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63, 249–294.
- Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts & P. R. Pintrich (Eds.), *Handbook of self-regulation* (pp. 532–566). San Diego, CA: Academic Press.
- Wong, B. Y. L. (1997). Research on genre-specific strategies for enhancing writing in adolescents with learning disabilities. *Learning Disability Quarterly*, 20, 140–159.
- Wong, B. Y. L., Butler, D. L., Ficzero, S. A., & Kuperis, S. (1996). Teaching low achievers and students with learning disabilities to plan, write, and revise opinion essays. *Journal of Learning Disabilities*, 29, 197–212.
- Wong, B. Y. L., Harris, K. R., Graham, S., & Butler, D. L. (2003). Cognitive strategies instruction research in learning disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 383–402). New York: Guilford Press.
- Zech, L., Gause-Vega, C. L., Bray, M. H., Secules, T., & Goldman, S. R. (2000). Content-based collaborative inquiry: A professional development model for sustaining educational reform. *Educational Psychologist*, 35, 207–217.
- Zimmerman, B. J. (1998). Academic studying and the development of personal skill: A self-regulatory perspective. *Educational Psychologist*, 33, 73–86.
- Zimmerman, B. J. (1999). Commentary: toward a cyclically interactive view of self-regulated learning. *International Journal of Educational Research*, 31, 545–551.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner. *Theory into practice*, 41, 64–70.
- Zimmerman, B. J., & Bandura, A. (1994). Impact of self-regulatory influence on writing course attainment. *American Educational Research Journal*, 31, 845–862.
- Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: the role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, 29, 623–676.
- Zimmerman, B. J., & Kitsantas, A. (1997). Developmental phases in self-regulation: Shifting from process goals to outcome goals. *Journal of Educational Psychology*, 89, 29–36.

- Zimmerman, B. J., & Kitsantas, A. (1999). Acquiring writing revision skill: shifting from process to outcome self-regulatory goals. *Journal of Educational Psychology, 91*, 241–250.
- Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology, 94*, 660–668.
- Zimmerman, B. J., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology, 80*, 284–290.
- Zimmerman, B. J., & Risemberg, R. (1997). Research for the future. Becoming a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology, 22*, 73–101.

Anhang¹⁶

A1 Lehrerbefragung

A2 Testungen

A3 Schreibanlässe im Training

A4 Pilotstudie

A4.1 Gruppenarbeiten

A4.2 Trainingsmanuale

A4.3 Hausaufgaben

A4.4 Übungen

A5 Studie I

A5.1 USAT

A5.1.1 Partnerarbeiten

A5.1.2 Trainingsmanuale

A5.2 WW

A5.2.1 Gruppenarbeiten

A5.2.2 Phantasiereisen

A5.2.3 Trainingsmanuale

A5.3 Implementationsgenauigkeit

A5.3.1 Beobachtungsbögen

A5.3.2 Checklisten für Trainer

A5.3.3 Differenzierung zwischen den Bedingungen

A5.3.4 Interviewleitfäden für Lehrer

¹⁶ Zur Einsicht in den Anhang wenden Sie sich bitte an die Autorin (Sandra.Budde@psychol.uni-giessen.de).

A6 Studie II**A6.1 USAT**

- A6.1.1 Trainingsmanuale kurz
- A6.1.2 Trainingsmanuale lang
- A6.1.3 Übungen
- A6.1.4 Hausaufgaben
- A6.1.5 Rückmeldung für Lehrer

A6.2 LUG

- A6.2.1 Trainingsmanuale
- A6.2.2 Übungen
- A6.2.3 Merkhilfen
- A6.2.4 Hausaufgaben
- A6.2.5 Rückmeldung für Lehrer

A6.3 Implementationsgenauigkeit

- A6.3.1 Beobachtungen
- A6.3.2 Checklisten für Lehrer
- A6.3.3 Evaluation

A7 Schreibstrategien**A7.1 Inhalt****A7.2 Ausdruck****A7.3 Zusammenhang****A7.4 Planen****A7.5 Überarbeiten**

Eidesstattliche Erklärung

„Ich erkläre:

Ich habe die vorgelegte Dissertation selbständig und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht.“

Gießen, den 24. Mai 2010

Sandra Budde