Web-based Training in der Hochschullehre –

Bedingungsanalyse lern- und motivationsförderlicher Gestaltungsmerkmale

Inaugural-Dissertation
zur
Erlangung des Doktorgrades
der Philosophie des Fachbereiches Psychologie und Sportwissenschaft
der Justus-Liebig-Universität Gießen

vorgelegt von
Nina Müller

aus Butzbach Kirch-Göns

2017

Dekan: Prof. Dr. Dr. Jürgen Hennig
1. Berichterstatter: Prof. Dr. Ulrich Glowalla
2. Berichterstatter: Prof. Dr. Marko Ennemoser
Tag der Disputation: 14. Juni 2017

Zusammenfassung

E-Learning spielt im Rahmen der Hochschullehre eine zunehmend bedeutendere Rolle. Ziel der vorliegenden Dissertation war es deshalb zu untersuchen, inwieweit Akzeptanz und Lernwirksamkeit von E-Learning-Anwendungen für das Studium gefördert werden können. Um dies zu untersuchen, wurden im Rahmen der ersten Studie die Lernwirksamkeit und Akzeptanz der beiden prominentesten E-Learning-Methoden E-Lecture und Web-based Training (WBT) im Vergleich zur traditionellen Präsenzvorlesung evaluiert. Für die zweite Studie wurden Web-based Trainings hinsichtlich verschiedener lernförderlicher Gestaltungsmerkmale nach dem ARCS-Modell von Keller (2010) variiert und vergleichend evaluiert.

Bei beiden Studien handelt es sich um umfangreiche Felduntersuchungen, an denen insgesamt etwa 2500 Studierende teilnahmen. Außerdem folgten beide Untersuchungen einem Versuchsplan mit Messwiederholung. Dies ermöglichte jeweils mehrere Replikationen sowie inhalts-, person- und materialinvariante Vergleiche. Die Arbeit mit einer großen studentischen Zielgruppe im Kontext regulärer Lehrveranstaltungen stellt darüber hinaus eine hohe ökologische Validität der Ergebnisse sicher.

Die Ergebnisse aus beiden Studien sprechen für eine gute Lernwirksamkeit und Akzeptanz der eingesetzten Lehrmethoden. Alle Lehrmethoden können folglich zukünftig mit gutem Gewissen in der Lehre eingesetzt werden. Wenn signifikante Unterschiede hinsichtlich der Lernwirksamkeit und Akzeptanz auftraten, fielen diese im Vergleich zwischen E-Learning und Präsenzlehre zumeist zugunsten der E-Learning-Anwendungen aus. Im Vergleich zwischen E-Lectures und WBTs schnitten die Web-based Trainings in vielen Bereichen signifikant besser ab.

Es zeigte sich, dass die Studierenden die Stärken der zur Verfügung gestellten Lehrmethoden und -materialien gut einschätzen können. Sie setzten sie je nach Lernziel unterschiedlich ein bzw. wünschten sie sich für unterschiedliche Zwecke. Je nach Lernziel sollte in der Hochschullehre folglich die jeweils geeignetste Methode ausgewählt werden. Dabei wünschten sich die Studierenden eine Fortsetzung der traditionellen Vorlesung und zusätzlich eine Unterstützung durch E-Learning-Maßnahmen. Dies spricht für ein Blended Learning-Konzept in der Hochschullehre, in dem Vorteile von Präsenzlehre und E-Learning miteinander kombiniert werden.

Im Hinblick auf die Frage, was bei der Erstellung von lernförderlichen und gut akzeptierten E-Learning-Anwendungen beachtet werden muss, sprechen die Ergebnisse von Studie II für den Einsatz des ARCS-Modells (Keller, 2010). Auf jeden Fall ist es sinnvoll, die Komponenten „Erfolgszuversicht“ und „Zufriedenheit“ umzusetzen. Die Komponente „Relevanz“ sollte immer dann zum Einsatz kommen, wenn sich Übertragbarkeit und praktische Anwendbarkeit nicht direkt aus den Inhalten heraus ergeben. Die Komponente „Aufmerksamkeit“ erwies sich als die am wenigsten einflussreiche Komponente und kann in vielen Lernsituationen getrost vernachlässigt werden.

Summary

Today, e-learning plays a significant role in higher education. And, as many experts say, the impact of e-learning is still increasing. Consequently, the main objective of the present dissertation was to find out whether it is possible to increase both the acceptance and the learning effectiveness of the two most prominent forms of e-learning, namely e-lectures and web-based trainings (wbts). In particular, study 1 compared the learning effectiveness and acceptance of both of these e-learning applications to a traditional university lecture on psychology for student teachers. For study 2, different versions of the same wbt content were created based on the ARCS-model proposed by Keller (2010). Study 2 focussed on the learning effectiveness and acceptance of these different wbt versions.

Both studies were extensive field tests with a total of 2.500 participating students approximately. Both studies followed stringent experimental designs with repeated measurement and multiple replications of the experimental variations. Consequently, the comparions of the different instructional methods were possible invariant of teaching content and subjects. Of course, the examination of a large group of students participating in a regular university course guarantees a high ecological validity of the findings.

The results of both studies indicate that both traditional university lectures as well as e-learning were very well received by the students and lead to good learning results. Hence, all of these teaching methods can be recommend for higher education. However, whenever significant differences were observed between e-learning and lectures with respect to learning effectiveness and acceptance, these comparisons were in favor of the e-learning applications. Moreover, when comparing e-lectures to web-based trainings, the latter outperform the e-lectures in the majority of instances.

In addition, it was found that students are very good at identifying and utilizing the benefits and strenghts of each teaching method depending on their current learning objective. Therefore, it makes a lot of sense in designing higher education to provide the most appropriate teaching method for any particular learning objective. As far as students are concerned universities should continue to offer traditional lectures and enrich this learning experience by providing a variety of e-learning-applications. This is further evidence for blended learning in higher education in order to make use of the advantages of both lectures and e-learning applications.

With respect to the key design features of learning efficient and highly accepted e-learning-applications, the findings of study 2 provide strong evidence for the use of the ARCS model (Keller, 2010). In each case, Keller's components „confidence" and „satisfaction" should be implemented. The component „relevance" should be used whenever the teaching content does not provide sufficient information about its practical relevance and applicability. Finally, the component „attention" seems to be the least important variable in Keller's model and can therefore be neglected in many learning situations.

Inhalt

1 THEORETISCHE GRUNDLAGEN 1

1.1 EINLEITUNG 1
1.2 E-LEARNING 3
1.2.1 DEFINITION UND ENTWICKLUNG 3
1.2.2 E-LEARNING IN SCHULE, HOCHSCHULE UND BETRIEBLICHER BILDUNG 4
1.2.3 VOR- UND NACHTEILE VON E-LEARNING 13
1.3 TRADITIONELLE LEHRE: DIE VORTRAGSMETHODE 19
1.4 VERSCHIEDENE FORMEN VON E-LEARNING 22
1.4.1 PROMINENTE E-LEARNING-FORMEN IN STUDIUM UND BERUFLICHER WEITERBILDUNG 27
1.4.2 E-LECTURE 29
1.4.3 WEB-BASED TRAINING (WBT) 32
1.5 EVALUATION VON E-LEARNING-ANWENDUNGEN 37
1.5.1 NOTWENDIGE VORÜBERLEGUNGEN 37
1.5.2 EVALUATIONSKRITERIEN 38
1.5.3 EVALUATIONSMETHODEN 40
1.6 FRAGESTELLUNGEN DER ERSTEN STUDIE 41
1.6.1 E-LEARNING VS. PRÄSENZLEHRE – EIN BLICK IN DIE LITERATUR 41
1.6.2 FRAGESTELLUNGEN DER ERSTEN STUDIE 47

2 STUDIE I: METHODE 55

2.1 STICHPROBE 55
2.1.1 BESCHREIBUNG DER GESAMTSTICHPROBE 56
2.1.2 TEILSTICHPROBEN FÜR EINZELNE AUSWERTUNGEN 62
2.2 DURCHFÜHRUNG 63
2.3 MATERIAL 65
2.3.1 LEHRMATERIALIEN: VORLESUNG, E-LECTURES UND WEB-BASED TRAININGS 66
2.3.2 ABHÄNGIGE VARIABLEN 67

3 STUDIE I: ERGEBNISSE UND DISKUSSION 75

3.1 LERNWIRKSAMKEIT 77
3.1.1 KURZFRISTIGE LERNEFFEKTIVITÄT UND -EFFIZIENZ 77
3.1.2 LÄNGERFRISTIGE LERNEFFEKTIVITÄT UND -EFFIZIENZ 80
3.1.3 LANGFRISTIGE LERNWIRKSAMKEIT 83
3.1.4 ZUSAMMENFASSUNG UND DISKUSSION 86
3.2 AKZEPTANZ 90
3.2.1 KURZFRISTIGE EINSCHÄTZUNGEN 90
3.2.2 LÄNGERFRISTIGE EINSCHÄTZUNGEN IM RAHMEN DER PROBEKLAUSUR 95
3.2.3 ABSCHLIEßENDE EINSCHÄTZUNGEN 100
3.2.4 VERHALTENSAKZEPTANZ 103

3.2.5 Zusammenfassung und Diskussion 109
3.3 Verhalten vor, während und nach Bearbeitung/Besuch des Lehrangebotes 118
3.3.1 Zusammenfassung und Diskussion 120
3.4 Übungsaufgaben 121
3.4.1 Beurteilung der Übungsaufgaben durch die Studierenden 121
3.4.2 Vergleich der WBTs mit und ohne Übungsaufgaben 122
3.4.3 Zusammenfassung und Diskussion 124

4 STUDIE I: FAZIT 127

4.1 Lernwirksamkeit 127
4.2 Akzeptanz 129
4.3 Limitationen von Studie I 130
4.4 Ausblick 133

5 STUDIE II: THEORETISCHE GRUNDLAGEN 135

5.1 Überblick über verschiedene Instruktions- und Motivationsdesignmodelle 135
5.2 Vergleich zu Kellers „Macro Model of Motivation and Performance“ (2010) 137
5.2.1 Einführung zum „Macro Model of Motivation and Performance“ (Keller, 2010) 137
5.2.2 Vergleich zu den anderen Instruktions- und Motivationsdesignmodellen 138
5.3 Das „Macro Model of Motivation and Performance“ nach Keller (2010) 139
5.3.1 Aufmerksamkeit 141
5.3.2 Relevanz 142
5.3.3 Erfolgszuversicht 142
5.3.4 Zufriedenheit 143
5.3.5 Empirische Evidenz für das ARCS-Modell 144
5.4 Fragestellungen der zweiten Studie 147
5.4.1 Lernleistung, Nutzungszeit und Lerneffizienz 149
5.4.2 ARCS-Skala 152
5.4.3 Akzeptanz 154
5.4.4 Freiwillige Nutzung im Folgesemester 154

6 STUDIE II: METHODE 159

6.1 Stichprobe 159
6.1.1 Allgemeine Angaben 159
6.1.2 Interesse und Vorkenntnisse hinsichtlich der Vorlesung und des Projekts 162
6.1.3 Computerbezogene Fragen 163
6.1.4 Personenvariablen 165
6.2 Durchführung 167
6.2.1 Planung und Vorbereitung 167
6.2.2 Datenerhebung 168
6.3 Material 172
6.3.1 Lehrmaterial: E-Lectures und Web-based Trainings 172

6.3.2 Abhängige Variablen 181

7 STUDIE II: ERGEBNISSE UND DISKUSSION 191

7.1 Nutzungszeit, Lerneffektivität und -effizienz 192
7.1.1 Kurz- und längerfristige Lerneffektivität 192
7.1.2 Nutzungszeit und -häufigkeit 201
7.1.3 Kurz- und längerfristige Lerneffizienz 206
7.1.4 Zusammenfassung und Diskussion 211
7.2 ARCS-Skala 219
7.2.1 Skala „Aufmerksamkeit“ 220
7.2.2 Skala „Relevanz“ 222
7.2.3 Skala „Erfolgszuversicht“ 224
7.2.4 Skala „Zufriedenheit“ 226
7.2.5 Gesamtskala „ARCS“ 228
7.2.6 Zusammenfassung und Diskussion 231
7.3 Akzeptanz 237
7.3.1 Skala „Akzeptanz der Lernumgebung“ 237
7.3.2 Schulnote 240
7.3.3 Zusammenfassung und Diskussion 242
7.4 Freiwillige Nutzung im Folgesemester 244
7.4.1 Nutzungszeit und -häufigkeit 247
7.4.2 Lernleistung 251
7.4.3 Einschätzung zentraler Eigenschaften von E-Learning-Anwendungen 257
7.4.4 Zusammenfassung und Diskussion 260

8 STUDIE II: FAZIT 265

8.1 Lernleistung und Lerneffizienz 265
8.2 Akzeptanz 267
8.3 ARCS-Skala: Wahrnehmung der Lernunterstützung 268
8.4 Betrachtung der Ergebnisse in Bezug auf die vier Komponenten 268
8.5 Abschließendes Fazit zur Studie II 271
8.6 Limitationen von Studie II und Anregung für weitere Studien 272

9 ABSCHLIEßENDE DISKUSSION UND FAZIT 277

10 LITERATURVERZEICHNIS 283

11 ANHANG 299

EIDESSTATTLICHE ERKLÄRUNG 399

Abbildungsverzeichnis

Abbildung 1. Weltweite Anzahl an Internetnutzern insgesamt und in Bezug auf 100 Einwohner (International Telecommunication Union, 2015) 1
Abbildung 2. Der "Gartner E-Learning Hype Cycle 2000" (nach Dittler & Jechle, 2011) 12
Abbildung 3. Gründe für den Einsatz von E-Learning in den TOP 500-Unternehmen Deutschlands (Institut für Medien- und Kompetenzforschung, 2010) 16
Abbildung 4. E-Learning-Formen (Bundesministerium für Wirtschaft und Technologie, 2008). 23
Abbildung 5. Genutzte E-Learning-Formen in deutschen TOP 500-Unternehmen (Institut für Medien- und Kompetenzforschung, 2010) 27
Abbildung 6. Angabe der Nützlichkeit verschiedener digitaler Anwendungen für das Studium (nach Kleimann, Özkilic & Göcks, 2008) 28
Abbildung 7. Beispiel einer NetLecture-Produktion mit Darstellung ihrer Funktionen (vgl. Brömme, 2010) 29
Abbildung 8. Inhaltsverzeichnis des Web-based Trainings "Entwicklungspsychologie" 33
Abbildung 9. Eine typische Inhaltsseite des Web-based Trainings "Entwicklungspsychologie" 33
Abbildung 10. Übungsaufgabe mit Rückmeldung nach falscher Beantwortung aus dem WBT "Entwicklungspsychologie" 34
Abbildung 11. Abschließende Rückmeldungen am Ende der Übungstests 34
Abbildung 12. Verteilung der Studierenden auf die fünf Lehramtsstudiengänge sowie die Kategorie "Anderer Studiengang" (Angabe der gültigen Prozente) 57
Abbildung 13. Verteilung der Abiturnote innerhalb der Stichprobe 58
Abbildung 14. Gültige prozentuale Angaben der Studierenden auf die Frage: "Wo haben Sie Ihr Abitur erworben?" 58
Abbildung 15. Eingesetzte Fragebögen und Wissensüberprüfungen innerhalb der Studie I 74
Abbildung 16. Durchschnittliche Punktzahl der vier Nutzergruppen „Keine Nutzung“, „E-Lecture-Nutzung“, „WBT-Nutzung“ und „Kombinierte Nutzung“ in der Probeklausur sowie der Modulabschlussprüfung 84
Abbildung 17. Mittlere Ränge der Lehrmethoden E-Lecture, WBT und Vorlesung zu den Variablen „Größter Nutzen“, „Motivation“, „Spaß“ und „Konzentration“ 97
Abbildung 18. Prozentuale Häufigkeit des generellen Nutzungswunsches sowie des Nutzungwunsches zur Klausurvorbereitung von Vorlesung, E-Lecture, WBT, Übungsaufgaben, Vorlesungsskript und Literatur 99
Abbildung 19. Mittlere Einschätzungen zur Häufigkeit der Nutzung, Motivation und Spaß für Vorlesungsskript, Mitschrift, E-Lecture, WBT, Literatur und „Sonstiges“ 101
Abbildung 20. Mittlere Note für das Vorlesungsskript, die Mitschriften, die E-Lecture, das WBT, die Literatur und Sonstiges 102
Abbildung 21. Einteilung der Studierenden in die vier Nutzergruppen „Keine Nutzung“, „E-Lecture-Nutzer“, „WBT-Nutzer“ und „kombinierte Nutzung“ (in Prozent) 104
Abbildung 22. Durchschnittliche Nutzungszeit der E-Lectures und WBTs zu allen zwölf Themen 108
Abbildung 23. Kellers Macro Model of Motivation and Performance (nach Keller, 2010) 140
Abbildung 24. Altersverteilung der Stichprobe (n = 1163) 160
Abbildung 25. Prozentuale Verteilung der Studierenden auf die Lehramtsstudiengänge 160
Abbildung 26. Gültige prozentuale Angaben der Studierenden auf die Frage: "Wo haben Sie Ihr Abitur erworben? 161
Abbildung 27. Kategorisierte Vortestergebnisse der Studierenden 163
Abbildung 28. Prozentuale Häufigkeit der studentischen Angaben zu der Frage „Wie oft nutzen Sie Ihren Computer?“ 164
Abbildung 29. Kategorisierte Ergebnisse der Studierenden im CUSE-D-r 165

Abbildung 30. Überblick über die Datenerhebung zu Studie II 168
Abbildung 31. Beispielseiten aus dem 0-WBT zum Thema "Konditionierung" 173
Abbildung 32. Beispielseiten aus dem A-WBT zum Thema "Konditionierung" 175
Abbildung 33. Beispielseiten aus dem R-WBT zum Thema "Konditionierung" 176
Abbildung 34. Beispielseiten aus dem C-WBT zum Thema "Konditionierung" 178
Abbildung 35. Beispielseiten aus dem S-WBT zum Thema "Konditionierung" 180
Abbildung 36. Beispielanwendungsaufgabe aus der ersten Nachbefragung („Konditionierung“) 185
Abbildung 37. Überblick über die genutzten Fragebögen in Studie II 189
Abbildung 38. Durchschnittlicher Lernzuwachs durch die Bearbeitung der verschiedenen E-Learning-Varianten zu allen drei Themenbereichen 194
Abbildung 39. Längerfristiger Lernzuwachs der E-Learning-Versionen für alle drei Themen 197
Abbildung 40. Mittlere Nutzungsdauer (in Sek.) der E-Learning-Gruppen zu allen drei Themen 202
Abbildung 41. Mittlere Nutzungsdauer der E-Learning-Gruppen zu allen drei Themen 204
Abbildung 42. Kurzfristige Lerneffizienz der E-Learning-Gruppen für alle drei Themen 207
Abbildung 43. Längerfristige Lerneffizienz der E-Learning-Gruppe für alle drei Themen 209
Abbildung 44. Gewichtete Mittelwerte der Skala "Aufmerksamkeit" für die E-Learning-Gruppen zu allen drei Themen 220
Abbildung 45. Gewichtete Mittelwerte der Skala "Relevanz" für die E-Learning-Gruppen zu allen drei Themen 222
Abbildung 46. Gewichtete Mittelwerte der Skala "Erfolgszuversicht" für die E-Learning-Gruppen zu allen drei Themen 224
Abbildung 47. Gewichtete Mittelwerte der Skala "Zufriedenheit" für die E-Learning-Gruppen zu allen drei Themen 227
Abbildung 48. Gewichtete Mittelwerte zur gewichteten Gesamt-ARCS-Skala für die E-Learning-Gruppen zu allen drei Themen 229
Abbildung 49. Mittelwerte der Skala "Akzeptanz der Lernumgebung" für die E-Learning-Gruppen zu allen drei Themen 238
Abbildung 50. Mittlere „Schulnote“ der E-Learning-Versionen zu allen drei Themen 241
Abbildung 51. Durchschnittliche Nutzungszeit der E-Learning-Anwendungen für die 12 Themen 249
Abbildung 52. Durchschnittliche Nutzungszeit der E-Learning-Anwendungen getrennt für die Kalenderwochen 21 bis 30 relativiert durch die Anzahl der möglichen Nutzungstage 250
Abbildung 53. Durchschnittliches Gesamtergebnis der sechs Nutzergruppen in der Probeklausur 253
Abbildung 54. Durchschnittliche Punktzahl der Nutzergruppen in den beiden Klausurteilen der Modulabschlussprüfung 255
Abbildung 55. Einschätzung der Wichtigkeit der Eigenschaften von E-Learning-Anwendungen 259

Tabellenverzeichnis

Tabelle 1	Zusammenfassung der im Text genannten Vor- und Nachteile von Vorträgen	22
Tabelle 2	Vorteile von Präsenzlehre und E-Learning (nach Dittler & Jechle, 2009)	25
Tabelle 3	Akzeptanz des Projekts "Lehrinnovation" vor und nach der Durchführung	59
Tabelle 4	Mittlere Vorkenntnisse der Studierenden bezüglich des bearbeiteten Themas sowie Erfahrung mit der genutzten Lehrmethode	60
Tabelle 5	Abhängige Varianzanalysen mit Greenhouse-Geisser-Korrektur (ANOVA) zu den Variablen "Vorkenntnisse zum gestrigen Thema" sowie "Erfahrung mit der genutzten Lehrmethode“	60
Tabelle 6	Angaben der Studierenden zu computerbezogenen Items und zur Einschätzung von E-Learning	61
Tabelle 7	Übersicht über die Teilstichproben	62
Tabelle 8	Größe der zwölf Projektgruppen	62
Tabelle 9	Versuchsplan zu Studie I	63
Tabelle 10	Eingesetzte Software sowie Learning Management Systeme (LMS)	66
Tabelle 11	Aufteilung der Klausurfragen auf die elf Vorlesungsthemen bzw. sechs Themenkomplexe in den vier Klausurversionen A bis D sowie Angabe der Studiengänge der Klausurgruppe	70
Tabelle 12	Deskriptive Statistik sowie Mittelwertsanalysen zum Vergleich der Klausurversionen	71
Tabelle 13	Durchschnittliche Abiturnoten der Studierenden der fünf Lehramtsstudiengänge	71
Tabelle 14	Deskriptive Statistik zur kurzfristigen Lernleistung der Vorlesungs-, E-Lecture-, und WBT-Gruppe	78
Tabelle 15	Ergebnisse der Mittelwertsanalysen zum Vergleich der kurzfristigen Lernleistung der Vorlesungs-, E-Lecture-, und WBT-Gruppe	78
Tabelle 16	Deskriptive Statistik zur kurzfristigen Lernleistung, Lernzeit sowie der aus diesen Variablen berechneten kurzfristigen Lerneffizienz der Vorlesungs- und WBT-Gruppe für die Themen 7 bis 12	79
Tabelle 17	Ergebnisse der Mittelwertsanalysen zum Vergleich der kurzfristigen Lernleistung, Lernzeit sowie der hieraus abgeleiteten kurzfristigen Lerneffizienz der Vorlesungs- und WBT-Gruppe für die Themen 7 bis 12	80
Tabelle 18	Deskriptive Statistik zur längerfristigen Lernleistung der Vorlesungs-, E-Lecture- und WBT-Gruppe	81
Tabelle 19	Ergebnisse der Mittelwertsanalysen zum Vergleich der längerfristigen Lernleistung der Vorlesungs-, E-Lecture- und WBT-Gruppe	81
Tabelle 20	Deskriptive Statistik zur längerfristigen Lernleistung, Lernzeit sowie der längerfristigen Lerneffizienz der Vorlesungs- und WBT-Gruppe für die Themen 7 bis 12	82
Tabelle 21	Ergebnisse der Mittelwertsanalysen zum Vergleich der längerfristigen Lernleistung, Lernzeit sowie der längerfristigen Lerneffizienz der Vorlesungs- und WBT-Gruppe für die Themen 7 bis 12	82
Tabelle 22	Deskriptive Statistik zur langfristigen Lernleistung (Probeklausur vor der MAP sowie MAP selbst) der Nutzergruppen „Keine Nutzung“, „E-Lecture-Nutzer“, „WBT-Nutzer“ sowie „Kombinierte Nutzung“	84
Tabelle 23	Deskriptive Statistik zur langfristigen Lernleistung (Probeklausur vor der MAP sowie MAP selbst) für drei verschiedene Zufallsstichproben aus den Nutzergruppen „Keine Nutzung“, „E-Lecture-Nutzer“, „WBT-Nutzer“ sowie „Kombinierte Nutzung“	85
Tabelle 24	Ergebnisse der Mittelwertsanalysen zum Vergleich der langfristigen Lernleistung (Probeklausur vor der MAP sowie MAP selbst) der	

	Nutzergruppen „Keine Nutzung“, „E-Lecture-Nutzer“, „WBT-Nutzer“ sowie „Kombinierte Nutzung“ (jeweils drei Zufallsziehungen)	85
Tabelle 25	Deskriptive Statistik zur Lernförderlichkeit der Lehrmethoden Vorlesung, E-Lecture und WBT	91
Tabelle 26	Ergebnisse der Mittelwertsanalysen zum Vergleich der Lernförderlichkeitseinschätzung zur Vorlesung, E-Lecture und dem WBT	91
Tabelle 27	Deskriptive Statistik zur Qualität der Lernumgebung, Flexibilität sowie Gesamtakzeptanz von Vorlesung, E-Lecture und WBT	92
Tabelle 28	Ergebnisse der Mittelwertsanalysen zum Vergleich der Qualität der Lernumgebung, Flexibilität sowie Gesamtakzeptanz von Vorlesung, E-Lecture und WBT	92
Tabelle 29	Deskriptive Statistik zur Note für die Vorlesung, die E-Lecture und das WBT	93
Tabelle 30	Ergebnisse der Mittelwertsanalysen zum Vergleich der Note für die Vorlesung, die E-Lecture und das WBT	93
Tabelle 31	Deskriptive Statistik zu den in der Abschlussbefragung erfragten Variablen für die Vorlesung, die E-Lecture, das WBT, die Übungsaufgaben sowie die Literatur	94
Tabelle 32	Ergebnisse der Mittelwertsanalysen zum Vergleich der in der Abschlussbefragung erfragten Variablen für die Vorlesung, die E-Lecture, das WBT, die Übungsaufgaben sowie die Literatur	95
Tabelle 33	Deskriptive Statistik zur Note für die Vorlesung, die E-Lecture, das WBT, die Übungsaufgaben, das Vorlesungsskript sowie die Literatur	96
Tabelle 34	Ergebnisse der Mittelwertsanalysen zum Vergleich der Note von Vorlesung, E-Lecture, WBT, Übungsaufgaben, Vorlesungsskript und Literatur	96
Tabelle 35	Deskriptive Statistik zu den in der längerfristigen Befragung erhobenen Akzeptanzvariablen zu Vorlesung, E-Lecture und WBT	97
Tabelle 36	Ergebnisse der Mittelwertsanalysen zum Vergleich der in der längerfristigen Befragung erhobenen Akzeptanzvariablen zu Vorlesung, E-Lecture und WBT	98
Tabelle 37	Deskriptive Statistik zum generellen Nutzungswunsch bzw. dem Nutzungswunsch zur Klausurvorbereitung von Vorlesung, E-Lecture, WBT, Übungsaufgaben, Vorlesungsskript sowie Literatur	98
Tabelle 38	Ergebnisse der Mittelwertsanalysen zum Vergleich des generellen Nutzungwunsches bzw. des Nutzungswunsches zur Klausurvorbereitung von Vorlesung, E-Lecture, WBT, Übungsaufgaben, Vorlesungsskript sowie Literatur	99
Tabelle 39	Deskriptive Statistik zu den abschließenden Akzeptanzvariablen zu Vorlesungsskript, Mitschrift, E-Lecture, WBT, Literatur sowie „Sonstiges“	101
Tabelle 40	Ergebnisse der Mittelwertsanalysen zum Vergleich der in der abschließenden Befragung erhobenen Akzeptanzvariablen zu Vorlesungsskript, Mitschrift, E-Lecture, WBT und Literatur	102
Tabelle 41	Nutzungen der Nutzergruppen „Keine Nutzung“, „E-Lecture-Nutzer“, „WBT-Nutzer“ und „Kombinierte Nutzung“ für alle zwölf Themen sowie insgesamt	104
Tabelle 42	Durchschnittliche Nutzungshäufigkeit der E-Lectures und WBTs für alle zwölf Themen sowie insgesamt (Summe und durchschnittliche Nutzung)	106
Tabelle 43	Ergebnisse der Mittelwertsanalysen zum Vergleich der mittleren Nutzungshäufigkeit der E-Lecture- und WBT-Nutzer (drei Zufallsziehungen)	107
Tabelle 44	Durchschnittliche Nutzungszeit der E-Lectures und WBTs getrennt für alle zwölf Themen sowie insgesamt (Summe und durchschnittliche Nutzungszeit)	107
Tabelle 45	Ergebnisse der Mittelwertsanalysen zum Vergleich der mittleren Nutzungszeit der E-Lecture- und WBT-Nutzer	108
Tabelle 46	Durchschnittliche Gesamtnote in der kurzfristigen, längerfristigen sowie abschließenden Evaluation zu den insgesamt sieben verschiedenen Lehrmaterialien/-methoden	109

Tabelle 47	Deskriptive Statistik zur Vor- und Nachbereitungszeit, Pausenhäufigkeit und -zeit sowie dem Machen von Notizen während der Vorlesung, der Bearbeitung der E-Lecture und dem WBT	119
Tabelle 48	Ergebnisse der Mittelwertsanalysen zum Vergleich der Vor- und Nachbereitungszeit, Pausenhäufigkeit und -zeit sowie Notizenhäufigkeit während der Vorlesung, der Bearbeitung der E-Lecture und dem WBT	119
Tabelle 49	Deskriptive Statistik der Variablen zur Beurteilung der Übungsaufgaben	121
Tabelle 50	Deskriptive Statistik und Ergebnisse der Mittelwertsanalysen zu den Lernleistungsvariablen sowie den signifikanten Akzeptanzvariablen zum Vergleich der WBTs mit und ohne Übungsaufgaben	122
Tabelle 51	Deskriptive Statistik sowie Ergebnisse der Mittelwertsanalysen zum Vergleich der Lernzeit und der kurz- und längerfristigen Lerneffizienz von WBTs mit und ohne Übungsaufgaben	123
Tabelle 52	Deskriptive Statistik sowie Ergebnisse der Mittelwertsanalysen zum Vergleich der anhand der Inhaltsseiten relativierten Lernzeit und Lerneffizienz von WBTs mit und ohne Übungsaufgaben	124
Tabelle 53	Deskriptive Statistik der Personen-Skalen für die Gesamtstichprobe	166
Tabelle 54	Experimentalgruppen der Studie II mit Variationen über die drei Erhebungen	169
Tabelle 55	Zeitraum der Evaluationswellen sowie Anzahl der Studierenden, die jeweils teilgenommen haben	170
Tabelle 56	Überblick über die Vortestitems	182
Tabelle 57	Grundlegender Aufbau der drei kurzfristigen Nacherhebungen	182
Tabelle 58	Beispielitems und deskriptive Statistiken zu den vier Skalen des ARCS-Fragebogens sowie dem Fragebogen insgesamt	183
Tabelle 59	Korrelationen der ARCS-Skalen mit den Validierungsskalen	184
Tabelle 60	Deskriptive Statistik zu den Vor-, Nachtestergebnissen und dem Lernzuwachs für die E-Learning-Gruppen getrennt für die drei Themenbereiche sowie insgesamt	193
Tabelle 61	Ergebnisse der Mittelwertsanalysen zum Vergleich der Vor- und Nachtestergebnisse der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt	195
Tabelle 62	Ergebnisse der Kovarianzanalysen zum Vergleich der Nachtestergebnisse der E-Learning-Gruppen unter Einbezug der Kovariate „Vortestergebnis“ getrennt für die drei Themen sowie insgesamt	196
Tabelle 63	Deskriptive Statistik zur längerfristigen Lernwirksamkeit der E-Learning-Gruppen für die drei Themen sowie insgesamt	197
Tabelle 64	Ergebnisse der Mittelwertsanalysen zum Vergleich der längerfristigen Lernwirksamkeit der E-Learning-Versionen getrennt für die drei Themen sowie insgesamt	198
Tabelle 65	Ergebnisse der Kovarianzanalysen zum Vergleich der Probeklausurergebnisse der E-Learning-Gruppen unter Einbezug der Kovariate „Vortestergebnis“ getrennt für die drei Themen sowie insgesamt	199
Tabelle 66	Ergebnisse der Mittelwertsanalysen zum Vergleich des Vortestergebnisses, der längerfristigen Lernwirksamkeit sowie des längerfristigen Lernzuwachses zu den Themen, die durch die Vorlesung und den Themen, die durch E-Learning-Anwendungen gelehrt wurden	200
Tabelle 67	Deskriptive Statistik zur Nutzungszeit (in Sek.) aller E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt	201
Tabelle 68	Ergebnisse der Mittelwertsanalysen zum Vergleich der Nutzungszeit der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt	202
Tabelle 69	Ergebnisse der Mittelwertsanalysen zum Vergleich der Nutzungszeit aller Gruppen zwischen den Themen	203

Tabelle 70 Deskriptive Statistik zur Nutzungshäufigkeit der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 204
Tabelle 71 Ergebnisse der Mittelwertsanalysen zum Vergleich der Nutzungshäufigkeit der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 205
Tabelle 72 Ergebnisse der Mittelwertsanalysen zum Vergleich der Nutzungshäufigkeit aller Gruppen zwischen den Themen 205
Tabelle 73 Deskriptive Statistik zur kurzfristigen Lerneffizienz der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 206
Tabelle 74 Ergebnisse der Mittelwertsanalysen zum Vergleich der kurzfristigen Lerneffizienz der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 207
Tabelle 75 Deskriptive Statistik zur längerfristigen Lerneffizienz der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 208
Tabelle 76 Ergebnisse der Mittelwertsanalysen zum Vergleich der längerfristigen Lerneffizienz der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 209
Tabelle 77 Ergebnisse der Mittelwertsanalysen zum Vergleich der längerfristigen Lernwirksamkeit und -effizienz zu den Themen, die durch die Vorlesung und den Themen, die durch E-Learning-Anwendungen gelehrt wurden 210
Tabelle 78 Deskriptive Statistik zur gewichteten Skala "Aufmerksamkeit" für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 220
Tabelle 79 Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Aufmerksamkeit“ zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 221
Tabelle 80 Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Aufmerksamkeit“ über alle Gruppen zwischen den Themen 221
Tabelle 81 Deskriptive Statistik zur gewichteten Skala "Relevanz" für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 222
Tabelle 82 Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Relevanz“ zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 223
Tabelle 83 Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Relevanz“ über alle Gruppen zwischen den Themen 223
Tabelle 84 Deskriptive Statistik zur gewichteten Skala "Erfolgszuversicht" für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 224
Tabelle 85 Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Erfolgszuversicht“ zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 225
Tabelle 86 Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Erfolgszuversicht“ über alle Gruppen zwischen den Themen 226
Tabelle 87 Deskriptive Statistik zur gewichteten Skala "Zufriedenheit" für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 226
Tabelle 88 Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Zufriedenheit“ zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 227
Tabelle 89 Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Zufriedenheit“ über alle Gruppen zwischen den Themen 228
Tabelle 90 Deskriptive Statistik zur gewichteten Gesamt-ARCS-Skala für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 229
Tabelle 91 Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Gesamt-ARCS-Skala zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 230

Tabelle 92 Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Gesamt-ARCS-Skala über alle Gruppen zwischen den Themen 230
Tabelle 93 Deskriptive Statistik zur Skala "Akzeptanz der Lernumgebung" für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 238
Tabelle 94 Ergebnisse der Mittelwertsanalysen zum Vergleich der Skala „Akzeptanz der Lernumgebung“ zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 239
Tabelle 95 Ergebnisse der Mittelwertsanalysen zum Vergleich der Skala „Akzeptanz der Lernumgebung“ über alle Gruppen zwischen den Themen 240
Tabelle 96 Deskriptive Statistik zur Variable „Schulnote“ für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 240
Tabelle 97 Ergebnisse der Mittelwertsanalysen zum Vergleich der Variable „Schulnote“ zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt 241
Tabelle 98 Ergebnisse der Mittelwertsanalysen zum Vergleich der Variable „Schulnote“ über alle Gruppen zwischen den Themen 242
Tabelle 99 Hauptsächliche Nutzung der Lernmaterialien zur Vorbereitung auf die MAP 244
Tabelle 100 Deskriptive Statistik zur Nutzungszeit der E-Learning-Anwendungen für die verschiedenen Selbstauskunft-Gruppen 245
Tabelle 101 Ergebnisse der Mittelwertsanalysen zum Vergleich der Nutzungszeiten der drei E-Learning-Anwendungen zwischen der Vorlesungs-, der E-Lecture- und den beiden WBT-Gruppen 246
Tabelle 102 Deskriptive Statistik sowie Summe zur Gesamtnutzungszeit und Nutzungshäufigkeit der E-Learning-Nutzergruppen 247
Tabelle 103 Deskriptive Statistik zu den Testleistungen der Nutzergruppen bezüglich der zwölf Themen der Probeklausur 252
Tabelle 104 Deskriptive Statistik zur Lernleistung der Nutzergruppen in der Probeklausur 253
Tabelle 105 Ergebnisse der Mittelwertsanalysen zum Vergleich der Gesamttestleistung in der Probeklausur der Vorlesungs-, der E-Lecture- und der beiden WBT-Gruppen 254
Tabelle 106 Deskriptive Statistik zu den Leistungen der Nutzergruppen in den beiden Teilen der MAP sowie der MAP insgesamt 255
Tabelle 107 Ergebnisse der Mittelwertsanalysen zum Vergleich der Leistungen der Vorlesungs-, E-Lecture- und der beiden WBT-Gruppen in den beiden Teilen der MAP 256
Tabelle 108 Ergebnisse der Mittelwertsanalysen zum Vergleich der Gesamtpunkte sowie der Note der Vorlesungs-, E-Lecture- und der beiden WBT-Gruppen in der MAP 257
Tabelle 109 Durchschnittliche Einschätzung der Wichtigkeit von 20 Eigenschaften von E-Learning-Anwendungen sowie die Angabe, zu welcher ARCS-Komponente diese Eigenschaft nach Keller (2010) gehört 258

Abkürzungsverzeichnis

%	Prozentualer Anteil
0-WBT	WBT, in dem keine der Strategien nach Keller (2010) umgesetzt wurden
ANOVA	Einfaktorielle Varianzanalyse
A-WBT	WBT, in dem die Aufmerksamkeitsstrategien nach Keller (2010) umgesetzt wurden
ARCS-WBT	WBT, in dem alle Strategien nach Keller (2010) umgesetzt wurden
C-WBT	WBT, in dem die Erfolgszuversichtsstrategien nach Keller (2010) umgesetzt wurden
df	Freiheitsgrade
EL	E-Lecture
F	*F*-Wert (Prüfstatistik der Varianzanalyse bzw. des Brown-Forsythe-Tests)
IMMS	Instructional Material Motivational Survey von Keller (2006)
Lit	Literatur
LMS	Lehr-Management-System
M	Mittelwert
MAP	Modulabschlussprüfung
Mit	Mitschrift aus der Vorlesung
N/n	Stichprobengröße
p	Signifikanz
R-WBT	WBT, in dem die Relevanzstrategien nach Keller (2010) umgesetzt wurden
S-WBT	WBT, in dem die Zufriedenheitsstrategien nach Keller (2010) umgesetzt wurden
SD	Standardabweichung
t	*t*-Wert (Prüfstatistik des *t*- bzw. des Welch-Tests)
VL	Vorlesung
VLSk	Vorlesungsskript
WBT	Web-based Training

1 Theoretische Grundlagen

1.1 Einleitung

Das Internet und der Computer spielen eine immer gewichtigere Rolle in unserer Gesellschaft. Wir zahlen Rechnungen online, kaufen unsere Weihnachtgeschenke bei *Amazon*, planen Autoreisen mit *Google Maps*, *skypen* mit Freunden oder erledigen unsere Steuererklärung mit *ELSTERWEB*[1] (elektronische Steuererklärung). Diese Aufzählung könnte beliebig fortgeführt werden.

Heute verfügen 82% aller europäischen Haushalte über einen Internetanschluss und auf 100 Europäer fallen etwa 120 Mobiltelefone. Die Beobachtung der letzten Jahre lässt vermuten, dass die Nutzung von Internet und Computer – aber insbesondere von Tablets und Smartphones – weiterhin steigen wird. Abbildung 1 zeigt die weltweite Entwicklung der Internetnutzung insgesamt sowie auf 100 Einwohner bezogen. Mittlerweile nutzen etwa 43% der Weltbevölkerung – vornehmlich selbstverständlich in den Industrienationen – das Internet (International Telecommunication Union, 2015).

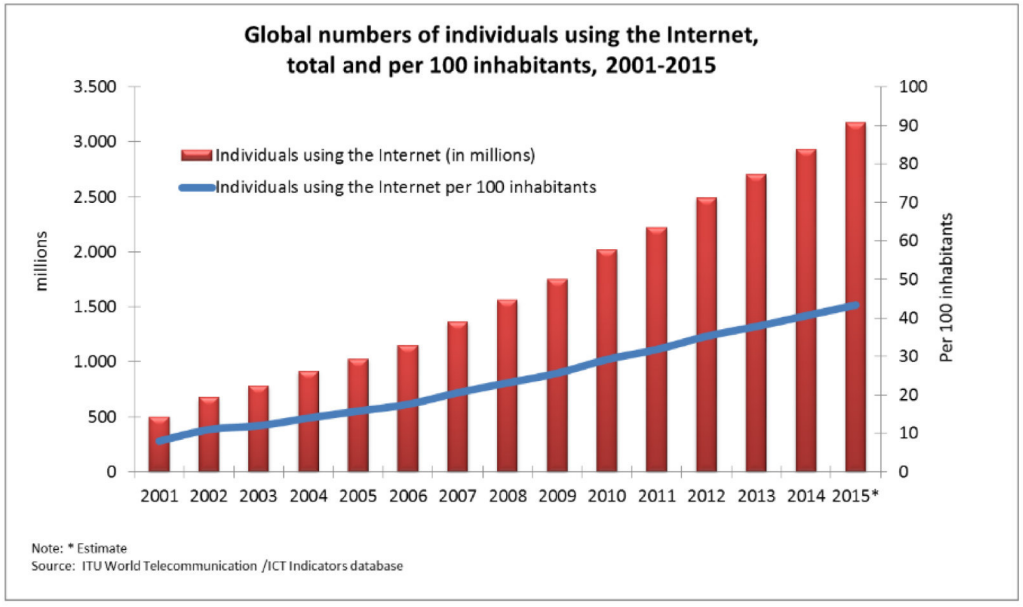

Abbildung 1. Weltweite Anzahl an Internetnutzern insgesamt und in Bezug auf 100 Einwohner (International Telecommunication Union, 2015)

So ist es nicht überraschend, dass Internet und Computer auch in der Aus- und Weiterbildung eine immer wichtigere Rolle spielen – auch wenn die Lehre in Schule, Hochschule und betrieblicher Bildung noch immer vornehmlich durch Präsenzlehre sowie über Selbststudium mit Literatur vonstattengeht. Die neuen Informations- und Kommunikationsmittel sind zu einem maßgeblichen Auslöser für einen enormen

[1] **EL**ektronische **ST**euer**ER**klärung, siehe https://www.elster.de/

Qualifizierungsbedarf geworden, da sie heute in den meisten Berufen fester Bestandteil der täglichen Arbeit sind. Eine gute Medienbildung ist folglich eine wesentliche Voraussetzung für die berufsqualifizierende Bildung. Die neuen Medien können allerdings auch dazu beitragen, den zusätzlichen Qualifizierungsbedarf effektiv und effizient zu unterstützen (Köhler, Kahnwald & Reitmeier, 2008; Weinmann, 1998; siehe auch Glowalla & Schoop, 1996).

Nicht nur deshalb wächst gerade im Bereich der Hochschulen der Bedarf an E-Learning. Aktuell wurden durch die Verkürzung der Schulzeit aufgrund von G8[2] Doppeljahrgänge auf die Hochschulen gespült – und das in einer Zeit, in der schon demographisch eine größere Nachfrage nach Studienplätzen herrscht. Dadurch werden in der Spitzenzeit über 400.000 Studienplätze mehr benötigt. Dies setzt die Universitäten unter Druck, Maßnahmen zu ergreifen, die geeignet sind, die Diskrepanz zwischen Angebot und Nachfrage zu reduzieren (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2005).

Das Internet spielt folglich nicht nur als Informations- und Kommunikationsmedium eine bedeutende Rolle, sondern wird auch zum Anbieter von Lernangeboten (Beer, Hamburg, Lindecke & Terstriep, 2003), da es für die neuen Anforderungen neue Möglichkeiten zur Aufbereitung und Präsentation von Lerninhalten sowie zur Gestaltung der Vermittlungsprozesse bietet (Bundesministerium für Bildung und Forschung, 2000). Dies wird durch E-Learning aber nicht per se erreicht. Hierzu ist es notwendig, sinnvolle, lernförderliche und möglichst preiswerte E-Learning-Anwendungen zu erstellen und diese mit Mehrwert in bestehende Curricula zu integrieren.

Aus diesem Grunde beschäftigt sich die vorliegende Dissertation mit der lernförderlichen Gestaltung von E-Learning-Angeboten – im Rahmen von Studie II insbesondere von Web-based Trainings, die eine sehr prominente Form von E-Learning-Anwendungen darstellen und ausgesprochen flexibel gestaltet werden können.

Um einen theoretischen Rahmen zu geben, soll zunächst auf E-Learning sowie seine Bedeutung für Schule, Hochschule und betriebliche Aus- und Weiterbildung eingegangen werden. Danach werden die in der vorliegenden Dissertation untersuchten Lehrformen Vorlesung, E-Lecture und Web-based Training beschrieben sowie das Untersuchungsdesign der ersten Studie hergeleitet.

[2] Auch „achtjähriges Gymnasium“

1.2 E-Learning

In der Literatur finden sich einige Begriffe, die das Lernen mit Computer und Internet beschreiben. Der Begriff „E-Learning“ (für „Electronic Learning“ oder „Internet-Enabled Learning“) stellt vermutlich einen der meist genutzten Begriffe dar. Google liefert hierzu in nur einer halben Sekunde 314.000.000 Treffer. Weitaus häufiger genutzt werden jedoch die Begriffe „Online Learning“ (1.150.000.000 Treffer) oder auch „Net Learning“ (726.000.000 Treffer), die jedoch genauso wie „netzbasiertes Lernen“, „Cyber Learning“ oder „Online Lernen“ lediglich die Internet- oder Online-Formen des Lernens mit neuen Medien umschreiben. Weitere weniger häufig genutzte Begriffe stellen zum Beispiel das „Virtual Learning“ oder „virtuelle Lernen“, das „Telelernen“ oder das „technologiegestützte Lernen“ dar (Döring, 2002; Wache, 2003).
Im weiteren Verlauf dieser Dissertation wird der Begriff „E-Learning“ genutzt, der sich sowohl im Deutschen als auch im Englischen als gängiger Begriff herauskristallisiert hat und auch Offline-Formen des Lernens mit neuen Medien umspannt. Dieser Begriff soll im Folgenden definiert und die Entwicklung von E-Learning erläutert werden.

1.2.1 Definition und Entwicklung

In der Literatur finden sich viele verschiedene – und doch sehr ähnliche Definitionen – für den Begriff „E-Learning“. Beispielhaft sollen im Folgenden drei Definitionen gegeben werden:
Kaltenbaek (2009, S. 370) fasst sich ganz kurz. Ihm zufolge versteht man unter E-Learning ein „Lernen unterstützt durch elektronische Medien (v. a. Computern) in Offline- und Online- (netzgestützten) Settings“.
Für Kleimann und Schmid (2006, S. 12) ist „E-Learning“ „[…] die Gesamtheit aller Lehr- und Lernformen […], die Fern- oder Präsenzstudium durch synchrone oder asynchrone digitale Medien unterstützen.“
Nach Döring (2002, S. 247) werden „Lernprozesse, die sich soziotechnisch ganz oder teilweise auf das Internet stützen, […] summarisch oft als „E-Learning“ oder „Online-Lernen“ bezeichnet.“
Allen gemeinsam ist „der Einsatz von elektronischen Medien zu Lehr- und Lernzwecken“. Diese Definition soll im Folgenden zugrundegelegt werden.

Die Idee des „E-Learning“ ist nicht neu. Schon vor über 100 Jahren wurden Kurse angeboten, die eine Art „Fernlernen“ ermöglichen sollten (Hamburg & Terstriep, 2003). Weitere

Vorläufer des heutigen „E-Learning“ waren Lernangebote auf Audiokassetten oder das Bildungsfernsehen (Wache, 2003).
Erst in den frühen 90er Jahren des 20. Jahrhunderts bekam die Idee des E-Learning durch die Verbreitung des Computers einen immensen Schub. E-Learning wurde in Form von Computer-based Trainings (CBTs) realisiert, die den Lernenden allerdings zumeist einen strikten Trainingsablauf vorgaben und ausschließlich auf Selbstlernprozesse beschränkt waren (Hamburg & Terstriep, 2003). Das enttäuschte Viele, die auf diese neue Form des Lernens gesetzt hatten. Nach Connolly und Stansfield (2007) lag dies auch daran, dass in dieser ersten „E-Learning-Generation“ traditionelle Materialien unbedacht in ein Online-Format gebracht wurden, anstatt die neuen Möglichkeiten auszuschöpfen. Die zweite Generation der E-Learning-Anwendungen, die etwa zur Jahrtausendwende entstand, konnte auf immer höhere Bandbreiten und eine bessere Technik zurückgreifen. Dadurch entstand eine Bewegung, die sich bemühte, virtuelle Lernumgebungen zu erschaffen, die Zugang zu Lernmaterialien, Kommunikationsmöglichkeiten und eine onlinebasierte Administration ermöglichen. Dies gestattete nun sehr flexible Formen des elektronischen Lernens (Connolly & Stansfield, 2007; Hamburg & Terstriep, 2003).
Dadurch, dass die technische Entwicklung immer weiter voranschreitet, entstehen immer wieder neue Möglichkeiten zur Entwicklung von E-Learning-Anwendungen. So befinden wir uns nach Connolly und Stansfield (2007) derzeit in der dritten „E-Learning-Generation“, welche beispielsweise E-Portfolios, Wikis, Blogs, Game-based Learning oder Online-Simulationen – zunehmend aber auch Mobile Learning (mLearning) über PDAs, Tablets oder Smartphones – einsetzt.
Generell lässt sich folglich zusammenfassen, dass sich das Verständnis von „E-Learning“ in den letzten Jahren stark verändert hat. Verstand man Anfang des Jahrtausends darunter zumeist noch den Einsatz von PowerPoint Folien oder Videos in Präsenzseminaren, spricht man heute vermehrt über Web 2.0-Tools.
Im folgenden Abschnitt sollen nun die Entwicklung sowie der aktuelle Stand des E-Learning-Einsatzes in den drei Bildungskontexten Schule, Hochschule sowie betriebliche Aus- und Weiterbildung dargestellt werden, um in der Folge anhand dieser einen Überblick über die Vor- und Nachteile des E-Learning-Einsatzes zu geben.

1.2.2 E-Learning in Schule, Hochschule und betrieblicher Bildung

Auch wenn ein Großteil der aktuellen Schülergeneration hinsichtlich der Nutzung von Smartphones, Computern und Tablets sicherlich schon sehr geübt ist, heißt das noch lange nicht, dass sie auch das Lernen und Arbeiten mit diesen neuen Medien beherrschen. Nach Knop, Hefner, Schmitt und Vorderer (2015) können 64% der 8- bis 14-Jährigen und 86%

der 13-/14-Jährigen über ein Smartphone auf das Internet zugreifen. Sie nutzen diese neuen Medien hauptsächlich zur Kommunikation, zum Spielen oder zum Musik hören. Vergleichbar besitzen laut Statista (2015) 76% der 12- bis 19-Jährigen Deutschen auch einen eigenen Computer oder Laptop. Eine Umfrage zu Online-Aktivitäten am PC bzw. am Laptop in der Schweiz (Statista, 2014) zeigte, dass ein Großteil diese Medien nutzt, um Websites wie z.B. Newsportale zu besuchen oder Filme, Fernsehen sowie Radio zu konsumieren. Folglich nutzen nur wenige die neuen Medien im privaten Bereich, um zu lernen oder zu arbeiten. Junge Menschen müssen deshalb auch heute noch in ihrer Schul- und Hochschulzeit bzw. der beruflichen Ausbildung (Bundesministerium für Bildung und Forschung, 2000) auf die neuen Anforderungen der Wissensgesellschaft vorbereitet werden. Denn die Anforderungen in unserer Arbeitswelt haben sich im Übergang von der Industrie- zur Informationsgesellschaft stark verändert. Wissen und Qualifikation spielen eine immer größere Rolle. Gleichzeitig ist die „Halbwertszeit" von Wissen stark gesunken (Glowalla & Schoop, 1996, Weinmann, 1998). Aus dem „Lernen auf Vorrat" ist ein „Lernen nach Bedarf" – „just in time" und „just in place" – geworden (Bundesministerium für Bildung und Forschung, 2000; Ihringer, 2000; Kaltenbaek, 2009; Keller & Back, 2004). Dies erhöht auch für Arbeitnehmer den Bildungsbedarf. Sie werden in die Rolle versetzt, lebenslang zu lernen (Bundesministerium für Bildung und Forschung, 2000). Für Unternehmen heißt dies, dass Mitarbeiter „on demand[3]" zielbezogen auf bevorstehende Aufgaben vorbereitet werden müssen. Mit traditionellen Medien und Präsenzlehre ist dies nur noch schwer zu bewältigen (Beer et al., 2003; Bundesministerium für Bildung und Forschung, 2000; Dittler & Jechle, 2009; Keller & Back, 2004).

Speziell in der Hochschule muss darüber hinaus auch deshalb ein Strukturwandel stattfinden, um die internationale Wettbewerbsfähigkeit der Hochschulen und Studierenden zu erhöhen (Bundesministerium für Bildung und Forschung, 2000; siehe auch „Bologna-Prozess"). Dies ist unter anderem durch den Einsatz neuer Medien in der Schule und Hochschule möglich.

An dieser Stelle sollen deshalb nun die Entwicklung sowie der aktuelle Stand des E-Learning-Einsatzes in Schule, Hochschule und betrieblicher Bildung kurz dargestellt werden, bevor in der Folge die Vor- und Nachteile dieses Einsatzes thematisiert werden.

1.2.2.1 E-Learning in der Schule

Im Jahr 2000 finanzierte die Bundesregierung durch das Handlungskonzept „IT in der Bildung – Anschluss statt Ausschluss" verschiedene Maßnahmen, wie z.B. die Programme

[3] Dies bezeichnet das spontane Lernen zu jeder Zeit und an jedem Ort – dann, wenn es notwendig ist. Wissen wird für Mikro-Lernprozesse dann bereitgestellt, wenn es gebraucht wird. Mittlerweile ist dies auch mobil möglich (dann spricht man von „M-Learning").

„Neue Medien in der Bildung“ oder „Schulen ans Netz“, um den Einsatz neuer Medien in der Schule zu fördern (Bundesministerium für Wirtschaft und Arbeit & Bundesministerium für Bildung und Forschung, 2006). Dadurch konnten unter anderem bis zum Herbst 2001 alle bundesdeutschen Schulen mit einem Internetanschluss versorgt und die Anzahl der SchülerInnen pro Computer im Durchschnitt halbiert werden (Bundesministerium für Wirtschaft und Arbeit & Bundesministerium für Bildung und Forschung, 2006).

Deutschland nimmt aber weiterhin im internationalen Vergleich keinen Spitzenrang hinsichtlich der Nutzung neuer Medien in der Schule ein (Bundesministerium für Bildung und Forschung, 2000). Die Gründe für den immer noch eher zögerlichen Einsatz sind offensichtlich: Zum einen ist es in Deutschland für jedes Kind möglich, eine Schule zu erreichen. Der Druck, E-Learning einsetzen zu *müssen*, ist folglich nicht gegeben. Dies ist in anderen Ländern, in denen Lernende weit verstreut wohnen, sicherlich nicht in der Form gegeben, weshalb diese vermehrt auch im Schulbereich auf E-Learning-Szenarien setzen. Zum anderen hat sich die technische Hard- und Software in den letzten Jahren rasant weiterentwickelt. Dieses Tempo können viele Schulen nicht mithalten. Es fehlt an finanziellen Mitteln und an Know-how. So kommt es innerhalb der Schulen zu einer Überalterung der ohnehin schon spärlichen technischen Ausstattung (Bundesministerium für Bildung und Forschung, 2000).

Hinzu kommen Defizite in der Lehreraus- und -fortbildung. Lehrer sind noch immer nicht darauf vorbereitet, neue Medien im Unterricht mit didaktischem Mehrwert einzusetzen (Bundesministerium für Bildung und Forschung, 2000). So zeigte eine Lehrerbefragung von Breiter, Aufenanger, Averbeck, Welling und Wedjelek (2013), dass die selbsteingeschätzte Medienkompetenz von Lehrkräften insbesondere hinsichtlich der neuen Medien (Computer und Internet) noch ausbaufähig ist – wobei zu beachten ist, dass sie sich im Vergleich zu einer ähnlichen Befragung aus dem Jahre 2000 bereits deutlich gebessert hat. Es ist folglich notwendig, die Medienkompetenz des Lehrpersonals weiter zu stärken. Breiter et al. (2013) fordern deshalb (neben anderen Aspekten zur Förderung der Medienbildung in der Schule) eine verbindliche Integration von Medienbildung in die Ausbildung von Lehrkräften und Schulleitungen.

Allerdings darf – wie oben bereits erläutert – auch nicht davon ausgegangen werden, dass alle SchülerInnen in der Lage sind, neue Medien kompetent zu nutzen und E-Learning-Anwendungen selbstgesteuert zu bearbeiten. Neben den oben genannten Einschränkungen wird es immer noch SchülerInnen geben, die wenig Erfahrung mit Computer und Internet gesammelt haben oder die durch das Internet überfordert oder sogar gelangweilt sind (Döring, 2002).

Insgesamt geht die Entwicklung an den deutschen Schulen allerdings in die richtige Richtung. Durch Patenschaften von Unternehmen mit Schulen (*Public-Private Partnerships*;

Bundesministerium für Bildung und Forschung, 2000) und das Angebot der *Deutschen Telekom AG*, allen Schulen den Netzzugang und die Internetnutzung entgeltfrei zu ermöglichen, ist im Bereich der Ausstattung der Schulen schon Einiges passiert. Hinsichtlich der Hard- und Software-Ausstattung existiert jedoch Nachholbedarf. Außerdem fehlt es immer noch an sinnvollen Konzepten zur Einbettung der neuen Medien in die Schule und auch die Maßnahmen zur Pflege, technischen Unterstützung, Beratung und Qualifizierung der Lehrkörper müssen weiter ausgebaut werden (Bundesministerium für Bildung und Forschung, 2000).

1.2.2.2 E-Learning an Hochschulen

Die Wunschvorstellung des Bundesministeriums für Bildung und Forschung (BMBF) in Bezug auf die Hochschullehre der Zukunft sah im Jahr 2000 so aus: Studierende können sich die besten Kurse verschiedener Anbieter aussuchen. Diese sind so zertifiziert, dass sie von allen beteiligten Hochschulen anerkannt werden. Das Abschlussexamen kann dann an einer beliebigen Hochschule abgelegt werden. Dafür müssen die Hochschulen verstärkt miteinander kooperieren. E-Learning-Module gibt es für einzelne Lehrinhalte oder für ganze Lernveranstaltungen. Nach Bedarf können hieraus Lehrveranstaltungen zusammengestellt werden. HochschullehrerInnen erstellen ihre Materialien ganz selbstverständlich digital und veröffentlichen sie im hochschuleigenen Intranet. Auch werden diese Materialien mit kooperierenden Kollegen ausgetauscht. Selbst dem BMBF war schon damals klar, dass eine solche Hochschulwelt noch entwickelt werden muss (Bundesministerium für Bildung und Forschung, 2000).

Die Realität sieht nach wie vor anders aus. Die Lehre an den deutschen Hochschulen steht und fällt immer noch mit den Präsenz-Vorlesungen – und das nicht ganz ohne Berechtigung (siehe auch Abschnitt „1.3 Traditionelle Lehre: die Vortragsmethode“). Vorlesungen haben sich schon in der vorschriftlichen Zeit vor über 2000 Jahren bewährt. Fast ausschließlich war man damals auf die mündliche Weitergabe von Wissen angewiesen. Etwa ab dem 8. Jahrhundert standen handgeschriebene Bücher und Manuskripte zur Verfügung und mit der Erfindung des Buchdrucks durch Gutenberg 1450 konnte Wissen dann endlich relativ kostengünstig vervielfältigt werden. Neben der Vorlesung trat auch die Bearbeitung von Literatur in den Vordergrund (Kaltenbaek, 2009).

Durch Wilhelm von Humboldt wurden der Hochschullehre zusätzlich Seminare zugefügt. An den Hochschulen sollte nicht mehr nur gelehrt werden – Hochschullehrer sollten nun gemeinsam mit ihren Studierenden forschen (Kaltenbaek, 2009).

Die Bereitstellung von Kopiergeräten in den 80er Jahren des 20. Jahrhunderts erleichterte dann die Vervielfältigung von eigenen Manuskripten. Zudem konnte man nun anstelle von Diavorträgen Videopräsentationen zeigen (Kaltenbaek, 2009).

Anfang der 90er Jahre begann schließlich der Einsatz von Multimedia und Internet an den Hochschulen. Seitdem entstanden verschiedene virtuelle Hochschulen und an den Präsenzhochschulen wurden Anwendungsprogramme wie z.B. Microsoft PowerPoint™ innerhalb von Präsenzveranstaltungen eingesetzt (Kaltenbaek, 2009). Ab 2002 begannen die Hochschulen, ihren Internetzugang auszubauen. Im Jahr 2006 verkündeten Kleimann und Schmid als Fazit einer von ihnen durchgeführten Befragung von Mitgliedern aller deutschen Hochschulen zum Einsatz von neuen Medien in Administration und Lehre, dass sich die deutschen Universitäten „auf dem Weg zur E-Readiness" befänden (näheres siehe Kleimann & Schmid, 2006 bzw. Kaltenbaek, 2009). Der Weg zur „eUniversity", in welcher alle universitären Prozesse – von der Verwaltung bis hin zur Lehre – durch neue Technologien unterstützt werden (Kerres & Stratmann, 2007), sei demnach offen (Kaltenbaek, 2009). Sicherlich hat sich in dem Bereich seither einiges getan. Von einer durchgängigen Abbildung aller universitären Prozesse durch neue Technologien kann jedoch auch heute noch nicht gesprochen werden.

Vergleicht man die Lehre zu Platos Zeiten mit der Hochschullehre in der heutigen Zeit, zeigt sich der deutliche Kontrast: Heute steht den Hochschuldozenten eine breite Fülle an Möglichkeiten für die Lehre zur Verfügung. Neben Vorlesungen, Seminaren, Übungen, Praktika etc. steht fast überall eine Tafel, ein Whiteboard oder sogar ein interaktives Whiteboard oder ein Tablet-PC zur Verfügung. Zudem kann auf vielfältige Printmedien wie Bücher, Zeitschriften und selbst erstellte Manuskripte zurückgegriffen werden. Neben der Nutzung von Overhead-Folien können beispielsweise Dias oder Videos vorgeführt sowie Präsentationen erstellt und gezeigt werden. Zudem können Hochschuldozenten in einem Intranet der Hochschule Informationen bereitstellen und austauschen. Hier ergibt sich eventuell auch die Möglichkeit, weitreichendere E-Learning-Anwendungen für Studierende zur Verfügung zu stellen. Aus all diesen Mitteln entsteht eine Unmenge an Kombinationsmöglichkeiten. Diese bilden jeweils verschiedene Formen des Blended Learning ab – von der reinen Präsenzlehre bis hin zu vollvirtuellen Lehrangeboten.

Welche dieser Möglichkeiten in der Realität von den Lehrenden jedoch wirklich genutzt werden, zeigt die umfangreiche HISBUS-Erhebung von Kleimann, Özkilic und Göcks (2008): Lehrveranstaltungsbegleitende Materialien wurden entsprechend den Ergebnissen dieser Studierendenbefragung bereits von 86% aller Hochschulen angeboten. Dies liegt vermutlich vor allem an dem relativ geringen technischen und organisatorischen Erstellungsaufwand und an dem direkten Mehrwert für die Lehrenden, die ihre Inhalte leicht und schnell bereitstellen können. Auch die Studierenden sehen die elektronische Bereitstellung von Material mittlerweile als selbstverständlich an (83% laut HISBUS-Studie, 2005, nach Kleimann & Schmid, 2006). Interaktive Lehrangebote wurden von 35% aller Hochschulen genutzt – und somit in deutlich geringerer Form als die lehrveranstaltungs-

begleitenden Materialien. Dies ist in dem wesentlich höheren technischen und didaktischen Aufwand zu begründen (Kleimann, Özkilic & Göcks, 2008). Virtuelle Seminare, Tutorien, Televorlesungen, virtuelle Praktika oder Labore hatten nur noch ca. 5 – 12% der befragten Studierenden an ihrer Hochschule kennengelernt. Nach den Autoren sind Lehrende dieser Form des E-Learning häufig skeptisch gegenüber eingestellt, vermuten einen erheblichen Zusatzaufwand oder kennen sie nicht einmal. Häufig verfügen sie aber auch nicht über die notwendige mediendidaktische Kompetenz. Zum Teil liegt es aber auch an dem eingeschränkten Einsatzzweck der E-Learning-Anwendungen, der eine Nutzung nur in speziellen Lernumgebungen möglich macht (Kleimann & Schmid, 2006; Kleimann et al., 2008).

Online-Studiengänge wurden nach Kleimann und Schmid (2006) von 16% der Hochschulen angeboten. Ihre Einführung erfordert zwar den größten zusätzlichen Aufwand, sie werden jedoch immer wichtiger – gewinnt doch das netzgestützte, berufsbegleitende Studium im Zuge des Bologna-Prozesses immer mehr an Bedeutung. Durch sie können Hochschulen neue Studierendengruppen gewinnen, Fächer durch Verbundstudienangebote stärken oder die Sichtbarkeit der Hochschule durch ein internetbasiertes Bildungsangebot erhöhen.

Insgesamt war (und ist) folglich die Verbreitung von lehrveranstaltungsbegleitenden Materialien und virtuellen Lernformen relativ groß, während die übrigen medien-unterstützten Lehrformen nur selten zum Einsatz kommen (Kleimann & Schmid, 2006).

Innerhalb der Erhebung von Kleimann et al. (2008) wurden zudem über 4000 Studierende online dazu befragt, wie sie den Einsatz von E-Learning in den deutschen Hochschulen einschätzen. Die Studierenden gaben an, dass sie E-Learning-Angebote der Hochschule, wie zum Beispiel Video-Aufzeichnungen von Veranstaltungen, Wikis, Web-based Trainings oder E-Portfolios, hauptsächlich zur Vor- und Nachbereitung von Lehrveranstaltungen sowie zur Prüfungsvorbereitung nutzen. Dabei werden von den Studierenden – wie weiter unten genauer erläutert – insbesondere Video-Aufzeichnungen von Veranstaltungen als sehr nützlich wahrgenommen.

Auch Buzzetto-More (2008) befragte Studierende der University of Maryland Eastern Shore (UMES) zu ihren Einschätzungen von E-Learning-Anwendungen. Es zeigte sich, dass sie die begleitenden E-Learning-Anwendungen zu ihren Veranstaltungen als sehr hilfreich wahrnahmen, um die Kursinhalte besser zu verstehen. Zudem schätzten sie die Rolle solcher E-Learning-Anwendungen für die Zukunft der Hochschullehre als hoch ein. Ihrer Meinung nach förderte die Online-Umgebung das Gefühl, eine Gemeinschaft zu sein und erhöhte die Interaktion zwischen Lernenden sowie Lernenden und Lehrenden.

In einer Umfrage im Rahmen des k-MED[4]-Projektes zeigte sich ebenfalls deutlich die Bereitschaft der Studierenden, elektronische Medien zum Lernen zu nutzen (Glowalla, 2003). Auf Dozentenseite sah es ähnlich positiv aus. Bereits damals nutzten fast 100% aller Hochschullehrer und -studierenden eine eigene E-Mail-Adresse und ca. 60% der Dozierenden stellten ergänzend zu ihren Veranstaltungen Materialien im Intranet zur Verfügung.

Die Erwartungen an die Einführung von E-Learning-Anwendungen in die Hochschullehre sind folglich von allen Seiten hoch. Durch den Einsatz neuer Medien verspricht man sich, Studierende mehr zufrieden zu stellen und die eigene Attraktivität für (neue) Studierende zu steigern (Kleimann und Schmid, 2006). Der Einsatz neuer Medien wird zu einem wettbewerbsprägenden Faktor an den Hochschulen (Bundesministerium für Bildung und Forschung, 2000). Je größer die Hochschule, desto größer der Verwaltungs- und Servicebedarf und entsprechend größer auch das Engagement, elektronische Medien einzusetzen (Kleimann und Schmid, 2006). Man erwartet, die Qualität der Lehre zu verbessern, indem man das Präsenzstudium durch den Einsatz neuer Medien ergänzt – nicht ersetzt (Bundesministerium für Bildung und Forschung, 2000, Kaltenbaek, 2009). Dadurch könnten der Anteil eines begleiteten Selbststudiums erhöht, neue Fernstudienangebote und neue Kombinationen von Präsenzlehre und Selbst-/Fernstudienanteilen entwickelt sowie neue Angebote für die Weiterbildung geschaffen werden (Bundesministerium für Bildung und Forschung, 2000). Weniger wird erwartet, dass durch E-Learning die Lehre erleichtert oder Kapazitätsprobleme gelöst werden können. Lehre wird nach Meinung der Hochschulverantwortlichen attraktiver, aber nicht weniger aufwändig oder kostengünstiger (Kleimann und Schmid, 2006).
Die Hochschulen erkennen folglich, dass sie durch die Medien neue Möglichkeiten haben und auch gesellschaftlich gezwungen sind, neue Medien einzusetzen. Für den sinnvollen Einsatz von neuen Medien in der hochschulischen Lehre müssen allerdings noch immer die Voraussetzungen verbessert werden. Zwar stehen mittlerweile in den Hochschulen ausreichend viele Computer für die Nutzung durch die Studierenden zur Verfügung (Kaltenbaek, 2009): Alle Hochschulen sind mit Hochgeschwindigkeitszugängen ausgestattet (Bundesministerium für Bildung und Forschung, 2000) und nahezu jede Universität in Deutschland verfügt über eine Lernplattform sowie ein Zentrum für die Administration und Betreuung der technischen Systeme und ihrer Nutzer (nach Kaltenbaek,

[4] In dem Projekt k-MED (Knowledge-based Multimedia Medical Education) werden E-Learning-Angebote für Humanmedizin-Studierende hessischer Universitäten entwickelt, bereitgestellt und evaluiert. Näheres siehe www.k-med.org.

2009). Dennoch muss die technische Ausstattung und ihre Betreuung weiter ausgebaut werden – insbesondere, da der technische Fortschritt sehr schnell vonstattengeht.

Darüber hinaus muss beachtet werden, dass Hochschullehrende zumeist nicht didaktisch ausgebildet sind, Lehre für Studierende anzubieten (Hochschuldidaktisches Kompetenzzentrum der JLU, 2016). Hier gibt es folglich schon per se einen Nachholbedarf. Es ist naheliegend, dass Lehrende darüber hinaus qualifiziert und beraten werden müssen, wie sie neue Medien gewinnbringend in der Lehre einsetzen können. Es müssen didaktische und methodische Konzepte für den Einsatz der Medien entwickelt werden. Dafür sind vielfach auch fächerübergreifende Kompetenzen und interdisziplinäre Zusammenarbeit gefordert.

Neben der Weiterbildung der Lehrenden müsste jedoch auch gewährleistet werden, dass es Anreize für E-Learning-Leistungen gibt oder diese auf die Lehrverpflichtung angerechnet werden können. Dies ist für Dozierende naturgemäß wichtig, aber leider noch immer nur in begrenztem Maße möglich. So besagt die „Verordnung über den Umfang der Lehrverpflichtung des wissenschaftlichen und künstlerischen Personals an den Hochschulen des Landes (Lehrverpflichtungsverordnung)“ (Hessische Staatskanzlei, 2013) in §2, Absatz 3: „Die Erstellung und Betreuung von E-Learning-Angeboten kann bis zu einem dem Zeitaufwand entsprechenden Umfang auf die Lehrverpflichtung angerechnet werden, jedoch höchstens bis zu 25 Prozent der festgelegten Lehrverpflichtung.“ Sicherlich ist diese Regelung bereits ein guter Anfang, die Erstellung von umfangreichen E-Learning-Anwendungen wird hierdurch allerdings vermutlich nicht ausreichend gefördert. Dadurch bleibt es zumeist bei der elektronischen Bereitstellung von lehrveranstaltungsbegleitenden Materialien (wie durch Kleimann & Schmid (2006) beobachtet).

Aber auch die Regelung von Urheber- und Nutzungsrechten darf nicht vergessen werden. Ohne eine klare Regelung hierfür ist der Einsatz von E-Learning an Hochschulen undenkbar (Bundesministerium für Wirtschaft und Arbeit & Bundesministerium für Bildung und Forschung, 2006; Kaltenbaek, 2009, Kleimann & Schmid, 2006, siehe auch Overmann, 2002).

Vielen Hochschulen fehlt es zudem an einer Gesamtstrategie. Gegenwärtig werden hauptsächlich Einzelprojekte („Insel- oder Leuchtturmprojekte“, nach Görlich, 2006) vorangetrieben, die durch Kurzlebigkeit gekennzeichnet sind und deren längerfristiger Erhalt zumeist schwierig ist. Dies liegt hauptsächlich daran, dass sie häufig von einzelnen Personen getragen und mit begrenzter Förderung durchgeführt wurden. Wenn eine Institutionalisierung innerhalb der Projektzeit ausbleibt, werden die Anwendungen meist nicht mehr gepflegt und weiter in der Lehre eingesetzt (Kaltenbaek, 2009).

1.2.2.3 E-Learning in der beruflichen Weiterbildung

In den 1990er Jahren waren es zunächst die Großunternehmen, die die Einsatzmöglichkeiten von E-Learning in der beruflichen Weiterbildung testeten, um ihre steigenden Weiterbildungskosten einzudämmen. Insbesondere die Finanzdienstleistungsunternehmen spielten hier eine große Rolle, da sie durch ihre häufig dezentrale Struktur und hohe Durchdringung mit PC-Arbeitsplätzen eine Möglichkeit sahen, Kosten (z.B. Reise-/Abwesenheitskosten) einzusparen. Auch Unternehmen aus der EDV-, Telekommunikations- und Elektronik-Branche nutzten frühzeitig E-Learning für ihre Fortbildungen – hier galt es, die hohe IT-Kompetenz der Mitarbeiter auszunutzen. In anderen Branchen setzte sich das Lernen mit neuen Medien langsamer oder bis zum heutigen Tage gar nicht durch, weil z.B. die notwendige Infrastruktur am Arbeitsplatz nicht gegeben ist oder die Inhalte sich nicht für eine E-Learning-Anwendung eignen (Dittler & Jechle, 2009).

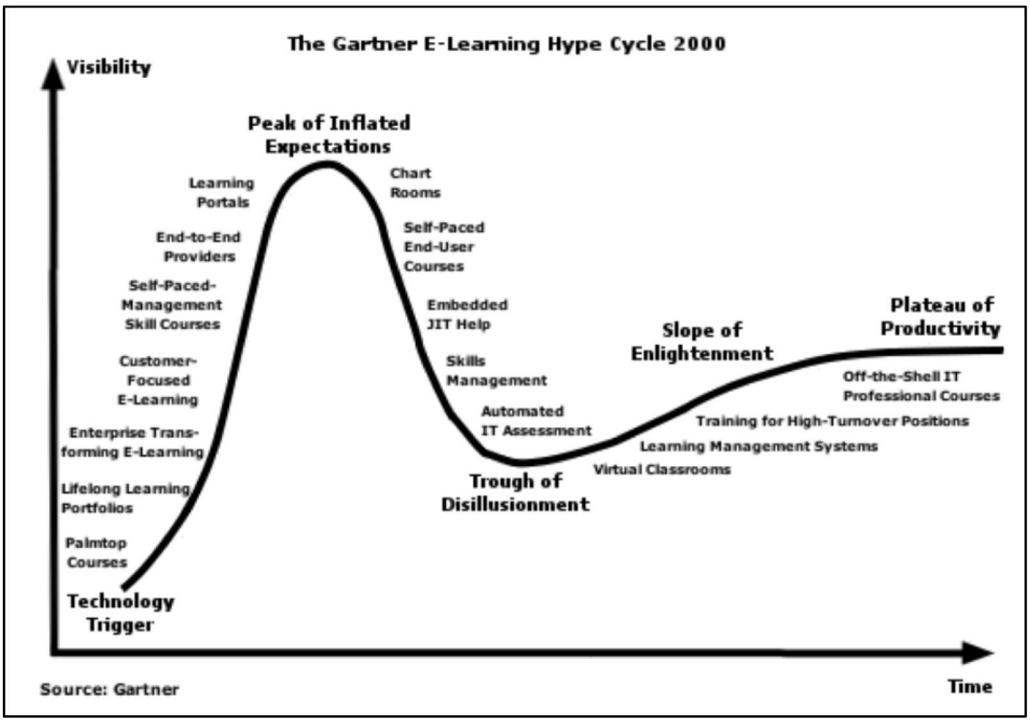

Abbildung 2. Der "Gartner E-Learning Hype Cycle 2000" (nach Dittler & Jechle, 2011)

Um die Jahrtausendwende wurde vielen Unternehmen klar, dass die hohen Erwartungen an die Einsparung durch E-Learning nicht realistisch sind. Zudem kam es häufig zu Akzeptanzproblemen und viele Lernende vermissten das Lernen in einer sozialen Umgebung. Hinzu kam, dass die Produktion der Anwendungen in vielen Fällen weitaus mehr Vorlauf benötigte als die Vorbereitung einer Präsenzschulung (Dittler & Jechle, 2009). Gartner (nach Dittler & Jechle, 2009 bzw. 2011) bildete diese Entwicklung in seinem „E-Learning Hype Cycle 2000“ (siehe Abbildung 2) als „Gipfel der überzogenen Erwartungen“ ab. Danach folgte der Weg in das „Tal der Enttäuschungen“. Die Unternehmen

verstanden nun, dass eine vollständige Umstellung der Weiterbildung auf E-Learning-Angebote nicht realistisch ist. Über den „Pfad der Erleuchtung“ hin entwickelte sich der Prozess allerdings zum „Plateau der Produktivität“. Wollte man vorher durch E-Learning lediglich Kosten einsparen, hatte man sich nun die Verbesserung der Aus- und Weiterbildung zum Ziel gesetzt. Hieraus resultierten Ansätze wie „Blended Learning“, die Einbindung von Tele-Tutoren oder die Gestaltung von interaktiven Lernanwendungen (Dittler & Jechle, 2009).

Heute ist E-Learning für die Mehrheit der deutschen Großunternehmen zur Selbstverständlichkeit geworden. Dabei werden die Zielgruppen von E-Learning und die Themen, die mit E-Learning bearbeitet werden sollen, immer vielfältiger (Institut für Medien- und Kompetenzforschung, 2010; siehe auch Schüler, 2002). Auch nach einer Befragung des Bundesministeriums für Wirtschaft und Technologie (2008) äußern sich die meisten Mitarbeiter mittlerweile positiv über E-Learning. Auch sie sehen E-Learning nicht mehr nur als flüchtige Modeerscheinung an, sondern nehmen es als festen Bestandteil der betrieblichen Weiterbildung wahr.

1.2.3 Vor- und Nachteile von E-Learning

Die Erwartungen an das Lehren und Lernen mit neuen Medien sind hoch. Ob diese jedoch erfüllt werden können, soll im Folgenden erläutert werden. Dabei werden zunächst die möglichen Vorteile von E-Learning-Szenarien beleuchtet, bevor die denkbaren Nachteile diskutiert werden.

1.2.3.1 Mögliche Vorteile des E-Learning-Einsatzes

Döring (2002) unterteilt die möglichen Vorteile von E-Learning in ökonomische, politische, didaktische und inhaltliche Aspekte. Ihre Auflistung soll im Folgenden ergänzt durch Ausführungen verschiedener anderer Autoren (Beer et al., 2003; Connolly & Stansfield, 2007; Dalsgaard, 2005; Institut für Medien- und Kompetenzforschung, 2010; Kaltenbaek, 2009; Keller & Back, 2004; Klauser, Kim & Born, 2002; Köhler, Kahnwald & Reitmeier, 2008; Norton, 2000; Overmann, 2002; Schlageter & Feldmann, 2002; Vogt, 2002; Wache, 2003 sowie Weidenmann, 2000) dargestellt werden.

Ökonomische Aspekte

Der Einsatz von E-Learning-Anwendungen ist zumeist verbunden mit der Möglichkeit, orts- und zeitflexibel zu lernen. Zeitunabhängigkeit sowie die Selbstbestimmung des Lerntempos werden in der Befragung von Klauser et al. (2002) auch von Studierenden als Antwort auf

die Frage genannt, welche Vorteile der E-Learning-Einsatz in der Hochschullehre ihrer Meinung nach bietet.

Des Weiteren können Inhalte einfacher und schneller ausgetauscht werden und durch die Möglichkeit, Lernenden Wissen permanent zur Verfügung zu stellen, ist es ihnen möglich, jederzeit und so häufig sie mögen auf das Lernangebot zuzugreifen. Je nach Art der Anwendung können Inhalte zudem ggf. schnell und kostengünstig aktualisiert sowie wieder- und weiterverwendet werden. Dies kann die Nachhaltigkeit von E-Learning-Anwendungen erhöhen, was Kosteneinsparungen ermöglicht. Im Rahmen der betrieblichen Weiterbildung nennt Norton (2000) konkret den Wegfall von Reisekosten, von Kosten für Lehrmittel und -räume oder die Verminderung von administrativen Kosten und Fehlzeiten auf dem Arbeitsplatz. Solche ökonomischen Vorteile sind selbstverständlich nicht per se für jedes Lernszenario gegeben und die ggf. Mehrkosten für die Erstellung der E-Learning-Anwendungen muss dabei mitbedacht werden.

Politische Aspekte

Im Sinne des „Open Learning“ kann Bildung für alle möglich gemacht werden – auch für Berufstätige, in ihrer Mobilität Behinderte, in abgelegenen Gebieten Wohnende oder Lernende, die Familienmitglieder betreuen müssen. So ermöglicht E-Learning nach Schlageter und Feldmann (2002) erstmals berufsbegleitendes, bedarfsorientiertes Lernen. Die Schwierigkeiten, die solche reinen E-Learning-Szenarien allerdings mit sich bringen, wie Motivationsverluste oder ein hoher Drop-out, werden noch im Laufe der Arbeit diskutiert.

Didaktische Aspekte

Einige Studien zeigen, dass Lernende sich Lernerfolge in kürzerer Zeit erarbeiten können. Lehren und Lernen wird effizienter, wenn Lernszenarien so beschaffen sind, dass Wissen reflektiert, rezipiert, produziert, kommuniziert und kollaborativ verarbeitet wird. Lernen kann „Just-in-Time“ und „Just-in-Place“ stattfinden. Dies ist in Offline-Szenarien häufig so nicht möglich.

Hinzu kommt eine Erhöhung der Lernmotivation durch die Möglichkeit, die Lerninhalte durch die Multimedialität attraktiver und spielerischer zu gestalten. Komplexe, unsichtbare, abstrakte oder numerische Lerninhalte können durch Visualisierungen und Animationen anschaulich und dynamisch gezeigt, langfristige Prozesse im Zeitraffer dargestellt, Experimente mit z.B. Tieren oder Menschen mit dem Computer durchgeführt, Lernobjekte interaktiv bearbeitet und Übungen wirklichkeitsnah und interaktiv gestaltet werden.

Durch die neuen Kommunikations- und Kooperationsmöglichkeiten können neue teamorientierte Lernszenarien geschaffen und die Kommunikation zwischen allen

Beteiligten verbessert werden. Dadurch kann sich auch die Betreuung der Lernenden zwischen Präsenzveranstaltungen vertiefen. Besonders hilfreich bzw. sogar notwendig sind solche Möglichkeiten vor allem in Regionen, in denen einzelne Lernende sehr weit entfernt voneinander leben.

Ebenso ist eine bessere Anpassung an die Bedürfnisse der Nutzer möglich (Individualisierung), Lernende können ihren Lernprozess selbständig gestalten, ihre Lerngeschwindigkeit selbst bestimmen und sich bedarfsgerecht (z.B. je nach Vorwissen) weiterbilden. Wenn zusätzlich gute Übungsmöglichkeiten und Wissensdiagnosen bereitgestellt werden, kann der Lernende sein Wissen sichern und Lernfortschritte selbst kontrollieren.

Inhaltliche Aspekte

Neben den attraktiveren Lehrinhalten können auch die technische und soziale Medienkompetenz erweitert und Lernende zudem für medienbezogene Fragen sensibilisiert werden (Medienbildung).

Sicherlich sind noch nicht alle der von den verschiedenen Autoren genannten Vorteile ausreichend wissenschaftlich belegt. Hierzu werden weitere Studien vonnöten sein. Des Weiteren ist zu ergänzen, dass es sich hier um eine Auflistung von Vorteilen handelt, die durch E-Learning erreicht werden *können* und nicht zwangsläufig *müssen*. E-Learning-Anwendungen, die unbedacht erstellt und eingesetzt werden, können diese Vorteile sicherlich nicht mit sich bringen.

Darüber hinaus kann die Optimierung des einen Vorteils manchmal auf Kosten des anderen gehen. Insbesondere ökonomische und didaktische Ziele lassen sich häufig nur schwer vereinbaren. Will man einen didaktischen Mehrwert durch den Einsatz von E-Learning erreichen, muss man unweigerlich mit Zusatzaufwand rechnen. Auch die Betreuung der Anwendungen und der Lernenden sind mit Aufwand verbunden – aber dringend notwendig, um den Lernprozess zu fördern (Döring, 2002).

Außerdem spielt das jeweilige Lernszenario eine immense Rolle. Wenn eine E-Learning-Anwendung mit hohen Kosten für einen einzelnen Einsatz und nur wenige Lernende erstellt wurde, kann sie zum Beispiel keine ökonomischen Vorteile im Vergleich zur Präsenzlehre mit sich bringen. Die gleiche Anwendung hätte diese Vorteile allerdings, wenn sie häufig eingesetzt und von vielen Lernenden genutzt würde.

So zeigt Abbildung 3 die Ergebnisse einer Befragung des Instituts für Medien- und Kompetenzforschung (2010, siehe auch Scheer, 2009), bei der Personalverantwortliche der TOP-500-Unternehmen Deutschlands unter anderem zu den Gründen für den Einsatz von E-Learning in ihren Unternehmen befragt wurden. Es zeigt sich, dass sich die genannten Gründe gut mit den oben erläuterten Vorteilen decken.

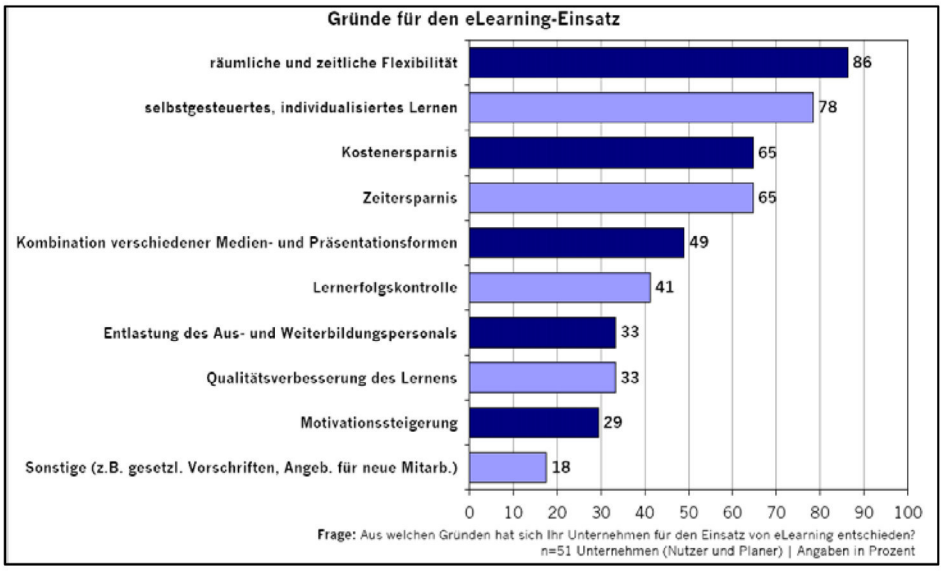

Abbildung 3. Gründe für den Einsatz von E-Learning in den TOP 500-Unternehmen Deutschlands (Institut für Medien- und Kompetenzforschung, 2010)

Insbesondere Großunternehmen können die Mehrwertpotentiale von E-Learning nutzen, da sie auf viele Standorte verteilt sind und sich auf kurzen Produktlebenszyklen und hohen Innovationsraten befinden (Wache, 2003). In den klein- und mittelständischen Unternehmen (KMU) wird E-Learning weniger umfangreich eingesetzt. Ihre Weiterbildungsetats sind deutlich geringer und es findet häufig gar keine strategische Personalentwicklung statt (Bundesministerium für Wirtschaft und Technologie, 2008), da sich die Unternehmen keinen wirklichen „Return on Investment“ (ROI, Wache, 2003) hiervon versprechen. In öffentlichen Verwaltungen sieht es noch schlechter aus. Hier sind es insbesondere die unerfreuliche Lage der öffentlichen Haushalte, die sehr restriktiven Regelungen bezüglich der Nutzung von Internet am Arbeitsplatz sowie die mangelnde Motivation der Mitarbeiter, die die Einführung von E-Learning erschweren (Wache, 2003). Hier wird erneut deutlich, wie wichtig die jeweilige Lernumgebung für die Beurteilung der einzelnen Vor- und Nachteile von E-Learning-Szenarien ist.

E-Learning-Angebote sind also – genauso wie alle anderen Lernangebote – nicht per se lernförderlich und kostensparend und der Einsatz neuer Medien in Lernprozessen ist nicht per se über längere Zeit motivierend. Die modernsten Technologien reichen folglich nicht aus, um Lernprozesse zu fördern, wenn ihrem Einsatz keine didaktischen Konzepte zugrunde liegen. Außerdem eignen sich bestimmte Medien und Methoden zum Erreichen von bestimmten didaktischen Zielen besser als andere (Kaltenbaek, 2009). Es ist folglich wichtig, dass die Anwendungen kognitiv, ergonomisch und instruktional zweckmäßig gestaltet (Glowalla, 2003), auf die Voraussetzungen, Bedürfnisse und Interessen der Lernenden abgestimmt sind (Klauser, Kim & Born, 2002) sowie didaktisch sinnvoll in den Lernprozess/das Curriculum eingebunden werden. Erst dann sind sie geeignet,

Lernprozesse zu optimieren und das Erreichen von Lernzielen zu erleichtern (Glowalla, 2003, siehe auch Schlageter & Feldmann, 2002) damit die oben genannten Vorteile eintreten können.

1.2.3.2 Mögliche Nachteile des E-Learning-Einsatzes

Neben den beschriebenen Vorteilen sind allerdings auch klare Nachteile zu nennen, die durch den E-Learning-Einsatz auftreten *können*. So können Lernende sehr hohe – zum Teil nicht erfüllbare – Erwartungen an die Inhalte, Materialien, Betreuung und mediale Integration stellen und dann womöglich schnell enttäuscht werden. Es gibt darüber hinaus immer noch Lernende und Lehrende, die E-Learning aus emotionalen Gründen ablehnen (Kaltenbaek, 2009). Die generelle Akzeptanz, Motivation und Medienkompetenz aller Beteiligten muss deshalb immer bedacht werden. Treten in diesen Bereichen Defizite auf, ist der erfolgreiche Einsatz von E-Learning gefährdet.

Die Betreuung, Entwicklung und Aktualisierung der Kurse kann zudem, wie oben bereits erwähnt, mit erheblicher Arbeit und immensen Kosten verbunden sein (Connolly & Stansfield, 2007), wobei diese sich mit steigender Lernerzahl amortisieren (Vogt, 2002).

Befragt man Studierende selbst, sehen diese besonders große Nachteile von E-Learning-Anwendungen in der ihrer Meinung nach eingeschränkten Kommunikation und dem fehlenden Druck zum Lernen. Die Gelegenheit zum selbstorganisierten Lernen durch E-Learning wird von Studierenden selbst folglich sogar als nachteilig eingeschätzt (Klauser et al., 2002). Dies widerspricht sicherlich den oben von Studierenden genannten Vorteilen, zeigt aber, dass Lernende zwar die Flexibilität durch E-Learning gutheißen, vermutlich vielen Lernenden jedoch die notwendigen Fähigkeiten[5] zum effizienten selbstorganisierten Lernen fehlen. Auch Lehrende müssen die notwendigen Fähigkeiten[5] für die Erstellung und Betreuung lerneffizienter E-Learning-Anwendungen häufig erst noch erlernen. E-Learning kann auch nur dann einen Mehrwert haben, wenn es gelingt, Lernende und Lehrende ausreichend hinsichtlich solcher Fähigkeiten zu qualifizieren (Wache, 2003).

Soziale Komponenten der Präsenzlehre lassen sich darüber hinaus durch reines E-Learning zumeist nicht abdecken. Die Lernenden klagen eventuell über fehlendes „Involvement“, über das Gefühl des Isoliertseins oder sie fühlen sich vernachlässigt („I am neglected“-Effekt, Helic, Maurer & Scherbakov, 2000). Dies wiederum kann in hohe Abbruchquoten oder niedrige Lernleistung münden (Connolly & Stansfield, 2007; Görlich, 2006). Lernende möchten gerne mit anderen Lernenden und Lehrenden zusammenarbeiten. Deshalb ist es sinnvoll, reine E-Learning-Anwendungen durch Möglichkeiten des sozialen Austausches zu ergänzen, wie z.B. Emails, Chats oder Foren. Zu beachten ist allerdings, dass diese

[5] Nähere Ausführungen hierzu siehe Wache (2003)

Kommunikationsmöglichkeiten nicht von selbst funktionieren. Es bedarf eines Experten, der diese Angebote betreut.
Auch über diese Wege kann jedoch die ganzheitliche Kommunikation, die im Rahmen der Präsenzlehre stattfinden würde, nicht vollständig ersetzt werden. Wichtige non- und paraverbale Signale stehen in der Kommunikation über Chats, Foren oder Emails nicht zur Verfügung. Die fehlende Gestik und Mimik kann dazu führen, dass dem Lehrenden z.B. keine Rückmeldung über das Verstehen eines Sachverhaltes durch die Lernenden zukommt. Auch kann im Rahmen von Diskussionen die Koordination erschwert sein (Wache, 2003), was dazu führen kann, dass Fragen von Lernenden offen bleiben (Lindecke & Beer, 2003).
Hinzu kommt, dass sich das Setting der Präsenzlehre durch eine besondere Atmosphäre auszeichnet (Kaltenbaek, 2009), welche innerhalb von E-Learning-Anwendungen auch nicht über sogenannte „pädagogische Agenten[6]“ aufzufangen ist. Immerhin konnten Schilbach, Wohlschläger, Krämer, Newen, Shah, Fink & Vogeley (2006) zeigen, dass virtuelle Figuren, die über Gestik verfügten, beim Beobachter ähnliche Hirnregionen aktivieren können wie eine Mensch-zu-Mensch-Interaktion. Den persönlichen Austausch zwischen realen Menschen können sie dennoch nicht vollständig ersetzen.
Hinzu kommt, dass E-Learning – wie jede andere Lehrmethode auch – nicht für jedes Lernziel geeignet ist. Nicht jeder Wissensbereich und jede Fähigkeit kann über E-Learning-Anwendungen gelehrt werden (Glowalla & Häfele, 1997). Motorische Ziele, wie das Treffen eines Basketballkorbes oder das Fahren eines Fahrrades, werden beispielsweise durch E-Learning unmöglich zu erreichen sein. Und auch für das Bilden von Meinungen sind sicherlich andere Methoden – wie die Diskussionsmethode oder Rollenspiele – besser geeignet, als ein reines E-Learning-Szenario.
In der betrieblichen Weiterbildung kommen weitere Nachteile hinzu, denn eine erfolgreiche Implementation von E-Learning (oder Blended Learning) erfordert auch immer starke Veränderungen im Umgang mit Weiterbildungen innerhalb des Unternehmens. So müssen zum Beispiel Mitarbeiter motiviert werden teilzunehmen und es muss genaue Regelungen für das Lernen am Arbeitsplatz und die Anerkennung von durch E-Learning erworbene Qualifikationen geben. Findet E-Learning darüber hinaus außerhalb der Arbeitszeit statt, vermittelt es den TeilnehmerInnen nicht dasselbe Prestige wie ein Präsenzseminar (Lindecke & Beer, 2003). Dem muss entgegengewirkt werden.

Nach der ausführlichen Vorstellung von E-Learning in verschiedenen Kontexten mit seinen Vor- und Nachteilen soll im Vergleich dazu die traditionelle Lehre vorgestellt werden. Da die Studien der vorliegenden Dissertation im Rahmen der Hochschullehre stattfinden und

[6] Pädagogische Agenten sind anthropomorph gestaltete Figuren, die dem Lernenden Wissen vermitteln und ihn durch die Lernanwendung begleiten sollen (Bendel, 2004).

hier die Vorlesung bzw. der Vortrag als prominenteste Form der traditionellen Lehre angesehen werden kann, wird im Folgenden auf diese Lehrmethode fokussiert.

1.3 Traditionelle Lehre: die Vortragsmethode

Obwohl das angeblich „passive Lernen“ des Frontalunterrichts häufig kritisiert wird, ist der Vortrag als älteste Lehrmethode noch immer die dominante Lehrform innerhalb der Universität und vielen anderen Aus- und Weiterbildungsbereichen (siehe u.a. Görlich, 2006; Vogt, 2002,). Damit ist er der prominenteste Vertreter der traditionellen Präsenzlehre.
Nach Glowalla (2004) liegt dies daran, dass Vorträge sehr ökonomisch sind. Ohne großen Mehraufwand kann gleichzeitig vielen Zuhörern Wissen vermittelt werden. Dabei kann der Vortragende seine Zuhörer gut überblicken. Und auch seine Vorbereitungszeit und sein Aufwand, um den Vortrag zu aktualisieren, sind minimal. Zumeist kann der Experte, der den Vortrag halten soll, schon auf Anschauungsmaterial oder eine Präsentation zurückgreifen, die er dann lediglich auf die Gegebenheiten oder/und unter Berücksichtigung neuer Erkenntnisse überarbeiten muss. Dadurch ist ein Vortrag auch einfach aktualisierbar – sogar Informationen, die noch nicht in gedruckter Version vorliegen, können bereits in den Vortrag eingebracht werden. So können zum Beispiel in der Hochschullehre Forschung und Lehre elegant verknüpft werden. Innerhalb des Vortrags kann der Vortragende dann den Inhalt (häufig Materialien aus unterschiedlichen Quellen) prägnant zusammenfassen, verdichten und für die Lernenden vorstrukturieren. So kann er erreichen, dass Zuhörer in ein Thema eingeführt werden oder sie ihr Wissen erweitern oder vertiefen können. Am besten geeignet ist der Vortrag dafür, dass dem Lernenden eine Einführung oder ein Überblick in ein Thema gegeben wird. Außerdem bietet er immer die Gelegenheit, etwas praktisch zu demonstrieren, was es den Zuhörern ermöglicht, am Modell zu lernen. Je nach seinen Fähigkeiten kann der Vortragende die Lernenden zudem durch seinen Enthusiasmus anstecken und für seine Inhalte motivieren (Glowalla, 2004; Görlich, 2006; Soliman, 1999). So kann der Vortrag zu einem lebendigen Geschehen werden, wenn Vortragender und Inhalt stimmig zusammenspielen.
Als weiterer Pluspunkt kommt hinzu, dass Studierende dieser Lehrmethode positiv gegenüber stehen. In einer Untersuchung von Glowalla (siehe Glowalla, 2004) hielten 70% der Studierenden die Vorlesung (als bekannteste Einsatzmöglichkeit für die Vortragsmethode innerhalb der Hochschullehre) für eine geeignete Form der Wissensvermittlung.

Die Vortragsmethode hat jedoch auch Nachteile. Zum einen wird häufig eine fehlende Lerneraktivität genannt, welche durch die zumeist unidimensionale Kommunikation entstehe. Das könne dazu führen, dass die Zuhörer von der Flut an Inhalten, die präsentiert werden, überwältigt würden und nicht wirklich wüssten, welche Inhalte nun wirklich gelernt werden sollen (Soliman, 1999). Hinzu komme, dass innerhalb eines Vortrages dem Lernenden keine Rückmeldung über seinen Lernfortschritt geboten würde. Hesse und Mandl (2000, S. 32) sprechen in diesem Zusammenhang auch von „trägem Wissen", das durch den „Wissenstransport vom Lehrenden zum Lernenden" entstehe, da Lernen vornehmlich „rezeptiv" erfolge. Über den Punkt der fehlenden Lerneraktivität kann man allerdings streiten, da der Zuhörer zumindest kognitiv aktiv sein muss, um dem Vortrag zu folgen. So ist die Lerneraktivität während eines Vortrages sicherlich als geringer einzuschätzen als zum Beispiel während einer Unterrichtsdiskussion, von „fehlender Lerneraktivität" kann jedoch auch nicht ausgegangen werden, da die Vorlesung sonst keinerlei Lernwirksamkeit aufweisen dürfte.

Bligh (2000) argumentiert allerdings, dass die Teilnehmer an einem Vortrag sehr häufig unaufmerksam sind. Dabei erfordert die komprimierte Form der Wissenspräsentation innerhalb eines Vortrages (oben als Vorteil genannt!) eine gute Konzentration auf Seiten der Zuhörer. Diese Konzentration kann allerdings nur selten über die gesamte Vortragszeit aufrechterhalten werden. So konnte McLeish (1968) zeigen, dass mit zunehmender Länge eines Vortrages die Menge an Wissen sank, die die Zuhörer aufnehmen konnten. Direkt nach einem 15-minütigem Vortrag konnten die Lernenden noch 41% der Inhalte wiedergeben, während sie nach einem 45-minütigem Vortrag lediglich 20% erinnern konnten. Auch eine Befragung im Rahmen des EVALIS-Projektes (Glowalla, Glowalla & Kohnert, 2001) zeigte, dass viele Studierende im Verlauf einer Vorlesung abgelenkt sind. 55% geben an, mit ihren Gedanken abzuschweifen, 49% reden mit ihren Kommilitonen und 10% lesen vortragsfremde Inhalte.

Hinzu kommt, dass es für den einzelnen Lernenden innerhalb eines großen Auditoriums schwierig ist, sich aktiv mit Fragen oder anderen Beiträgen zu beteiligen. Zwar ist die Größe des Auditoriums nicht direkt mit der Lernwirksamkeit des Vortrages korreliert, aber sie beeinflusst die Möglichkeit für persönliche Beiträge Einzelner bzw. die individuelle Zuwendung des Vortragenden (Bligh, 2000; McKeachie, Pintrich, Lin & Smith, 1986). Nach Glowalla, Glowalla und Kohnert (2001) stellen lediglich 4% aller Studierenden in einer vollbesetzten Vorlesung eine Frage. Zudem kann der Vortragende selbst nicht auf individuelle Unterschiede der Lernenden eingehen. Viele Lernende fühlen sich deshalb anonym oder entfremdet (Soliman, 1999). Motivation und Aufmerksamkeit sind in der Folge stark beeinträchtigt, so dass der Lernstoff zumeist nur unvollständig erworben wird (u.a. Glowalla, Glowalla & Görlich, 2004; Scheele, Wessels & Effelsberg, 2004). Je mehr

Lernende anwesend sind, desto mehr stört den Einzelnen zudem die Unruhe im Raum und desto mehr fällt jeder Einzelne in eine rezeptive Haltung. Dadurch wird der Kontakt zwischen Lernendem und Lehrendem noch zusätzlich minimiert und eine Rückkopplung zum Lehrenden ist nur begrenzt möglich.

Nach Bligh (2000) können Lernende durch die Vorlesung genauso viel Wissen erwerben wie durch andere Lehrmethoden auch – allerdings beachtet er hierbei die Nachhaltigkeit dieses Wissens nicht (Vogt, 2002). Denn ein weiterer Nachteil von Vorträgen ist in ihrer längerfristigen Lernwirksamkeit zu sehen. Diese ist eher gering, da die Inhalte schnell wieder vergessen werden. In einer Studie von Glowalla (Glowalla 2004) war die Behaltensleistung für die in der Vorlesung erläuterten Inhalte schon unmittelbar nach der Vorlesung auf 40-60% begrenzt. Und das Vergessen geht schnell voran: McLeish (1968) konnte zeigen, dass bereits eine Woche nach einer Unterrichtsstunde nur noch ca. 17% der Inhalte erinnert werden können, wohingegen es etwa 41% direkt nach dem Vortrag waren. Durch das Anfertigen von Notizen oder die Nutzung von Folienausdrucken kann das Behalten von Vorlesungsinhalten zwar verbessert werden. Durch die Dichte an Informationen ist es allerdings zumeist unmöglich, neben kurzen Notizen auch noch eigene Folgerungen zu notieren (Vogt et al., 2006). Zudem muss der Zuhörer seine Aufmerksamkeit aufteilen, wenn er während des Vortrages Notizen machen möchte. Um die Nachhaltigkeit eines Vortrages zu steigern, ist es darüber hinaus notwendig, die Notizen möglichst zügig nach dem Besuch der Vorlesung nachzubearbeiten (Glowalla, 2004). Zwar geben 80% der Studierenden an, ihre Veranstaltungen nachzubereiten – 50% jedoch lesen in diesem Rahmen lediglich noch einmal ihre Notizen. Aktivere Formen der Nachbereitung, wie zum Beispiel die Diskussion über den Stoff mit Kommilitonen, die Überarbeitung der Mitschriften oder das Lesen von Skript oder Literatur, werden nur selten genutzt (Glowalla et al., 2001). Aber es ist gerade diese aktive Auseinandersetzung mit dem Lernstoff, die zu guten Lernleistungen führt (Glowalla et al., 2004). Wenn die eigenen Notizen darüber hinaus erst wieder nach Wochen oder Monaten zur Klausurvorbereitung genutzt werden, sind sie für die Lernenden häufig nicht mehr hilfreich.

Zudem eignet sich die Vorlesung hauptsächlich für die Demonstration und Illustration eines Sachverhaltes (Bligh, 2000; Soliman, 1999). Diese Vermittlung von deklarativem Wissen reicht jedoch nicht immer aus (Görlich, 2006). Für Lernziele wie Analyse, Synthese, Evaluation, Einstellungsänderung oder kritisches Urteilen ist die Vorlesung allein eher nicht geeignet. Nach Vogt et al. (2006) sollte sie deshalb durch stärker lernerzentrierte Methoden ergänzt werden, um aktives Lernen und die praktische Anwendung von Lerninhalten zu fördern.

Tabelle 1 fasst die genannten Vor- und Nachteile von Vorträgen in einer Tabelle zusammen.

Tabelle 1
Zusammenfassung der im Text genannten Vor- und Nachteile von Vorträgen

Vorteile	Nachteile
• Ökonomie • Hohe Aktualität • Experte als Wissensrepräsentant • Soziale Interaktion • Hohe Akzeptanz bei Lernenden	• Rezeptives Lernverhalten • Begrenzte soziale Interaktion und störende Umgebungsfaktoren • Begrenzte Nachhaltigkeit des Gelernten • Keine Lernerfolgsrückmeldung • Nicht geeignet für Lernziele, wie Analyse, Synthese, Evaluation etc.

Insgesamt ist der Vortrag folglich eine geeignete Lehrmethode zur Vermittlung von Wissen. Aktive Verarbeitung oder praktische Anwendung des Wissens, um tieferes Verständnis oder die Anwendung des Wissens zu erreichen, ist allein durch die Vorlesung jedoch nur bedingt möglich. In den deutschen Hochschulen werden deshalb zusätzlich Seminare, Praktika oder Übungen angeboten. Leider werden diese Lehrformen von den Dozierenden häufig auch mit eigenen Vorträgen oder Referaten der Studierenden gefüllt. Diese erneute Nutzung der Vortragsmethode ist vermutlich ein Zeichen für die oben erwähnte häufig nichtvorhandene didaktische Ausbildung der Lehrenden an den Universitäten.
Eine weitere Möglichkeit, Nachteile der Vorlesung aufzufangen, ist die von Vogt et al. (2006) genannte Erweiterung des Vortrages durch mehr lernerzentrierte Lehrmethoden. Ein Beispiel hierfür stellen interaktive Vorlesungen dar. Hierbei werden Präsenzvorlesungen mit motivations- und interaktionsfördernden Elementen ergänzt, wie z.B. Übungsaufgaben oder Meinungsumfragen (zur Umsetzung und Evaluation siehe Scheele et al., 2004). Interaktive Vorlesungen erfordern allerdings einen erheblichen technischen Aufwand und bergen nicht die Vorteile, die zum Beispiel eine zeit- und ortsabhängige E-Learning-Anwendung mit sich bringen könnte. Eine weitere Möglichkeit, die Nachteile der Präsenzvorlesung auszugleichen, wird in Abschnitt „1.4.2 E-Lecture“ vorgestellt. Bevor allerdings auf diese spezielle E-Learning-Anwendung eingegangen wird, soll im folgenden Abschnitt zunächst ein Überblick über die verschiedenen Formen von E-Learning-Anwendungen gegeben werden.

1.4 Verschiedene Formen von E-Learning

Vor ca. fünfzehn Jahren bedeutete „E-Learning“, sich Texte am Rechner durchzulesen und anschließend Übungsaufgaben zu beantworten (Computer-based Training). Heute gibt es viele verschiedene Variationen und Möglichkeiten für E-Learning-Anwendungen. Für ihre Beschreibung haben sich bisher keine verbindlichen Konventionen durchgesetzt (Wache, 2003). In den letzten Jahren wurden allerdings immer wieder neue Klassifikationen von

E-Learning-Anwendungen erstellt. Eine davon stammt zum Beispiel von Köhler, Kahnwald und Reitmeier (2008), die E-Learning-Anwendungen in sechs Gruppen klassifizieren:

1 Software zur Präsentation und Visualisierung (z.B. *Microsoft PowerPoint™*),
2 Drill- und Test-Software (z.B. Vokabeltrainer),
3 tutorielle Systeme (Web-based Trainings),
4 intelligente tutorielle Systeme (im Prinzip adaptive Web-based Trainings),
5 Simulationssysteme (z.B. *SimCity*) und
6 Hypermediasysteme (z.B. *Wikipedia*).

Weitere Klassifikationen finden sich zum Beispiel in Krapp und Weidenmann (2001), Niegemann, Domagk, Hessel, Hein, Hupfer & Zobel (2008) oder Klimsa und Issing (2011). An dieser Stelle soll allerdings die Klassifikation des Bundesministeriums für Wirtschaft und Technologie (Bundesministerium für Wirtschaft und Technologie (BMWT), 2008) genutzt werden, da sie eine etwas feinere Gliederung auf zwei Dimensionen erlaubt: formelle Methoden, mit denen offiziell gelernt werden soll vs. informelle Methoden, die eher arbeitsbegleitend zum „Learning on demand" genutzt werden sowie individuelle vs. kollaborative Methoden. Die Klassifikation des BMWT ist in Abbildung 4 dargestellt.

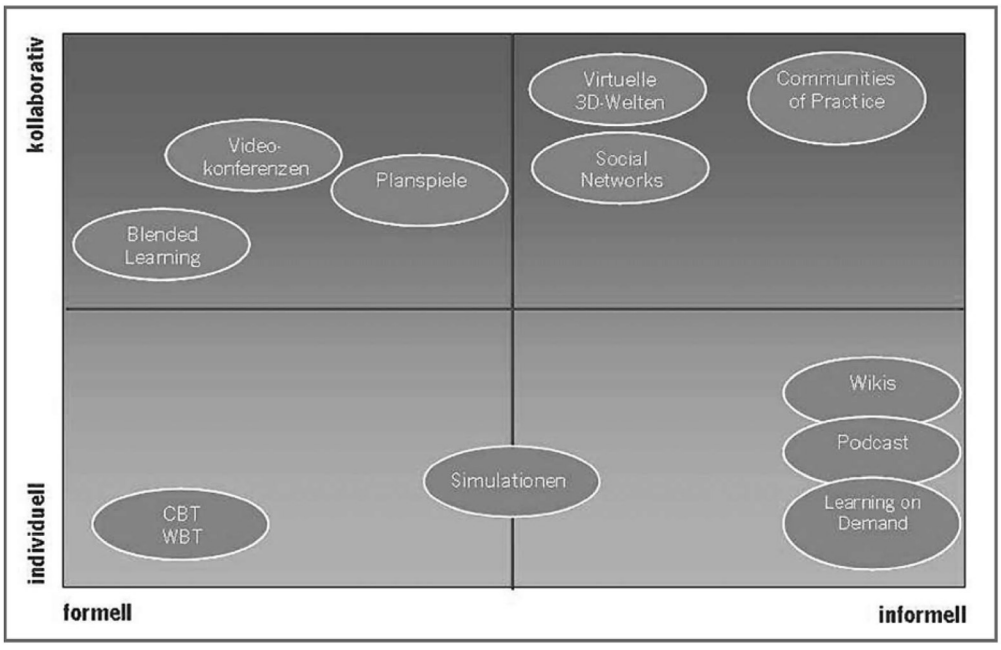

Abbildung 4. E-Learning-Formen (Bundesministerium für Wirtschaft und Technologie, 2008).

Die abgebildeten E-Learning-Formen werden im Folgenden kurz erläutert. Zusätzlich werden Televorlesungen und E-Lectures erläutert, da diese für die vorliegende Dissertation eine besondere Rolle spielen.

Computer-based (CBT) und Web-based Training (WBT)

Nach dem BMWT gewährleisten diese E-Learning-Anwendungen ein eher formelles und individuelles Lernen. Ein CBT ist ein E-Learning-Kurs, der auf einer CD-ROM oder DVD enthalten ist. Wie andere E-Learning-Anwendungen hat es den Vorteil der räumlichen und zeitlichen Flexibilität, allerdings auch den Nachteil, dass Kommunikationsmöglichkeiten zwischen Lernenden oder Lernenden und Lehrenden sowie eine tutorielle Unterstützung häufig fehlen.

Das WBT (auch „Web-based Teaching"), welches über das Internet verfügbar gemacht wird, stellt eine Weiterentwicklung des CBTs dar. Dies ermöglicht im Gegensatz zum CBT eine schnelle Aktualisierung und Anpassung der Inhalte. Für den Lernenden kommt hinzu, dass er so mit einer Software lernen kann, die ihm vertraut ist – einem Internetbrowser. Auch die Kommunikation und tutorielle Betreuung kommen nicht zu kurz, wenn zusätzlich Emails, Chats, Foren oder ähnliches angeboten werden (Bundesministerium für Wirtschaft und Technologie, 2008). Ausführlichere Informationen zum Web-based Training folgen in Abschnitt „1.4.3 Web-based Training (WBT)".

Blended Learning

Unter „Blended Learning"[7] versteht man eine Kombination aus Präsenzlehre und E-Learning-Angeboten. Keller und Back (2004) definieren Blended Learning folgendermaßen: Es „[…] bezeichnet die didaktisch und wirtschaftlich sinnvolle Verknüpfung von verschiedenartigen Methoden und Medien im Rahmen von Lehr-/Lernkonzepten, wobei stets E-Learning-Ansätze einbezogen werden; die Kombination geschieht zielorientiert und im Sinne der optimalen Gestaltung bzw. Optimierung einer Lernmaßnahme sowie deren Lerneffekt." Dabei können die Anteile der verschiedenen Lernformen jeweils frei bestimmt werden.

Durch Blended Learning können Vorteile von E-Learning und Präsenzlehre genutzt (siehe Tabelle 2) und ihre Nachteile minimiert werden. So ist es Lernenden möglich, Inhalte selbständig und flexibel zu lernen, ohne dabei auf Kontakte zu anderen Lernenden oder dem Lehrenden verzichten zu müssen.

[7] Im Deutschen „vermengt, vermischt", man spricht auch von „hybridem Lernen", „Distributed oder Integrated Learning" etc. (Keller & Back, 2004).

Tabelle 2
Vorteile von Präsenzlehre und E-Learning (nach Dittler & Jechle, 2009)

Präsenzlehre	E-Learning
• **Direkter Kontakt** zwischen allen Beteiligten beschleunigt und vereinfacht Kommunikation; Kontakt zwischen Teilnehmern auch in der Pause • **Gewohnte** Lernform • **Soziale Lernformen** wie z.B. das Rollenspiel sind einfach zu realisieren • Bestimmte **Interaktionen** wie z.B. Vor- und Nachmachen sind nur hier realisierbar	• **Individuelle** Lernvoraussetzungen, Lerngewohnheiten und -bedürfnisse können berücksichtigt werden • Lernort, -zeit und -geschwindigkeit können **flexibel** bestimmt werden • Lernmaterialien können einfach **aktualisiert** werden • Lernmaterial ist räumlich und zeitlich **unbegrenzt verfügbar**

Blended Learning-Szenarien erreichen nach Buzzetto-More (2008) sowie Buzzetto-More und Sweat-Guy (2006) hohe Zufriedenheitswerte auf Seiten der Lernenden. Auch die Erfolgsraten sind gleich gut oder häufig besser als reine E-Learning-Lösungen oder traditionelle Lehre (siehe u.a. Dziuban & Moskal, 2001; Rivera, McAlister & Rice, 2002).

Videokonferenzen

Ein sehr kollaboratives Lernarrangement wird durch Videokonferenzen erzeugt. Hierbei handelt es sich um ins Internet verlagerte Gruppenarbeits- und/oder Präsenzunterrichtsphasen. Innerhalb von virtuellen Klassenräumen oder durch Dienste wie *Skype™*[8] oder *Connect™*[9] treffen sich Lernende zeitgleich zu einer Videokonferenz. Hier können Lernende und Lehrende z.B. einem Vortrag folgen, sich selbst zu Wort melden und über Chats oder Emails kommunizieren. Über „Application Sharing" können sich die Teilnehmer wechselseitig ihren Bildschirm freigeben und zum Beispiel gemeinsam an einer Präsentation arbeiten (Bundesministerium für Wirtschaft und Technologie, 2008). Sicherlich verbreiteter als im Bildungsbereich ist der Einsatz von Videokonferenzen jedoch im Rahmen von beruflichen Konferenzen, an denen sich Personen unterschiedlicher Standorte beteiligen.

Televorlesungen und E-Lectures

Televorlesungen sind den Videokonferenzen sehr ähnlich und werden vielfach genutzt, um Vorträge an Hochschulen live zu anderen Orten zu übertragen. Zum Teil wird auch ein Audio- oder Videorückkanal bereitgestellt. Der didaktische Mehrwert der Televorlesungen gegenüber einer Präsenzvorlesung ist allerdings gering, da die Schwierigkeiten, die in einer

[8] Kostenlose Software von Microsoft, welche Instant-Messaging, Dateiübertragung und Videotelefonie ermöglicht (Wikipedia, 2016).
[9] Web-Kommunikationssystem der Firma Adobe für Schulungen, Web-Konferenzen und Online-Zusammenarbeit (Adobe Systems GmbH, 2012).

normalen Vorlesung herrschen (Erhalt der Aufmerksamkeit im Plenum etc.) sich hier noch verstärken, da häufig kein Lehrender im „Empfangshörsaal“ anwesend ist.
Flexibler als Televorlesungen sind E-Lectures: Vorträge, die zunächst aufgezeichnet werden und danach den Studierenden per CD-ROM/DVD oder über das Internet zur Verfügung gestellt werden. Da sie im Rahmen der vorliegenden Dissertation genutzt werden, werden sie in Abschnitt „1.4.2 E-Lecture“ ausführlich vorgestellt.

Simulationen und Planspiele, Virtuelle 3D-Welten

Simulationen, Planspiele und virtuelle 3D-Welten gelten als weniger formell als die bereits erläuterten E-Learning-Formen. Diese Anwendungen können auch unter dem Begriff „Game-based Learning“ angesiedelt werden. Ein Prozess wird am Computer nachgebildet, damit der Lernende durch eigenes Erleben selbständig nachvollziehen kann, wie er funktioniert. Bei Simulationen – beispielsweise Flugsimulatoren – arbeiten Lernende eher für sich. Planspiele, bei denen zum Beispiel Arbeitsprozesse innerhalb einer Fabrik durchgespielt werden können, bearbeiten sie zumeist innerhalb einer Gruppe. Die Lernenden treffen Entscheidungen, verändern bestimmte Aspekte innerhalb der Anwendung und bekommen durch die resultierenden Veränderungen innerhalb der Anwendung eine Rückmeldung, ob ihre Entscheidungen korrekt waren. Die Erstellung solcher Anwendungen ist sehr aufwendig und kostspielig (Bundesministerium für Wirtschaft und Technologie, 2008).

Social Networks und Communities of Practice

Soziale Netzwerke dienen zwar nicht in erster Linie als Lernangebot, können aber helfen, passende Ansprechpartner für bestimmte Probleme zu finden. Sie ermöglichen „social learning“. Auch besteht hier die Möglichkeit, Lerngruppen zu bilden, innerhalb denen sich Lernende austauschen können – sogenannte „Communities of Practice“ (Bundesministerium für Wirtschaft und Technologie, 2008). Zum Beispiel gaben Studierende in der Befragung von Kleimann et al. (2008) an, die Social Communities neben Freizeitaktivitäten auch zum Austausch über Studienangelegenheiten, wie z.B. das Klären von Fragen für das Selbststudium, die Prüfungsvorbereitung oder den Austausch von Dokumenten zu nutzen.

Web 2.0: Wikis und Podcasts, Learning on Demand

Hierunter werden internetbasierte Computeranwendungen gefasst, die ein eher informelles vernetztes gemeinsames Arbeiten von Lernenden ermöglichen. Dennoch arbeitet der einzelne Lernende überwiegend individuell mit den angebotenen Inhalten. Die „Web 2.0“- bzw. „Social Software“-Tools sind einfach zu bedienen, so dass Lernende ohne Schwierigkeiten eigene Inhalte einstellen und mit anderen austauschen können.

Ein Beispiel für ein Wiki ist das bekannte Internet-Lexikon *Wikipedia*[10], Podcasts sind downloadbare Audiodateien und unter „Learning on Demand" versteht man das spontane Lernen zu jeder Zeit und an jedem Ort (Bundesministerium für Wirtschaft und Technologie, 2008).

1.4.1 Prominente E-Learning-Formen in Studium und beruflicher Weiterbildung

Innerhalb der oben bereits erwähnten Umfrage des Instituts für Medien- und Kompetenzforschung (2010) wurden Personalverantwortliche der Top 500-Unternehmen in Deutschland nach dem E-Learning-Einsatz in ihrer Aus- und Weiterbildung befragt. Es zeigte sich, dass von den meisten Unternehmen Web- oder Computer-based Trainings, Blended Learning oder eine Kombination aus WBT und CBT genutzt werden. Ein Drittel der Unternehmen nutzt zudem Wikis oder Weblogs (siehe Abbildung 5).

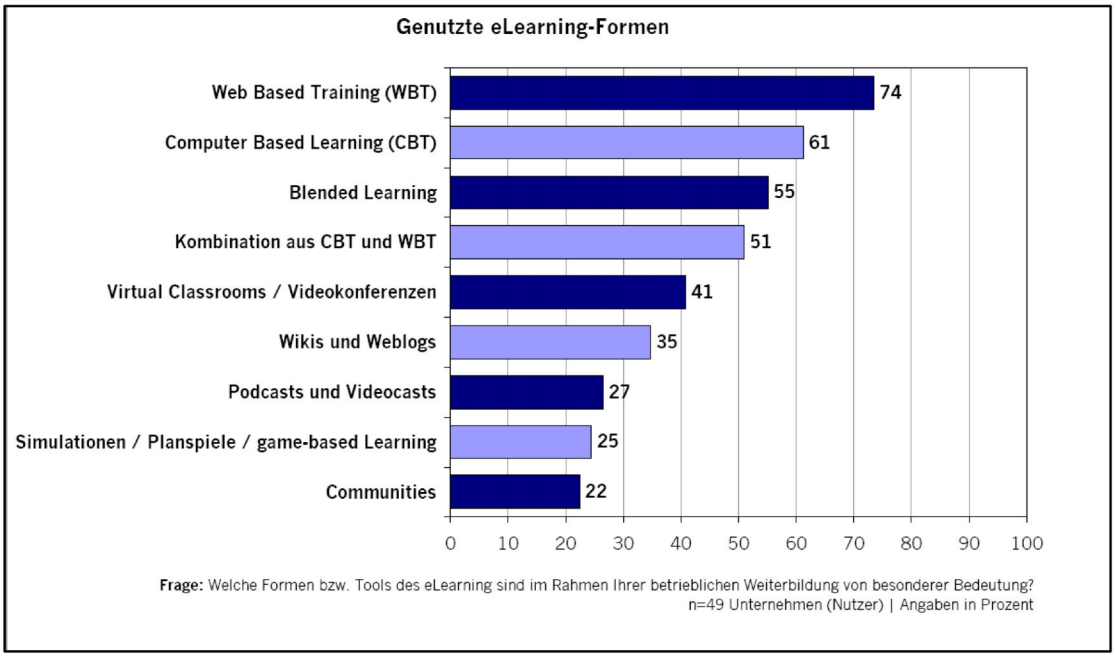

Abbildung 5. Genutzte E-Learning-Formen in deutschen TOP 500-Unternehmen (Institut für Medien- und Kompetenzforschung, 2010)

In der HISBUS-Befragung 2008 wurden hierzu im Vergleich deutsche Studierende zum Einsatz von E-Learning in ihrer Hochschule befragt (Kleimann et al., 2008). Wie in Abschnitt „1.2.2 E-Learning in Schule, Hochschule und betrieblicher Bildung" bereits erläutert, ergab sich hier, dass der Einsatz von E-Learning an den Hochschulen zumeist bedeutet, dass digitale lehrveranstaltungsbegleitende Materialien zur Verfügung gestellt werden. Dies ist sehr einfach für die Dozierenden und wird von den Studierenden mittlerweile als

[10] www.wikipedia.de

selbstverständlich angenommen. Weitere medienunterstützte Lehrformen kommen hingegen eher selten zum Einsatz (Kleimann & Schmid, 2006; Kleimann et al., 2008).
Betrachtet man allerdings, welche E-Learning-Formen deutsche Studierende zur Unterstützung ihres Studiums als sinnvoll erachten (siehe Abbildung 6, nach Kleimann et al., 2008), zeigt sich, dass 58% von ihnen insbesondere elektronische Video-Aufzeichnungen/Videopodcasts von Veranstaltungen als „sehr nützlich“ oder „nützlich“ einschätzen. Gleich darauf folgen Online-Tests und -übungen (52%), Wikis (46%) und Web-based Trainings (45%).
Ein Grund für die häufige Nennung von E-Lectures könnte nach Kleimann et al. (2008) die Vertrautheit der Lehrsituation sein, da die E-Lectures der Frontalvorlesung am ähnlichsten sind. Online-Tests bzw. -übungen sind für Studierende zudem nützlich, da es an den Hochschulen häufig nur wenige Möglichkeiten zur Wissensüberprüfung gibt und viele Studierende einen sehr prüfungsorientierten Lernstil verfolgen.

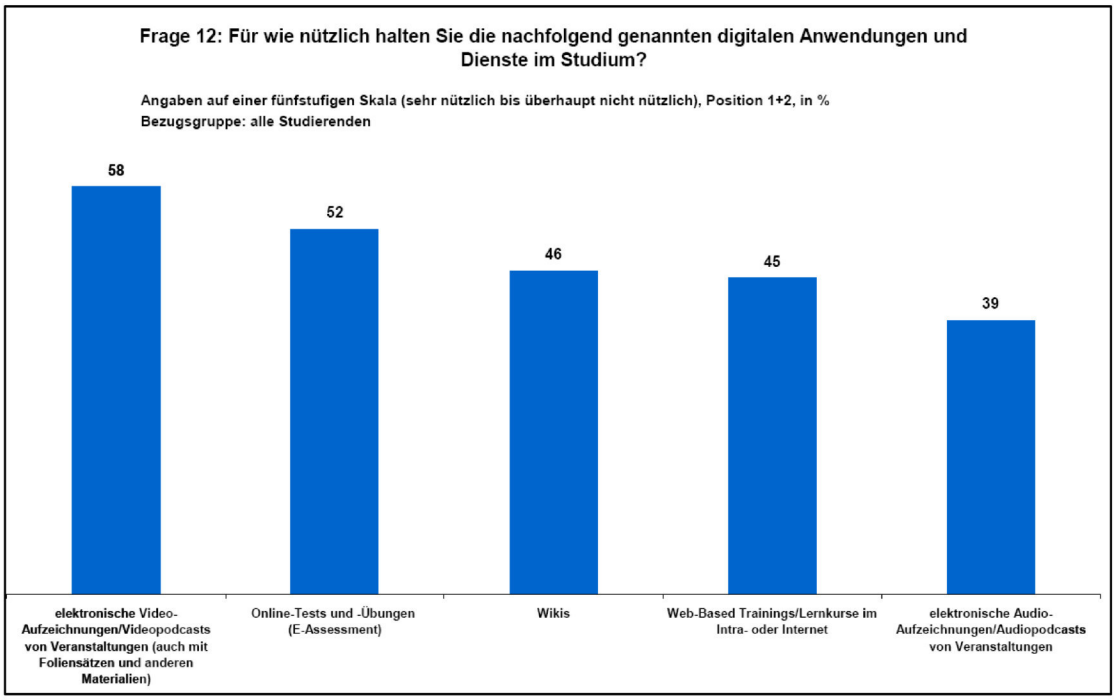

Abbildung 6. Angabe der Nützlichkeit verschiedener digitaler Anwendungen für das Studium (nach Kleimann, Özkilic & Göcks, 2008)

Etwa die Hälfe der befragten Studierenden innerhalb der HISBUS-Befragung 2008 schätzten WBTs für ihr Studium folglich als nützlich ein und auch in der betrieblichen Weiterbildung werden WBTs mit Abstand am häufigsten genutzt. Da die WBTs zudem viele weitere Vorteile aufweisen (siehe Abschnitt „1.4.3 Web-based Training (WBT)“), wurden sie für die vorliegende Untersuchung zum Vergleich mit der prominentesten Form der Präsenzlehre – der Vorlesung – ausgewählt. Daneben wurden auch E-Lectures genutzt, da sie neben ihrer Beliebtheit bei Studierenden der Vorlesung sehr ähnlich sind und darüber hinaus sehr

kostengünstig und schnell produziert werden können. Im nächsten Abschnitt wird aus diesem Grunde zunächst die E-Lecture näher erläutert, bevor ausführlich auf das Web-based Training – und in diesem Rahmen auch auf Online-Übungstests – eingegangen wird.

1.4.2 E-Lecture

Die digitalisierte Form von Vorlesungen sind die „Lecture-on-demand-Anwendungen" oder kurz „E-Lectures". Wie oben bereits erwähnt, handelt es sich hierbei um die Videoaufzeichnung einer Vorlesung bzw. eines Vortrages, welche durch synchron gezeigte Folien, schriftliche Bemerkungen des Vortragenden oder ähnliches ergänzt wird. Diese folglich hauptsächlich audiovisuellen Medien sind geeignet, viele Vorteile eines Vortrages zu erhalten und dabei einige seiner Nachteile auszugleichen.
Abbildung 7 zeigt einen Screenshot aus einer E-Lecture[11], die mit dem Autorentool *NetLecture* erstellt und in den Untersuchungen dieser Dissertation genutzt wurde.

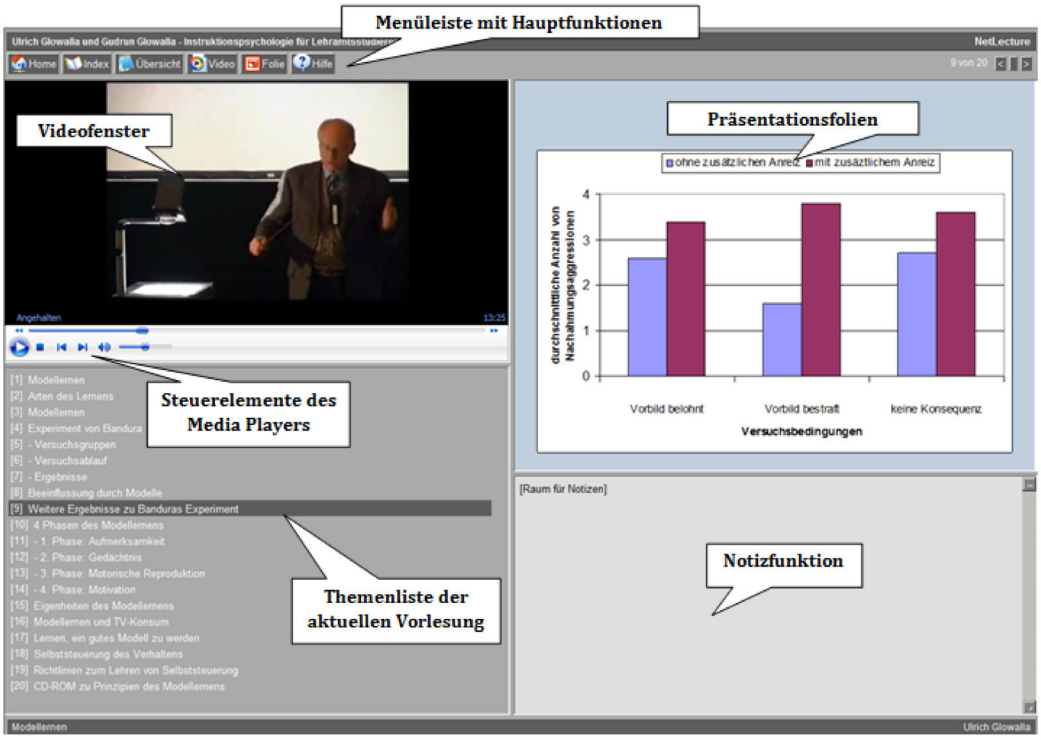

Abbildung 7. Beispiel einer *NetLecture*-Produktion mit Darstellung ihrer Funktionen (vgl. Brömme, 2010)

In der einfachsten Fassung beinhalten E-Lectures (nach Glowalla, 2003; Glowalla, 2004 sowie Görlich, 2006) eine Aufzeichnung eines Vortrags oder einer Vorlesung, welche am Stück oder in Ausschnitten abgespielt werden kann sowie die parallel zur Aufzeichnung

[11] Es handelt sich um eine der E-Lectures aus der Reihe „Instruktionspsychologie für Lehramtsstudierende." von Glowalla und Glowalla (2005).

gezeigten Folien oder vergleichbare Begleitmaterialien, wie Kommentare des Vortragenden o.ä.

Hinzu können je nach genutztem Autorentool[12] weitere Funktionen kommen (nach Glowalla, 2003; Glowalla, 2004 sowie Görlich, 2006):

- Ein Inhaltsverzeichnis, über welches bestimmte Themen angesteuert werden können. Der gewählte Aufzeichnungsausschnitt sowie die dazugehörige Folie werden automatisch angezeigt. So kann jedes Thema individuell und so häufig wie gewünscht bearbeitet werden.
- Ein Indexverzeichnis, welches ebenfalls eine gezielte Suche und Anwahl bestimmter Themen nach Schlagworten ermöglicht.
- Ein Notizfeld, welches den Lernenden die Möglichkeit bietet, Notizen zum Lehrinhalt zu machen und später weiterzuverarbeiten. Die Notizen werden gespeichert und können zusammen mit den Folien ausgedruckt werden.
- Eine Verknüpfung zu einem Learning-Management-System[13], in dem weitere Materialien wie z.B. Literaturverweise oder Übungsaufgaben zur Verfügung gestellt werden und Lernende miteinander oder mit dem Dozenten kommunizieren können.

Die Vorteile von E-Lectures liegen auf der Hand: Zum einen motivieren sie die Lernenden ähnlich wie ein realer Vortrag, zum anderen werden Nachteile des realen Vortrages ausgeglichen. So können sich die Lernenden während des Zuhörens ganz auf das Gesagte konzentrieren. Sie müssen nicht hektisch mitschreiben, um Inhalte für sich verfügbar zu halten, da sie den Vortrag jederzeit erneut aufrufen und dabei individuell Pausen machen können – das einmalige Ereignis des Vortrages wird folglich „konserviert“ (Frenger, 2010). Zudem können sich Lernende im Rahmen ihrer Prüfungsvorbereitung gezielt die Passagen ansehen, die sie nicht gut verstanden haben und noch einmal wiederholen möchten. Dies ist insbesondere dann ein großer Vorteil, wenn der Lerninhalt komplex ist oder der Vortrag in einer Fremdsprache gehalten wird. Eine E-Lecture ist demnach nachhaltiger als ein Präsenzvortrag.

Aber auch aus organisatorischen Gründen bietet eine E-Lecture klare Vorteile im Vergleich zu einer Vorlesung: Verpasst ein Studierender eine Vorlesung – sei es aus gesundheitlichen

[12] Programme, mit denen E-Lectures produziert werden können, sind zum Beispiel *IMC Lecturnity, TechSmith Camtasia, Lerndesign NetLecture, Articulate Presenter, Microsoft Producer* oder *Adobe Breeze Presenter* (Vogt et al., 2006).

[13] Learning-Management-Systeme (LMS) sind Software-Pakete, durch die Lerninhalte bereitgestellt, Kommunikation zwischen Lernenden und Lehrkräften ermöglicht sowie administrative Aufgaben (z.B. Verwaltung von Lernenden etc.) übernommen werden können. Gerade in großen Veranstaltungen ist eine angemessene Betreuung der Lernenden ohne ein LMS nur mit erheblichem Aufwand möglich (Bundesministerium für Wirtschaft und Technologie, 2008).

Gründen oder weil er/sie arbeiten oder Kinder betreuen muss – kann diese ganz bequem von zuhause aus nachgearbeitet werden – in der Zeit, die hierfür zur Verfügung steht (Glowalla, 2003; Glowalla, 2004). E-Lectures haben also klare Vorteile hinsichtlich ihrer Flexibilität im Einsatz.

Auch bezüglich der Ökonomie können E-Lectures Vorteile bieten, weil durch sie Lerninhalte ohne immensen zusätzlichen Aufwand einer größeren Menge an Lernenden verfügbar gemacht werden können. Verglichen mit anderen E-Learning-Anwendungen ist ihre Herstellung – zumindest in der einfachsten Form ohne Sonderfunktionen – relativ kostengünstig. Mittlerweile ist die Produktion von E-Lectures so vereinfacht, dass eine E-Lecture noch am gleichen Tag des Real-Vortrages im Internet oder auf DVD veröffentlicht werden kann (Frenger, 2010).

Allerdings müssen auch die Nachteile von E-Lectures hier Erwähnung finden[14]: So sind Kontakte zwischen den Lernenden und zwischen Lernendem und Lehrendem sowie eine Rückmeldung über den Lernerfolg nur über Zusatzprogramme und zusätzlichen personellen Aufwand möglich (Vogt, 2002). Darüber hinaus ist eine Anpassung der Inhalte auf verschiedene Gruppen von Lernenden nur schwer realisierbar (Frenger, 2010; Glowalla, 2004; Vogt et al., 2006).

Bell, Cockburn, McKenzie und Vargo (2001) beobachteten in ihrer Studie darüber hinaus, dass die Verfügbarkeit von E-Lectures zu einem selteneren Besuch der Präsenzveranstaltungen durch die Studierenden führte, diese in der Folge die E-Lectures dann jedoch nicht nutzten, um das Verpasste nachzuholen. In der Studie von Glowalla (2004) ergab sich allerdings, dass nur wenige Studierende (zwischen 13% und 25%) bei Verfügbarkeit von Vorlesung und E-Lecture nur mit der E-Lecture arbeiteten. Alle anderen besuchten weiterhin die Vorlesung und nutzen die E-Lecture als Ergänzung.

Zudem hängt die Qualität einer E-Lecture sehr stark von den Fähigkeiten des Vortragenden sowie von der Qualität der Produktion ab (Vogt et al., 2006). Bezüglich der Aufzeichnung von E-Lectures muss zum Beispiel zwischen „in vivo-“ und „in vitro-“Aufzeichnungen unterschieden werden. „In vivo“-E-Lectures werden live vor Studierenden aufgezeichnet, „in vitro“-E-Lectures in einem Studio ohne Publikum. Es leuchtet ein, dass der Vortragende in der „in vitro“-E-Lecture weniger die Rolle eines Lehrers einnimmt als im Rahmen der Aufzeichnung einer „in vivo“-E-Lecture. Dies kann seine Wirkung auf die Studierenden beeinflussen (Demetriadis & Pombortsis, 2007).

Damit nicht weitere Nachteile hinzukommen, ist es nach Demetriadis und Pombortsis (2007) notwendig, dass E-Lectures thematisch fokussiert und so nicht zu lang sind.

[14] Hier werden nur diejenigen Nachteile genannt, die über die in Abschnitt „1.2.3 Vor- und Nachteile von E-Learning“ hinausgehen und speziell für die E-Lecture zutreffen.

Außerdem sollten sie Themen behandeln, die nicht regelmäßig aktualisiert werden müssen, da für jede Aktualisierung der gesamte Vortrag neu aufgezeichnet werden muss. Auch ist es notwendig, dass man den Vortragenden und das, was er gerade zeigt, zur gleichen Zeit sehen kann und es eine Möglichkeit gibt, bestimmte Inhalte gezielt zu suchen und auszuwählen. Zudem sollten die Anwendungen unbedingt in Blended Learning-Szenarien eingebaut werden, um Nachteile wie z.B. hohe Drop-out-Quoten zu vermeiden. So könnte eine Vorlesung durch die E-Lecture ersetzt und die nun frei gewordene Vorlesungszeit für vertiefende Diskussionen mit Studierenden genutzt werden (vgl. Glowalla, 2004). Weitere Studien zur Evaluation von E-Lectures werden in Abschnitt „1.6.1 E-Learning vs. Präsenzlehre – Ein Blick in die Literatur“ erläutert.

1.4.3 Web-based Training (WBT)

Ein „Web-based Training“, „Web-based Teaching“ oder eine „Web-based Instruction“ (Döring, 2002; Driscoll, 1998) zeichnet sich im Gegensatz zur Vorlesung und E-Lecture durch eine stärker „buchartige“ und visuelle Darbietung des Lehrstoffs aus. Zusätzlich zu Texten und Bildern enthalten sie häufig dynamische Medien wie Animationen, Simulationen oder Videos. Durch Übungstests können Lernende ihren Lernfortschritt überprüfen. Aufwendigere WBTs bieten zudem einen Vortest, der die Lernenden zu Beginn einstuft und von unterschiedlichen Ausgangspunkten aus sowie auf unterschiedlichen Lernwegen durch den Lernstoff leitet. Zudem können in Abhängigkeit von den Testergebnissen differenzierte Feedbacks und Vertiefungen angeboten werden.
Darüber hinaus können WBTs durch Kommunikationswerkzeuge wie Chats, Emails, News, Foren etc. ergänzt werden, über die Lernende die Möglichkeit erhalten, mit anderen Lernenden oder dem Lehrenden zu interagieren oder zu kommunizieren. Diese Interaktions- und Kommunikationsmöglichkeiten sind allerdings nicht Teil des WBTs selbst, sondern werden wie bei den E-Lectures durch Learning-Management-Systeme angegliedert.
Die Abbildungen 8 bis 11 zeigen Screenshots aus dem Web-based Training zum Thema „Entwicklungspsychologie“, das im Rahmen der ersten Untersuchung dieser Dissertation genutzt wurde. Es handelt sich um das Inhaltsverzeichnis, eine typische Textseite, eine Übungsaufgabe mit einer Rückmeldung sowie die abschließende Rückmeldung am Ende eines Übungstests.



Abbildung 8. Inhaltsverzeichnis des Web-based Trainings "Entwicklungspsychologie"



Abbildung 9. Eine typische Inhaltsseite des Web-based Trainings "Entwicklungspsychologie"



Abbildung 10. Übungsaufgabe mit Rückmeldung nach falscher Beantwortung aus dem WBT "Entwicklungspsychologie"



Abbildung 11. Abschließende Rückmeldungen am Ende der Übungstests

Vorteile von Web-based Trainings sind, dass die Lerninhalte multimedial großen Nutzergruppen jederzeit und auf einfachem Wege zur Verfügung gestellt werden können. Sie bieten folglich sowohl für die Lehrenden als auch für die Lernenden Zeit- und Ortsunabhängigkeit. Lernen kann dezentralisiert und flexibel stattfinden – und dass ohne aufwendige Betreuung, da die Handhabung von WBTs für die Nutzer zumeist sehr intuitiv ist. WBTs können viele verschiedene Medien enthalten und darüber hinaus mühelos mit anderen Anwendungen verknüpft werden. Sie sind nahezu beliebig skalierbar, bieten transparente Lernerfolgskontrollen und können einfach und schnell aktualisiert werden (Scheer, 2009; Wache, 2003).

Die Erstellung eines multimedialen WBTs mit didaktischem Mehrwert kann allerdings sehr aufwendig sein. Ein gutes WBT sollte nach Döring (2002) qualitativ hochwertige Inhalte bieten, welche aktuell sind und einen Praxisbezug herstellen. Diese Inhalte müssen strukturiert, interaktiv aufbereitet und mit Visualisierungen verdeutlicht werden. Der Lernende soll durch den Lernstoff begleitet werden und die Möglichkeit haben, seinen Lernprozess jederzeit zu unterbrechen und an derselben Stelle wieder einzusteigen. Das System sollte ihm Bearbeitungszeiten, Testergebnisse oder ähnliches zurückmelden und so

seinen selbstgesteuerten Lernprozess unterstützen. Dass die Erstellung einer solchen Lernanwendung aufwendig ist, liegt auf der Hand. Darüber hinaus zeigen Erfahrungswerte, dass die Produktion eines einfachen Web-based Trainings aufwendiger ist als die einer schlichten E-Lecture. Für die Durchführung der Studien zur vorliegenden Dissertation wurden insgesamt 48 WBTs erstellt. Die inhaltsgleichen E-Lectures lagen zwar bereits zu Beginn der Studien vor, im Rahmen von anderen Projekten der Forschungsgruppe wurden jedoch einige E-Lectures mit den Autorentools *NetLecture* oder *Camtasia* erstellt. Hierbei wurde sehr schnell deutlich, dass die Erstellung einer E-Lecture deutlich weniger Ressourcen in Anspruch nimmt als die eines WBTs. Im Gegenzug dazu haben die WBTs klare Vorteile hinsichtlich der Aktualisierung von Lerninhalten. Während Inhalte in WBTs durch den Austausch einzelner Seiten, Abbildungen oder gar die Überarbeitung einzelner Texte erfolgen kann, muss eine E-Lecture von Grunde auf neu erstellt werden.
Studien zur Evaluation von Web-based Trainings werden in Abschnitt „1.6.1 E-Learning vs. Präsenzlehre – Ein Blick in die Literatur“ erläutert.

1.4.3.1 Exkurs: Online-Übungstests

Eine wichtige Rolle innerhalb von Web-based Trainings spielen Online-Übungstests. Sie sind wichtig, um den Lernerfolg der Lernenden zu erfassen und durch gezieltes Feedback die Aktivität und die Lernfortschritte der Lernenden zu fördern. Nach Bearbeitung einer Übungsfrage kann der Lernende in Abhängigkeit von dem Ergebnis sein Lernverhalten anpassen. Hat er die Aufgabe korrekt beantwortet, fährt er mit dem nächsten Lerninhalt fort. Hat er sie jedoch falsch beantwortet, muss er noch einmal „zurückblättern“ – am besten gezielt an die Stelle, die er noch nicht verstanden hat – und den Lernstoff erneut bearbeiten. Dabei hat die computerbasierte Wissensdiagnose klare Vorteile gegenüber traditioneller Wissensüberprüfung: Ergebnisse können automatisch gespeichert und ausgewertet werden. Der Lernende erhält eine unmittelbare Rückmeldung über sein Ergebnis – gegebenenfalls mit Musterlösung und Link zu relevanten Informationen (Glowalla, Kohnert & Schneider, 2003). Dies geschieht, ohne dass ein Lehrender die Geduld verliert oder sich der Lernende für sein Nichtwissen schämen müsste.
Dass Übungsfragen positive Effekte auf die Lernwirksamkeit haben, ohne dabei auch unweigerlich die Lernzeit zu erhöhen, konnte bereits die Metaanalyse von Hamaker (1986) zeigen (siehe auch Jacobs, 2001). Nungester und Duchastel (1982) ließen Studierende nach dem Lesen eines Textes entweder Übungsaufgaben oder eine irrelevante Aufgabe bearbeiten oder den Text erneut lesen. Zwei Wochen später zeigte sich, dass sich diejenigen, die die Übungsaufgaben bearbeitet hatten, besser an die Inhalte aus dem Text erinnern konnten als die Studierenden aus den beiden Kontrollbedingungen. Auch Roedinger III und Karpicke (2006) konnten zeigen, dass Tests das Lernen gegenüber dem erneuten Lesen der

Inhalte deutlich verbessern können („testing effect“) – und das sogar, wenn die Tests kein Feedback boten. Ähnliche Ergebnisse finden sich auch bei Larsen, Butler und Roediger III (2009), welche die Lernenden mehrfach mit Übungsaufgaben oder dem erneuten Bearbeiten der Inhalte konfrontierten. Diejenigen, die wiederholt (mit Feedback) getestet wurden, erreichten nach sechs Monaten ein höheres Lernergebnis als diejenigen, die das Material mehrfach bearbeitet hatten. Vergleichbare Ergebnisse zeigten sich bei Butler (2010).

Siegert, Schlieck, Jäger, Gotthardt und Glowalla (2003) fanden allerdings keine Unterschiede zwischen der Akzeptanz oder der Lernwirksamkeit von Web-based Trainings mit und ohne Übungsaufgaben. Alle Untersuchungsgruppen bewerteten ihr Web-based Training überwiegend als sehr effektiv und zufriedenstellend und unterschieden sich nicht hinsichtlich ihrer Lernzeit und Lerneffektivität. Bezogen auf die überraschend nicht-signifikanten Lernwirksamkeitsunterschiede argumentieren die Autoren, dass die Lernwirksamkeit der WBTs mit 74% bereits sehr hoch lag. Augenscheinlich wiesen die WBTs mit Übungsaufgaben hier Vorteile auf (77% / 78% vs. 70%). Durch den insgesamt sehr hohen Lernerfolg der Teilnehmer ergab sich jedoch eine eingeschränkte Varianz, welche es unwahrscheinlicher macht, dass Signifikanzen auftreten. Hinzu kommt eine zum Teil sehr kleine Stichprobe, die diesen Effekt noch verstärkt. Die Ergebnisse dieser Studie müssen folglich mit Achtsamkeit interpretiert werden.

Bezüglich der Darbietungsform des Übungstests zeigte Görlich (2006), dass sowohl Papier- und Bleistift- als auch Online-Übungstests von Studierenden gut angenommen werden. Die Studierenden beurteilten sie als nützlich für die Klausurvorbereitung und wünschten sich häufiger solche Tests. Dabei schnitten die Studierenden nach Bearbeitung der Online-Tests deutlich besser ab als nach Bearbeitung von Papier- und Bleistift-Tests. Dies ist von Seiten der Ersteller erfreulich, da für die onlinebasierten Tests weniger administrativer Aufwand nötig ist. Vergleichbares zeigte die Studie von DeSouza und Fleming (2003). Auch hier schnitten die Nutzer des Online-Quizzes besser ab als die Nutzer des Papier- und Bleistift-Tests. Dabei zeigten sich die Lernenden sehr zufrieden mit den Online-Quizzes.

Studierende begrüßen folglich alle Angebote, mit denen sie ihr Wissen überprüfen können, ohne gleich geprüft und benotet zu werden (Glowalla et al., 2004).

Auch in der Studie von Klauser, Kim und Born nannten Studierende die Merkmale Übersichtlichkeit, klare Lernzielformulierung, Lernerfolgskontrolle und Übungsaufgaben als besonders wichtige Kriterien für gute E-Learning-Anwendungen (Klauser, Kim & Born, 2002). Dabei ist der Lernerfolg umso größer, je besser die Lernverhaltenssteuerung durch den Einsatz von Übungstests, Feedback und Lernempfehlungen gelingt (Glowalla et al., 2003).

Nach Blick auf die Literatur kann folglich zusammengefasst werden, dass Übungsaufgaben wichtige Bestandteile von Web-based Trainings sind und auch sein sollten.

Bevor nun auf die Fragestellungen der ersten Studie eingegangen wird, soll zunächst erläutert werden, welche Möglichkeiten zur Evaluation von E-Learning-Anwendungen bestehen und bei der Planung einer solchen Studie zu bedenken sind.

1.5 Evaluation von E-Learning-Anwendungen

Wie bereits erläutert kann die Entwicklung und Bereitstellung von E-Learning-Anwendungen relativ kostenintensiv sein. Aber auch in den Fällen, in denen E-Learning aus Gründen der Kostenersparnis eingesetzt wird, ist es notwendig, die Anwendungen sinnvoll zu evaluieren[15] und ihre Qualität zu beurteilen, um ihre Lernwirksamkeit sicherzustellen. Die korrekte Auswahl und Anwendung der hierfür notwendigen Evaluationskriterien und -methoden können den Mehrwert und die Nachhaltigkeit von E-Learning-Anwendungen substantiell steigern.

1.5.1 Notwendige Vorüberlegungen

Um die Qualität einer E-Learning-Anwendung zu untersuchen, gibt es verschiedene Möglichkeiten. Deshalb ist es notwendig, gleich zu Beginn des Evaluationsprozesses festzulegen, welches Ziel man mit der Evaluation erreichen möchte. Will man zum Beispiel lediglich feststellen, ob die Anwendung dem geplanten Ziel dient, gilt es selbstverständlich anders vorzugehen, als wenn man untersuchen möchte, ob eine Anwendung besser ist als eine andere (Glowalla, Herder, Süße & Koch, 2011). Der Evaluator muss folglich zunächst entscheiden, ob neben der Experimentalgruppe auch eine Vergleichsgruppe erhoben werden soll, die mit einer anderen Lernmethode als der zu evaluierenden arbeitet.
Des Weiteren gilt es zu überlegen, ob es für den gegebenen Fall sinnvoller ist, eine Feld- oder eine Laboruntersuchung durchzuführen. Eine Laboruntersuchung hat den Vorteil, dass man für alle Lernenden eine kontrollierte Lernsituation herstellen kann, was den Einfluss möglicher Störfaktoren minimiert. Man benötigt allerdings ausreichend Computerarbeitsplätze mit gleicher Ausstattung oder muss über einen relativ langen Zeitraum hinweg einzelne Lernende an den Rechnern arbeiten lassen. Die Untersuchung ist orts- und zeitgebunden und die Anzahl an Probanden wird durch den Aufwand stark eingeschränkt. Eine Felduntersuchung hingegen hat den Vorteil, dass ohne großen Aufwand

[15] Unter Evaluation verstehen Glowalla et al. (2011, S. 310) die „systematische Kontrolle von Qualität, Funktionalität und Wirkung eines Produktes oder Verfahrens“.

eine große Anzahl an Probanden in die Untersuchung einbezogen werden kann. Sie ist folglich ökonomischer und flexibler und bietet zudem eine realistischere Lernsituation. Allerdings ist hier zu beachten, dass bei Befragungen relativ niedrige Rücklaufquoten in Kauf zu nehmen sind. Dies führt häufig auch dazu, dass sich die Stichprobe nicht mehr zufällig zusammensetzt, da sich der einzelne Lernende aus bestimmten Gründen für oder gegen die Teilnahme entscheidet. Eventuell nehmen nur diejenigen teil, die die Anwendung besonders gut oder besonders schlecht fanden. Hier müssen folglich geeignete Maßnahmen zur Erhöhung der Rücklaufquote ergriffen werden. Zudem kann die Lernsituation nicht umfassend kontrolliert werden. Gerade die Zeit- und Ortsunabhängigkeit von E-Learning-Anwendungen kann dazu führen, dass Nutzer in einer ganz anderen Lernumgebung lernen als von den Entwicklern der Anwendung geplant. Dieser Aspekt kann einen starken Einfluss auf den Lernerfolg und die Akzeptanz des Angebotes haben (Glowalla et al., 2011).

1.5.2 Evaluationskriterien

Auch die Auswahl der geeigneten Evaluationskriterien spielt im Rahmen der Planung eines Evaluationsprozesses eine wichtige Rolle. So kann die Erfassung der Akzeptanz einer Lernanwendung durchaus interessant sein, von einer hohen Akzeptanz kann jedoch nicht unmittelbar auf eine gute Lernwirksamkeit geschlossen werden. Akzeptanzbeurteilungen von Lernenden sind immer subjektiv und damit Urteilsfehlern unterworfen. Lernende selbst sind zumeist nicht in der Lage zu beurteilen, ob eine Lernanwendung lernwirksam ist oder nicht – oder lernwirksamer ist als eine andere (Glowalla & Häfele, 1997). Da eine hohe Akzeptanz allerdings häufig mit einer intensiveren Nutzung der Anwendung einhergeht, muss man die Beurteilung der Lernenden sehr ernst nehmen (Glowalla et al., 2011). Zudem können spezifische Akzeptanzdaten dazu genutzt werden, Schwächen einer E-Learning-Anwendung genauer zu identifizieren und zu korrigieren (Vogt, 2002).
In der Literatur wird zwischen Einstellungs- und Verhaltensakzeptanz unterschieden. Nach Müller-Böling und Müller (1986) setzt sich die Einstellungsakzeptanz aus einer affektiven, motivational-emotionalen und einer kognitiven (Kosten-Nutzen-) Komponente zusammen. Sie kann lediglich erfragt werden. Die Verhaltensakzeptanz meint hingegen das konkret beobachtbare Verhalten, wie z.B. die Nutzung einer E-Learning-Anwendung. Die Unterscheidung zwischen Einstellungs- und Verhaltensakzeptanz ist sinnvoll, da eine hohe (subjektive) Einstellungsakzeptanz nicht zwangsläufig bedeutet, dass die Lernenden das Lehrmedium wirklich zum Lernen nutzen. Die Einstellungsakzeptanz ist allerdings Hauptprädiktor für die objektivere Verhaltensakzeptanz.

Die Qualität einer Lernanwendung kann nicht alleine durch die Akzeptanz belegt werden. Deshalb sollte unbedingt auch die Lernleistung erfasst werden. Nur sie kann Auskunft darüber geben, ob und wie viel die Lernenden durch eine Anwendung gelernt haben. Allerdings eignet sich nicht jedes Maß zur Lernerfolgsmessung gleichermaßen. Zum Beispiel eignen sich Klausurnoten nur bedingt, um die Qualität einer Lernanwendung zu beurteilen. Eine schlechte Lehrstoffpräsentation wird von Lernenden häufig kompensiert, indem sie beispielsweise den Lernstoff intensiver mit anderen Quellen nachbearbeiten (Glowalla, Glowalla & Kohnert, 2002). Glowalla et al. (2011) unterscheiden deshalb die Erhebung der Lernleistung auch von der Erhebung der Lernwirksamkeit – wobei die Lernwirksamkeit durch das Erreichen von externen Zielen, wie z.B. dem Bestehen einer Prüfung, belegt wird, während die Lernleistung durch extra dafür vorgesehene Tests gemessen wird.

Entscheidend ist ebenfalls, welches Ziel ein Lernender durch seine Lernhandlung verfolgt. Je nachdem, welches Lernziel erreicht werden soll, wird ein anderes Lernverhalten an den Tag gelegt. Erfasst man also nicht, ob jemand sein Lernziel erreicht hat, verbietet sich streng genommen auch eine Interpretation des Lernverhaltens (Glowalla & Häfele, 1997).

Notwendig ist zudem, nicht nur die Lernleistung *nach* dem Lernen, sondern auch das Vorwissen der Lernenden *vor* dem Lernen zu erfassen. Erst die Differenz dieser beiden Messzeitpunkte kann als Maß für den Wissenszuwachs gewertet werden. Um eine Aussage über die längerfristige Lernleistung zu machen, ist es außerdem sinnvoll, die Lernleistung nach einer bestimmten Zeit erneut zu messen (Glowalla et al., 2011).

Ergänzend zur Lernleistung sollte immer auch die Lernzeit erhoben werden. So kann es sein, dass zwei Lernangebote zu vergleichbaren Lernleistungen führen – eine sich jedoch zusätzlich durch eine viel kürzere Lernzeit auszeichnet. Die Lernangebote sind folglich vergleichbar effektiv unterscheiden sich jedoch hinsichtlich ihrer Effizienz.

Eine weitere sinnvolle Variable ist die Erhebung des Lernverlaufs/Lernweges bzw. des Lernverhaltens. Evaluiert man mittels den oben genannten Kriterien eine E-Learning-Anwendung im Ganzen, betrachtet man hier wesentlich spezifischer einzelne Komponenten innerhalb einer E-Learning-Anwendung. E-Learning-Anwendungen können von den Lernenden sehr unterschiedlich genutzt werden. Der eine bearbeitet sie vollständig, der andere schaut sich nur das an, was er zum Schließen seiner Wissenslücken bearbeiten muss. Genauso werden bestimmte Elemente innerhalb einer E-Learning-Anwendung (z.B. Animationen, einzelne Passagen, Übungsaufgaben) intensiv genutzt oder völlig ignoriert. Erst durch die genaue Analyse des Lernverhaltens und -weges können Schwierigkeiten der Lernenden mit dem Lernangebot sowie dessen Gebrauchstauglichkeit (Usability) aufgedeckt werden. Darüber hinaus wird sichtbar, welche Komponenten innerhalb der

Anwendung besonders gut akzeptiert werden bzw. lernwirksam sind und für welche Komponenten Optimierungsbedarf besteht. So kann das E-Learning-Angebot gezielt verbessert werden. Dies lässt sich durch die Erhebung der vorher genannten Aspekte allein nicht bewerkstelligen. Dabei reicht in den meisten Fällen die Befragung der Lernenden, in anderen Fällen muss das Nutzerverhalten genauer gemessen werden (Glowalla & Häfele, 1997; Glowalla et al., 2011).

Zu guter Letzt ist zudem eine Betrachtung des Kosten-Nutzen-Verhältnisses sinnvoll. Diese in der Universität eher vernachlässigte Variable spielt in Unternehmen eine bedeutende Rolle, will man doch wissen, ob sich die Kosten, die man für die Produktion und Bereitstellung der E-Learning-Anwendung aufgewendet hat, auch gelohnt haben (Glowalla et al., 2011).

1.5.3 Evaluationsmethoden

Um die oben genannten Variablen zu erfassen, stehen unterschiedliche Möglichkeiten zur Verfügung. Zum einen kann man mit Befragungen arbeiten, die zumeist die subjektive Meinung oder Einschätzung der Nutzer erfassen. Dies erlaubt – besonders bei der Nutzung von Online-Fragebögen – die Befragung von sehr vielen Teilnehmern ohne großen Aufwand. Die Urteile müssen allerdings nicht unbedingt mit objektiv erhobenen Daten übereinstimmen.
Eine objektivere Erhebungsmethode ist die Beobachtung. Hierbei wird das Verhalten der Lernenden im Umgang mit der E-Learning-Anwendung beobachtet. Dies ist über eine direkte Beobachtung, eine Auswertung der Videoaufzeichnung, eine Beobachtung von Denkprozessen, eine Wissensdiagnostik als besondere Beobachtungsform oder eine Blickbewegungsmessung möglich. Insbesondere die Blickbewegungsmessung wird gerne für Usability-Untersuchungen eingesetzt.
Eine besondere Erhebungsform, die nur im Rahmen von Untersuchungen mit computerbasiertem Material möglich ist, ist das „User Tracking“. Dabei zeichnet der Computer das Verhalten des Nutzers im Umgang mit dem jeweiligen Programm auf. Hierdurch können eine Vielzahl an Variablen aufgezeichnet werden, wie zum Beispiel die Nutzungszeiten, Aufrufe bestimmter Funktionen (z.B. Hilfe, Suche), der Lernweg, der genutzt wurde, Pausen und Fehlaktionen des Lernenden, Antworten der Lernenden auf die Übungsaufgaben sowie die dazugehörigen Antwortzeiten etc. (Glowalla et al., 2011).

In den meisten Evaluationen werden mehrere der vorgenannten Variablen auf unterschiedliche Art erhoben, um eine tragfähige Aussage zur Qualität der Anwendung(en)

machen zu können. Führen unterschiedliche Methoden zum gleichen Ergebnis, verstärkt dies die Zuverlässigkeit des Evaluationsergebnisses („Converging Evidence“). Es ist allerdings auch möglich, dass hierdurch Widersprüche aufgedeckt werden. So kann die subjektive Lernförderlichkeit und Akzeptanz einer Methode ganz anders ausfallen als die objektiv gemessene Lernleistung (Glowalla et al., 2011).

1.6 Fragestellungen der ersten Studie

Da das Hauptaugenmerk der ersten Studie auf dem Vergleich zwischen den prominenten Lehrformen Vorlesung, E-Lecture und Web-based Training liegt, soll im Folgenden zunächst ein Überblick über verschiedene Studien gegeben werden, die ebenfalls einen Akzeptanz- oder/und Lernwirksamkeitsvergleich zwischen einer computervermittelten Lehrform und einer Präsenzlehrform vorgenommen haben (weitere Studien siehe z.B. Schulmeister, 2007). In der Folge soll erläutert werden, wie sich die vorliegende Studie von diesen unterscheidet und was ihren Mehrwert ausmacht, bevor die Fragestellungen hergeleitet werden.

1.6.1 E-Learning vs. Präsenzlehre – Ein Blick in die Literatur

Zunächst lassen sich in der Literatur einige Studien finden, die sich generell dem Vergleich von Präsenzlehre und E-Learning widmen, ohne die Form des E-Learning konkret einzugrenzen. Hierzu gehört zunächst die Studie von Kulik und Kulik (1991), die eine Metaanalyse über 254 Evaluationsstudien zum Vergleich zwischen computerbasierter Instruktion (CBI) und herkömmlicher Lehre durchführten. Hierbei zeigte sich, dass Lernende eine bessere Lernleistung (0.3 Standardabweichungen) erreichten, wenn sie computerbasierte Instruktion genutzt hatten.

Auch Schutte (1997) verglich internetbasierte mit traditioneller Lehre. Die Lernenden, die internetbasiert gelehrt wurden, zeigten u.a. bessere Lernleistungen (20%) und positivere Einstellungen gegenüber dem Fach und berichteten über ein besseres Verständnis der Inhalte.

Carey (2001) fand in ihrer Studie hingegen keine Unterschiede bezüglich Lernerfolg und Zufriedenheit, jedoch eine höhere Dropout-Rate in der E-Learning-Gruppe.

In der Studie von Glowalla (2003) wurden E-Learning-Anwendungen hinsichtlich ihrer Effektivität und Konzentrationserhaltung von den Studierenden als deutlich überlegen gegenüber der inhaltsgleichen Vorlesung eingeschätzt. Etwa 50% der Studierenden erlebte das E-Learning-Szenario als motivierender. Ein Vergleich zwischen der E-Learning-Anwendung und dem Lehrbuch fiel ähnlich aus.

Lyon Jr. et al. (1992) konnten durch den Einsatz einer E-Learning-Maßnahme sowohl für die Studierenden als auch für die Lehrenden Effizienzgewinne erreichen.
In der Studie von Wernet, Olliges und Delicath (2000) gaben die Lernenden an, das zur Verfügung gestellte Online-Lernmaterialien als hilfreich zu empfinden.
Sanders und Morrison-Shetlar (2001) evaluierten einen E-Learning-Kurs und beobachteten, dass die Bereitstellung vielfältiger Onlinematerialien positive Effekte auf die Lernleistung, die Problemlösefähigkeiten sowie das kritische Denken der Lernenden hatte.
In der Studie von Maki, Maki, Patterson und Whittaker (2000) führte die Nutzung einer E-Learning-Anwendung zu einem besserem Ergebnis im Nachtest. Darüber hinaus nutzten die Lernenden in der Folge häufiger das Internet und den Computer für Studienbelange und ihre Computerangst verringerte sich deutlich. Die Teilnehmer an der Vorlesung zeigten allerdings eine bessere Akzeptanz als die Teilnehmer am E-Learning-Szenario.
Wofford, Spickard III und Wofford (2001) betrachteten acht Studien zum Vergleich zwischen Vorlesung und E-Learning. In sechs der Studien ergab sich kein signifikanter Unterschied hinsichtlich der Lerneffektivität, in zwei schnitt das E-Learning-Angebot signifikant besser ab.
Shomaker, Ricks und Hale (2002) verglichen ein E-Learning- mit einem Präsenz- und einem hieraus gebildeten Blended Learning-Szenario. Bezüglich aller drei Varianten ergaben sich eine generell gute Akzeptanz, jedoch keine Signifikanzen hinsichtlich der Lernwirksamkeit. Die Studierenden in der E-Learning-Gruppe verbrachten signifikant weniger Zeit mit den Lernmaterialien als die Studierenden in der Blended Learning-Gruppe.

Da sich die vorliegende Studie im Gegensatz zu den bisher vorgestellten gezielt mit bestimmten Formen von E-Learning-Anwendungen auseinandersetzt, sollen in der Folge darüber hinaus solche Studien betrachtet werden, die einen Vergleich zwischen E-Lectures (bzw. ihren Vorgängerformen) und/oder Web-based Trainings und der Präsenzlehre unternommen haben.
Zuerst zu nennen ist hier die Studie von Schramm (1962), der einen Vergleich von TV-gestützten Vorlesungen und Präsenzvorlesungen unternahm. Es zeigte sich, dass Lernende schnell und effizient durch TV-gestützte Vorlesungen lernen können und dabei zumindest gleich gut, manchmal jedoch sogar besser lernen als durch Präsenzlehre. Schramm (1962) diskutiert als Ursache hierfür die häufig bessere Vorbereitung von TV-gestützten Vorlesungen.
Chu und Schramm (1967) betrachteten 207 Studien zum Vergleich zwischen Vorlesung und TV-Übertragung. In der Mehrheit der betrachteten Studien zeigten sich keine signifikanten Unterschiede.

LaRose, Gregg und Eastin (1998) verglichen eine auf Tonband aufgenommene Vorlesung mit einer Präsenzvorlesung und fanden ebenfalls keine signifikanten Unterschiede zwischen den beiden Lehrformen bezüglich Lernleistung und Akzeptanz.

Ellis und Mathis (1985) konnten ebenfalls keinen signifikanten Unterschied feststellen hinsichtlich der Lernleistung oder der Abwesenheit von Studierenden, die eine Präsenzvorlesung und Studierenden, die die Übertragung dieser Vorlesung via Fernseher sehen konnten.

Johnson, Aragon, Shaik und Palma-Rivas (2000) verglichen eine E-Lecture (lediglich mit Audiokanal) und eine Präsenzvorlesung. Es zeigten sich keine Unterschiede bezüglich der Lernleistung und lediglich geringe Unterschiede hinsichtlich der Akzeptanz beider Experimentalgruppen. Die Lernenden waren hier etwas zufriedener mit der Präsenzvorlesung.

Spickard III, Alrajeh, Cordray und Gigante (2002) stellten eine Präsenzvorlesung einer Online-Vorlesung gegenüber. Beide Teilnehmergruppen schnitten vergleichbar gut ab und waren ähnlich zufrieden mit dem Lehrangebot, die Online-Vorlesung erreichte jedoch eine deutlich höhere Lerneffizienz.

Demetriadis und Pombortsis (2007) verglichen eine Präsenzvorlesung mit anschließender Diskussion mit der Nutzung einer E-Lecture mit anschließender Präsenzdiskussion. Sie konnten keinen Unterschied hinsichtlich der Lernleistung der Lehrformen feststellen, jedoch beteiligten sich die Studierenden, die vorab die E-Lecture genutzt hatten, weniger stark an der Diskussion. Generell betonten die Studierenden, dass sie die E-Lectures gut fanden – allerdings erschien ihnen auch die Kommunikation zu den Lehrenden als sehr wichtig.

In der Studie von Glowalla (2003) stand es Studierenden frei, sich nach einer Präsenzvorlesung mit oder ohne eine E-Lecture-CD-ROM, welche diese Vorlesung inhaltlich abdeckte, auf die abschließende Klausur vorzubereiten: 72% derjenigen, die die E-Lectures zum Lernen genutzt hatten, bestanden die Klausur, während dies nur für 49% derjenigen zutraf, die auf die E-Lectures verzichtet hatten. Dabei stieg die in der Klausur erreichte Punktzahl mit der Nutzungsdauer der E-Lecture. Glowalla (2004) schränkt ein, dass dieses Ergebnis evtl. auch dadurch zustande kommt, dass besonders motivierte Studierende, die wegen ihrer insgesamt umfangreicheren Vorbereitung ganz natürlich in der Klausur besser abschneiden, auch eher die E-Lecture für ihre Klausurvorbereitung genutzt hatten. In einer weiteren Variation zeigte sich, dass Studierende, die anstatt eines Dozenten-Videos nur ein Standbild des Dozenten sahen, ihren Lernprozess ähnlich effektiv einschätzten wie innerhalb einer Vorlesung. Studierende, die jedoch ein Video dargeboten bekamen, schätzten ihre Arbeit als effektiver ein als in der Vorlesung. Auch waren sie der Meinung, sich besser konzentrieren zu können.

Glowalla, Glowalla und Görlich (2004) befragten Studierende, die mit einer E-Lecture gearbeitet hatten. Diese erlebten die Arbeit mit der E-Lecture als effektiver als den Besuch der Vorlesung – sie schätzten den Besuch einer Vorlesung jedoch als motivierender ein. 98% der Studierenden wünschten sich mehr E-Lectures für ihr Studium, 79% möchten diese allerdings nur als Ergänzung zur Vorlesung nutzen, nicht als Ersatz.

Glowalla (2004) konnte im Rahmen einer Usability-Untersuchung zeigen, dass Lernende mit der Bedienung von E-Lectures gut zurechtkommen. Zudem beurteilten die meisten das Arbeiten mit der E-Lecture als positiv und effektiv. 90% der Studierenden machte die Arbeit mit der E-Lecture Spaß. Die E-Lectures wurden sehr intensiv bearbeitet – jede Vorlesungsstunde wurde zumindest einmal betrachtet. Darüber hinaus wurden bestimmte Themen wiederholt. Die E-Lectures motivierten die Studierenden folglich, die Vorlesungen nachzubearbeiten und dem Vergessen entgegenzuwirken.

Görlich (2006) konnte dies bestätigen. Auch in seiner Studie stellten sich die E-Lectures als zum Arbeiten und Lernen gebrauchstauglich heraus und wurden ebenfalls von den Studierenden positiv bewertet. Zudem konnte gezeigt werden, dass die Studierenden das Arbeiten mit einer E-Lecture sehr ähnlich zum Besuch der realen Vorlesung wahrnehmen. Auch der Dozent wurde nicht kühler oder unpersönlicher wahrgenommen, wobei es anscheinend eine Rolle spielte, dass die E-Lecture „in vivo“ aufgezeichnet wurde. Die Arbeit mit der E-Lecture wurde als effektiver und konzentrierter eingestuft als der Besuch der Vorlesung und zum Lernen besser eingeschätzt als viele andere Lehrmedien wie Skript, Vorlesung, Bücher oder Folien. Jedoch ergab sich auch in der Studie von Görlich (2006) kein Unterschied zwischen Vorlesung und E-Lecture in Bezug auf die Lernwirksamkeit. Dabei arbeiteten die Studierenden länger mit der E-Lecture als sie die Vorlesung besucht hätten. Studierende sahen die E-Lectures eher als Ergänzung zur Vorlesung für ihre konkrete Prüfungsvorbereitung – nicht als Ersatz der Vorlesung. Die Zusatzfunktionen der E-Lectures (Notizenfeld, Indexsuche) wurden nur selten genutzt (siehe auch Glowalla, Glowalla & Kohnert, 2002). Stattdessen wurden die E-Lectures von den Studierenden zumeist von vorne nach hinten durchgesehen, ohne dass häufig gespult oder pausiert wurde.

In der Studie von Hasebrook (2002) ergab ein Vergleich zwischen Web-based Training und Präsenzlehre keinen Unterschied hinsichtlich der Lerneffektivität – allerdings lernten die WBT-Nutzer die gleichen Inhalte in nur 70% der Zeit. Die WBTs waren damit lerneffizienter als die Präsenzlehre. Zudem schätzten die Teilnehmer an der Präsenzlehre ihre Lernergebnisse höher ein als die WBT-Nutzer, obwohl dies objektiv nicht der Fall war.

Olson und Wisher (2002) fassten 47 WBT-Evaluationen aus den Jahren 1996 bis 2002 zusammen. 15 der Studien boten Effektstärken zum Vergleich der Lernwirksamkeit von WBT und Präsenzlehre. Acht dieser Vergleiche fielen zugunsten der WBTs aus, sieben

zugunsten der Präsenzlehre. Die Effektstärken variierten dabei sehr stark von -.4 zu 1.6, so dass sich über alle 15 Studien hinweg eine Effektstärke von .24 zugunsten der Web-based Trainings ergab.

In der Studie von Vogt et al. (2006) wurde ein Vergleich zwischen Web-based Training, E-Lecture und einer Präsenzveranstaltung gezogen. Die Teilnehmer in allen drei inhaltsgleichen Experimentalgruppen beurteilten den Inhalt, die Nützlichkeit und die Verständlichkeit des jeweiligen Lehrangebotes als sehr gut. Die Teilnehmer der Präsenz-Gruppe favorisierten in einem hypothetischen Vergleich die Präsenzsitzung, während die Teilnehmer der computervermittelten Lehre sich demgegenüber zweigeteilt zeigten. Die Einstellung der Lernenden gegenüber computervermittelter Lehre konnte durch die Teilnahme an einem E-Learning-Lernszenario folglich etwas verbessert werden.

Auch in der Studie von Vogt et al. (2006) nutzten die Lernenden die computerbasierten Lehrformen hauptsächlich linear. Ausnahmen bestanden lediglich darin, dass in der E-Lecture einzelne Stellen wiederholt und innerhalb der WBTs die Übungstests häufiger genutzt wurden. Die Möglichkeiten, mit anderen Lernenden oder mit den Lehrenden zu interagieren, wurden in der Präsenzveranstaltung am besten beurteilt.

Hinsichtlich der Lernleistung ergab sich lediglich zum ersten Testzeitpunkt ein signifikanter Unterschied zwischen der E-Lecture- und WBT-Gruppe zugunsten des WBTs, langfristig unterschieden sich die Gruppen nicht mehr voneinander. Die computerbasierten Lernanwendungen waren allerdings durchgängig lerneffizienter: Um den gleichen Lerneffekt zu erreichen, benötigten die Lernenden nur 60% der Zeit, die sie für die Präsenzveranstaltung hätten aufwenden müssen.

1.6.1.1 Zusammenfassung und Diskussion

Die vorgestellten Studien zum Vergleich von E-Learning-Anwendungen und Präsenzlehre zeigen, dass sich zwar häufig kein Lernwirksamkeitsvorteil einer Methode ergibt, dass allerdings in den Fällen, in denen ein Unterschied hinsichtlich der Lerneffektivität vorliegt, zumeist die E-Learning-Anwendungen besser abschneiden. In mehreren Studien mussten die Lernenden für die Bearbeitung der E-Learning-Anwendung zudem weniger Zeit investieren, um ein vergleichbares Lernergebnis zu erzielen wie die Lernenden in den Präsenzveranstaltungen (siehe z.B. Vogt et al., 2006). Insgesamt kann hieraus geschlossen werden, dass E-Learning generell – aber auch die spezifischen Formen E-Lecture und Web-based Training – mit gutem Gewissen in der Lehre eingesetzt werden können. Ein genereller Nachteil der damit Lernenden ist nicht zu befürchten.

Des Weiteren finden sich signifikante Unterschiede bezüglich verschiedener Akzeptanzvariablen. Auch hier schnitten des Öfteren die E-Learning-Anwendungen besser ab – insbesondere, was Variablen wie Verständnis, Konzentrationserhaltung oder

Motivation betrifft. Dies ist sicherlich hauptsächlich der Möglichkeit geschuldet, die Lerngeschwindigkeit selbst zu bestimmen. Studien wie zum Beispiel die von Maki et al. (2000) zeigen allerdings, dass in manchen Fällen auch die Präsenzlehre von den Lernenden besser akzeptiert wird. Hinsichtlich der Akzeptanz zeigt sich folglich ein noch weniger einheitlicheres Bild als für die Lernwirksamkeit. Dies liegt aber vermutlich daran, dass innerhalb der Studien sehr unterschiedliche Akzeptanzaspekte erfragt wurden.
Interessant ist zudem, dass in verschiedenen Studien eine lineare Bearbeitung der E-Learning-Anwendungen beobachtet wurde. So wird die E-Lecture von den Lernenden (nach Görlich, 2006) sehr ähnlich zur Vorlesung wahrgenommen und in der Folge hauptsächlich von vorne nach hinten bearbeitet. Falls überhaupt spulten sie ein paar Mal zurück, um nicht-verstandene Inhalte noch einmal zu wiederholen. Auch die Zusatzfunktionen der E-Lecture (Notizfunktion, Indexfunktion etc.) wurden nur selten von den Lernenden genutzt. Ein ähnliches Vorgehen beobachteten Vogt et al. (2006) bei der Bearbeitung von Web-based Trainings. Hier wurden lediglich die Übungsaufgaben häufiger wiederholt – die Inhaltsseiten wurden stringent von vorne nach hinten bearbeitet.

Die Nützlichkeit von Studien zum Vergleich von E-Learning-Anwendungen und Präsenzlehre – oder von verschiedenen Lehrformen im Allgemeinen – wurden häufig kritisiert (siehe z.B. Carey, 2001 oder Schulmeister, 2007). Kritiker sind der Meinung, dass solche Studien zumeist Neuheitseffekte, instruktionale Methode sowie Lehrinhalte „in einen Topf werfen“. Nach den vorliegenden Studien könne häufig nicht eindeutig gesagt werden, ob es wirklich die Wahl des Lehrmediums war, die einen Einfluss auf das Lernen hatte. Hierzu seien nach Carey (2001) umfangreiche Metaanalysen über viele Kurse vonnöten. Da sich in vielen Fällen auch keine signifikanten Effekte oder nur sehr uneinheitliche Ergebnisse zeigen (Carey, 2001), könne daraus gefolgert werden, dass der Effekt der Lehrform relativ klein sei. Shomaker, Ricks und Hale (2002) fügen dem hinzu, dass bei der Interpretation solcher Studien immer im Auge behalten werden muss, dass es eine ganze Reihe von nicht-kontrollierbaren Störfaktoren gibt, die die Aussagekraft von Vergleichen zwischen E-Learning-Anwendungen und traditionellem Unterricht einschränken (Clark, 1983). Dabei spielt sicherlich auch die Neuheit der Anwendungen (Kerres, 2009) bzw. die Abwechslung, die durch ihren Einsatz entsteht, eine Rolle. Nach Schulmeister (2007) seien sogar so viele verschiedene konfundierende Variablen zu kontrollieren, wie z.B. Erfahrung der Lernenden mit neuen Medien oder das Vorwissen der Lernenden, dass „[…]kein vernünftiges Design für experimentelle Evaluationsstudien mehr möglich“ sei.
Diese Kritikpunkte treffen sicherlich für viele Studien zum Vergleich von Lehrmethoden zu. Diese Einschränkungen können allerdings durchaus vermieden werden, wenn der Versuchsplan so gestaltet ist, dass durch Messwiederholung inhalts-, person- und

methodeninvariante Vergleiche vorgenommen werden können. Genau das wird in Studie I dieser Arbeit realisiert.

1.6.2 Fragestellungen der ersten Studie

Im Gegensatz zu den meisten der oben berichteten Studien soll in der ersten Studie der vorliegenden Dissertation ein Vergleich zwischen zwei sehr spezifischen E-Learning-Anwendungen – E-Lecture und Web-based Training – und einer Präsenzvorlesung vorgenommen werden. Die beiden E-Learning-Anwendungen unterscheiden sich hinsichtlich der Darstellung der Lehrinhalte deutlich: Während die E-Lecture der Präsenzvorlesung sehr ähnlich ist, werden die Inhalte im Web-based Training hauptsächlich durch Texte und Bilder erläutert. Damit unterscheidet sich das Web-based Training deutlich sowohl von der Vorlesung als auch von der E-Lecture. Da gerade diese beiden E-Learning-Anwendungen vermehrt die traditionelle Lehre ersetzen oder zumindest ergänzen, ist es notwendig zu wissen, welche Vor- und Nachteile hieraus für die Lernenden entstehen. Dabei ist es auch von Interesse zu untersuchen, welche Faktoren innerhalb der E-Learning-Szenarien für die Vorteile verantwortlich sind – ein Aspekt, der im Fokus der zweiten Studie dieser Dissertation steht.

Durch Vogt et al. (2006, siehe Abschnitt „1.6.1 E-Learning vs. Präsenzlehre – Ein Blick in die Literatur“) wurde bereits ein Vergleich zwischen E-Lecture, Web-based Training und Präsenzlehre durchgeführt. Diese Studie umfasste allerdings sehr kleine und damit nur bedingt aussagekräftige Stichproben (3x10 Studierende). In dem Umfang der hier vorliegenden Untersuchung (mit einer kohärenten Stichprobe von mehr als 1.300 Studierenden – genauere Informationen siehe Abschnitt „2 Studie I: Methode“) ist ein Vergleich der drei Lehrangebote bisher noch nicht durchgeführt worden. Im Gegensatz zu Vogt et al. (2006), die die Studierenden die verschiedenen Lehrformen im Laborsetting nutzen ließen, wird in Studie I eine Felduntersuchung durchgeführt. Eine Laboruntersuchung hat – wie oben bereits erläutert – den Vorteil, dass Störvariablen besser kontrolliert werden können, sie lässt allerdings keinen Blick auf die reale Nutzung und den wirklichen Lernerfolg durch diese zu. Die vorliegende Studie versucht, diesen Nachteil auszugleichen, indem die Lernenden hier innerhalb einer Felduntersuchung die Lehrmethoden in dem Lernumfeld nutzen, die den realen Bedingungen entsprechen. Dieses Studiendesign dürfte daher eine wesentlich höhere ökologische Validität besitzen. Dadurch können zudem ohne großen Aufwand gleichzeitig viele Studierende an der Untersuchung teilnehmen, da den einzelnen Studierenden keine Arbeitsplätze zur Verfügung gestellt werden müssen. Dies macht es auch möglich, die Studie innerhalb eines Semesters

mehrfach zu replizieren und sie insgesamt so anzulegen, dass die Vergleiche durch Messwiederholung sowohl inhalts-, methoden- als auch personinvariant möglich sind.

Im Folgenden werden die Fragestellungen erläutert, die im Rahmen der vorliegenden Studie untersucht werden. Dabei werden zunächst diejenigen Fragestellungen dargestellt, die sich auf die Lernwirksamkeit der Lehrmethoden beziehen. Danach folgen die Fragestellungen zur Akzeptanz und zum Verhalten während der Bearbeitung/des Besuch der Lehrangebote. Zu guter Letzt werden die Fragestellungen beschrieben, die sich für die freiwillige Nutzung der E-Learning-Anwendungen im Rahmen der Klausurvorbereitung ergeben.

1.6.2.1 Lernwirksamkeit

Wie oben dargestellt, existieren hinsichtlich der Lernwirksamkeit sowohl Studien, die einen Unterschied zwischen E-Learning und Präsenzlehre zeigen als auch Studien, die diese Unterschiede nicht belegen konnten. Dennoch gilt: Wenn innerhalb der betrachteten Studien Lernwirksamkeitsunterschiede auftraten, fielen diese zugunsten der E-Learning-Anwendungen aus (siehe u.a. Glowalla, 2003; Kulik & Kulik, 1991; Maki et al., 2000; Sanders & Morrison-Shetlar, 2001; Schutte, 1997; Wofford et al., 2001).

Zum konkreten Vergleich der beiden E-Learning-Anwendungen E-Lecture und WBT existieren nur wenige Studien. Die Studie von Vogt et al. (2006) konnte lediglich kurzfristig einen Lernwirksamkeitsunterschied zwischen den beiden Anwendungen zugunsten des WBTs nachweisen. Hierfür könnte allerdings die sehr geringe Stichprobe innerhalb der Studie (mit)verantwortlich sein. Betrachtet man die Eigenschaften der E-Learning-Anwendungen genauer, sollte sich ein Vorteil der WBTs abzeichnen: Innerhalb der WBTs werden die Inhalte weniger flüchtig in geschriebenem Text darstellt. Diese Darstellung erleichtert es den Studierenden, die Inhalte im eigenen Lerntempo zu bearbeiten. Schon Görlich (2006) konnte zeigen, dass die Studierenden E-Lectures hauptsächlich linear nutzen, ohne häufig zurückzuspulen oder gezielt über den Index Schlagworte nachzuschauen. Demzufolge nutzen sie häufig nicht die Funktionen, die es ihnen ermöglichen würden, die Anwendung an ihr Lerntempo anzupassen. Zudem können innerhalb des WBTs spezifische Inhalte auch wesentlich einfacher nachgeschlagen und nachgelesen werden als innerhalb einer E-Lecture.

Darüber hinaus konnte Salomon (1984) zeigen, dass Lernende weniger Anstrengung investieren, je geringer sie die Anforderungen eines Mediums einschätzen[16]. Es ist möglich, dass sie die Anforderungen der Vorlesung und insbesondere der E-Lecture geringer

[16] Der Titel seines Artikels hierzu lautet bezeichnend „Television is easy, and print is tough".

einschätzen, da beides ihnen ermöglicht, sich rezipierend zurückzulehnen, während sie im Rahmen der WBT-Bearbeitung Anstrengung investieren müssen, um die Texte zu lesen.
All dies sollte zu einem höheren Lernergebnis durch die Bearbeitung der WBTs im Gegensatz zu den E-Lectures führen. Im Rahmen der vorliegenden Studie wurde die Lernwirksamkeit kurz- und längerfristig erfasst. Die kurzfristige Lernwirksamkeit bezieht sich auf die Erfassung des Lernstandes einen Tag nach der Bearbeitung der E-Learning-Anwendungen bzw. dem Besuch der Vorlesung, die längerfristige auf die Erfassung des Lernstandes am Ende des Semesters. Aus den oben genannten Punkten werden für die kurz- und längerfristige Lernwirksamkeit folgende Hypothesen hergeleitet:

- *Hypothese 1a: Kurzfristig weisen die E-Lectures eine höhere Lerneffektivität auf als die Vorlesung.*
- *Hypothese 1b: Kurzfristig weisen die WBTs eine höhere Lerneffektivität auf als die Vorlesung.*
- *Hypothese 1c: Kurzfristig weisen die WBTs eine höhere Lerneffektivität auf als die E-Lectures.*
- *Hypothese 2a: Längerfristig weisen die E-Lectures eine höhere Lerneffektivität auf als die Vorlesung.*
- *Hypothese 2b: Längerfristig weisen die WBTs eine höhere Lerneffektivität auf als die Vorlesung.*
- *Hypothese 2c: Längerfristig weisen die WBTs eine höhere Lerneffektivität auf als die E-Lectures.*

Da auf eine Auswertung der E-Lecture-Nutzungszeiten aus technischen Gründen verzichtet werden muss (siehe Abschnitt „2.3.2 Erhobene Variablen“ – Unterkapitel „User Tracking“), resultiert für die Lerneffizienz lediglich eine vergleichende Hypothese zur WBT-Nutzung und dem Vorlesungsbesuch. Angelehnt an die Ergebnisse der Studien von Hasebrook (2002), Lyon Jr. et al. (1992), Spickard III et al. (2002) sowie Vogt et al. (2006), die bei zum Teil gleicher Lerneffektivität häufig kürzere Bearbeitungszeiten und eine dementsprechend höhere Lerneffizienz der E-Learning-Anwendungen beobachten konnten, wird auch hinsichtlich der Lerneffizienz ein Vorteil für die E-Learning-Anwendungen (die WBTs) angenommen:

- *Hypothese 1d: Kurzfristig erreichen die WBTs eine höhere Lerneffizienz als die Vorlesung.*
- *Hypothese 2d: Längerfristig erreichen die WBTs eine höhere Lerneffizienz als die Vorlesung.*

Nungester und Duchastel (1982), Roedinger III und Karpicke (2006) sowie Larsen et al. (2009) konnten die Lernwirksamkeit von Übungsaufgaben verglichen zum erneuten Studieren der Inhalte belegen. Dabei konnten Siegert et al. (2003) zeigen, dass Web-based

Trainings mit Übungsaufgaben nicht signifikant länger bearbeitet werden mussten als WBTs ohne Übungsaufgaben. Auch die Metaanalyse von Hamaker (1986) belegte, dass die Lernwirksamkeitsvorteile von Übungsaufgaben erhalten blieben, wenn die Lernzeit durch die Versuchsleitung kontrolliert wurde. Insgesamt sollten WBTs mit Übungsaufgaben deshalb durch die höhere Lernwirksamkeit und gleichbleibende Lernzeit insgesamt lerneffizienter sein. Daher werden für den Unterschied zwischen WBTs mit und ohne Übungsaufgaben folgende Hypothesen aufgestellt:

- *Hypothese 3a: Kurzfristig erreichen WBTs mit Übungsaufgaben eine höhere Lerneffektivität als WBTs ohne Übungsaufgaben.*
- *Hypothese 3b: Längerfristig erreichen WBTs mit Übungsaufgaben eine höhere Lerneffektivität als WBTs ohne Übungsaufgaben.*
- *Hypothese 3c: Kurzfristig erreichen WBTs mit Übungsaufgaben eine höhere Lerneffizienz als WBTs ohne Übungsaufgaben.*
- *Hypothese 3d: Längerfristig erreichen WBTs mit Übungsaufgaben eine höhere Lerneffizienz als WBTs ohne Übungsaufgaben.*

1.6.2.2 Akzeptanzvariablen

Basierend auf den Ergebnissen der Studien von Glowalla (2003), Glowalla et al. (2004), Görlich (2006) sowie Schutte (1997), nach denen Lernende unter anderem die Effektivität und Konzentrationserhaltung von E-Learning-Anwendungen höher einschätzten als von der Präsenzlehre, wird angenommen, dass die E-Learning-Anwendungen auch in der vorliegenden Studie als lernförderlicher eingeschätzt werden als die Vorlesung. Im Vergleich zwischen E-Lecture und WBT wird zudem angenommen, dass das WBT von den Studierenden als lernförderlicher wahrgenommen wird, da es die Inhalte (wie oben bereits erwähnt) weniger flüchtig in geschriebenem Text darstellt und ein einfacheres Nachschlagen spezifischer Inhalte erlaubt. Es resultieren zur Lernförderlichkeit der drei Lehrmethoden aus diesem Grunde folgende Hypothesen:

- *Hypothese 4a: Studierende schätzen die Lernförderlichkeit der E-Lecture höher ein als die Lernförderlichkeit der Vorlesung.*
- *Hypothese 4b: Studierende schätzen die Lernförderlichkeit des WBTs höher ein als die Lernförderlichkeit der Vorlesung.*
- *Hypothese 4c: Studierende schätzen die Lernförderlichkeit des WBTs höher ein als die Lernförderlichkeit der E-Lecture.*

Da Studien zum Akzeptanzvergleich zwischen traditionellen und E-Learning-Lehrmethoden häufig keinen Unterschied (z.B. Carey, 2001), manchmal einen Vorteil für die E-Learning-Anwendungen (z.B. Glowalla, 2003) und selten einen Vorteil für die

traditionelle Lehrmethode ergaben (z.B. Maki et al., 2000) kann dieser Hintergrund nicht genutzt werden, um eine Hypothese für die Akzeptanzeinschätzung der drei Lehrmethoden abzuleiten.

Angelehnt an den von Kerres (2009) erläuterten Neuigkeitseffekt kann jedoch angenommen werden, dass die E-Learning-Anwendungen kurzfristig schon deshalb besser von den Studierenden beurteilt werden, da ihre Nutzung eine Abwechslung zum „üblichen" Vorlesungsszenario der Hochschullehre darstellt. An dieser Stelle soll ganz klar von einem „Ausnahme-" oder „Abwechslungseffekt" gesprochen werden, nicht von einem „Neuigkeitseffekt". Schon 2007 schrieb Schulmeister (S. 368), dass sich der Neuigkeitseffekt „abschleife", da die meisten Menschen regelmäßig Zugang zu einem Computer haben. Seitdem sind fast zehn Jahre vergangen und es ist davon auszugehen, dass sich dieser Effekt noch verstärkt hat. Die Studierenden sollten nun zwar an die Nutzung des Computers gewöhnt sein, die „übliche" Lehre innerhalb der Universität hat sich jedoch seither nicht stark verändert, so dass der Einsatz von E-Learning in der Hochschullehre zumindest in einem Gefühl der Abwechslung/Ausnahme resultieren und sich dies wiederum in der Akzeptanzbeurteilung niederschlagen sollte.

Da die E-Lecture hier sowohl in der Darstellung der Inhalte (Vortragsaufzeichnung) als auch hinsichtlich der Nutzung (Anschauen des Vortrages) sehr der Vorlesung ähnelt und das WBT damit eine größere „Abwechslung" für die Studierenden darstellen sollte, resultieren folgende Hypothesen:

- *Hypothese 5a: Im kurzfristigen[17] Vergleich wird die E-Lecture von den Studierenden besser akzeptiert als die Vorlesung.*
- *Hypothese 5b: Im kurzfristigen Vergleich wird das WBT von den Studierenden besser akzeptiert als die Vorlesung.*
- *Hypothese 5c: Im kurzfristigen Vergleich wird das WBT von den Studierenden besser akzeptiert als die E-Lecture.*

Längerfristig sind zudem Vergleiche der drei eingesetzten Lehrmethoden mit den ebenfalls zur Verfügung gestellten Übungsaufgaben, dem Skript zur Vorlesung und der angegebenen Literatur geplant. Hierbei soll erhoben werden, inwieweit diese **sechs Lehrmaterialien/-methoden** hinsichtlich ihrer Akzeptanz und ihrem Nutzungswillen eingeschätzt werden und welche Gesamtnote den Lehrmethoden gegeben wird. Vergleichbar zur kurzfristigen Akzeptanz wird hier folgende Hypothese angenommen:

[17] Auch hinsichtlich der Akzeptanzbefragung wurde sowohl kurzfristig (ein Tag nach Bearbeitung der E-Learning-Anwendungen bzw. dem Besuch der Vorlesung) als auch längerfristig (am Ende des Semesters) eine Beurteilung durch die Studierenden erfragt.

- *Hypothese 6: Die E-Learning-Anwendungen (E-Lecture, Web-based Training, Übungsaufgaben) erreichen eine höhere Akzeptanz als die „traditionellen Lehrmethoden" der Hochschule (Vorlesung, Vorlesungsskript, Literatur).*

1.6.2.3 Verhalten vor, während und nach Bearbeitung/Besuch des Lehrangebotes

Da Inhalte in der Vorlesung nur einmalig dargeboten werden und die Inhalte in der E-Lecture flüchtiger erscheinen als im WBT, sollten die Studierenden insbesondere in der Vorlesung Notizen machen und auch bei der Bearbeitung der E-Lecture vermehrt das Bedürfnis haben, sich etwas zu notieren als bei Bearbeitung der Web-based Trainings. Daraus ergeben sich für das Notizenmachen folgende Hypothesen:

- *Hypothese 7a: In der Vorlesung machen die Studierenden mehr Notizen als bei Bearbeitung der E-Lecture.*
- *Hypothese 7b: In der Vorlesung machen die Studierenden mehr Notizen als bei Bearbeitung des Web-based Trainings.*
- *Hypothese 7c: Bei Bearbeitung der E-Lecture machen die Studierenden mehr Notizen als bei Bearbeitung des Web-based Trainings.*

Für die Pausenhäufigkeit sowie die Pausenlänge wird angenommen, dass die Studierenden sicherlich häufiger im Rahmen der E-Learning-Nutzung eine Pause einlegen, da sie bei Pausen während der Vorlesung Inhalte verpassen würden. Die E-Learning-Anwendungen bieten die Möglichkeit, eine Pause einzulegen. Gleiches gilt auch für die Länge der Pausen. Es resultieren folgende Hypothesen:

- *Hypothese 8a: Während der Nutzung der E-Learning-Anwendungen legen die Studierenden häufiger eine Pause ein als während des Vorlesungsbesuchs.*
- *Hypothese 8b: Während der Nutzung der E-Learning-Anwendungen legen die Studierenden längere Pausen ein als während des Vorlesungsbesuchs.*

1.6.2.4 Freiwillige Nutzung im Sommersemester

Neben den subjektiv erhobenen Akzeptanzeinschätzungen (Einstellungsakzeptanz) soll auch die Verhaltensakzeptanz (Müller-Böling & Müller, 1986) erhoben werden, indem den Studierenden die E-Lectures und Web-based Trainings mehrere Monate lang für ihre Klausurvorbereitung zur Verfügung gestellt werden. Hier soll betrachtet werden, welche der E-Learning-Anwendungen häufiger und länger freiwillig genutzt werden.

Da kurz vor der Klausur erneut eine Probeklausur zur Verfügung gestellt wird, ist es möglich festzustellen, welche Vorgehensweise der Studierenden zum besten Lernerfolg führt (Nutzung der E-Lectures, Nutzung der WBTs, Nutzung von beiden Anwendungen, Nutzung von keiner der Anwendungen). Auch ist es in diesem Rahmen möglich,

abschließend die Akzeptanz der Lernenden zu erfassen und zu betrachten, wie sich diese durch die freiwillige – und vor allem womöglich längerfristigere – Nutzung der Anwendungen verändert hat.
Da die Einstellungsakzeptanz nach Bürg und Mandl (2005) Hauptprädiktor für die Verhaltensakzeptanz ist, ergibt sich folgende Hypothese:

- *Hypothese 9: Die E-Learning-Anwendung mit der höheren Einstellungsakzeptanz am Ende des Wintersemesters wird im Sommersemester häufiger freiwillig genutzt.*

Vergleichbar zu den Ergebnissen der Studie von Glowalla (2003), in der die Nutzer der E-Learning-Anwendung in der abschließenden Klausur besser abschnitten als die Nicht-Nutzer, erschließen sich für die Lernleistung der Studierenden folgende Hypothesen:

- *Hypothese 10a: Studierende, die die E-Learning-Anwendungen genutzt haben, schneiden besser in der Probeklausur ab als Studierende, die die Anwendungen nicht genutzt haben.*
- *Hypothese 10b: Studierende, die die E-Learning-Anwendungen genutzt haben, schneiden besser in der Abschlussklausur (MAP) ab als Studierende, die die Anwendungen nicht genutzt haben.*

Zum Zeitpunkt der abschließenden Akzeptanzbefragung sollte der oben diskutierte „Abwechslungs-/Ausnahmeeffekt“ (angelehnt an Kerres (2009) „Neuigkeitseffekt“) für die E-Learning-Anwendungen nach der potentiell häufigen Nutzung im Sommersemester nicht mehr existieren. Es besteht allerdings die Annahme, dass die am meisten genutzte Anwendung auch diejenige ist, die hinsichtlich der Akzeptanz nun erneut am besten abschneidet. Dadurch resultiert folgende Hypothese:

- *Hypothese 11: Die E-Learning-Anwendung, die im Sommersemester häufiger freiwillig genutzt wurde, wird hinsichtlich der Akzeptanz kurz vor der MAP am besten beurteilt.*

2 Studie I: Methode

Beide Studien dieser Dissertation wurden im Rahmen der jeweils im Wintersemester stattfindenden Vorlesung „Instruktionspsychologie für Lehramtsstudierende A“ von Professor Glowalla durchgeführt. Die Vorlesung ist Teil des Grundmoduls „Psychologie für Lehramtsstudierende“, das von allen Lehramtsstudierenden der Justus-Liebig-Universität Gießen vornehmlich in den ersten beiden Semestern des Studiums besucht werden sollte. Die Studien fanden in den Jahrgängen 2009/2010 (Studie I) sowie 2010/2011 (Studie II) statt.

Neben der Vorlesung im Wintersemester gehört zu dem Modul eine weitere Vorlesung, die im Sommersemester stattfindet („Psychologie für Lehramtsstudierende B“ von Professor Stiensmeier-Pelster) sowie ein Proseminar zur Vertiefung eines thematischen Bereiches der beiden Vorlesungen. Um an der Modulabschlussprüfung (MAP) teilzunehmen, müssen die Studierenden an beiden Vorlesungen teilgenommen und das Proseminar bestanden haben (z.B. durch Halten eines Referates, Anfertigen einer Seminararbeit etc.). Die Modulabschlussprüfung beinhaltet dann zu gleichen Teilen die Inhalte der beiden Vorlesungen und findet nach dem Sommersemester statt. Es ist folglich davon auszugehen, dass die Studierenden durch die anstehende Klausur von sich aus motiviert sind, die Inhalte beider Vorlesungen zu lernen. Damit die Studierenden zusätzlich möglichst vollständig an der Untersuchung teilnehmen, wurden im Rahmen der Studie I zudem – gestaffelt nach dem Umfang der Beteiligung – bis zu drei Bonuspunkte für die MAP vergeben[18].

2.1 Stichprobe

Im Folgenden soll zunächst die Gesamtstichprobe der Studierenden beschrieben werden, die an der ersten Studie im Wintersemester 2009/2010 teilgenommen haben. Da die Studierenden nicht zu allen Fragen Angaben machten, unterscheiden sich die Stichprobengrößen für einzelne Variablen. Die Stichprobengröße wird deshalb jeweils berichtet. Bei den Prozentangaben werden die gültigen Prozente nach Abzug der fehlenden Antworten angegeben.

Nachdem die Gesamtstichprobe vorgestellt wurde, folgt ein kurzer Überblick über die Teilstichproben, die sich durch die Einteilung der Studierenden in die Projektgruppen sowie die Beteiligung der Studierenden an den einzelnen Evaluationen und Leistungserhebungen ergaben und den Auswertungen im Ergebnisteil zugrunde liegen.

[18] Insgesamt konnten ca. 70 Punkte in der Klausur erreicht werden.

2.1.1 Beschreibung der Gesamtstichprobe

An der Vorlesung im Wintersemester 2009/10 nahmen insgesamt 1357 Studierende teil. Aufgrund der Menge an Studierenden wurde die Vorlesung mittwochs mittags und abends gehalten – 821 Studierende (60.5%) besuchten die Mittagsvorlesung, 524 (38.6%) die Abendvorlesung. Elf Studierende (0.8%) meldeten sich im Laufe des Semesters von der Vorlesung ab, eine Person (0.1%) konnte keiner der beiden Vorlesungen zugeordnet werden, da sie auf der Lernplattform der Justus-Liebig-Universität Gießen (*Stud.IP*[19]) zu keiner der Veranstaltungen angemeldet war.

2.1.1.1 Allgemeine Angaben

Zu ihrem Geschlecht gaben uns insgesamt 1093 Studierende Auskunft (264 ohne Angabe). 67.1% von diesen fast 1100 Studierenden sind weiblich, 32.9% männlich.

Die 1096 Studierenden (261 ohne Angabe), die über ihr Alter Auskunft gaben, sind zwischen 18 und 47 Jahre alt. Dabei ist ein Großteil der Stichprobe (72.9%) zwischen 18 und 21 Jahre alt. Im Schnitt beträgt das Alter der Stichprobe dementsprechend circa M = 21.5 Jahre.

Die Stichprobe setzt sich aus Studierenden aus fünf unterschiedlichen Lehramtsstudiengängen zusammen: L1 (Lehramt an Grundschulen), L2 (Lehramt an Haupt- und Realschulen), L3 (Lehramt an Gymnasien), L4 bzw. BaBBB (Lehramt an Berufsschulen bzw. Bachelor-Studiengang "Berufliche und betriebliche Bildung“) sowie L5 (Lehramt an Förderschulen). Von den 1100 Studierenden (257 ohne Angabe), die eine Auskunft über Ihren Studiengang machten, studieren die meisten Lehramt an Gymnasien (L3, 46.2%). Danach folgt der Studiengang „Lehramt an Haupt- und Realschulen“ (L2, 29.6%). Die Studiengänge L1 und L5 werden jeweils von 9.5% der Studierenden angegeben. Das Lehramt an Berufsschulen bzw. den neuen Bachelor-Studiengang „Berufliche und betriebliche Bildung“ studieren 5.1%. Nur 0.1% (ein/e Studierende/r) geben einen anderen Studiengang an, nämlich „MA/BBB“. Hierbei handelt es sich um den Masterstudiengang zur „Beruflichen und betrieblichen Bildung“. Abbildung 12 veranschaulicht die Verteilung der Studierenden auf die Studiengänge.

[19] Stud.IP steht für „Studienbegleitender Internetsupport von Präsenzlehre“ und dient an der Justus-Liebig-Universität Gießen der Veranstaltungsadministration (Justus-Liebig-Universität Gießen, 2012).

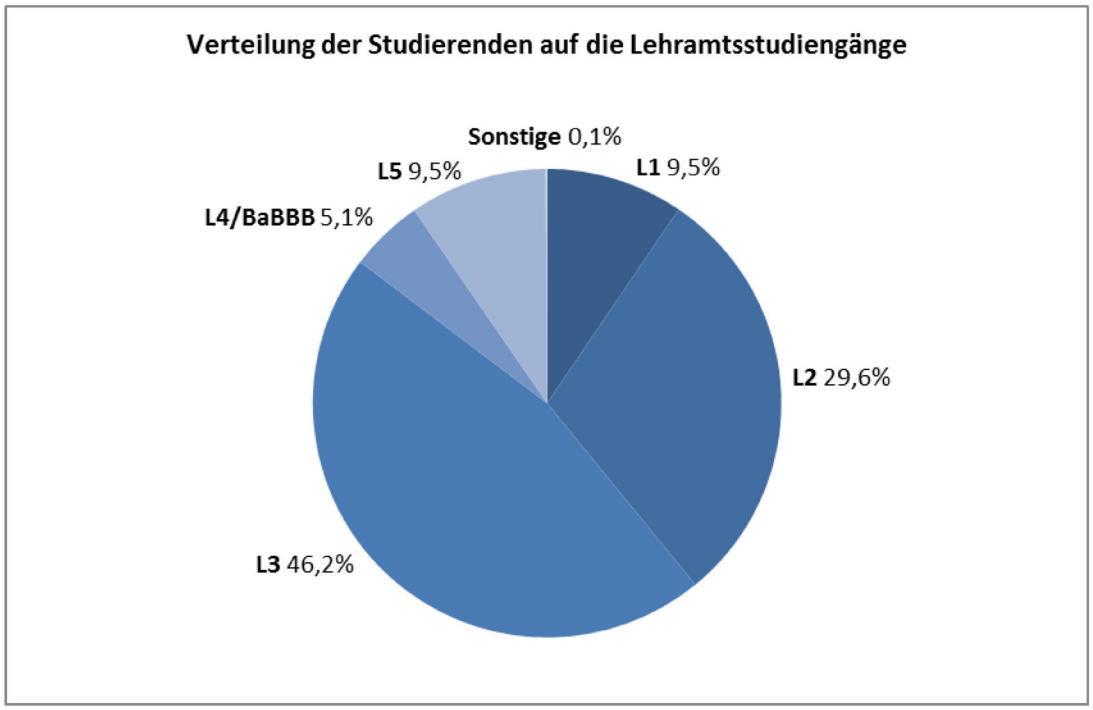

Abbildung 12. Verteilung der Studierenden auf die fünf Lehramtsstudiengänge sowie die Kategorie "Sonstiges" (Angabe der gültigen Prozente)[20]

1098 Studierende machten eine Angabe zu ihrem Fachsemester (259 ohne Angabe). Wie erwartet befinden sich die meisten Studierenden (96.1%) im ersten Fachsemester. 1.9% sind bereits im zweiten oder dritten, weitere 2.1% in einem höheren Semester.

Die meisten der 1101 Studierenden, die hierüber berichteten (99.3%, 256 ohne Angabe), besuchen die Vorlesung „Psychologie für Lehramtsstudierende A“ zum ersten Mal. Von den übrigen acht Personen (0.7%) gaben fünf an, die Vorlesung bereits einmal besucht zu haben. Jeweils eine Person gab an, die Vorlesung zwei-, drei- oder viermal besucht zu haben[21].

Befragt nach ihrer Abiturnote machten 1065 Studierende eine Angabe (292 ohne Angabe). Die Abiturnoten dieser Studierenden verteilen sich über die maximal mögliche Notenskala – die besten haben ihr Abitur mit 1.0 bestanden, die schlechtesten mit 4.0. Dabei ergibt sich ein Notendurchschnitt von M = 2.6 (SD = 0.55). Abbildung 13 zeigt die Verteilung der

[20] L1 = Lehramt an Grundschulen, L2 = Lehramt an Haupt- und Realschulen, L3 = Lehramt an Gymnasien, L4 bzw. BaBBB = Lehramt an Berufsschulen bzw. Bachelor-Studiengang "Berufliche und betriebliche Bildung“, L5 = Lehramt an Förderschulen

[21] Die Frage wurde von einzelnen Studierenden falsch verstanden. Diese verstanden wohl, dass sie angeben sollten, wie viele Vorlesungs*sitzungen* sie bereits besucht hatten. Daher wurden die Angaben zu dieser Variablen vor der Auswertung manuell bereinigt: Studierende, die vorher angegeben hatten, erst im ersten Semester zu sein, konnten die Vorlesung nicht schon einmal besucht haben. Andere Angaben bei diesen Studierenden wurden deshalb korrigiert. Nur noch bei den beschriebenen acht Studierenden resultierte eine plausible Kombination aus Fachsemesterzahl und Anzahl der angegebenen Vorlesungsbesuche.

Abiturnote innerhalb der Stichprobe. Es resultiert eine fast normalverteilte – etwas rechtsschiefe – Verteilung.

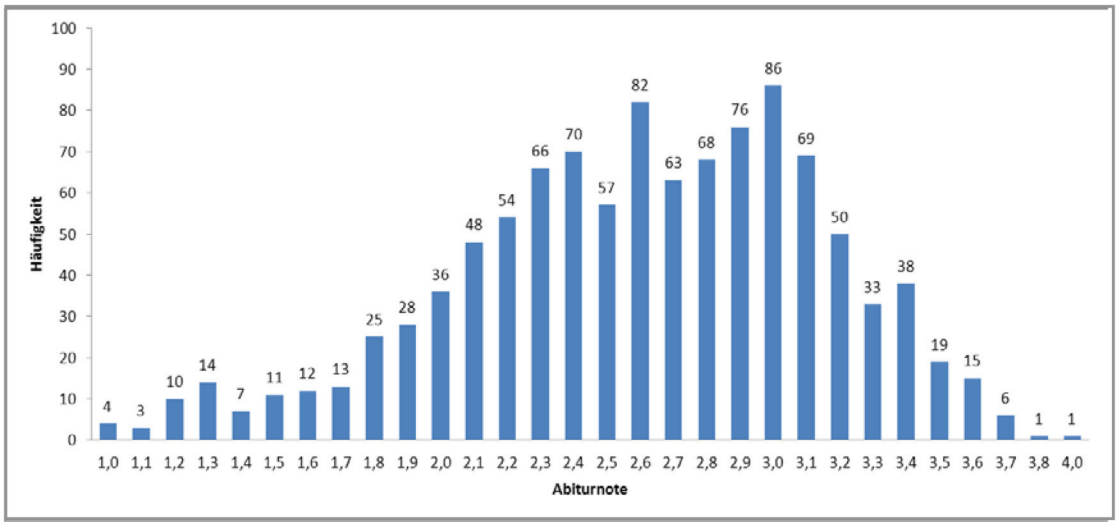

Abbildung 13. Verteilung der Abiturnote innerhalb der Stichprobe

Abbildung 14 zeigt zudem, wo die 1102 Studierenden, die hierzu eine Angabe machten, ihr Abitur erworben haben (255 fehlend). Häufigkeitsangaben von 5% und höher sind rot markiert. Naturgemäß haben die meisten Befragten ihr Abitur in Hessen erworben.

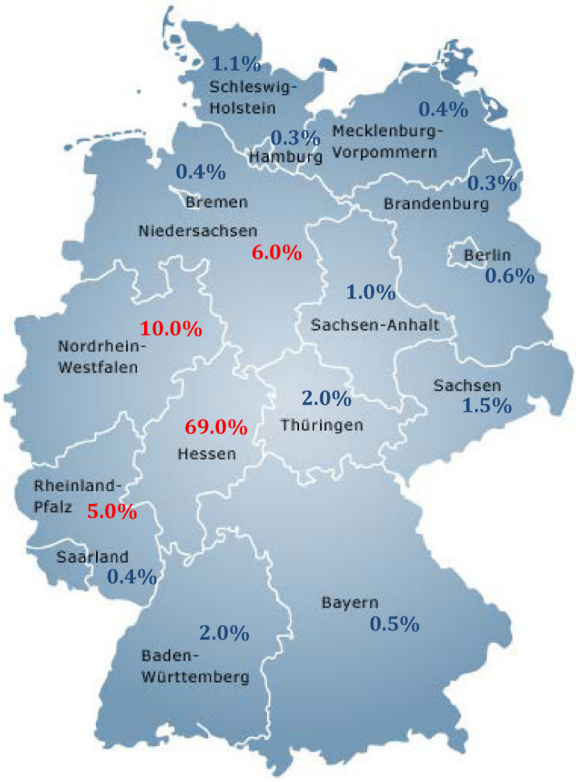

Abbildung 14. Gültige prozentuale Angaben der Studierenden auf die Frage: "Wo haben Sie Ihr Abitur erworben?[22]"

[22] Karte der Bundesländer: Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz (2016)

Vier Studierende gaben zudem an, in einem anderen Staat ihr Abitur erworben zu haben. Genannt wurden hier zweimal die Ukraine und jeweils einmal Mazedonien und Tunesien.

2.1.1.2 Interesse und Vorkenntnisse im Hinblick auf die Vorlesung und das Projekt

Die Studierenden sollten des Weiteren Angaben darüber machen, wie sie es finden, an dem Projekt teilzunehmen. Zu Beginn des Wintersemesters gaben 91.1% der Studierenden (n = 1096, 261 ohne Angabe) an, die Teilnahme an dem Projekt gut zu finden. Nur 8.9% fanden dies nicht gut. Eine ähnliche Frage wurde noch einmal nach Beendigung des Projekts gestellt. Hier gaben 941 Studierende eine Antwort (416 ohne Angabe). Von diesen fanden 92.3% die Teilnahme gut, 7.7% lehnten sie ab.

Wie in Tabelle 3 zu sehen ist, fehlen in der Vorbefragung etwa 19% der Angaben, in der Nachbefragung allerdings fast 31%. Hier könnte argumentiert werden, dass das abschließende Urteil weiterhin so gut ausfiel, da sich nur noch diejenigen an der Umfrage beteiligten, denen das Projekt per se gefallen hat – sich folglich die Ablehnung des Projekts zur Nachbefragung hin verstärkt hat. Die 31% fehlenden Urteile kamen allerdings nicht ausschließlich dadurch zustande, dass die Studierenden, die das Projekt nicht mochten, am Ende nicht mehr teilnahmen, sondern hauptsächlich dadurch, dass bei der Nachbefragung technische Schwierigkeiten auftraten, durch die einzelne Daten nicht gespeichert wurden. Deshalb kann aus den vergleichbaren gültigen Prozentwerten (91% positive Einschätzungen in der Vorbefragung und 92% in der Nachbefragung) auf eine in etwa gleichbleibend hohe Akzeptanz des Projektes über die Laufzeit hinweg geschlossen werden.

Tabelle 3
Akzeptanz des Projekts "Lehrinnovation" vor und nach der Durchführung

	Gut	Schlecht	Fehlend
Wie finden Sie es, dass Sie an unserem Projekt „Lehrinnovation“ teilnehmen? (n = 1096)	73.6% (91.1%)	7.1% (8.9%)	19.2%
Wie hat Ihnen die Teilnahme an unserem Projekt „Lehrinnovation“ gefallen? (n = 941)	64.0% (92.3%)	5.3% (7.7%)	30.7%

Anmerkung. Angegeben sind jeweils die Prozentangaben ohne und (in Klammern) mit Anrechnung der fehlenden Werte; n = Stichprobengröße

Auf die Frage nach ihrem Interesse an den Vorlesungsthemen gaben die meisten der Studierenden die Auskunft, sehr (66.8%) oder zumindest etwas (31.7%) an den Themen der Vorlesung interessiert zu sein (n = 1101; 256 ohne Angabe). Nur insgesamt 1.5% gaben an, dass sie die Themen der Vorlesung wenig oder gar nicht interessieren.

Ihre Vorkenntnisse hinsichtlich der Inhalte der Vorlesung schätzten die Studierenden (n = 1101, 256 ohne Angabe) dabei eher gering ein. Insgesamt 91.4% von ihnen berichteten keine oder geringe Vorkenntnisse in Bezug auf die Inhalte der Vorlesung. 8.1% meinten, gute Vorkenntnisse zu besitzen, lediglich 0.5% gaben sogar sehr gute Vorkenntnisse an.

Zudem sollten die Vorkenntnisse bezogen auf die einzelnen Vorlesungsthemen erfasst werden. Dazu wurden die Studierenden direkt nach Bearbeitung eines Themas befragt, inwieweit sie bereits Vorkenntnisse zum bearbeiteten Thema oder Erfahrungen mit der Lehrmethode hatten.

Tabelle 4
Mittlere Vorkenntnisse der Studierenden bezüglich des bearbeiteten Themas sowie Erfahrung mit der genutzten Lehrmethode

	E-Lecture			WBT			Vorlesung		
	n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
Vorkenntnisse	1031	1.86	0.66	1007	1.91	0.64	981	1.88	0.65
Erfahrung mit der Lehrmethode	1028	1.50	0.68	1007	1.43	0.70	983	2.51	0.74

Anmerkung. Skala „Vorkenntnisse“: Keine Vorkenntnisse = 1, sehr gute Vorkenntnisse = 4; Skala „Erfahrung“: Keine Erfahrung = 1, sehr viel Erfahrung = 4; *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Wie Tabelle 4 und Tabelle 5 zeigen, gaben die Studierenden aller drei Experimentalgruppen an, eher wenige Vorkenntnisse zu dem bearbeiteten/gehörten Thema gehabt zu haben. Hinsichtlich der Vorerfahrungen mit der genutzten Lehrmethode ergaben sich allerdings signifikante Unterschiede zwischen den Gruppen: Naturgemäß hatten die Studierenden mit der Vorlesung bereits mehr Erfahrung gesammelt als mit der E-Lecture oder dem WBT.

Tabelle 5
Abhängige Varianzanalysen mit Greenhouse-Geisser-Korrektur (ANOVA) zu den Variablen "Vorkenntnisse zum gestrigen Thema" sowie "Erfahrung mit der genutzten Lehrmethode“

	n	Vorlesung		E-Lecture		WBT		ANOVA				Posthoc
		M	***SD***	***M***	***SD***	***M***	***SD***	***F***	***df***	***p***	***Eta²***	
Vorkenntnisse	532	1.93	0.68	1.88	0.67	1.93	0.63	1.12	2.00	.33	.00	-
Erfahrung	534	2.53	0.72	1.49	0.69	1.44	0.70	497.81	1.91	<.01*	.48	VL>EL,WBT

Anmerkung. α = .05; ANOVA wurde anhand der Stichprobe derjenigen Studierenden berechnet, die alle drei Lehrmethoden beurteilt haben; *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke

Computerbezogene Fragen, Einschätzung von E-Learning

Um die Computeraffinität der Studierenden zu erheben, wurde gefragt, welches Betriebssystem sie hauptsächlich verwenden, wie häufig sie mit einem Computer arbeiten und wie sie ihre Fähigkeiten im Umgang mit dem Computer einschätzen. Zudem wurden 14 Items zum Thema „Computerselbstwirksamkeit“ erhoben, von denen nach einer Faktorenanalyse (Hauptkomponentenanalyse) und inhaltlichen Überlegungen sechs Items ausgewählt und zu einer Skala zusammengefasst wurden. Diese Skala weißt ein Cronbachs α = .85 auf und beinhaltet zum Beispiel das Item „Im Umgang mit dem Computer fühle ich mich sicher.“ Die vollständige Skala sowie die Ergebnisse der Skalenerstellung befinden sich in Anhang A.

Darüber hinaus wurden die Studierenden in drei Items nach Ihrer Einschätzung von E-Learning-Anwendungen in der Bildung befragt. Auch hier wurde eine Faktorenanalyse (Hauptkomponentenanalyse) durchgeführt, die für zwei der drei Items eine Skala

(Cronbachs α = .60) ergab. Eines der beiden Items lautet „E-Learning wird eine immer wichtigere Rolle im Bildungsbereich spielen." Die vollständige Skala sowie die zugehörigen Auswertungen befinden sich in Anhang A.

Die gültigen prozentualen Häufigkeiten zu diesen Skalen sind in Tabelle 6 abgetragen. Es zeigt sich, dass die meisten Studierenden täglich mit dem Computer arbeiten. Dabei nutzen sie zumeist das Betriebssystem *Windows*® von *Microsoft*®. Im Umgang mit dem Computer schätzen sich die Studierenden überwiegend als fortgeschritten ein. Dies spiegelt sich auch in ihrer Computerselbstwirksamkeit wider. Hier erreichen die meisten Studierenden einen Skalenwert von 16 bis 20 Punkten (bei einem möglichen Skalenwertebereich von 0 bis 24 Punkten) – es resultiert ein Skalenmittelwert von M = 17.4.

In Bezug auf die Einschätzung von E-Learning erreichen die meisten Studierenden hingegen einen Wert von 3 oder 4 Punkten (bei einem möglichen Skalenwertebereich von 0 bis 8 Punkten). Hier resultiert ein Skalenmittelwert von M = 3.8. Die Bedeutung von E-Learning für die Bildung wird folglich durch die Studierenden eher gering eingeschätzt.

Tabelle 6
Angaben der Studierenden zu computerbezogenen Items und zur Einschätzung von E-Learning

Welches Betriebssystem verwenden Sie hauptsächlich? (n = 1101)		**Wie häufig arbeiten Sie mit dem Computer?** (n = 1101)	
Betriebssystem	**Häufigkeit in %**	**Häufigkeit**	**Häufigkeit in %**
Windows	96.0%	Nie	0.1%
Linux	1.0%	Seltener	0.8%
MacOS	2.9%	1-3 Mal/Monat	0.4%
Anderes[23]	0.1%	1-3 Tage/Woche	4.7%
		4-6 Tage/Woche	19.6%
		Jeden Tag	74.4%

Wie schätzen Sie Ihre Fähigkeiten im Umgang mit dem Computer ein? (n = 1102)	
Kategorie	**Häufigkeit in %**
Anfänger	216 (19.6%)
Fortgeschrittener	802 (72.8%)
Experte	84 (7.6%)

Computerselbstwirksamkeit (n = 463)					
Punktzahl	**Häufigkeit in %**	***M***	***SD***	***Min***	***Max***
<10	1.9%	17.37	3.59	6	24
10-15	29.1%				
16-20	46.5%				
>20	22.5%				

Einschätzung zu E-Learning (n = 463)					
Punktzahl	**Häufigkeit in %**	***M***	***SD***	***Min***	***Max***
1 oder 2	14.0%	3.83	1.21	2	8
3 oder 4	60.0%				
5 oder 6	23.3%				
7 oder 8	2.6%				

Anmerkung. n = Stichprobengröße, M = Mittelwert, *SD* = Standardabweichung, *Min* = Minimum, *Max* = Maximum

[23] Angabe: *Win, Linux* und *Openso*

2.1.2 Teilstichproben für einzelne Auswertungen

Wie oben bereits erwähnt nahmen nicht alle Studierende an allen Teilen der Evaluation teil. Tabelle 7 gibt deshalb einen Überblick über die einzelnen den Auswertungen zugrundeliegenden Teilstichproben.

Tabelle 7
Übersicht über die Teilstichproben

Evaluationsabschnitte	Teilgenommen	Nicht teilgenommen
Kurzfristige Evaluation der E-Lecture	1077	280
Kurzfristige Evaluation des Web-based Trainings	1029	328
Kurzfristige Evaluation der Vorlesung	1009	348
Vollständige kurzfristige Evaluation (Evaluation aller drei Lehrmethoden)	819	538
Vollständige kurzfristige sowie längerfristige Evaluation	697	660
Abschließende Evaluation nach dem Sommersemester	329	1028
Modulabschlussprüfung	927	430

Dabei wurden die Studierenden zu Beginn des Projekts zufällig in zwölf verschiedene ursprünglich gleich große Gruppen eingeteilt, die zu unterschiedlichen Zeitpunkten an dem Projekt teilnehmen sollten (siehe Abschnitt „2.2 Durchführung“). Dies sollte sicherstellen, dass innerhalb des einen Wintersemesters mehrere Replikationen durchgeführt werden können.

Aufgrund dessen, dass insbesondere zu Beginn des Semesters einige Studierende die Teilnahme vergaßen oder technische Probleme auftraten, mussten diese in spätere Projektgruppen umgruppiert werden, damit sie weiterhin teilnehmen und sich die Bonuspunkte für die abschließende Klausur „verdienen“ konnten. Daher resultieren die im Endeffekt ungleichmäßig großen Projektgruppen, wie in Tabelle 8 zu sehen. Allerdings wurde darauf geachtet, dass die zu vergleichenden Gruppen (1 – 3, 4 – 6, 7 – 9 und 10 – 12, siehe Tabelle 9) in etwa gleich groß blieben.

Tabelle 8
Größe der zwölf Projektgruppen

Projektgruppe	Anzahl Studierende	Prozentualer Anteil der Gesamtstichprobe
Gruppe 1	79	5.8%
Gruppe 2	66	4.9%
Gruppe 3	89	6.6%
Gruppe 4	112	8.3%
Gruppe 5	114	8.4%
Gruppe 6	117	8.6%
Gruppe 7	117	8.6%
Gruppe 8	114	8.4%
Gruppe 9	116	8.5%
Gruppe 10	145	10.7%
Gruppe 11	145	10.7%
Gruppe 12	143	10.5%
Gesamt	**1357**	**100%**

2.2 Durchführung

Die vorliegende Studie fand im Zeitraum von Oktober 2009 bis August 2010 statt. Tabelle 9 gibt einen Gesamtüberblick über den Versuchsplan. Wie zu sehen ist, war es durch die Versuchsanordnung möglich, die experimentellen Variationen innerhalb des Wintersemesters viermal zu erheben. Durch diese mehrfache Replikation wurde die Material-, Inhalts- und Personinvarianz der Daten sichergestellt.

Die einzelnen Phasen der Durchführung werden im Folgenden näher erläutert.

Tabelle 9
Versuchsplan zu Studie I

Oktober 2009: Vorbesprechung/Organisatorisches												
Oktober 2009 bis Februar 2010: Kurzfristige Akzeptanzerhebung und Wissensüberprüfung nach Besuch der Vorlesung oder Nutzung von E-Lecture oder Web-based Training												
Vorlesungsthemen	**12 Projektgruppen, 4 Replikationen**											
	1	2	3	4	5	6	7	8	9	10	11	12
1 Gedächtnis im Überblick	VL	EL	WBT									
2 Kurzzeitgedächtnis	EL	WBT	VL									
3 LZG: Enkodierung[24]	WBT	VL	EL									
4 LZG: Erinnern & Vergessen, Verbesserung der Gedächtnisleistung [25,26]				VL	EL	WBT						
5 Klassisches/operantes Konditionieren				EL	WBT	VL[27]						
6 Lernen am Modell				WBT	VL[24]	EL						
7 Denken, Problemlösen, Intelligenz[22]						VL	VL	EL	WBT			
8 Entwicklungspsychologie					VL		EL	WBT	VL			
9 Unterrichtsplanung, -orga., -methoden							WBT	VL	EL			
10 Lehren und Lernen mit Medien										VL	EL	WBT
11 Leistungsdiagnostik[23]										EL	WBT	VL
12 Rückblick										WBT	VL	EL
Februar/März 2010: Längerfristige Akzeptanzerhebung und Wissensüberprüfung												
März bis August 2010: Bereitstellung aller E-Lectures und Web-based Trainings zur Klausurvorbereitung												
August 2010: Abschließende Akzeptanzerhebung und Wissensüberprüfung												
23./24. August 2010: **Modulabschlussprüfung**												

Anmerkung. VL = Vorlesungs-, EL = E-Lecture-, WBT = WBT-Gruppe

Vorbesprechung/Organisatorisches

In den ersten beiden Vorlesungssitzungen im Oktober 2009 wurden die Studierenden über die organisatorischen Abläufe der Vorlesung und des Projekts informiert und bekamen eine Einführung in das Thema „Effektives Lernen“. Auch wurden Sie instruiert, sich möglichst

[24] Professor Glowalla wurde in dieser Vorlesung aufgrund einer Erkrankung von der Verfasserin dieser Arbeit vertreten.

[25] Das Thema wurde in der Vorlesung nicht vollständig bearbeitet – dies wird bei der Erstellung bzw. Auswertung der Wissensüberprüfungen berücksichtigt.

[26] Die E-Lecture zu diesem Thema wurde von Frau Dr. Glowalla gehalten.

[27] Diese Vorlesungen fielen leider aufgrund eines Studierendenstreiks aus, so dass für die Auswertung dieser Themen keine Vorlesungsgruppen zur Verfügung stehen. Die Studierenden konnten die Evaluation zwei Wochen später nachholen. Die Inhalte wurden gestrafft in der Sitzung „Rückblick“ nachgeholt.

bald in dem Lehrmanagementsystem der Justus-Liebig-Universität Gießen *Stud.IP* zur Vorlesung anzumelden.

Anhand der Teilnehmerliste in *Stud.IP* wurden die Studierenden in der Folge zufällig in zwölf etwa gleich große Projektgruppen eingeteilt. Diese Gruppen sollten später sukzessive in vier Untersuchungsblöcken à jeweils drei Gruppen an dem Projekt teilnehmen (siehe Tabelle 9). Dabei wurde die Bearbeitung der Lehrmethoden innerhalb jedes Untersuchungsblockes rotiert, um Reihenfolgeeffekte vorzubeugen. Jeder Studierende evaluierte so im besten Falle jede Lehrmethode einmal und zu jedem Vorlesungsthema wurden jeweils alle drei Lehrmethoden von unterschiedlichen Studierendengruppen evaluiert. Dies ermöglicht sowohl einen querschnittlichen als auch einen längsschnittlichen Vergleich.

Nach der Gruppenzuordnung wurde jede Untersuchungsgruppe per Email über die Termine ihrer Projektteilnahme informiert. Auch wurden die Studierenden instruiert, wie sie sich zu dem Lehrmanagementsystem *Moodle*[28] der Forschungsgruppe IIM registrieren konnten, da die Bereitstellung der E-Lectures, Web-based Trainings und Evaluationsbefragungen sowie die Speicherung der Daten über diese Plattform erfolgte. Der Zugangscode, der jeweils an die Gruppen versandt wurde, sollte gewährleisten, dass die einzelnen Studierenden innerhalb von *Moodle* ebenfalls der richtigen Gruppe zugeordnet waren und so nur diejenigen Lehrangebote und Befragungen sehen und nutzen konnten, die jeweils für sie freigeschaltet wurden.

Kurzfristige Akzeptanzerhebung und Wissensüberprüfung

Die kurzfristige Evaluation der Lehrmethoden fand sukzessive in vier Untersuchungsblöcken zwischen Oktober 2009 und Februar 2010 statt (siehe Tabelle 9). Dafür wurden die teilnehmenden Studierendengruppen jeweils zwei Tage vor Bereitstellung der E-Lecture/des WBTs bzw. der Vorlesung per Email an ihre Teilnahme erinnert. Mittwochs sollten die drei teilnehmenden Projektgruppen die Vorlesung besuchen, die E-Lecture ansehen oder das WBT bearbeiten. Die aktuelle Vorlesungsgruppe musste hierbei auf einer Anwesenheitsliste unterschreiben, die Teilnahme der E-Lecture- und WBT-Gruppe wurde automatisch durch das System erfasst. Studierende, die in der jeweiligen Woche nicht am Projekt teilnahmen, besuchten die reguläre Vorlesungssitzung, ohne auf einer Anwesenheitsliste zu unterschreiben, da die Vorlesung an sich keine Teilnahmepflicht beinhaltete.

Ein Tag später sollten die teilnehmenden Studierenden in der Folge das entsprechende Lehrangebot evaluieren und einen Wissenstest bearbeiten. Dabei war der

[28] „Modular Object-Oriented Dynamic Learning Environment“ – weit verbreitete Open Source Lernplattform (Weitere Informationen unter http://moodle.de).

Evaluationsfragebogen jeweils der Wissensüberprüfung vorgeschaltet, damit die Studierenden ihn vollständig bearbeiteten. Die genutzten Evaluationsfragebögen und Wissensüberprüfungen werden in Abschnitt „2.3 Material“ erläutert.

Längerfristige Akzeptanzerhebung und Wissensüberprüfung

Im Februar/März 2010 fanden die längerfristige, vergleichende Abschlussbefragung sowie daran anschließend eine Wissensüberprüfung durch eine alle Themen umfassende „Probeklausur“ statt (siehe Abschnitt „2.3 Material“). Auch hierüber wurden die Studierenden im Vorfeld per Email informiert.

Leider fehlen von dieser Erhebung einige Datensätze, da es zu technischen Schwierigkeiten innerhalb des erfassenden Systems kam. Zwar zeigte das Systemprotokoll verlässlich an, dass die allermeisten der Studierenden an der Erhebung teilgenommen hatten, die Trackingdaten mancher Studierender wurden jedoch nicht vom Server gespeichert. Die Auswertungen hierzu werden dennoch ohne Einschränkung mit der reduzierten Stichprobe durchgeführt. Da die technischen Schwierigkeiten sich zufällig über die Stichprobe verteilten, ist davon auszugehen, dass diese Datenreduktion die Verteilung der Studierenden nicht beeinflusst hat.

Abschließende Akzeptanzerhebung und Wissensüberprüfung

Um nach Erfragung der subjektiven Einstellungsakzeptanz auch die Verhaltensakzeptanz zu erheben, wurden den Studierenden von März 2010 bis zur Modulabschlussprüfung im August 2010 alle zwölf E-Lectures und Web-based Trainings zur freien Nutzung online gestellt. Etwa zehn Tage vor der Modulabschlussprüfung wurden zudem eine letzte abschließende Kurzevaluation sowie erneut die Probeklausur zur Verfügung gestellt (siehe ebenfalls Abschnitt „2.3 Material“), die gemeinsam mit der anschließenden Modulabschlussprüfung als letztes Lernleistungsmaß in die Auswertung einging.

2.3 Material

Zur Durchführung des Projekts wurden verschiedene Lehrmaterialien und Evaluationsfragebögen eingesetzt, die im Folgenden näher beschrieben werden. Die hierfür verwendete Software sowie die genutzten Lehrmanagementsysteme werden übersichtshalber in Tabelle 10 dargestellt.

Tabelle 10
Eingesetzte Software sowie Learning Management Systeme (LMS)

Software	**Verwendungszweck**
Microsoft PowerPoint®	Erstellung der Präsentationsfolien für die Vorlesung
Lectora Publishing Suite®	Autorenwerkzeug der Firma Trivantis, das zur Erstellung der Web-based Trainings sowie der Evaluationsfragebögen genutzt wurde.
NetLecture	Autorensoftware der Firma Lerndesign GmbH, mit dem die E-Lectures erstellt wurden.
Lehrmanagementsystem	**Verwendungszweck**
Stud.IP	Lehrmanagementsystem der JLU-Gießen, das zur Veranstaltungsadministration genutzt wird und in der Studie insbesondere dafür genutzt wurde, die Studierendengruppen gezielt anzuschreiben.
Moodle	Open Source Lernplattform, welche zusätzlich genutzt wurde, um alle elektronischen Materialien zur Verfügung zu stellen und Nutzerdaten zu erfassen.

2.3.1 Lehrmaterialien: Vorlesung, E-Lectures und Web-based Trainings

Die drei genutzten und evaluierten Lehrmethoden Vorlesung, E-Lecture und Web-based Training wurden bereits ausführlich im theoretischen Abschnitt dieser Dissertation erläutert (siehe Abschnitte „1.3 Traditionelle Lehre: die Vortragsmethode“, „1.4.2 E-Lecture“ und „1.4.3 Web-based Training (WBT)“). In den Abschnitten zur E-Lecture und zum Web-based Training wurden zur Veranschaulichung Beispiele aus der vorliegenden Studie genutzt. Aus diesem Grunde soll auf eine erneute ausführliche Beschreibung der Lehrmethoden an dieser Stelle verzichtet und lediglich auf die Entstehung der elektronischen Lehrmaterialien eingegangen werden.

Die E-Lecture entstand aus einer Aufzeichnung der Vorlesung aus den Jahren 1999/2000. Die hieraus entstanden Videos wurden mittels der Software *NetLecture* mit den zugehörigen Präsentationsfolien zusammengefügt. Im Jahre 2005 wurde sie erneut als DVD aufgelegt (Glowalla & Glowalla, 2005) und im Rahmen der vorliegenden Studie genauso wie die Web-based Trainings im Lehrmanagementsystem *Moodle* online bereitgestellt. An der E-Lecture wurden keine Veränderungen vorgenommen, da sie inhaltlich noch immer vergleichbar zur Vorlesung war. Kleinere Änderungen wurden bei der Erstellung der Testaufgaben zur Wissensüberprüfung berücksichtigt.

Um für die zwölf Vorlesungsthemen Web-based Trainings zu erstellen, wurden im Herbst 2009 die E-Lectures transkribiert. Die hieraus resultierenden Redetexte wurden in der Folge in Lesetexte überführt, welche inhaltlich sinnvoll auf einzelnen WBT-Seiten angeordnet und durch Bilder, Videos etc. ergänzt wurden. Inhaltlich sind die Web-based Trainings damit vergleichbar zu den E-Lectures – und damit zur Vorlesung.

Für die Themen 1 bis 3 sowie 7 bis 9 (siehe Tabelle 9) wurden die Web-based Trainings zudem mit Übungsaufgaben bestückt. Die Web-based Trainings zu den Themen 4 bis 6 sowie 10 bis 12 enthielten keine Übungsaufgaben, um zu überprüfen, ob eine sich eventuell ergebende bessere Lerneffektivität oder Akzeptanz der Web-based Trainings ggf. nur auf die Übungsaufgaben zurückzuführen ist.

2.3.2 Abhängige Variablen

Während des gesamten Projekts wurden mehrere Fragebögen und Wissensüberprüfungen erstellt und eingesetzt, welche im Folgenden vorgestellt werden sollen. Abbildung 15 am Ende dieses Abschnittes gibt einen Überblick über die Erhebungen. Die vollständigen Fragebögen befinden sich in Anhang A.
Die Items für die Akzeptanzerhebungen entstammen hauptsächlich einer Fragensammlung aus früheren Studien der Forschungsgruppe „Instruktion und Interaktive Medien“. Im Rahmen der vorliegenden Studie wurden Fragen zu einzelnen Themenkomplexen mittels Faktorenanalyse auf Eindimensionalität/Mehrdimensionalität überprüft und eine Reliabilitätsberechnung für die entstandenen Skalen durchgeführt. Die Ergebnisse dieser Berechnungen werden ebenfalls in Anhang A berichtet. Weitere Items wurden auf Einzel-Item-Ebene ausgewertet.

Allgemeiner Fragebogen

Der allgemeine Fragebogen wurde jeweils in der ersten Woche der Teilnahme an der kurzfristigen Akzeptanzerhebung dem Evaluationsfragebogen zur E-Lecture/dem WBT oder der Vorlesung (je nach Gruppenzugehörigkeit, siehe Tabelle 9) vorgeschaltet. Folgende Variablen wurden erhoben:

- *Demographische Variablen*: Geschlecht, Alter, Studiengang, Fachsemester, Abiturnote und Ort des Erwerbs des Abiturs
- *Allgemeine Fragen*: Einstellung zum Projekt, Häufigkeit der Teilnahme an der Vorlesung, Vorkenntnisse und Interesse an der Vorlesung
- *Computerbezogene Fragen*: Häufigkeit der Computernutzung, genutztes Betriebssystem, Einschätzung der eigenen Computerfähigkeiten

Die Ergebnisse der Auswertung dieses Fragebogens wurden bereits in Abschnitt „2.1 Stichprobe“ berichtet.

Kurzfristige Akzeptanzerhebung und Wissensüberprüfung

Evaluation von Vorlesung, E-Lecture und Web-based Training

Nach zwei Vorabfragen, in denen die Vorkenntnisse der Studierenden bezüglich des bearbeiteten Themas und ihre bisherigen Erfahrungen mit der am Vortag genutzten Lehrmethode abgefragt wurden, sollten die Studierenden die Lernförderlichkeit der jeweiligen Lehrmethode einschätzen. Zudem wurden die Akzeptanz der Lehrmethode sowie das Verhalten vor, während und nach Bearbeitung/des Besuchs der Lehrmethode (Vor-/Nachbereitung, Pausen und Notizen) erhoben.

Evaluation der „Übungstests", Vergleich der Lehrmaterialien

In der dritten und damit letzten Woche der kurzfristigen Akzeptanzerhebung wurde dem Evaluationsfragebogen zur E-Lecture/dem WBT oder der Vorlesung (je nach Gruppenzugehörigkeit, siehe Tabelle 9) ein weiterer kurzer Fragebogen nachgeschaltet. Dieser enthielt sechs Items zur Beurteilung der Wissensüberprüfungen („Übungstests") sowie fünf Items zum Vergleich der Lehrmethoden „Vorlesung", „E-Lecture", „WBT", „Übungsaufgaben" und „Literatursammlung". Darüber hinaus gab er den Studierenden zusätzlich die Möglichkeit, ein offenes Feedback zum Projekt zu geben.

Wissensüberprüfung – „Übungstests"

Im Anschluss an jede kurzfristige Akzeptanzerhebung sollten die Studierenden mit einem sogenannten „Übungstest" ihr Wissen zu dem am Vortag bearbeiteten Thema überprüfen. Mittels dieser Übungstests wurde die kurzfristige Lernleistung der Studierenden erhoben. Sie enthielten Lückentext-, Mehrfach- bzw. Einfachauswahl- sowie Wahr-/Falsch-Aufgaben. Nach jeder Aufgabe erhielten die Studierenden ein informatives Feedback über die Richtigkeit ihrer Antwort sowie (im Falle der Falschbeantwortung) die richtige Lösung. Am Ende des Tests erhielten sie zudem eine Rückmeldung zu ihrer Gesamtlernleistung (ähnlich der in Abbildung 11).

Die Fragen für die Wissensüberprüfung wurden für die vorliegende Studie neu entwickelt. Sie werden gemeinsam mit ihren Item-, Skalenkennwerten (Schwierigkeit, korrigierte Trennschärfe, Cronbachs α) und Besonderheiten in Anhang A im Überblick dargestellt. Dies gilt auch für die im nächsten Abschnitt beschriebene Probeklausur.

Längerfristige Akzeptanzerhebung und Wissensüberprüfung

Abschlussfragebogen

Vorgeschaltet zur Probeklausur am Ende des Wintersemesters wurden eine Skala zu den Computerfähigkeiten sowie eine Skala zur Einstellung zu E-Learning erfasst. Beide Skalen

sowie ihre Ergebnisse wurden bereits in Abschnitt „2.1.1 Beschreibung der Gesamtstichprobe“ erläutert.
In der Folge sollten die Studierenden zudem die Lehrmaterialien/-methoden „Vorlesung“, „E-Lecture“, „Web-based Training“, „Übungsaufgaben“, „Vorlesungsskript“ und „Literatur“ in fünf Items hinsichtlich ihres Nutzungswunsches und einer Gesamtnote beurteilen.
Hinzu kamen vier vergleichende Fragen zu den untersuchten Lehrmethoden „Vorlesung“, „E-Lecture“ und „Web-based Training“: Welche Lehrmethode erbrachte für sie den größten Nutzen, welche war am meisten motivierend, welche machte ihnen am meisten Spaß und im Rahmen welcher konnten sie sich am besten konzentrieren?

Wissensüberprüfung – „Probeklausur“
Die Wissensüberprüfung am Ende des Wintersemesters bestand aus einer „Probeklausur“, bei der zu allen zwölf Veranstaltungsthemen je vier Fragen gestellt wurden. Zwei Fragen wurden bereits in den vorherigen kurzfristigen Wissensüberprüfungen genutzt, zwei weitere wurden für die Probeklausur neu erstellt. Sie enthielt damit insgesamt 48 Fragen in verschiedenen Formaten (siehe „Übungstests“), maximal sechs Fragen von diesen hatten die Studierenden durch die kurzfristigen Wissensüberprüfungen bereits kennengelernt. Auch hier erhielten die Studierenden nach jeder Frage sowie am Ende der Probeklausur eine Rückmeldung zu ihrem Lernstand.

Abschließende Akzeptanzerhebung und Wissensüberprüfung
Kurzevaluation
Innerhalb von vier Items sollten die Studierenden kurz vor der Modulabschlussprüfung sechs Lehrmethoden/-materialien miteinander vergleichen, die bei ihrer Klausurvorbereitung zum Einsatz gekommen sein könnten: das Vorlesungsskript, ihre Mitschriften aus der Vorlesung, die E-Lectures, die Web-based Trainings, die Literatur und eine Kategorie „Sonstiges“, die die Studierenden frei angeben konnten. Zu diesen sechs Materialien sollten sie angeben, wie häufig sie diese für ihre Klausurvorbereitung genutzt hatten, wie motivierend sie diese empfunden haben und wie viel Spaß ihnen die Nutzung gemacht hatte. Außerdem sollten sie die Lehrmaterialien abschließend auf der Notenskala von 1 bis 6 bewerten.

Wissensüberprüfung – „Probeklausur“
Die langfristige Wissensüberprüfung wurde im Anschluss an die abschließende Kurzevaluation durchgeführt. Hier wurde erneut die Probeklausur genutzt, die bereits als längerfristige Wissensüberprüfung am Ende des Wintersemesters zum Einsatz gekommen war. Da fast ein halbes Jahr zwischen den beiden Erhebungen lag, wurde davon

ausgegangen, dass diese erneute Nutzung der Probeklausur sich nur in sehr geringem Maße auf die Lernleistungen der Studierenden auswirken sollte.

Modulabschlussprüfung

In die Auswertung der vorliegenden Arbeit geht lediglich der Teil der Modulabschlussprüfung ein, der die Inhalte der Vorlesung „Psychologie für Lehramtsstudierende A“ abprüft. Der zweite Teil der Prüfung bezieht sich auf die ebenfalls zum Modul gehörende Sommersemestervorlesung, die nicht Bestandteil der Studie war.

Die Modulabschlussprüfung wurde – aufgrund der großen Zahl an Studierenden – in mehreren Gruppen geschrieben. Deshalb kamen vier Versionen der Klausur zum Einsatz, bei denen die Fragen zudem jeweils in zwei Reihenfolgen dargeboten wurden. Hieraus resultierten folglich acht Klausurversionen. Jede Klausurversion bestand aus insgesamt 24 Fragen (14 Einfachauswahl-, zwei Mehrfachauswahl- und acht offene Fragen), die Studierenden konnten in diesem Teil der Klausur 34 Punkte erzielen (Einfachauswahl-Fragen ergaben einen Punkt, Mehrfachauswahl- und offene Fragen zwei Punkte).

Die Klausurfragen wurden so ausgewählt, dass in etwa gleich schwere Klausuren entstanden. Die Itemkennwerte, die hierfür zugrunde gelegt wurden, lagen bereits aus den Vorjahren vor. Da die Fragen in der einen oder anderen Form in den kommenden Studienjahren weiterhin verwendet werden sollen, werden sie in dieser Arbeit nicht aufgeführt. Die an der aktuellen Stichprobe neu berechneten Itemkennwerte der Klausurversionen befinden sich in Anhang A. Tabelle 11 zeigt die Verteilung der Klausurfragen auf die elf Themen[29] bzw. sechs Themenkomplexe der Vorlesung.

Tabelle 11

Aufteilung der Klausurfragen auf die elf Vorlesungsthemen bzw. sechs Themenkomplexe in den vier Klausurversionen A bis D sowie Angabe der Studiengänge der Klausurgruppe

Klausurversion		**Vorlesungsthemen**											**Studiengang der Klausurgruppe**
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	
A	Fragen	1	-	3	3	2	2	5	3	2	1	2	L3: 100%
	Punkte	1	0	5	4	2	4	8	3	3	1	3	
	Punkte Thema	10				6		8	3	4		3	
B	Fragen	1	1	2	3	3	2	4	2	4	-	2	L2: 98.4%; L3: 1.0% L1: 0.5%
	Punkte	2	1	2	5	4	3	6	2	6	0	3	
	Punkte Thema	10				7		6	2	6		3	
C	Fragen	-	2	2	4	2	2	4	2	4	-	2	L3: 45.7%; L5: 36.8%; L4: 20.0%; L2: 0.4%
	Punkte	0	2	4	6	3	3	6	2	5	0	3	
	Punkte Thema	12				6		6	2	5		3	
D	Fragen	-	-	2	6	2	1	6	2	3	-	2	L1: 48.4%; L3: 41.0%; L2: 9.6%; L4,L5: je 0.5%
	Punkte	0	0	2	9	3	1	8	2	6	0	3	
	Punkte Thema	11				4		8	2	6		3	
Punkte gesamt		**10-12**				**4-7**		**6-8**	**2-3**	**4-6**		**3**	

Anmerkung. Themenkomplexe: Gedächtnis (Thema 1-4), Lernen (Thema 5/6), Denken und Problemlösen (Thema 7), Entwicklungspsychologie (Thema 8), Unterrichten (Thema 9/10) sowie Pädagogische Diagnostik (Thema 11)

[29] Thema 12 ist hier irrelevant, da es einen Rückblick auf alle anderen elf Themen darstellte.

Es zeigt sich, dass die Vorlesungsthemen in den verschiedenen Klausurversionen stellenweise nicht mit dem gleichen Gewicht abgefragt wurden, die Punktevergabe bezogen auf die sechs Themenkomplexe ist jedoch in etwa vergleichbar. Die Punkte zu den einzelnen Themen wurden deshalb für die Auswertungen in Prozente umgerechnet, damit sie zwischen den Klausurversionen vergleichbar sind.

Um die Unterschiede zwischen den Klausurversionen auf Signifikanz zu überprüfen, wurden *t*-Tests oder einfaktorielle Varianzanalysen bzw. Brown-Forsythe-Tests berechnet. Die Ergebnisse sind in Tabelle 12 abgetragen. Wie zu sehen ist, unterscheiden sich die Klausurversionen für die vorliegende Stichprobe in Ihrer Schwierigkeit. Insbesondere die Klausurversion B scheint in einigen Bereichen sowie insgesamt schwieriger gewesen zu sein, die Klausurversion D leichter.

Tabelle 12
Deskriptive Statistik sowie Mittelwertsanalysen zum Vergleich der Klausurversionen

	A (*n* = 263)		**B (*n* = 213)**		**C (*n* = 247)**		**D (*n* = 204)**		**Gesamt**		**ANOVA/BF-Test**				**Posthoc**
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*t/F*	*df*	*p*	*Eta²*	
1	84.8	35.98	84.6	26.18	-	-	-	-	88.9	32.14	0.06	468.87	.95	-	-
2	-	-	85.4	35.35	90.3	21.31	-	-	88.0	28.74	-1.74	337.14	.08	-	-
3	71.3	24.47	55.6	33.87	71.1	25.96	68.9	28.50	67.1	28.80	15.18	809.68,3	<.01*	.05	B<A,C,D
4	66.8	23.67	46.2	24.53	61.1	20.57	57.6	18.56	58.5	23.25	36.20	877.62,3	<.01*	.10	B<D=C<A
5	71.7	31.59	77.6	22.82	61.5	24.91	70.5	25.54	70.1	27.25	14.72	896.28,3	<.01*	.05	C<A,B,D; D<B
6	61.6	25.12	52.3	27.31	63.6	26.36	87.7	32.87	65.7	30.42	61.67	923,3	<.01*	.17	B<A=C<D
7	46.6	24.94	38.8	24.42	45.9	24.75	69.1	18.83	49.6	25.88	66.39	901.32,3	<.01*	.18	B<A=C<D
8	84.8	23.77	69.7	34.15	91.7	21.19	89.7	22.57	84.3	26.98	31.95	729.25,3	<.01*	.10	B<A,C,D; A<C
9	58.5	25.73	57.1	19.24	73.4	24.34	69.9	22.75	64.6	24.36	29.17	907.52,3	<.01*	.09	B=A<D=C
10	96.2	19.2	-	-	-	-	-	-	96.2	19.2	-	-	-	-	-
11	63.0	37.70	60.9	35.05	70.8	26.92	72.6	26.49	66.7	32.46	7.27	857.22,3	<.01*	.02	B=A<D=C
Ges.	21.9	5.92	19.6	5.37	22.5	4.90	23.3	4.84	21.8	5.45	18.68	906.37,3	<.01*	.06	B<A,C,D; A<D

Anmerkung. α = .05; *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *t* = Prüfstatistik des *t*-Tests, *F* = Prüfstatistik der ANOVA bzw. des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke

Die signifikanten Unterschiede zwischen den Klausurversionen sind vermutlich allerdings dadurch zustande gekommen, dass die Klausurgruppen unter anderem anhand des Lehramtsstudiengangs gebildet wurden (siehe Tabelle 11). Die Klausurversion A wurde von L3-Studierenden geschrieben, Klausurversion B hauptsächlich von L2-Studierenden, Klausurversion C hauptsächlich von L3-, L4- und L5-Studierenden und Klausurversion D hauptsächlich von L1- und L3-Studierenden. Wie Tabelle 13 zeigt, unterscheiden sich diese Gruppen von Lehramtsstudierenden hinsichtlich ihres Abiturschnittes. Dieser Unterschied ist signifikant mit $F(4) = 31.40$ und $p < .01$ ($Eta^2 = .11$).

Tabelle 13
Durchschnittliche Abiturnoten der Studierenden der fünf Lehramtsstudiengänge

Studiengang	***n***	***M***	***SD***
L1	102	2.31	0.53
L2	315	2.85	0.50
L3	488	2.54	0.53
L4/BaBBB	51	2.71	0.60
L5	105	2.41	0.47

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Posthoc-Tests zeigten, dass insbesondere die L2-Studierenden ein schlechteres Abitur haben als die L1-, L3- und L5-Studierenden[30]. Dies kommt auch dadurch zustande, dass die Studiengänge L1 und L5 universitätsintern zulassungsbeschränkt sind. Studierende der Lehramtsstudiengänge L1, L3 und L5, die hauptsächlich die Klausurversionen A, C und D geschrieben haben, haben folglich ein besseres Abitur als Studierende des Lehramtsstudienganges L2, denen die Klausurversion B vorgelegt wurde. So könnten die Schwierigkeitsunterschiede der Klausurversionen auch durch unterschiedliche Eigenschaften (Intelligenz, Interesse, Leistungsbereitschaft etc.) der Studierendengruppen erklärbar sein.

Diese Überlegung wurde mittels einer Kovarianzanalyse überprüft, in welcher die Abiturnote als Kovariate für den Vergleich der vier Studierendengruppen (Klausurversionen A bis D) hinsichtlich ihres Abschneidens im Teil A der Klausur berücksichtigt wurde. Es ergab sich nur noch eine knappe Signifikanz mit geringer Effektstärke zwischen den Klausurversionen ($F(3) = 2.70$, $p = .045$, Partielles $Eta^2 = .01$ – die Ergebnisse der einfaktoriellen Varianzanalyse ohne Berücksichtigung der Abiturnote befinden sich in Tabelle 12, letzte Zeile). Innerhalb der Posthoc-Vergleiche ergab sich nunmehr lediglich eine Signifikanz zwischen den Klausurversionen B und D. Selbst nach Berücksichtigung der Abiturnote der Studierenden scheint die Klausurversion D insgesamt etwas leichter gewesen zu sein als die Klausurversion B.

User-Tracking

Des Weiteren wurde im Lehr-Management-System *Moodle* direkt erfasst, welcher Studierende wann und wie lange mit der E-Lecture oder dem Web-based Training gearbeitet hat. Diese Daten gehen in die Berechnung der Lerneffizienz der drei Lehrmethoden Vorlesung, E-Lecture und Web-based Training ein. Die Zeitmessung zur E-Lecture-Nutzung hat sich im Wintersemester allerdings als unzuverlässig herausgestellt. Für die kurz- und längerfristige Lerneffizienz kann dementsprechend lediglich ein Vergleich zwischen den WBTs und Vorlesungen angestellt werden.

Im Nachgang mussten zudem einzelne Studierende anhand ihrer Nutzungszeiten aus den Auswertungen genommen werden. Im Wintersemester wurden diejenigen WBT-Nutzungen eliminiert, die kürzer als fünf Minuten oder länger als sechs Stunden dauerten. Dies wurde dadurch begründet, dass bei beiden Extremen nicht mehr davon ausgegangen werden kann, dass sinnvoll mit den Lernanwendungen gelernt wurde. Insbesondere bei den sehr kurzen Nutzungen kann nicht davon ausgegangen werden, dass die Studierenden die Lernanwendungen gut genug kennengelernt hatten, um sie evaluieren zu können. Die sehr

[30] Weitere Signifikanzen in den Post-hoc-Tests: L4>L1 und L5; L3>L1

langen Nutzungszeiten lassen hingegen vermuten, dass die Lerneinheit lediglich geöffnet und dann ggf. vergessen wurde. Die Begrenzungen wurden dabei bewusst sehr extrem gesetzt, um möglichst nur wenige Studierende auszuschließen und die flexible Lernweise der Studierenden nicht zu sehr zu beschneiden.

Vergleichbar wurden diejenigen Studierenden aus den Auswertungen ausgeschlossen, deren Nutzungszeiten im Sommersemester unter einer Minute lagen. Die Dauer wurde hier niedriger gesetzt als im Wintersemester, da davon ausgegangen werden kann, dass in der Vorbereitung auf die Klausur in den E-Lectures oder den WBTs beispielsweise einzelne Begriffe kurz nachgeschlagen wurden. Um in der Summe zu den „E-Lecture-“ oder „WBT-Nutzern“ gezählt zu werden, mussten die Studierenden allerdings insgesamt zehn Minuten mit den WBTs oder E-Lectures gearbeitet haben. Das Maximum der Lernzeit wurde im Sommersemester bei acht Stunden begrenzt, da der Lernzeitraum in der Phase der Klausurvorbereitung noch freier gehalten werden sollte als im Wintersemester.

Die Dauer der Vorlesung wurde anhand der Netto-Laufzeit der E-Lecture (ohne Sprünge oder Vor- und Zurückspulen) festgelegt, da es nicht möglich war, sie aufzuzeichnen. Bei den Themen, die Professor Glowalla in der Vorlesung inhaltlich nicht beenden konnte (Thema „Gedächtnis IV“ und „Denken, Problemlösen und Intelligenz“), wurde dennoch die komplette E-Lecture-Laufzeit zugrunde gelegt, da die Studierenden in der WBT-Bedingung die Möglichkeit hatten, das gesamte Thema zu bearbeiten und die Bearbeitungsdauer der Lehrmethoden so relativ vergleichbar blieben. Die nicht-behandelten Inhalte wurden in den Übungstests/Probeklausuren nicht erfragt oder im Nachhinein aus den Auswertungen genommen, so dass alle Teilnehmer an der Befragung auch immer vorher alle Inhalte kennengelernt hatten.

Abbildung 15 gibt abschließend einen Überblick über die eingesetzten Fragebögen und Wissensüberprüfungen.

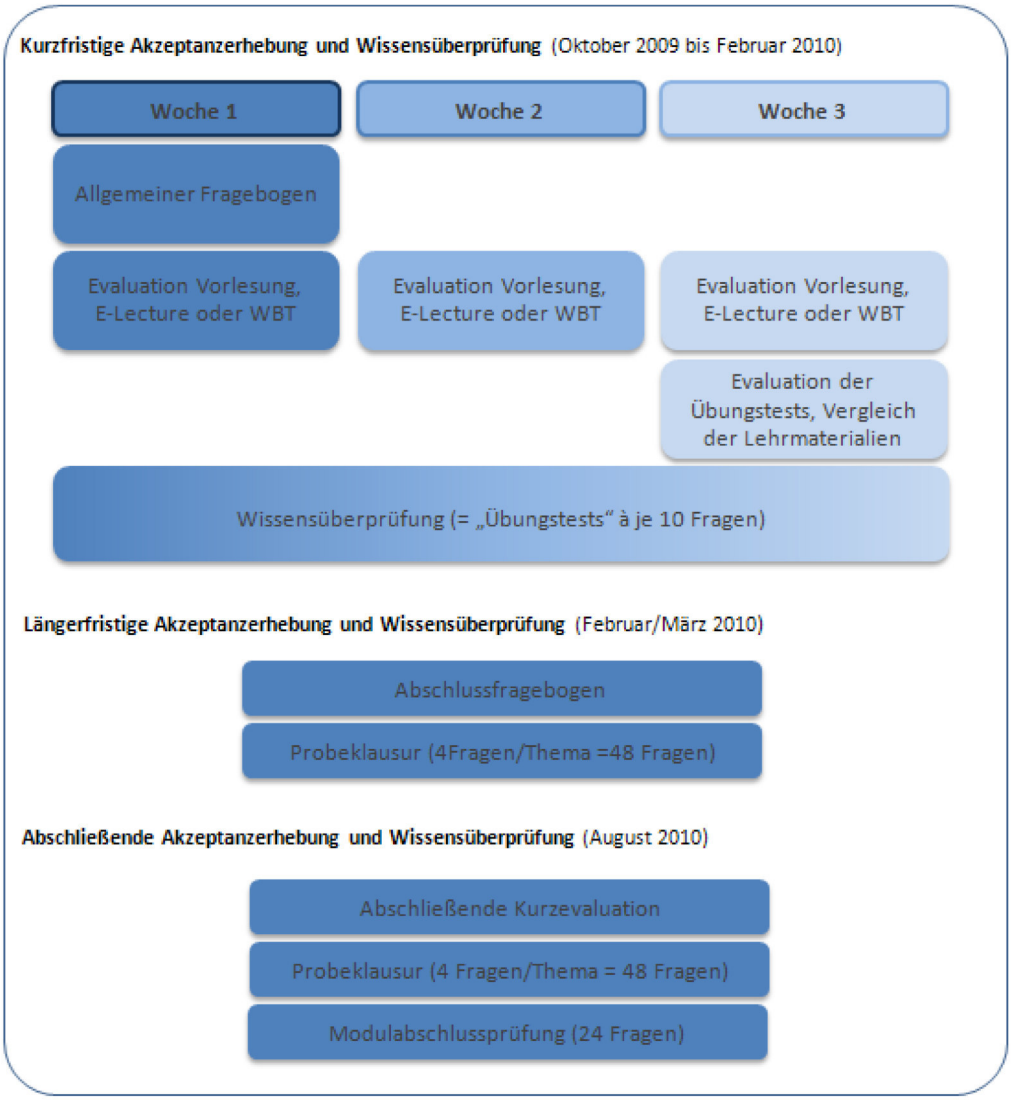

Abbildung 15. Eingesetzte Fragebögen und Wissensüberprüfungen innerhalb der Studie I

3 Studie I: Ergebnisse und Diskussion

Alle dem Ergebnisteil zugrundeliegenden Auswertungen wurden mit dem Statistikprogramm *SPSS 20* durchgeführt. Für die Prüfung von Hypothesen mittels Mittelwertsvergleichen wurde ein Signifikanzniveau von $\alpha = .05$ zugrunde gelegt.

Zur Hypothesenprüfung wurden hauptsächlich einfaktorielle Varianzanalysen (ANOVAs) berechnet, die voraussetzen, dass die zugrundeliegenden Daten normalverteilt und varianzhomogen sind. Die Überprüfung der Normalverteilung erfolgte deshalb mittels Kolmogorov-Smirnov-Test, die Überprüfung der Varianzhomogenität mittels Levene-Test. Ist Varianzhomogenität gegeben, Normalverteilung allerdings nicht, werden dennoch der *F*-Wert und die Ergebnisse der Post-Hoc-Tests (paarweise *t*-Tests mit Bonferroni-Adjustierung) angegeben. Nach Bortz und Schuster (2010) verlieren die Annahmen der Varianzanalyse mit steigender Stichprobengröße an Bedeutung. Zudem ist die ANOVA bei vergleichbar großen Stichproben relativ robust gegen Verletzungen ihrer Annahmen. Da in der vorliegenden Studie relativ große Stichproben vorliegen und diese vergleichbar groß sind (oder dies durch Ziehung von Zufallsstichproben aus großen Gruppen erzeugt wurde), kann die ANOVA berechnet werden. In Fällen, in denen jedoch auch die Annahme der Varianzhomogenität verletzt ist, werden Ergebnisse des Brown-Forsythe-Tests und Post-Hoc-Ergebnisse aus paarweisen Welch-Tests berichtet. Dies gilt auch in den Fällen, in denen nur zwei Gruppen miteinander verglichen werden. Bei nicht gegebener Varianzhomogenität werden hier anstelle von *t*-Tests die Ergebnisse des Welch-Tests angegeben. Die zugehörige Effektstärke wurde mittels des Programms *GPower 3.1* berechnet. Anhand der angegebenen Freiheitsgrade ist im konkreten Fall zu erkennen, welcher Test genutzt wurde.

In den Fällen, in denen die Berechnung einer abhängigen Varianzanalyse notwendig ist, wird generell das Verfahren von Box (Greenhouse-Geisser-Korrektur) angewandt, da die Messwiederholungs-Varianzanalyse anfällig für große Strichprobenumfänge und für die Verletzung der Normalverteilungsannahme ist. Deshalb raten Rasch, Friese, Hofmann und Naumann (2008) dazu, bei großen Stichproben in jedem Fall die Greenhouse-Geisser-Korrektur anzuwenden.

Da insbesondere für große Stichproben schon sehr kleine Effekte signifikant werden können, werden neben der Signifikanz stets auch die Effektstärken angegeben. Dabei wird vergleichbar zu Studie I nach Rasch, Friese und Hofmann (2009) ein Eta^2 von .01 als schwacher, ein Eta^2 von .06 als mittlerer und ein Eta^2 von .14 als starker Effekt interpretiert. Wenn auf die Effektstärke *d* zurückgegriffen wird, wird nach Cohen (1988) $d = .2$ als kleiner,

d = .5 als mittlerer und d = .8 als starker Effekt gedeutet. Darüber hinaus wird nach Bortz (2005) Cohens w = .1 als kleiner, w = .3 als mittlerer und w = .5 als großer Effekt bestimmt.

Die für die Berechnungen verwendete Stichprobe wird stets so gewählt, dass der für die jeweilige Fragestellung größtmögliche vollständige Datensatz genutzt wird. Die den Auswertungen zugrundeliegende Stichprobenzahl wird jeweils angegeben.
Die deskriptive Darstellung der Mittelwerte und Standardabweichungen erfolgt dabei für alle Variablen jeweils mit der verfügbaren Gesamtstichprobe. Für die kurzfristigen Akzeptanz- sowie alle Lernwirksamkeitsvergleiche zwischen den verschiedenen Lehrmethoden wurden allerdings die Erhebungswochen 5 und 6 zu den Themen „Konditionierung“ und „Lernen am Modell“ außen vor gelassen. In diesen Wochen konnte aufgrund des anhaltenden Studierendenstreiks keine Vorlesungsevaluation stattfinden und es ist davon auszugehen, dass es sich auch bei den Teilnehmern an der E-Lecture- und WBT-Evaluation in Woche 5 und 6 um eine ausgewählte Stichprobe handelt. Die langfristigen und abschließenden Akzeptanzvergleiche wurden hingegen mit der Gesamtstichprobe durchgeführt, da hier nicht mehr spezifisch nach einer bestimmten Lerneinheit gefragt wurde, sondern die Lehrmethoden insgesamt zu beurteilen waren. Da alle in die Auswertungen eingehenden Studierenden alle Lehrmethoden zu diesem Zeitpunkt kennengelernt hatten, ist davon auszugehen, dass Ihr Akzeptanzurteil relativ unabhängig von dem konkreten Thema ist, welches sie während der kurzfristigen Evaluation evaluiert hatten. Weitere ggf. notwendige Stichprobenkorrekturen werden an der jeweiligen Stelle angegeben.

Im Folgenden werden zunächst die Ergebnisse zu den kurzfristig (ein Tag nach der Nutzung der Anwendungen/Besuch der Vorlesung), längerfristig (am Ende des Wintersemesters) und langfristig (vor der MAP im Sommersemester) erhobenen Lernwirksamkeits-erhebungen dargestellt. Es erfolgt eine Zusammenfassung und Diskussion der Lernwirksamkeitsergebnisse.
Im zweiten Abschnitt werden die Akzeptanzeinschätzungen der Studierenden berichtet. Dabei werden vergleichbar zur Lernwirksamkeit die kurzfristigen, längerfristigen und abschließenden Einschätzungen betrachtet. In einem weiteren Unterkapitel wird die Verhaltensakzeptanz untersucht – die freiwillige Nutzung der E-Learning-Anwendungen im Laufe des Sommersemesters zur Vorbereitung auf die MAP. Erneut erfolgt eine Zusammenfassung und Diskussion aller Ergebnisse zur Akzeptanz der Lehrmethoden.
Im dritten Abschnitt wird über das Nutzerverhalten der Studierenden berichtet. Hier wird insbesondere die Vor- und Nachbereitungszeit, die Pausenzeiten sowie das Anfertigen von Notizen betrachtet.

Im vierten und letzten Abschnitt des Ergebnisteils werden noch einmal gesondert die zur Verfügung gestellten Übungsaufgaben in den Vordergrund gestellt. Dabei werden zunächst die Einschätzungen der Studierenden zur Qualität der Übungsaufgaben betrachtet. Danach werden die Web-based Trainings mit und ohne Übungsaufgaben hinsichtlich ihrer Akzeptanz und Lernwirksamkeit verglichen sowie die Ergebnisse zusammengefasst und diskutiert.

3.1 Lernwirksamkeit

Zunächst werden die kurz-, längerfristigen und langfristigen Ergebnisse zur Lerneffektivität und -effizienz der drei Lehrmethoden Vorlesung, E-Lecture und Web-based Training berichtet und diskutiert. Aufgrund dessen, dass die Bearbeitungszeit der E-Lectures unzuverlässig erfasst wurde (siehe Abschnitt „2.3.2 Erhobene Variablen“ – Unterkapitel „User Tracking“), kann für die kurz- und längerfristige Evaluation leider keine Auswertung der Lerneffizienz der E-Lectures erfolgen. Es wird deshalb lediglich ein Vergleich der Lernzeit und Lerneffizienz der WBT- und Vorlesungsgruppen angeführt. Die Dauer der Vorlesung wurde dabei anhand der Rohlaufzeit der E-Lecture geschätzt. Dies stellt eine eher strenge Schätzung dar, da die Rohlaufzeit der E-Lecture lediglich die Zeit abdeckt, innerhalb der konkret Inhalte vermittelt wurden. Die Zeit für z.B. organisatorische Aspekte wird dadurch ausgeklammert.
Da die Studierenden im Laufe des Sommersemesters auch andere Materialien nutzten als nur die E-Lectures und Web-based Trainings, ist eine Auswertung der Lerneffizienz hier nicht mehr sinnvoll. Auch die Ergebnisse zur Lerneffektivität müssen vor diesem Hintergrund mit Vorsicht interpretiert werden.

3.1.1 Kurzfristige Lerneffektivität und -effizienz

Um die kurzfristige Lernwirksamkeit zu erfassen, wurden am Tag nach der Bearbeitung der E-Learning-Anwendungen bzw. dem Besuch der Vorlesung kurze Übungstests durchgeführt, bei denen jeweils maximal zehn Punkte erreicht werden konnten. Da im Nachhinein bei einzelnen Tests Fragen nicht in die Auswertung einbezogen werden konnten (siehe „2.3 Material“ bzw. Anhang A), werden im Folgenden die Lernleistungen prozentual korrigiert angegeben, um eine Vergleichbarkeit zwischen den Tests zu gewährleisten.

3.1.1.1 Lerneffektivität

Hypothese 1a: Kurzfristig weisen die E-Lectures eine höhere Lerneffektivität auf als die Vorlesung.
Hypothese 1b: Kurzfristig weisen die WBTs eine höhere Lerneffektivität auf als die Vorlesung.
Hypothese 1c: Kurzfristig weisen die WBTs eine höhere Lerneffektivität auf als die E-Lectures.

Tabelle 14 zeigt über alle Erhebungswochen hinweg die deskriptiven Ergebnisse zur kurzfristigen Lernleistung. Augenscheinlich haben die Studierenden, die die WBTs bearbeitet haben, besser abgeschnitten als die Studierenden, die die E-Lecture sahen. Diese wiederum schnitten besser ab als diejenigen, die die Vorlesung besucht hatten.

Tabelle 14
Deskriptive Statistik zur kurzfristigen Lernleistung der Vorlesungs-, E-Lecture-, und WBT-Gruppe

Vorlesung			E-Lecture			WBT		
n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
988	53.95	24.02	1077	60.16	24.94	1011	**66.42**	22.50

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Um die Unterschiede zwischen den drei Gruppen hinsichtlich der kurzfristigen Lernleistung auf ihre Bedeutsamkeit hin zu überprüfen, wurde eine abhängige Varianzanalyse berechnet. Die Ergebnisse hierzu sind in Tabelle 15 abgetragen.

Tabelle 15
Ergebnisse der Mittelwertsanalysen zum Vergleich der kurzfristigen Lernleistung der Vorlesungs-, E-Lecture-, und WBT-Gruppe

n	Vorlesung		E-Lecture		WBT		ANOVA				Posthoc
	M	***SD***	***M***	***SD***	***M***	***SD***	***F***	***df***	***p***	***Eta²***	
551	55.78	23.79	59.37	23.16	**64.42**	22.33	25.51	1.99	<.01*	.04	WBT>EL>VL

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke

Über alle zehn Erhebungswochen hinweg (ohne die Wochen 5 und 6, in denen keine Vorlesung stattfand) schneiden die Studierenden, die am Vortag das WBT bearbeitet haben, in Bezug auf die Lernleistung signifikant besser ab als die Studierenden, die mit der E-Lecture gelernt haben. Beide E-Learning-Gruppen schneiden zudem insgesamt besser ab als die Studierenden, die die Vorlesung besucht haben. Die Effektstärke dieses Unterschiedes ist mit .04 als klein bis mittelgroß einzustufen. Die Hypothesen 1a bis 1c können demnach beibehalten werden.
Die getrennte Auswertung der zwölf Projektwochen (siehe Anhang B) unterstützt dieses Ergebnis – wenn auch nicht ganz so deutlich. Augenscheinlich erreichten die Studierenden nach Bearbeitung der WBTs in elf von zwölf Wochen die beste Testleistung. Signifikante Unterschiede ergeben sich allerdings nur in vier der zwölf Wochen. Diese fallen jeweils zugunsten des jeweiligen WBTs aus und sind begleitet von mittelgroßen Effektstärken.

3.1.1.2 Lerneffizienz

Hypothese 1d: Kurzfristig erreichen die WBTs eine höhere Lerneffizienz als die Vorlesung.

Wie oben bereits erwähnt, konnte im Rahmen der E-Lecture-Nutzung leider nicht auf verlässliche Daten zurückgegriffen werden. Aus diesem Grunde kann die Lerneffizienz lediglich für die WBTs und die Vorlesung verglichen werden. Für die Betrachtung der Lerneffizienz können zudem lediglich die Themen 7 bis 12 herangezogen werden, da die Erfassung der Lernzeit der WBT-Nutzer in den ersten vier Wochen aus technischen Gründen fehlschlug und zu den Wochen 5 und 6 aufgrund des Streiks keine Vorlesung – und damit auch keine Vorlesungsevaluation – stattfand.
Die Lerneffizienz wurde mittels der Formel „(Lernleistung * 1000) / Lernzeit in Sekunden" berechnet. Die Multiplikation mit 1000 erfolgte zur Verbesserung der Lesbarkeit der Ergebnisse. Je höher das Ergebnis, desto besser die Lerneffizienz der Lehrmethode. Die Lernzeit an sich wird zusätzlich getrennt von der Lerneffizienz ausgewertet, um eine bessere Interpretation der Lerneffizienz zu ermöglichen.

An dieser Stelle sollen zunächst erneut die Lernwirksamkeitsergebnisse für die betrachteten sechs Themen dargestellt werden, um die danach folgenden Lernzeit- und Lerneffizienzvergleiche zwischen den drei Lehrmethoden besser einordnen zu können. Dabei werden hier die Ergebnisse über alle sechs Themen insgesamt dargestellt, die getrennten Auswertungen zu den einzelnen Themen können in Anhang B eingesehen werden.

Tabelle 16
Deskriptive Statistik zur kurzfristigen Lernleistung, Lernzeit sowie der aus diesen Variablen berechneten kurzfristigen Lerneffizienz der Vorlesungs- und WBT-Gruppe für die Themen 7 bis 12

	Vorlesung			**WBT**		
	n	***M***	***SD***	***n***	***M***	***SD***
Lernleistung (in %)	591	52.20	24.96	433	**61.60**	23.49
Lernzeit (in Sek.)	601	4170.55	1631.95	387	**4644.12**	3870.96
Lerneffizienz	591	16.50	14.96	382	**25.16**	26.99

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Tabelle 16 gibt die deskriptiven Ergebnisse zur kurzfristigen Lernwirksamkeit, Lernzeit sowie Lerneffizienz der Vorlesung und der WBTs über die sechs betrachteten Themen wieder. Augenscheinlich erreichen die WBT-Nutzer auch für diese Themen die beste Testleistung. Dafür ist die Lernzeit der WBT-Nutzer auf den ersten Blick im Schnitt etwas länger. Die Lerneffizienz scheint dennoch für die WBT-Nutzer höher zu sein als für die Vorlesungsbesucher.

Um diese augenscheinlichen Unterschiede auf Signifikanz zu überprüfen, wurden für alle drei Variablen abhängige *t*-Tests berechnet. Die Ergebnisse dieser Analysen sind in Tabelle 17 abgetragen.

Tabelle 17
Ergebnisse der Mittelwertsanalysen zum Vergleich der kurzfristigen Lernleistung, Lernzeit sowie der hieraus abgeleiteten kurzfristigen Lerneffizienz der Vorlesungs- und WBT-Gruppe für die Themen 7 bis 12

	n	Vorlesung		WBT		*t*-Test			
		M	*SD*	*M*	*SD*	*t*	*df*	*p*	*d*
Lernleistung	424	52.61	24.68	**61.75**	23.29	-6.32	423	<.01*	.31
Lernzeit	383	3938.20	1577.92	**4656.78**	3869.95	-3.41	382	<.01*	.17
Lerneffizienz	373	17.74	16.07	**25.14**	27.13	-4.63	372	<.01*	.24

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *t* = Teststatistik des abhängigen *t*-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *d* = Effektstärke

Es zeigt sich erneut, dass die WBT-Nutzer auch für die betrachteten sechs Themen eine signifikant bessere Lernleistung erzielen als die Vorlesungsbesucher. Die Effektstärke dieses Unterschiedes ist allerdings als klein einzustufen und bestätigt somit das Ergebnis für die Gesamtprojektlaufzeit.

Für die Bearbeitung der WBTs benötigen die Studierenden bedeutsam mehr Lernzeit als für den Besuch der Vorlesung. Dieser Unterschied zwischen den beiden Lehrmethoden ist jedoch ebenfalls begleitet von einer geringen Effektstärke. In der Folge erreicht das WBT (trotz der strengen Zeitschätzung für die Vorlesung durch die Rohlaufzeit der E-Lecture!) eine höhere Lerneffizienz als die Vorlesung. Hierbei handelt es sich ebenfalls um einen kleinen Effekt.

Hypothese 1d ist folglich nach diesen Analysen über alle sechs Themen beizubehalten. Dieses Ergebnis zeigt sich auch in den für die Themen 7 bis 12 getrennten Auswertungen (siehe Anhang B).

3.1.2 Längerfristige Lerneffektivität und -effizienz

Um die längerfristige Lernwirksamkeit zu erfassen, wurden in der Probeklausur am Ende des Wintersemesters zu jedem Thema vier Fragen gestellt. Die Studierenden konnten folglich in jedem Bereich maximal vier Punkte erreichen. Betrachtet wird nun zum Beispiel, inwieweit die Studierenden, die mit dem WBT zum Thema 1, der Vorlesung zum Thema 1 oder der E-Lecture zum Thema 1 gelernt hatten, unterschiedlich bezüglich des Fragenkomplexes zum Thema 1 abgeschnitten haben. Da in die Auswertungen lediglich die Studierenden eingehen können, die während des Semesters an der jeweiligen Evaluation und am Ende des Semesters an der Probeklausur teilgenommen haben, resultieren zum Teil vergleichbar kleine Stichproben.

Analog zur kurzfristigen Lernwirksamkeit wird dabei zunächst die Lerneffektivität über alle zwölf Projektthemen betrachtet, danach werden die Lerneffektivität, die Lernzeit und die Lerneffizienz für die Themen 7 bis 12 für den Vergleich zwischen WBT und Vorlesung berichtet.

3.1.2.1 Lerneffektivität

Hypothese 2a: Längerfristig weisen die E-Lectures eine höhere Lerneffektivität auf als die Vorlesung.
Hypothese 2b: Längerfristig weisen die WBTs eine höhere Lerneffektivität auf als die Vorlesung.
Hypothese 2a: Längerfristig weisen die WBTs eine höhere Lerneffektivität auf als die E-Lectures.

Tabelle 18 gibt die durchschnittlichen längerfristigen Testleistungen getrennt für die Lehrmethoden Vorlesung, E-Lecture und WBT über alle Themen hinweg wieder. Auf den ersten Blick scheinen die drei Projektgruppen vergleichbar gut abgeschnitten zu haben.

Tabelle 18
Deskriptive Statistik zur längerfristigen Lernleistung der Vorlesungs-, E-Lecture- und WBT-Gruppe

Vorlesung			E-Lecture			WBT		
n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
291	1.96	1.18	285	1.99	1.21	292	**2.05**	1.29

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Dies wird auch durch eine einfaktorielle Varianzanalyse mit Greenhouse-Geisser-Korrektur deutlich (siehe Tabelle 19). Längerfristig unterscheiden sich die drei Projektgruppen folglich nicht mehr hinsichtlich ihrer Lernergebnisse. Demnach sind die Hypothesen 2a bis 2c zu verwerfen.

Tabelle 19
Ergebnisse der Mittelwertsanalysen zum Vergleich der längerfristigen Lernleistung der Vorlesungs-, E-Lecture- und WBT-Gruppe

n	Vorlesung		E-Lecture		WBT		ANOVA			
	M	***SD***	***M***	***SD***	***M***	***SD***	***F***	***df***	***p***	***Eta²***
253	1.97	1.19	1.99	1.23	**2.08**	1.27	0.85	2.00	.43	-

Anmerkung. *n* = Stichprobengröße; *M* = Mittelwert; *SD* = Standardabweichung; *F* = Teststatistik der Varianzanalyse; *df* = Freiheitsgrade; *p* = Signifikanz; *Eta²* = Effektstärke

Dieses Ergebnis bestätigt sich auch bei der getrennten Analyse der zwölf Projektwochen (siehe Anhang B). Hier ergibt sich lediglich für das Thema 3 („Gedächtnis III“) ein signifikanter Unterschied zwischen den WBT- und den E-Lecture-Nutzern zugunsten des WBTs.

Zudem wurde dem Verdacht nachgegangen, ob die Studierenden themenspezifisch unterschiedlich gut in der Probeklausur abgeschnitten haben, da die längerfristige Lernwirksamkeitsbefragung unterschiedlich lange Abstände zur eigentlichen E-Learning-Nutzung bzw. dem Vorlesungsbesuch aufwies. Es ist anzunehmen, dass die Studierenden deshalb innerhalb der Probeklausur insbesondere die Fragen zu den ersten Themen schlechter und die Fragen zu den letzten Themen besser beantworten können. Die Ergebnisse (siehe Anhang B, Tabelle B8) zeigen allerdings, dass dies nicht der Fall ist. Eine größere Rolle als der Abstand zum Vorlesungsbesuch/zu der E-Learning-Nutzung scheinen andere Faktoren zu spielen – wie z.B. die Komplexität der Inhalte oder die Schwierigkeit der Testfragen.

3.1.2.2 Lerneffizienz

Hypothese 2d: Längerfristig erreichen die WBTs eine höhere Lerneffizienz als die Vorlesung.

Tabelle 20 gibt die durchschnittliche längerfristige Lernwirksamkeit, Lernzeit und Lerneffizienz über die drei Projektgruppen in den Wochen 7 bis 12 wieder. Bei der Lernzeit handelt es sich naturgemäß um die gleichen Daten wie bereits bei der kurzfristigen Lerneffizienz besprochen, da die gleiche Nutzung im Wintersemester zugrundegelegt wird. Diese Ergebnisse werden hier nur noch einmal übersichtshalber aufgeführt.

Tabelle 20
Deskriptive Statistik zur längerfristigen Lernleistung, Lernzeit sowie der längerfristigen Lerneffizienz der Vorlesungs- und WBT-Gruppe für die Themen 7 bis 12

	Vorlesung			WBT		
	n	***M***	***SD***	***n***	***M***	***SD***
Lernleistung	254	1.81	1.17	190	**1.98**	1.30
Lernzeit (in Sek.)	601	4170.55	1631.95	387	**4644.12**	3870.96
Lerneffizienz	254	0.55	0.49	173	**0.93**	1.30

Anmerkung. n = Stichprobengröße; M = Mittelwert; SD = Standardabweichung

In Tabelle 21 werden die Ergebnisse der t-Tests berichtet, mittels denen die beiden Projektgruppen insgesamt hinsichtlich der drei Variablen auf signifikante Unterschiede untersucht wurden.

Tabelle 21
Ergebnisse der Mittelwertsanalysen zum Vergleich der längerfristigen Lernleistung, Lernzeit sowie der längerfristigen Lerneffizienz der Vorlesungs- und WBT-Gruppe für die Themen 7 bis 12

	n	Vorlesung		WBT		***t*-Test**			
		M	***SD***	***M***	***SD***	***t***	***df***	***p***	***d***
Lernleistung	188	1.87	1.21	**2.00**	1.30	-1.30	187	.20	-
Lernzeit	383	3938.20	1577.92	**4656.78**	3869.95	-3.41	382	<.01*	.17
Lerneffizienz	171	0.62	0.53	**0.94**	1.31	-3.26	170	<.01*	.25

Anmerkung. n = Stichprobengröße, M = Mittelwert, SD = Standardabweichung, t = Teststatistik des abhängigen t-Tests, df = Freiheitsgrade, p = Signifikanz, d = Effektstärke

Erneut ergibt sich kein signifikanter Unterschied zwischen den Gruppen, betrachtet man die Lernleistung alleine. Wie oben bereits berichtet, unterscheiden sich die beiden Projektgruppen allerdings signifikant hinsichtlich der aufgewendeten Lernzeit. Die WBTs werden länger bearbeitet als der Vorlesungsbesuch dauert. Bezogen auf die Lerneffizienz ergibt sich dennoch ein vergleichbares Ergebnis wie bereits bei der kurzfristigen Lerneffizienz beobachtet: Über alle sechs Themen hinweg erreichen die WBTs eine höhere Lerneffizienz als die Vorlesung. Die Effektstärke dieses Unterschiedes ist allerdings als klein einzustufen. Hypothese 2d ist demnach beizubehalten.

Diese Ergebnisse bestätigen sich allerdings nicht ganz, betrachtet man die Themen getrennt voneinander (siehe Anhang B): Zwar ergibt sich auch bei der spezifischeren Auswertung für keine der sechs betrachteten Themen ein bedeutsamer Lernwirksamkeitsunterschied zwischen den beiden Gruppen und auch hinsichtlich der Lernzeit ergeben sich für drei der sechs Themen längere Zeiten für die WBT-Gruppe als für die Vorlesungsgruppe. Hinsichtlich der Lerneffizienz kann das WBT allerdings für nur zwei der sechs Themen ein bedeutsam besseres Ergebnis erreichen als die Vorlesung. Hierbei handelt es sich jeweils um mittelstarke bis starke Effekte.

3.1.3 Langfristige Lernwirksamkeit

Hypothese 10a: Studierende, die die E-Learning-Anwendungen genutzt haben, schneiden besser in der Probeklausur ab als Studierende, die die Anwendungen nicht genutzt haben.

Hypothese 10b: Studierende, die die E-Learning-Anwendungen genutzt haben, schneiden besser in der Abschlussklausur (MAP) ab als Studierende, die die Anwendungen nicht genutzt haben.

Für die Einschätzung der langfristigen Lernwirksamkeit am Ende des Sommersemesters wurde die Probeklausur vor der Modulabschlussprüfung (MAP) sowie das Ergebnis der MAP selbst herangezogen. In der Probeklausur konnten insgesamt 48 Punkte erreicht werden, in der Modulabschlussprüfung 34 Punkte.

Unabhängige Variable war hier nun nicht mehr die Nutzung der Anwendungen im Wintersemester, sondern die freiwillige Nutzung der Anwendungen für die Klausurvorbereitung im Sommersemester (Einteilung der Studierenden auf die Auswertungsgruppen: siehe Abschnitt „3.2.4 Verhaltensakzeptanz“). Verglichen werden folglich die „E-Lecture-Nutzer“, „WBT-Nutzer“, „Kombi“-Nutzer und die „Nicht“-Nutzer der E-Learning-Anwendungen. Aufgrund dieser Gruppeneinteilung resultieren für den Lernwirksamkeitsvergleich sehr unterschiedliche Stichprobengrößen, was bei der varianzanalytischen Auswertung zu beachten war (siehe Tabelle 22).

Abbildung 16 und Tabelle 22 zeigen das Abschneiden der vier Nutzergruppen in der Probeklausur sowie in der MAP.

Tabelle 22
Deskriptive Statistik zur langfristigen Lernleistung (Probeklausur vor der MAP sowie MAP selbst) der Nutzergruppen „Keine Nutzung", „E-Lecture-Nutzer", „WBT-Nutzer" sowie „Kombinierte Nutzung"

	Keine Nutzung			E-Lecture-Nutzer			WBT-Nutzer			Kombi-Nutzer		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Probeklausur	123	34.75	7.47	16	37.07	6.16	212	35.53	6.50	38	**37.26**	6.26
MAP	376	20.60	6.15	49	22.14	5.85	418	22.47	4.57	84	**23.71**	4.86

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Auf den ersten Blick scheinen die vier Gruppen hinsichtlich beider Leistungserhebungen vergleichbar gut abgeschnitten zu haben. Tendenziell erreichen die Kombi-Nutzer bezüglich beider Lernwirksamkeitsmaße ein leicht besseres Ergebnis.

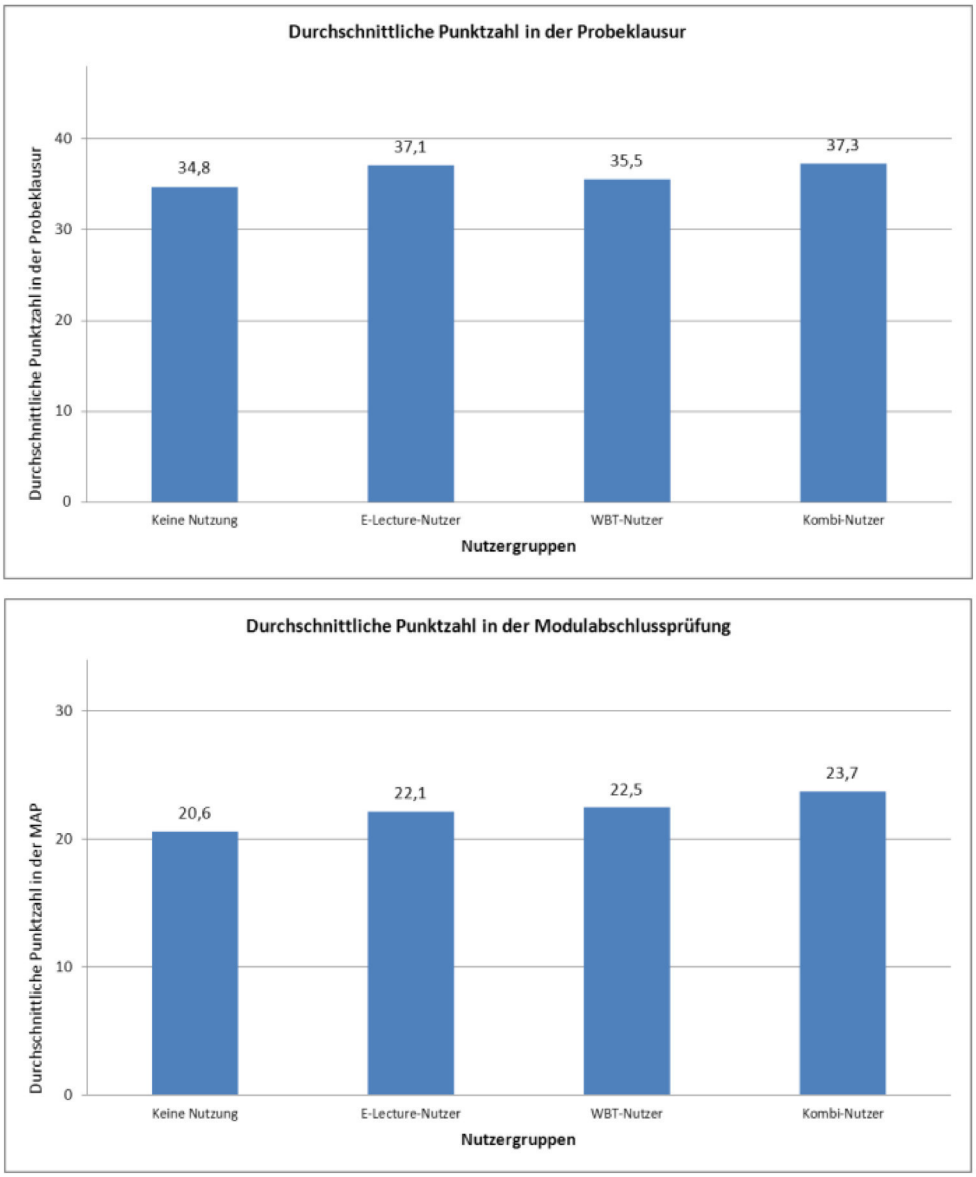

Abbildung 16. Durchschnittliche Punktzahl der vier Nutzergruppen „Keine Nutzung", „E-Lecture-Nutzung", „WBT-Nutzung" und „Kombinierte Nutzung" in der Probeklausur sowie der Modulabschlussprüfung

Um trotz der sehr unterschiedlichen Stichprobenanzahl die Leistungsunterschiede der Gruppen dennoch varianzanalytisch auf Unterschiede überprüfen zu können, wurden aus den größeren Gruppen „Keine Nutzer", „WBT-Nutzer" und „Kombinierte Nutzer" per Zufall mehrere Stichproben gezogen, deren Umfang für die Berechnungen angemessen sind (Maximales Verhältnis zur kleinsten Gruppe „E-Lecture-Nutzer" von 1 : 1.5). Um sicherzustellen, dass durch eine einmalige Zufallsauswahl kein falsches Ergebnis zustande kommt, wurde sowohl für den Lernwirksamkeitsvergleich bezüglich der Probeklausur als auch für den Vergleich des MAP-Ergebnisses dreimal eine Zufallsauswahl vorgenommen und mit dieser eine Varianzanalyse bzw. ein Brown-Forsythe-Test gerechnet. Die deskriptiven Ergebnisse dieser Analysen sind in Tabelle 23 abgetragen, die varianzanalytischen Ergebnisse befinden sich in Tabelle 24[31].

Tabelle 23
Deskriptive Statistik zur langfristigen Lernleistung (Probeklausur vor der MAP sowie MAP selbst) für drei verschiedene Zufallsstichproben aus den Nutzergruppen „Keine Nutzung", „E-Lecture-Nutzer", „WBT-Nutzer" sowie „Kombinierte Nutzung"

Probeklausur	Keine Nutzung			E-Lecture-Nutzer			WBT-Nutzer			Kombi-Nutzer		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Zufallsgruppen 1	24	36.24	5.95	16	**37.07**	6.16	24	37.03	7.62	19	36.48	7.19
Zufallsgruppen 2	24	33.67	9.55	16	**37.07**	6.16	23	36.97	5.49	19	36.48	7.19
Zufallsgruppen 3	25	33.90	8.20	16	37.07	6.16	23	35.60	6.77	19	**38.04**	5.24
MAP												
Zufallsgruppen 1	63	20.60	6.65	49	22.14	5.85	59	22.20	5.27	42	**23.45**	4.64
Zufallsgruppen 2	63	21.18	6.30	49	22.14	5.85	60	21.93	3.81	42	**23.98**	5.11
Zufallsgruppen 3	63	20.39	6.29	49	22.14	5.85	60	22.72	4.61	42	**23.98**	5.11

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Die mehrfachen Mittelwertsvergleiche zeigen ein ziemlich eindeutiges Bild: Hinsichtlich der Probeklausur ergibt sich in keiner der einfaktoriellen Varianzanalysen ein signifikantes Ergebnis. Alle vier Nutzergruppen schneiden folglich vergleichbar gut in der Probeklausur ab.

Tabelle 24
Ergebnisse der Mittelwertsanalysen zum Vergleich der langfristigen Lernleistung (Probeklausur vor der MAP sowie MAP selbst) der Nutzergruppen „Keine Nutzung", „E-Lecture-Nutzer", „WBT-Nutzer" sowie „Kombinierte Nutzung" (jeweils drei Zufallsziehungen)

Probeklausur	ANOVA/ Brown-Forsythe-Test				Posthoc
	F	*df*	*p*	*Eta²*	
Zufallsgruppen 1	0.08	3	.97	-	-
Zufallsgruppen 2	1.07	3	.37	-	-
Zufallsgruppen 3	1.51	3	.22	-	-
MAP					
Zufallsgruppen 1	2.26	202.92,3	.08	-	-
Zufallsgruppen 2	2.36	185.25,3	.07	-	-
Zufallsgruppen 3	3.89	197.32,3	.01*	.05	Kombi > Keine

Anmerkung. *F* = Teststatistik der Varianzanalyse bzw. des Brown-Forsythe-Tests; *df* = Freiheitsgrade; *p* = Signifikanz; *Eta²* = Effektstärke

[31] Ein Vergleich der vier Nutzergruppen getrennt für die zwölf Vorlesungsthemen war leider aufgrund der zum Teil sehr kleinen Stichproben (minimal *n* = 2) nicht möglich.

Bezogen auf das Abschneiden in der Modulabschlussprüfung ergibt sich lediglich einmal eine Signifikanz – hier schneidet die Gruppe, die sowohl E-Lectures als auch Web-based Trainings für ihre Vorbereitung auf die MAP genutzt hat (Kombi-Nutzer), signifikant besser ab als die Gruppe derer, die keine der E-Learning-Anwendungen nutzte. Der Effekt dieses Unterschiedes ist mit Eta^2 = .05 als klein bis mittelgroß einzustufen.
Die Mehrheit der Analysen spricht folglich dafür, dass die Nutzergruppen nicht signifikant hinsichtlich ihrer Lernleistung differieren. Die Hypothesen 10a und 10b müssen folglich verworfen werden.

3.1.4 Zusammenfassung und Diskussion

3.1.4.1 Kurzfristige Lerneffektivität und -effizienz

Kurzfristig (ein Tag nach Bearbeitung der E-Learning-Anwendungen bzw. nach dem Besuch der Vorlesung) ergab sich insgesamt über alle betrachteten Themen für die WBT-Nutzer mit durchschnittlich M = 64.4% richtigen Antworten eine signifikant bessere Lernwirksamkeit als für die E-Lecture-Nutzer (M = 59.4%) und die Vorlesungsbesucher (M = 55.8%). Die Lernleistung der Studierenden liegt damit ähnlich hoch wie in der Studie von Glowalla (2004), welcher unmittelbar nach einer Vorlesung eine Behaltensleistung von 40 – 60% nachweisen konnte.
Dabei schnitten neben den WBT-Nutzern auch die E-Lecture-Nutzer signifikant besser ab als die Vorlesungsbesucher. Betrachtet man die zwölf Themen getrennt, erreichten die WBT-Nutzer immerhin für vier Themen signifikant bessere Ergebnisse als die Vorlesungsbesucher, für drei Themen waren sie zudem besser als die E-Lecture-Nutzer. Zwischen den E-Lecture-Nutzern und den Vorlesungsbesuchern ergaben sich hier keine Signifikanzen. Dies bestätigt die Ergebnisse verschiedener Studien, wie z. B. Kulik und Kulik (1991), Maki et al. (2000), Sanders und Morrison-Shetlar (2001) oder Schutte (1997), die ebenfalls eine bessere Lernleistung nach E-Learning-Nutzung belegen und widerspricht Studien, die keine Unterschiede nachweisen konnten (z.B. Carey, 2001; Demetriadis & Pombortsis, 2007; Görlich, 2006; Hasebrook, 2002; Lyon Jr. et al., 1992; Spickard III et al., 2002; Vogt et al., 2006 oder Wofford et al., 2001 etc.).

Des Weiteren sollte ebenfalls die Lerneffizienz der Lehrmethoden betrachtet werden. In diesem Rahmen wurde auch die Lernzeit der Studierenden verglichen, um die Interpretation der Lerneffizienz zu unterstützen.
Für die Vorlesungsbesucher wurde die Rohlaufzeit der E-Lecture als Lernzeit zugrunde gelegt, da eine Angabe der regulären Vorlesungszeit von 90 Minuten die tatsächliche Zeit, welche für die Vermittlung der Inhalte genutzt wurde, häufig überschätzt. Hier wäre es

sicher sinnvoller gewesen, die tatsächliche Zeit exakt zu stoppen. Der Rückgriff auf die E-Lecture-Rohzeiten blendet jedoch zumindest die Zeit, die für Organisatorisches oder ähnliches aufgewandt wurde, automatisch aus und berücksichtigt so lediglich die Zeit für die Vermittlung der Inhalte. Dies erhöht die Vergleichbarkeit, da die Vorträge für die Erstellung der E-Lecture auf diese Inhalte zusammengeschnitten worden waren und dies auch die inhaltliche Grundlage für die WBTs darstellte. Da die E-Lecture zudem in den allermeisten Fällen vom gleichen Vortragenden[32] gehalten wurden, kann die Rohlaufzeit der E-Lecture als Schätzung für die Vorlesungszeit als relativ verlässlich eingeschätzt werden. Es handelt sich hierbei allerdings um eine sehr strenge Schätzung der Lernzeit für die Vorlesung.

Betrachtet man die Lernzeit, die die WBT-Nutzer und die Vorlesungsbesucher mit den Lerninhalten verbracht haben, zeigt sich, dass die WBT-Nutzer signifikant mehr Zeit investierten als die Vorlesungsbesucher. Dennoch erreichten die WBTe eine signifikant höhere Lerneffizienz als die Vorlesung. Dies zeigt sich auch im Einzelvergleich der sechs Themen: für fünf der sechs Themen ergibt sich eine signifikant bessere Lerneffizienz der WBTs.

Insgesamt kann also von einem klaren Vorteil der WBTs im Vergleich zur Vorlesung gesprochen werden – und das trotz der strengen Lernzeitschätzung für die Vorlesung. Der Vorteil des WBTs im Vergleich zur Vorlesung ergibt sich dabei klar aus der – wenn auch nicht immer signifikant – besseren Lernwirksamkeit, da die Lernzeit der WBT-Gruppen entweder genauso lang oder sogar weit länger war als die der Vorlesungsgruppen.

Dieser Vorteil von E-Learning-Anwendungen gegenüber der Präsenzlehre bestätigt die Ergebnisse einiger Studien, die ebenfalls Vorteile hinsichtlich der Lerneffizienz von E-Learning-Anwendungen belegen konnten (siehe z.B. Döring, 2002; Hasebrook, 2002; Lyon Jr. et al., 1992 oder Vogt et al., 2006). Die meisten dieser Studien führen dies jedoch klar darauf zurück, dass sich Lernende die Lernerfolge in kürzerer Zeit erarbeiten können. Im vorliegenden Fall scheint sich eher ein tendenzieller Lernwirksamkeitsvorteil durchzusetzen.

3.1.4.2 Längerfristige Lerneffektivität und -effizienz

Am Ende des Wintersemesters, nachdem die Studierenden alle drei Lehrmethoden genutzt und evaluiert hatten, wurde zu allen Themen der Vorlesung eine Probeklausur zur Verfügung gestellt, mittels der die Studierenden ihr Wissen testen konnten und zudem die

[32] Hinsichtlich der Themen, die für die E-Lecture- oder Vorlesungsgruppe von einem anderen Dozenten vertreten werden mussten (Themen 3 (Vorlesung), Themen 4 und 11 (E-Lecture)), erzielten die Lehranwendungen in keiner der Lernwirksamkeitsauswertungen auffällig abweichende Ergebnisse (siehe Anhang B), weshalb diese Variation vernachlässigt werden kann.

längerfristige Lernwirksamkeit erhoben werden konnte. Für die Lerneffektivität werden hier erneut alle zwölf Themen betrachtet, die Lernzeit und Lerneffizienz konnte wieder lediglich für die Themen 7 bis 12 und nur für WBT und Vorlesung analysiert werden.

Im Gegensatz zur kurzfristigen Lerneffektivität ergaben sich über alle Themen hinweg zwischen den drei Lerngruppen hinsichtlich der längerfristigen Lernleistung nun keine bedeutsamen Unterschiede mehr. Dies bestätigt Studien, die ebenfalls keine Unterschiede hinsichtlich der Lernwirksamkeit von E-Learning-Anwendungen nachweisen konnten (z.B. Carey, 2001; Demetriadis & Pombortsis, 2007; Görlich, 2006; Hasebrook, 2002; Lyon Jr. et al., 1992; Spickard III et al., 2002 oder Wofford et al., 2001). Insbesondere bestätigt wird die Studie von Vogt et al. (2006), bei welcher ebenfalls lediglich kurz-, jedoch nicht mehr längerfristig Lerneffektivitätsvorteile zugunsten des WBTs (hier aber im Vergleich zur E-Lecture) festgestellt werden konnten.
Unabhängig von der Lehrmethode erreichten die Studierenden allerdings etwa 50% der möglichen Punktzahl. Und auch getrennt für die zwölf Themen ergab sich nun lediglich einmal ein signifikanter Vorteil für die WBT-Gruppe im Vergleich zur E-Lecture-Gruppe. Während die Lehrmethode folglich einen kurzfristigen Effekt auf die Lerneffektivität hatte, trat dieser längerfristig nicht mehr auf.
Die Lernwirksamkeit aller drei Lehrmethoden ist insgesamt als positiv zu beurteilen. Die längerfristige Lernwirksamkeit wurde in der vorliegenden Studie zwischen einer Woche (Thema 12) und vier Monaten (Thema 1) nach Darbietung der Lerninhalte abgefragt. Dennoch resultiert eine so gute Lernwirksamkeit über alle drei Lehrmethoden – und das, obwohl die Klausur erst ein halbes Jahr später anstand. Dies zeigt, dass alle drei Lehrmethoden die Inhalte sehr lernförderlich vermitteln konnten.
An dieser Stelle ist allerdings zu diskutieren, dass die Probeklausur zur Hälfte aus Fragen bestand, die bereits bei der kurzfristigen Lernwirksamkeitserhebung zum Einsatz kamen. An diese Fragen könnten sich die Studierenden erinnert haben. Da die Studierenden jedoch nur zu drei der insgesamt zwölf Themenbereiche an einer kurzfristigen Lernwirksamkeitserhebung teilgenommen hatten, hatten sie lediglich sechs der insgesamt 48 Fragen schon einmal kennengelernt. Der Einfluss hieraus sollte folglich nicht zu einer bedeutsamen Verzerrung des Lernleistungsergebnisses geführt haben. Stuft man die wirkliche Lernleistung der Studierenden jedoch sicherheitshalber etwas geringer ein als die Probeklausur ergeben hat, ist sie dennoch im Vergleich zu anderen Studien als sehr erfreulich zu beurteilen.
Es ist erstaunlich, dass sich hierbei keine Unterschiede zwischen der Lernleistung zu den Themen in der Hinsicht zeigten, dass die ersten Themen wesentlich schlechter erinnert wurden als die letzten (siehe Abbildung B8 in Anhang B). Im Gegensatz dazu zeigte sich im

Vergleich zwischen den Themen, dass die Studierenden insbesondere mit den Fragen zum siebten Thema („Denken, Problemlösen und Intelligenz“) Schwierigkeiten hatten. Dieses Ergebnis ist nicht verwunderlich, da dieses Thema bei Weitem am umfangreichsten und komplexesten ist. Im Vergleich hierzu konnten die Studierenden die Fragen zu Thema 2 („Gedächtnis II“) über alle drei Projektgruppen hinweg am besten beantworten.

Das Ergebnis des längerfristigen Lernzeitvergleiches entspricht dem bereits oben im Rahmen der kurzfristigen Lerneffizienz erläuterten, da die gleiche Nutzung hierfür zugrunde gelegt wird. Die Studierenden lernen folglich mit den WBTs länger als sie die Vorlesung besuchen.
Auch für die Lerneffizienz ergibt sich ein vergleichbares Ergebnis: Das WBT zeichnet sich durch eine bedeutsam höhere Lerneffizienz aus als die Vorlesung. Dieser Unterschied ist vermutlich auf die – wenn auch nicht signifikant – etwas bessere Lernleistung der WBT-Gruppen zurückzuführen, da die Lernzeit der WBT-Gruppen signifikant länger ausfällt als die der Vorlesungsgruppen.

3.1.4.3 Vergleich der kurz- und längerfristigen Lerneffektivität und -effizienz

Die Lerneffektivität der drei Lehrmethoden verändert sich von der kurz- zur längerfristigen Messung. Kurzfristig schneiden die WBT-Nutzer deutlich besser ab als die E-Lecture-Nutzer und Vorlesungsbesucher und auch die E-Lecture-Nutzer erreichen ein besseres kurzfristiges Lernergebnis als die Vorlesungsbesucher. Längerfristig jedoch ergeben sich keine signifikanten Unterschiede zwischen den drei Projektgruppen. Der Vorteil der E-Learning-Anwendungen hinsichtlich der Lernwirksamkeit scheint folglich nur kurzfristig vorhanden zu sein.
Dabei nutzen die Studierenden aus der WBT-Gruppe das Lehrmaterial deutlich länger als die Studierenden aus der Vorlesungsgruppe.
Die Lerneffizienz ist in der Folge allerdings unverändert: Sowohl kurz- als auch längerfristig stellt sich das WBT als lerneffizienter heraus als die Vorlesung. Hinsichtlich der Lerneffizienz bleibt der Vorteil der WBTs also – trotz längerer Bearbeitungszeit – auch längerfristig bestehen. Das bessere Abschneiden des WBTs hinsichtlich der Lerneffizienz ist folglich mit dem – wenn auch nicht signifikantem, doch etwas besserem – Abschneiden der WBT-Gruppe in den Wissensüberprüfungen zu begründen.

3.1.4.4 Langfristige Lerneffektivität

Die langfristige Lernwirksamkeit betrifft andere Experimentalgruppen als die bisher betrachteten. Hier ging es nicht mehr um die Nutzung der E-Learning-Anwendungen und den Besuch der Vorlesung innerhalb des Wintersemesters, sondern um die freiwillige

Nutzung der E-Learning-Anwendungen im Rahmen der Klausurvorbereitung im Sommersemester. Deshalb wurden hier diejenigen Studierenden hinsichtlich ihrer Leistung in einer erneuten Probeklausur und der Modulabschlussprüfung verglichen, die keine der E-Learning-Anwendungen, nur die WBTs, nur die E-Lectures oder beide E-Learning-Anwendungen („Kombi-Nutzer“) zur Klausurvorbereitung genutzt hatten. Da die Studierenden zur Vorbereitung auf die Klausur sicherlich auch noch andere Materialien nutzten, verbot sich hier eine Auswertung der Lerneffizienz.

Hinsichtlich des Abschneidens in der Probeklausur ergaben sich keine Unterschiede zwischen den vier betrachteten Gruppen, in der Modulabschlussprüfung kam es lediglich in einer der drei Analysen mit Zufallsstichprobe zu einer Signifikanz mit kleinem bis mittleren Effekt zwischen der Gruppe der „Nicht-Nutzer“ und den „Kombi-Nutzern“ zugunsten der kombinierten Nutzung der E-Learning-Anwendungen. Die Mehrheit der Analysen spricht dementsprechend ganz klar gegen einen Lernleistungsunterschied zwischen den vier Gruppen.

Es ergab sich folglich kein Unterschied daraus, womit die Studierenden gelernt hatten – ob mit den E-Learning-Anwendungen, nur mit dem Vorlesungsskript oder doch ggf. mit der angegebenen Literatur: Im Schnitt erreichten die Studierenden in der Probeklausur und der MAP ein vergleichbar gutes Ergebnis. Dies zeigt, dass Lernende gegebene Lernmaterialien im Schnitt so für sich nutzen, dass sie das Lernziel erreichen können. Ggf. „schlechtes Material“ wird von guten Lernenden z.B. durch Ergänzung mit anderen Materialien oder durch Fleiß kompensiert (Glowalla et al., 2002). Genauso wird gutes Lernmaterial in der Form genutzt, dass auch „nur“ das (selbst gesteckte) Lernziel erreicht wird – z.B. wird es evtl. weniger intensiv bearbeitet.

3.2 Akzeptanz

3.2.1 Kurzfristige Einschätzungen

Die kurzfristigen Einschätzungen zur Akzeptanz wurden einen Tag nach der Nutzung der E-Learning-Anwendungen bzw. nach dem Besuch der Vorlesung erhoben. Wenn die Studierenden alle drei Lehrmethoden evaluiert hatten, wurde in einer weiteren kurzen Befragung noch einmal vergleichend nach ihrem Urteil zu den Methoden/Materialien Vorlesung, E-Lecture, WBT, Übungsaufgaben und Literatursammlung gefragt.

Der folgende Abschnitt beschäftigt sich zunächst mit den direkt am darauffolgenden Tag erhobenen Aspekten „Lernförderlichkeit“, „Akzeptanzskalen (Flexibilität, Qualität der Lernumgebung sowie Gesamtakzeptanz)“ und „Gesamtnote“. Zum Schluss werden die Variablen berichtet, die am Ende der kurzfristigen Erhebung die Studierenden veranlassen

sollten, vergleichende Urteile zu den Lehrmethoden Vorlesung, E-Lecture, WBT, Übungsaufgaben und Literatur zu fällen.

3.2.1.1 Lernförderlichkeit

Hypothese 4a: Studierende schätzen die Lernförderlichkeit der E-Lecture höher ein als die Lernförderlichkeit der Vorlesung.
Hypothese 4b: Studierende schätzen die Lernförderlichkeit des WBTs höher ein als die Lernförderlichkeit der Vorlesung.
Hypothese 4c: Studierende schätzen die Lernförderlichkeit des WBTs höher ein als die Lernförderlichkeit der E-Lecture.

Tabelle 25 zeigt die Einschätzung der Lernförderlichkeit zu den drei Lehrmethoden E-Lecture, WBT und Vorlesung durch die Gesamtstichprobe der Studierenden. Die Skala weist einen Wertebereich von 4 bis 16 auf. Mit einem Mittelwert von etwa 11 Punkten bescheinigen die Studierenden allen drei Lehrmethoden folglich eine relativ gute Lernförderlichkeit.

Tabelle 25
Deskriptive Statistik zur Lernförderlichkeit der Lehrmethoden Vorlesung, E-Lecture und WBT

Vorlesung			E-Lecture			WBT		
n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
979	10.80	2.42	1027	11.33	2.56	993	**11.50**	2.55

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Betrachtet man für den Vergleich zwischen den drei Lehrmethoden lediglich die Urteile der Studierenden, die alle drei Lehrmethoden korrekt evaluiert haben, resultiert eine wesentlich geringere Stichprobe – die Mittelwerte und Standardabweichungen bleiben trotz der Stichprobenreduktion jedoch vergleichbar (siehe Tabelle 26).

Tabelle 26
Ergebnisse der Mittelwertsanalysen zum Vergleich der Lernförderlichkeitseinschätzung zur Vorlesung, E-Lecture und dem WBT

n	Vorlesung		E-Lecture		WBT		ANOVA				Posthoc
	M	***SD***	***M***	***SD***	***M***	***SD***	***F***	***df***	***p***	***Eta²***	
524	10.83	2.45	11.35	2.53	**11.49**	2.52	11.74	2.00	<.01*	.02	WBT, EL>VL

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade; *p* = Signifikanz, *Eta²* = Effektstärke

In einer abhängigen Varianzanalyse mit Greenhouse-Geisser-Korrektur ergibt sich ein signifikanter Unterschied zwischen den drei Lehrmethoden, welcher mit Eta^2 = .02 allerdings einen nur geringen Effekt aufweist. Posthoc-Vergleiche zeigen, dass sowohl das WBT als auch die E-Lecture hinsichtlich ihrer Lernförderlichkeit etwas besser eingeschätzt wurden als die Vorlesung. Die Hypothesen 4a und 4b können nach diesem Ergebnis beibehalten werden, Hypothese 4c muss verworfen werden.

3.2.1.2 Akzeptanzskalen

Hypothese 5a: Im kurzfristigen Vergleich wird die E-Lecture von den Studierenden besser akzeptiert als die Vorlesung.
Hypothese 5b: Im kurzfristigen Vergleich wird das WBT von den Studierenden besser akzeptiert als die Vorlesung.
Hypothese 5c: Im kurzfristigen Vergleich wird das WBT von den Studierenden besser akzeptiert als die E-Lecture.

Tabelle 27 zeigt die deskriptiven Ergebnisse zu den beiden Akzeptanzskalen „Qualität der Lernumgebung" (Wertebereich von 5 bis 20) und „Flexibilität" (Wertebereich von 4 bis 16) sowie zur Gesamtakzeptanz (Wertebereich von 9 bis 36), welche aus diesen beiden Unterskalen resultiert. Die Qualität der Lernumgebung wird mit einem Mittelwert zwischen 13.6 und 16.2 bei allen drei Lehrmethoden mittelmäßig bis gut eingeschätzt, die Flexibilität wird bei den E-Learning-Anwendungen gut, bei der Vorlesung eher mittelmäßig beurteilt. Daraus resultiert für die E-Learning-Anwendungen eine gute, für die Vorlesung eine eher mittelmäßige Gesamtakzeptanz.

Tabelle 27
Deskriptive Statistik zur Qualität der Lernumgebung, Flexibilität sowie Gesamtakzeptanz von Vorlesung, E-Lecture und WBT

	Vorlesung			E-Lecture			WBT		
	n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
Qualität der Lernumgebung	976	13.60	2.60	1022	**16.19**	3.02	994	15.78	3.33
Flexibilität	976	10.25	2.93	1025	13.92	2.53	995	**14.09**	2.28
Gesamt	971	23.85	4.58	1017	**30.13**	4.89	987	29.87	4.74

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Eine abhängige Varianzanalyse mit Greenhouse-Geisser-Korrektur (siehe Tabelle 28) ergibt für alle drei Akzeptanzskalen einen signifikanten Unterschied zwischen den Lehrmethoden mit einer großen Effektstärke. Hinsichtlich aller drei Variablen schneiden die E-Learning-Anwendungen signifikant besser ab als die Vorlesung. Bezogen auf die Qualität der Lernumgebung erreicht die E-Lecture darüber hinaus auch ein signifikant besseres Ergebnis als das WBT. Die Hypothesen 5a und 5b können damit beibehalten werden, Hypothese 5c muss verworfen werden.

Tabelle 28
Ergebnisse der Mittelwertsanalysen zum Vergleich der Qualität der Lernumgebung, Flexibilität sowie Gesamtakzeptanz von Vorlesung, E-Lecture und WBT

	n	Vorlesung		E-Lecture		WBT		ANOVA				Posthoc
		M	***SD***	***M***	***SD***	***M***	***SD***	***F***	***df***	***p***	***Eta²***	
Qualität	520	13.67	2.68	**16.23**	2.97	15.75	3.45	106.89	1.96	<.01	.17	EL>WBT>VL
Flexibilität	520	10.22	2.91	13.89	2.59	**14.10**	2.29	342.57	1.41	<.01	.40	EL,WBT >VL
Gesamt	512	23.89	4.65	**30.11**	4.91	29.88	4.86	264.70	1.75	<.01	.34	EL,WBT >VL

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke

3.2.1.3 Gesamtnote

In Tabelle 29 werden die Mittelwerte der von den Studierenden vergebenen Noten für die drei Lehrmethoden über alle zwölf Themen berichtet. Die Noten zu allen drei Lehrmethoden bewegen sich im guten Bereich.

Tabelle 29
Deskriptive Statistik zur Note für die Vorlesung, die E-Lecture und das WBT

Vorlesung			E-Lecture			WBT		
n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
983	2.42	0.80	1032	2.41	0.86	1008	**2.37**	0.91

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

In einer Varianzanalyse mit Greenhouse-Geisser-Korrektur mit der eingeschränkten Stichprobe derer, die alle drei Lehrmethoden benotet haben, ergibt sich kein signifikanter Unterschied zwischen den Lehrmethoden (siehe Tabelle 30). Alle drei Lehrmethoden erreichen folglich eine vergleichbare Gesamtbewertung der Studierenden.

Tabelle 30
Ergebnisse der Mittelwertsanalysen zum Vergleich der Note für die Vorlesung, die E-Lecture und das WBT

n	Vorlesung		E-Lecture		WBT		ANOVA			
	M	***SD***	***M***	***SD***	***M***	***SD***	***F***	***df***	***p***	***Eta²***
535	2.42	0.77	2.44	0.87	**2.38**	0.96	0.82	1.96	.44	-

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke

3.2.1.4 Vergleich der Lehrmaterialien/-methoden

Nachdem die Studierenden alle drei Lehrmethoden genutzt bzw. besucht und evaluiert hatten, sollten Sie in einer zusammenfassenden kurzen Befragung die Lehrmethoden/-materialien Vorlesung, E-Lecture, WBT, Übungsaufgaben und Literatursammlung hinsichtlich ihrer Eignung zur Wissensvermittlung und Klausurvorbereitung, ihrer motivierenden Wirkung und ihrem Spaßfaktor einschätzen. Zudem sollten sie angeben, inwieweit sie diese Methoden für ihre Vorbereitung auf die Modulabschlussprüfung nutzen möchten. In die Auswertung der Angaben gehen nur diejenigen Studierenden ein, die alle drei Lehrmethoden (Vorlesung, E-Lecture und WBT) vorab korrekt evaluiert hatten und die nicht an den Terminen teilgenommen hatten, an denen aufgrund des Streiks die Vorlesung ausgefallen war. Tabelle 31 zeigt die Mittelwerte und Standardabweichungen zu den oben genannten Variablen getrennt für die fünf Lehrmethoden/-materialien.

Tabelle 31
Deskriptive Statistik zu den in der Abschlussbefragung erfragten Variablen für die Vorlesung, die E-Lecture, das WBT, die Übungsaufgaben sowie die Literatur

		Vorlesung	E-Lecture	WBT	Übungs-aufgaben	Literatur
Bitte geben Sie an, wie geeignet Sie folgende Lehrmaterialien/-methoden zur **Wissensvermittlung** finden. (1 = nicht geeignet, 4 = geeignet)	*n*	537	534	535	536	530
	M	3.31	**3.34**	3.27	3.31	2.82
	SD	0.74	0.76	0.86	0.79	0.83
Bitte geben Sie an, wie geeignet Sie folgende Lehrmaterialien/-methoden zur **Klausurvorbereitung** finden. (1 = nicht geeignet, 4 = geeignet)	*n*	533	534	534	533	529
	M	2.83	3.39	3.58	**3.71**	3.20
	SD	0.86	0.80	0.72	0.57	0.79
Bitte geben Sie an, wie **motivierend** Sie folgende Lehrmaterialien/-methoden finden. (1 = nicht motivierend, 4 = motivierend)	*n*	533	535	535	536	528
	M	2.95	2.93	2.76	**3.13**	2.10
	SD	0.86	0.89	1.03	0.84	0.82
Bitte geben Sie an, wie viel **Spaß** Ihnen das Lernen mit folgenden Lehrmaterialien/-methoden gemacht hat. (1 = kein Spaß, 4 = viel Spaß)	*n*	538	537	537	538	534
	M	**3.17**	2.93	2.75	2.91	2.01
	SD	0.80	0.88	1.02	0.84	0.77
Welche der Lehrmaterialien/-methoden würden Sie gerne für Ihre **Vorbereitung auf die MAP** nutzen (wenn sie alle jederzeit verfügbar wären)?	*n*	570	570	570	570	570
	M	0.34	0.68	0.72	**0.79**	0.32
	SD	0.47	0.47	0.45	0.41	0.47

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Dem Augenschein nach werden zur Wissensvermittlung alle Lehrmaterialien/-methoden als besser geeignet beurteilt als die Literatursammlung – zur Klausurvorbereitung wird die Vorlesung als am geringsten geeignet eingestuft. Auch hinsichtlich der Motivation und dem Spaßfaktor schneiden alle Lehrmaterialien/-methoden besser ab als die Literatur.

Dies zeigt sich auch bei der Frage nach dem Wunsch, die Materialien/Methoden zur Klausurvorbereitung nutzen zu wollen. Sowohl die Vorlesung als auch die Literatur werden am wenigsten für die Klausurvorbereitung gewünscht. Die E-Lectures, WBTs und Übungsaufgaben scheinen relativ gleich auf zu sein.

Im Großen und Ganzen wird dies auch durch die abhängigen Varianzanalysen bzw. (im Falle des Nutzungswunsches) durch den Cochrans Q-Test[33] sowie die darauffolgenden Posthoc-Vergleiche bestätigt (siehe Tabelle 32). Es zeigte sich hier, dass alle Lehrmaterialien/-methoden von den Studierenden als geeigneter zur **Wissensvermittlung** beurteilt werden als die Literatursammlung. Die Effektstärke dieses Unterschiedes ist als mittelgroß einzustufen.

Bei allen weiteren gefundenen Mittelwertunterschieden handelt es sich durchweg um starke Effekte. Hinsichtlich der **Klausurvorbereitung** unterscheiden sich alle Lehrmaterialien/-methoden signifikant voneinander: Die Übungsaufgaben werden am sinnvollsten für die Klausurvorbereitung eingeschätzt, danach folgt das WBT, dann die E-Lecture, die Literatur und zu guter Letzt die Vorlesung.

[33] Dieser wurde genutzt, um die Gleichverteilung abhängiger dichotomer Variablen zu überprüfen, wie es bei der Variable „Nutzen" der Fall ist.

Tabelle 32
Ergebnisse der Mittelwertsanalysen zum Vergleich der in der Abschlussbefragung erfragten Variablen für die Vorlesung, die E-Lecture, das WBT, die Übungsaufgaben sowie die Literatur

	n		Vorlesung (VL)	E-Lecture (EL)	WBT	Übungs-aufg. (ÜA)	Literatur (LIT)	ANOVA bzw. Cochran-Test				Posthoc
								F/Q	*df*	*p*	*Eta²/ w*	
Wissensver-mittlung	524	*M*	3.31	**3.34**	3.27	3.32	2.82	39.96	3.75	<.01*	.07	LIT<VL,EL,WBT,ÜA
		SD	0.74	0.76	0.86	0.79	0.83					
Klausur-vorbereitung	525	*M*	2.83	3.39	3.59	**3.71**	3.20	118.76	3.57	<.01*	.19	VL<LIT<EL<WBT<ÜA
		SD	0.86	0.80	0.72	0.57	0.79					
Motivation	523	*M*	2.95	2.93	2.76	**3.14**	2.10	112.52	3.45	<.01*	.18	LIT<WBT,VL,EL<ÜA; WBT<VL
		SD	0.86	0.89	1.02	0.83	0.82					
Spaß	529	*M*	**3.17**	2.94	2.76	2.92	2.00	151.57	3.49	<.01*	.22	LIT<WBT<EL,ÜA<VL
		SD	0.80	0.87	1.02	0.83	0.77					
Nutzen	570	*M*	0.34	0.68	0.72	**0.79**	0.32	486.76	4	<.01*	.92	VL,LIT<EL,WBT<ÜA
		SD	0.47	0.47	0.45	0.41	0.47					

Anmerkung. n = Stichprobengröße, M = Mittelwert, SD = Standardabweichung, F = Teststatistik der Varianzanalyse, Q = Teststatistik des Cochran-Tests, df = Freiheitsgrade, p = Signifikanz, Eta^2 bzw. w = Effektstärke

Als am **motivierendsten** werden von den Studierenden die Übungsaufgaben angegeben. Danach folgen die E-Lecture, die Vorlesung und das WBT, wobei das WBT signifikant schlechter abschneidet als die Vorlesung. Zwischen E-Lecture und WBT ergibt sich keine Signifikanz. Die Literatursammlung wird signifikant als am wenigsten motivierend eingeschätzt.

Beim **Spaß** ergibt sich ein anderes Bild: Hier erreicht die Vorlesung die beste Einschätzung. Danach folgen Übungsaufgaben und die E-Lecture. Signifikant weniger Spaß machte den Studierenden die Arbeit mit dem WBT – am schlechtesten schneidet allerdings auch hier die Literatur ab.

Für ihre Klausurvorbereitung möchten die Studierenden am häufigsten die Übungsaufgaben **nutzen**. Danach folgen die WBTs und E-Lectures. Die Vorlesung und die Literatur werden am signifikant wenigsten für die Klausurvorbereitung gewünscht.

3.2.2 Längerfristige Einschätzungen im Rahmen der Probeklausur

Wie oben bereits erläutert, wurden für die Auswertung der Bewertungen im Rahmen der Probeklausur am Ende des Wintersemesters nur die Einschätzungen derjenigen Studierenden herangezogen, die neben der Probeklausur auch an allen Evaluationen während des Semesters vollständig teilgenommen hatten. Dies reduziert die Stichprobe auf 697 Teilnehmer. Durch Datenverluste aufgrund technischer Schwierigkeiten während der Erhebung der Probeklausur konnte allerdings nur auf ca. 350 vollständige Befragungen zurückgegriffen werden.

Im Folgenden soll erneut zuerst auf das grobe Akzeptanzmaß „Gesamtnote“ eingegangen werden, bevor die Ergebnisse zu weiteren Akzeptanzfragen erläutert werden. Zu guter Letzt erfolgt ein Blick auf die von den Studierenden angegebenen Nutzungsbereiche.

3.2.2.1 Gesamtnote (Vorlesung, E-Lecture, WBT, Übungsaufgaben, Skript, Literatur)

> *Hypothese 6:* Die E-Learning-Anwendungen (E-Lecture, Web-based Training, Übungsaufgaben) erreichen eine höhere Akzeptanz als die „traditionellen Lehrmethoden“ der Hochschule (Vorlesung, Vorlesungsskript, Literatur).

In der längerfristigen Befragung im Rahmen der Probeklausur sollten die Studierenden die Lehrmethoden/-materialien „Vorlesung“, „E-Lecture“, „WBT“, „Übungsaufgaben“, „Vorlesungsskript“ und „Literatur“ vergleichend beurteilen. Hypothese 6 wurde hier mittels des groben Akzeptanzmaßes „Schulnote (1-6)“ überprüft. Tabelle 33 zeigt die mittleren Noten, die von den Studierenden vergeben wurden. Auf den ersten Blick wird deutlich, dass außer der Literatur alle genannten Lehrmaterialien eine Note im guten Bereich erhalten.

Tabelle 33
Deskriptive Statistik zur Note für die Vorlesung, die E-Lecture, das WBT, die Übungsaufgaben, das Vorlesungsskript sowie die Literatur

	n	*M*	*SD*
Vorlesung	343	2.27	0.94
E-Lecture	343	2.10	1.06
WBT	343	2.21	1.24
Übungsaufgaben	340	**1.99**	0.90
Vorlesungsskript	341	2.25	0.93
Literatur	327	3.05	1.23

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Eine abhängige Varianzanalyse, in die lediglich die Studierenden eingehen, die zu allen sechs Materialien eine Einschätzung abgegeben haben, ergibt einen signifikanten Unterschied zwischen den Lehrmethoden mit einer mittleren bis großen Effektstärke (siehe Tabelle 34). Die Literatur erreicht in den Posthoc-Vergleichen ein signifikant schlechteres Ergebnis als alle anderen Lehrmethoden. Zusätzlich erhalten die Übungsaufgaben eine signifikant bessere Note als die Vorlesung und das Vorlesungsskript.

Tabelle 34
Ergebnisse der Mittelwertsanalysen zum Vergleich der Note von Vorlesung, E-Lecture, WBT, Übungsaufgaben, Vorlesungsskript und Literatur

n		Vorlesung	E-Lecture	WBT	Übungs-aufgaben	VL-Skript	Literatur	ANOVA				Posthoc
								F	*df*	*p*	*Eta²*	
324	*M*	2.28	2.10	2.17	**1.98**	2.25	3.06	46.26	4.29	<.01*	.13	Lit > Alle; VL, VLSk > ÜA
	SD	0.94	1.08	1.23	0.90	0.94	1.22					

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* bzw. w = Effektstärke

Hypothese 6 muss nach Auswertung der studentischen Angaben zurückgewiesen werden. Das traditionelle Lehrmedium „Literatur“ erreicht zwar hinsichtlich dieses groben Akzeptanzmaßes ein signifikant schlechteres Ergebnis als alle anderen Lehrmaterialien/-methoden. Die beiden anderen „traditionellen“ Lehrmaterialien bzw. -methoden Vorlesung/Vorlesungsskript erhalten allerdings lediglich eine signifikant schlechtere Note als die Übungsaufgaben. Es ist folglich nicht so, dass die „traditionellen“ Materialien/Methoden generell schlechter akzeptiert werden.

3.2.2.2 Feinere Akzeptanzfragen (Vergleich Vorlesung, E-Lecture und WBT)

Etwas genauer sollte die Akzeptanz der drei im Vordergrund betrachteten Lehrmethoden „Vorlesung“, „E-Lecture“ und „WBT“ erfasst werden. Tabelle 35 gibt die erfragten Variablen sowie die mittleren Beurteilungen der Studierenden wieder. Die Ergebnisse zu den Variablen, bei denen die Studierenden die drei Lehrmethoden in eine Rangreihe von 1 bis 3 bringen sollten, werden zudem in Abbildung 17 veranschaulicht.

Tabelle 35
Deskriptive Statistik zu den in der längerfristigen Befragung erhobenen Akzeptanzvariablen zu Vorlesung, E-Lecture und WBT

Akzeptanzvariablen	Vorlesung			E-Lecture			WBT		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Welche der folgenden Lehrmaterialien erbrachte den größten Nutzen für Sie?	344	2.10	0.80	344	1.94	0.76	344	**1.91**	0.87
Welche der folgenden Lehrmaterialien waren für Sie am motivierendsten?	340	**1.74**	0.79	340	2.04	0.79	339	2.16	0.82
Mit welcher der folgenden Lehrmaterialien hatten Sie den meisten Spaß beim Lernen?	339	**1.60**	0.73	340	2.08	0.77	339	2.27	0.80
Mit welcher der folgenden Lehrmaterialien konnten Sie sich am besten beim Lernen konzentrieren?	340	2.33	0.76	339	1.86	0.76	340	**1.77**	0.83

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert bzw. mittlere Ränge, *SD* = Standardabweichung; Falls nicht anders angegeben, wurden die Fragen jeweils mit einem Rang von 1 bis 3 beurteilt.

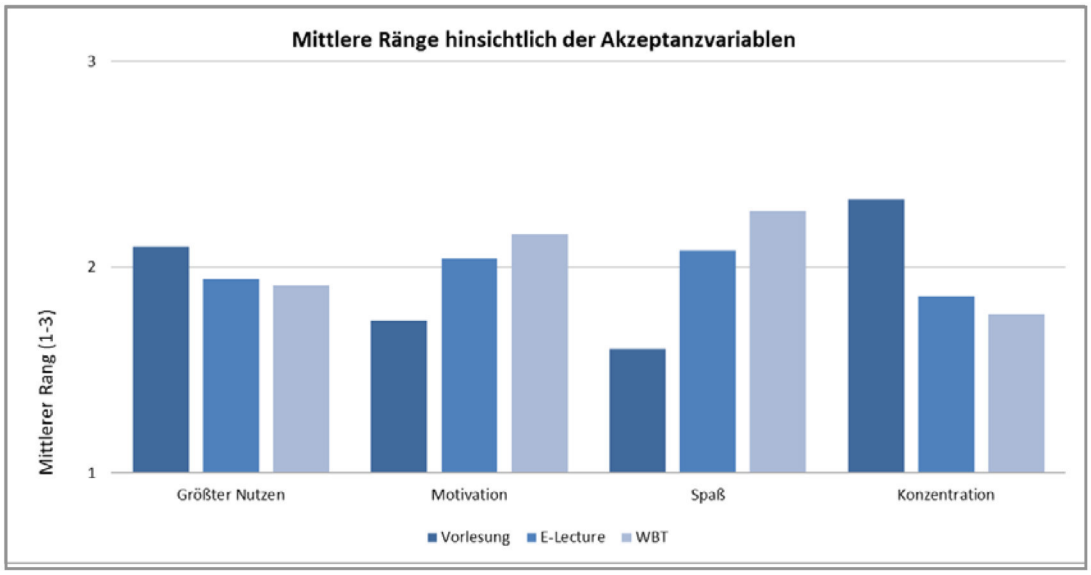

Abbildung 17. Mittlere Ränge der Lehrmethoden E-Lecture, WBT und Vorlesung zu den Variablen „Größter Nutzen“, „Motivation“, „Spaß“ und „Konzentration“

Erneut wurde mit abhängigen Varianzanalysen überprüft, inwieweit sich die Beurteilungen der Studierenden zu den drei Lehrmethoden signifikant unterscheiden. Die Ergebnisse dieser Analysen werden in Tabelle 36 wiedergegeben.

Tabelle 36
Ergebnisse der Mittelwertsanalysen zum Vergleich der in der längerfristigen Befragung erhobenen Akzeptanzvariablen zu Vorlesung, E-Lecture und WBT

	n	Vorlesung		E-Lecture		WBT		Friedman-Test				Posthoc
		M	*SD*	*M*	*SD*	*M*	*SD*	χ^2	*df*	*p*	w	
Größter Nutzen	344	2.10	0.80	1.94	0.76	**1.91**	0.87	8.01	2	.02*	.15	WBT<VL
Motivation	338	**1.75**	0.79	2.04	0.79	2.16	0.81	31.79	2	<.01*	.31	VL<EL,WBT
Spaß	339	**1.60**	0.73	2.09	0.77	2.27	0.80	82.46	2	<.01*	.50	VL<EL,WBT
Konzentration	338	2.33	0.75	1.87	0.76	**1.77**	0.83	64.60	2	<.01*	.44	EL,WBT<VL

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert bzw. mittlerer Rang, *SD* = Standardabweichung, χ^2 = Teststatistik des Friedman-Tests, *df* = Freiheitsgrade; *p* = Signifikanz, w = Effektstärke

Es zeigt sich, dass der Nutzen der Vorlesung signifikant schlechter eingeschätzt wird als der Nutzen des WBTs. Die Effektstärke dieses Unterschiedes ist allerdings als klein zu beurteilen.

In Bezug auf Motivation und Spaß werden sowohl die E-Lecture als auch das WBT von den Studierenden hingegen signifikant schlechter eingeschätzt als die Vorlesung. Dafür sind die Studierenden der Meinung, sich bei Bearbeitung der E-Lecture sowie des WBTs deutlich besser konzentrieren zu können als beim Besuch der Vorlesung. Diese Unterschiede weisen jeweils eine mittlere bis große Effektstärke auf.

3.2.2.3 Angegebene Nutzungsbereiche und -häufigkeit

Die Studierenden wurden befragt, wie häufig sie die Lehrmaterialien/-methoden nutzen (bzw. zur Vorbereitung auf die MAP nutzen) würden, wenn sie jederzeit zur Verfügung stünden. Tabelle 37 gibt die mittlere Beurteilung der Studierenden wieder.

Tabelle 37
Deskriptive Statistik zum generellen Nutzungswunsch bzw. dem Nutzungswunsch zur Klausurvorbereitung von Vorlesung, E-Lecture, WBT, Übungsaufgaben, Vorlesungsskript sowie Literatur

Genereller Nutzungswunsch	***N***	***M***	***SD***
Vorlesung	342	3.71	1.15
E-Lecture	342	3.95	1.05
WBT	340	3.84	1.19
Übungsaufgaben	339	3.92	1.05
Vorlesungsskript	341	**3.98**	0.92
Literatur	339	2.78	1.19
Nutzungswunsch zur Klausurvorbereitung	***n***	***M***	***SD***
Vorlesung	336	3.46	1.31
E-Lecture	339	4.22	1.03
WBT	338	4.20	1.10
Übungsaufgaben	338	**4.55**	0.74
Vorlesungsskript	340	4.25	0.82
Literatur	333	3.02	1.26

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung; Skala: 1 = gar nicht, 5 = häufig

Die Angaben werden durch Abbildung 18 veranschaulicht.

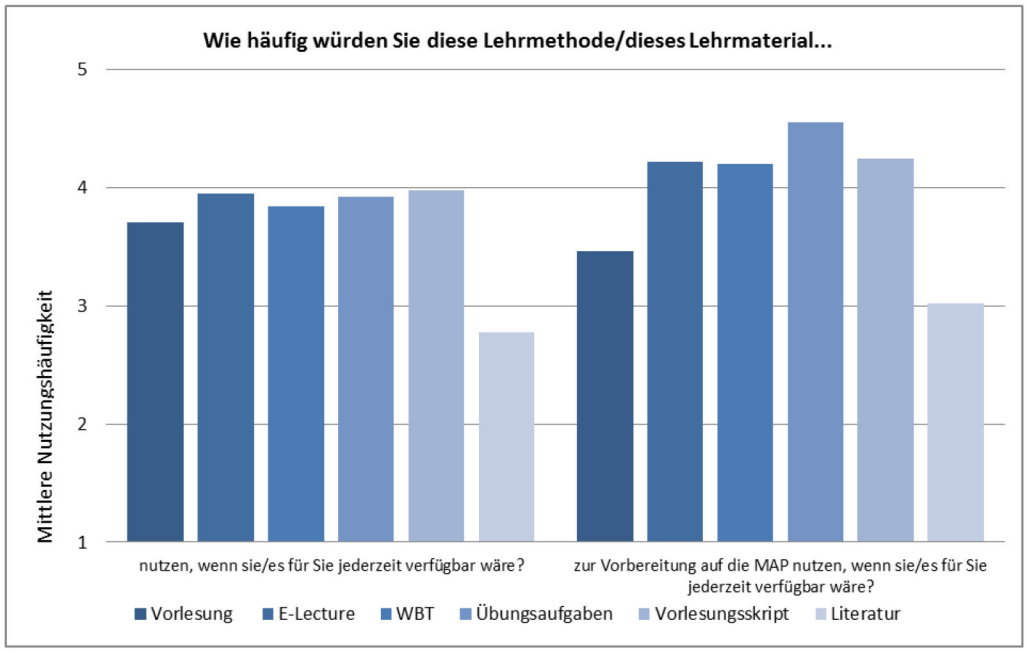

Abbildung 18. Prozentuale Häufigkeit des generellen Nutzungswunsches sowie des Nutzungwunsches zur Klausurvorbereitung von Vorlesung, E-Lecture, WBT, Übungsaufgaben, Vorlesungsskript und Literatur

Um zu überprüfen, inwieweit sich die Angaben der Studierenden signifikant unterscheiden, wurden erneut abhängige Varianzanalysen berechnet. Tabelle 38 zeigt die Ergebnisse dieser Analysen.

Tabelle 38
Ergebnisse der Mittelwertsanalysen zum Vergleich des generellen Nutzungwunsches bzw. des Nutzungswunsches zur Klausurvorbereitung von Vorlesung, E-Lecture, WBT, Übungsaufgaben, Vorlesungsskript sowie Literatur

	n		Vorlesung	E-Lecture	WBT	Übungs-aufgaben	Vorlesungs-skript	Literatur	ANOVA				Posthoc
									F	*df*	*p*	*Eta²*	
Generelle Nutzung	335	*M*	3.71	3.96	3.84	3.91	**3.98**	2.78	61.68	4.49	<.01*	.16	Lit<Alle; Vl<VlSk
		SD	1.15	1.04	1.19	1.05	0.92	1.19					
Nutzung zur Klausurvorb.	326	*M*	3.46	4.24	4.18	**4.56**	4.24	3.02	107.91	4.09	<.01*	.25	ÜA>EL,VLSk, WBT>VL>Lit
		SD	1.31	1.00	1.12	0.75	0.83	1.26					

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke

Generell würden die Studierenden die Literatur signifikant seltener nutzen als alle anderen Lehrmaterialien/-methoden. Die Vorlesung würde bedeutsam seltener genutzt als das Vorlesungsskript.

Zur Vorbereitung auf die MAP möchten die Studierenden am signifikant häufigsten die Übungsaufgaben nutzen. Signifikant seltener möchten sie die E-Lecture, das Vorlesungsskript sowie das WBT nutzen. Am seltensten werden auch hier die Vorlesung

sowie die Literatur genannt. Die Effektstärken zu beiden Mittelwertsunterschieden sind jeweils als stark einzustufen.

Des Weiteren wurde erfragt, für welche Nutzungsbereiche die Studierenden die einzelnen Lehrmethoden (Vorlesung, E-Lecture, WBT, Übungsaufgaben, Vorlesungsskript, Literatur) nutzen wollen würden. Die ausführlichen Ergebnisse hierzu sind in Anhang B zu sehen.
Zusammengefasst zeigte sich, dass sich die Studierenden für die Wissensaneignung am häufigsten die Vorlesung wünschen, für die Prüfungsvorbereitung am häufigsten die Übungsaufgaben. Zum Nachholen verpasster Vorlesungssitzungen möchten die Studierenden am liebsten die E-Lecture und zum Erarbeiten schwieriger Inhalte das WBT sowie die E-Lecture nutzen. Der Austausch mit Kommilitonen über die Inhalte sollte nach Meinung der Studierenden am besten innerhalb der Vorlesung erfolgen. Auf die Frage, welche der Lehrmethoden/-materialien sie gar nicht nutzen möchten, wird signifikant am häufigsten die Literatur genannt.
Betrachtet man diese Ergebnisse aus Sicht der Lehrmethoden, wird die Vorlesung am häufigsten für die Wissensaneignung und den Austausch mit Kommilitonen gewünscht, die E-Lecture am häufigsten für das Nachholen verpasster Sitzungen. In Bezug auf das WBT, die Übungsaufgaben und das Vorlesungsskript wünschen sich die Studierenden am häufigsten eine Nutzung zur Prüfungsvorbereitung, die Literatur möchten die Studierenden hauptsächlich zur Wissensaneignung oder zum Erarbeiten nicht-verstandener Inhalte nutzen.
Für andere studienrelevante Module wünschen sich die Studierenden am häufigsten die Übungsaufgaben, das Vorlesungsskript sowie die E-Lecture. Für interessante Themen außerhalb der Universität möchten sie am liebsten WBTs und E-Lectures nutzen.

3.2.3 Abschließende Einschätzungen

Hypothese 11: Die E-Learning-Anwendung, die im Sommersemester häufiger freiwillig genutzt wurde, wird hinsichtlich der Akzeptanz kurz vor der MAP am besten beurteilt.

Im abschließenden Fragebogen, welcher zur Probeklausur direkt vor der MAP gehörte, wurde den Studierenden die Frage gestellt, wie häufig sie die Lernmaterialien „Vorlesungsskript“, „Mitschrift“, „E-Lecture“, „WBT“ und „Literatur“ für die Vorbereitung auf die MAP genutzt hatten, wie motivierend sie diese für ihre bisherige Klausurvorbereitung fanden und wie viel Spaß ihnen das Arbeiten mit den Lernmaterialien bereitet hat. Zudem vergaben die Studierenden ein letztes Mal eine Note für die

Lernmaterialien als generelles Akzeptanzmaß. Tabelle 39 zeigt die deskriptive Auswertung dieser abschließenden Befragung.

Tabelle 39
Deskriptive Statistik zu den abschließenden Akzeptanzvariablen zu Vorlesungsskript, Mitschrift, E-Lecture, WBT, Literatur sowie „Sonstiges“

		VL-Skript	Mitschrift	E-Lecture	WBT	Literatur	Sonstiges*
Nutzungshäufigkeit	*n*	236	234	231	234	231	79
	M	**3.24**	2.67	1.57	3.08	0.96	0.57
	SD	1.10	1.33	1.42	1.39	1.12	1.17
Motivation	*n*	222	223	214	220	201	75
	M	2.79	2.67	2.59	**3.38**	1.74	1.60
	SD	0.86	0.90	0.99	0.87	0.84	0.94
Spaß	*n*	222	220	211	218	195	71
	M	2.58	2.37	2.43	**3.23**	1.65	1.55
	SD	0.81	0.86	1.02	0.93	0.80	0.92
Note (1-6)	*n*	215	215	207	212	172	63
	M	2.26	2.68	2.56	**1.50**	3.65	4.57
	SD	0.96	1.21	1.33	0.78	1.39	1.69

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung; *Die Kategorie „Sonstiges“ wird im Rahmen der Varianzanalyse nicht berücksichtigt.

Insgesamt gaben 18 Personen an, „Sonstige Materialien“ zum Lernen für die Klausur selten bis häufig genutzt zu haben. Von diesen 18 machten 13 eine Aussage, was unter „Sonstiges“ zu verstehen ist: Sieben gaben das „Internet“ an, jeweils einer „Google“, „Karteikarten“, „Gespräche“, „Forum in *Stud.IP*“ und „Material des Seminars“, eine Person gab „ert“ an – da es sich hier um eine Buchstabenfolge auf der Tastatur handelt, eine dementsprechend nicht ernst zu nehmende Antwort. Die Kategorie „Sonstiges“ wurde aufgrund der sehr kleinen und inhomogenen Stichprobe im Rahmen der varianzanalytischen Auswertungen außer Acht gelassen. Es werden lediglich die fünf Lernmaterialien „Vorlesungsskript“, „Mitschriften“, „E-Lecture“, „WBT“ und „Literatur“ miteinander verglichen.

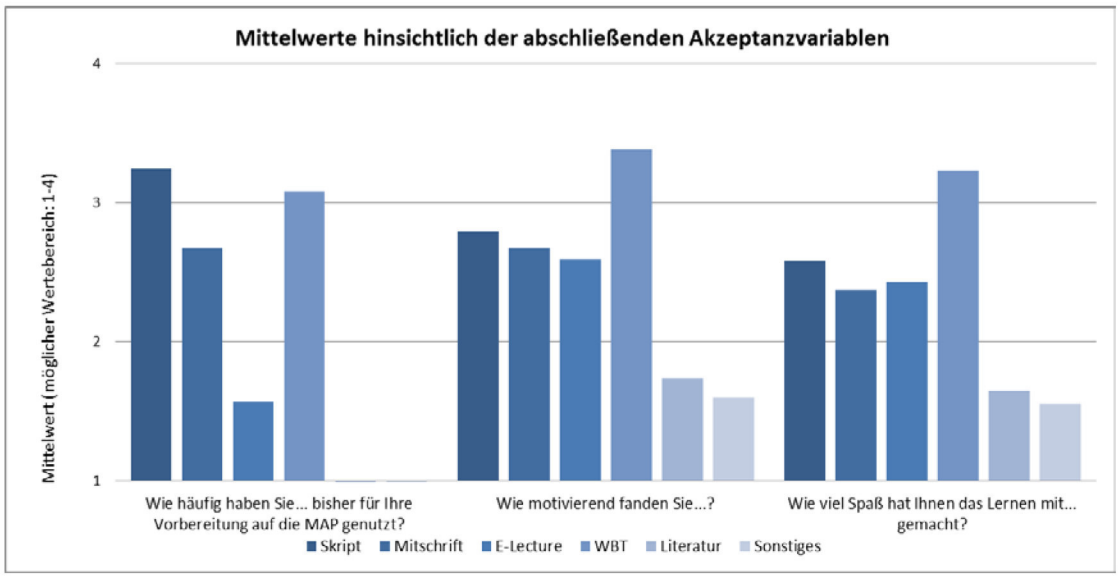

Abbildung 19. Mittlere Einschätzungen zur Häufigkeit der Nutzung, Motivation und Spaß für Vorlesungsskript, Mitschrift, E-Lecture, WBT, Literatur und „Sonstiges“

Zur Veranschaulichung stellen Abbildung 19 und 20 diese Ergebnisse noch einmal grafisch dar. Auf den ersten Blick scheinen das Skript und das WBT am häufigsten zur Klausurvorbereitung genutzt worden zu sein.

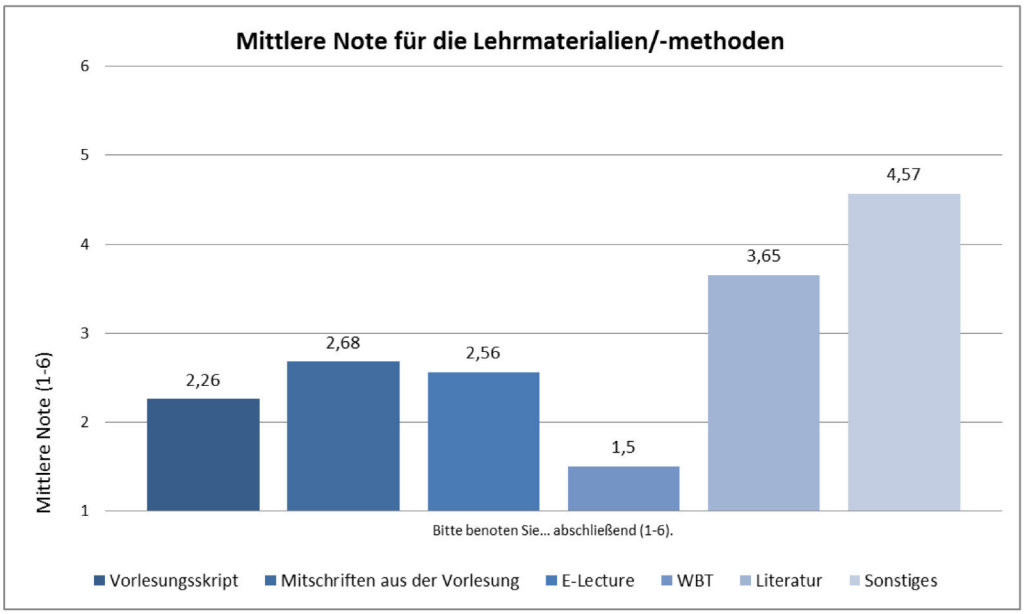

Abbildung 20. Mittlere Note für das Vorlesungsskript, die Mitschriften, die E-Lecture, das WBT, die Literatur und Sonstiges

Hypothese 11 („Die E-Learning-Anwendung, die im Sommersemester häufiger freiwillig genutzt wurde, wird hinsichtlich der Akzeptanz kurz vor der MAP am besten beurteilt.") muss folgendermaßen umformuliert werden: „Das WBT, welches im Sommersemester häufiger freiwillig genutzt wurde, wird hinsichtlich der Akzeptanz kurz vor der MAP besser beurteilt als die E-Lecture." Es zeigt sich bereits augenscheinlich, dass das WBT nun als am meisten motivierend und Spaß bringend wahrgenommen wird. Es erhält ebenfalls abschließend die deutlich beste Noteneinstufung.

Dies bestätigt sich durch die Varianzanalysen und Posthoc-Vergleiche (siehe Tabelle 40).

Tabelle 40

Ergebnisse der Mittelwertsanalysen zum Vergleich der in der abschließenden Befragung erhobenen Akzeptanzvariablen zu Vorlesungsskript, Mitschrift, E-Lecture, WBT und Literatur

	n		Vorlesungs-Skript	Mitschrift	E-Lecture	WBT	Literatur	ANOVA				Posthoc
								F	*df*	*p*	*Eta²*	
Häufigkeit der Nutzung	227	*M*	**3.25**	2.64	1.56	3.07	0.97	136.25	3.27	<.01*	.38	VLSk,WBT>Mit>EL>Lit
		SD	1.09	1.34	1.42	1.40	1.13					
Motivation	193	*M*	2.75	2.64	2.54	**3.36**	1.72	89.41	3.58	<.01*	.32	WBT>VLSk,Mit, EL>Lit
		SD	0.88	0.91	1.00	0.89	0.83					
Spaß	191	*M*	2.55	2.34	2.41	**3.21**	1.62	90.25	3.55	<.01*	.32	WBT>VLSk,EL, Mit>Lit (VLSk>Mit)
		SD	0.83	0.87	1.03	0.94	0.78					
Note	168	*M*	2.29	2.71	2.57	**1.52**	3.64	76.96	3.55	<.01*	.32	WBT<VLSk,EL, Mit<Lit (VLSk<Mit)
		SD	1.01	1.24	1.34	0.79	1.39					

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke

Die meisten Studierenden gaben folglich an, für die Vorbereitung auf die MAP das Skript sowie die WBTs genutzt zu haben. Signifikant etwas seltener wurden nach diesen Angaben die Mitschriften, noch seltener die E-Lectures und am seltensten die Literatur genutzt. Dieser und auch alle weiteren Mittelwertsunterschiede weisen eine sehr starke Effektstärke auf.

Am motivierendsten fanden die Studierenden in dieser abschließenden Befragung die WBTs. Signifikant weniger motivierend wurden das Skript, die E-Lecture und die Mitschriften wahrgenommen, als am wenigsten motivierend wurde die Literatur eingeschätzt. Das gleiche Bild ergab sich für die Einschätzung, wie viel Spaß das Lernen mit den Lernmaterialien gemacht hat sowie für die abschließend vergebene Gesamtnote.

Hypothese 11 kann folglich beibehalten werden: das häufiger genutzte WBT[34] schneidet hinsichtlich der Akzeptanz besser ab als die weniger genutzte E-Lecture.

3.2.4 Verhaltensakzeptanz

Hypothese 9: Die E-Learning-Anwendung mit der höheren Einstellungsakzeptanz am Ende des Wintersemesters wird im Sommersemester häufiger freiwillig genutzt.

Nachdem den Studierenden E-Lectures und WBTs zu allen zwölf Vorlesungsthemen für ihre Klausurvorbereitung zur Verfügung gestellt wurden, konnten sie über diese fast ein halbes Jahr lang frei verfügen. Nach der Klausur wurden die Studierenden anhand dieser freiwilligen Nutzung in vier Gruppen eingeteilt: Die erste Gruppe (Gruppe „Keine“) hatte für ihre Klausurvorbereitung weder eine E-Lecture noch ein WBT (in der Summe mehr als eine bzw. zehn Minuten[35]) genutzt. Die Gruppe „E-Lecture“ hatte mindestens eine E-Lecture genutzt (jedoch kein WBT), die Gruppe „WBT“ mindestens ein WBT (jedoch keine E-Lecture). Eine weitere Gruppe (die Gruppe „Kombiniert“ bzw. „Kombi-Nutzer“) nutzte sowohl mindestens eine E-Lecture als auch ein WBT für ihre Klausurvorbereitung. Tabelle 41 gibt einen Überblick über die Nutzergruppen getrennt für die zwölf Themen sowie über alle zwölf Themen insgesamt (für die Nutzungen, die tatsächlich in der Summe länger als zehn Minuten andauerten).

[34] Im folgenden Abschnitt „Verhaltensakzeptanz“ wird sich zeigen, dass diese Einschätzung der Studierenden auch mit der tatsächlichen Nutzung übereinstimmt.

[35] In der Gruppeneinteilung auf Ebene der zwölf Themen wurden alle Nutzungen berücksichtigt, die länger als eine Minute andauerten. In der Gruppeneinteilung über alle zwölf Themen wurden lediglich diejenigen Nutzungen berücksichtigt, die **in der Summe** länger als zehn Minuten bestanden.

Tabelle 41
Nutzungen der Nutzergruppen „Keine Nutzung“, „E-Lecture-Nutzer“, „WBT-Nutzer“ und „Kombinierte Nutzung“ für alle zwölf Themen sowie insgesamt

Nutzergruppen	Keine		E-Lecture		WBT		Kombiniert	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Thema 1	904	66.6	66	4.9	333	24.5	54	4.0
Thema 2	960	70.7	20	1.5	361	26.6	16	1.2
Thema 3	1153	85.0	28	2.1	160	11.8	16	1.2
Thema 4[36]	-	-	45	-	-	-	-	-
Thema 5	1033	76.1	29	2.1	276	20.3	19	1.4
Thema 6	1079	79.5	32	2.4	235	17.3	11	0.8
Thema 7	969	71.4	26	1.9	326	24.0	36	2.7
Thema 8	973	71.7	21	1.5	345	25.4	18	1.3
Thema 9	957	70.5	21	1.5	361	26.6	18	1.3
Thema 10	1081	79.7	18	1.3	254	18.7	4	0.3
Thema 11	1127	83.1	17	1.3	206	15.2	7	0.5
Thema 12	1138	83.9	21	1.5	189	13.9	9	0.7
Gesamt (> 10 Min.)	777	**57.3**	55	**4.1**	440	**32.4**	85	**6.3**

Anmerkung. n = Stichprobengröße, *%* = Prozentualer Anteil an der Gesamtstichprobe. Hinsichtlich der Gesamtauswertung wurden alle Nutzungen eingeschlossen, die insgesamt länger als zehn Minuten andauerten.

Betrachtet man die Gruppenverteilung der Studierenden über die zwölf Themen, zeigt sich, dass 57% der Studierenden die Anwendungen nicht oder in der Summe weniger als 10 Minuten genutzt haben. Fast ein Drittel der Studierenden hat darüber hinaus nur mit den Web-based Trainings gelernt, ca. 4% haben nur die E-Lectures genutzt und weitere 6% der Studierenden lernten sowohl mit Web-based Trainings als auch mit E-Lectures. Die E-Lectures wurden folglich insgesamt von etwa 10% der Studierenden genutzt, die WBTs von etwa 39% der Studierenden. Diese Verteilung wird durch Abbildung 21 noch einmal veranschaulicht.

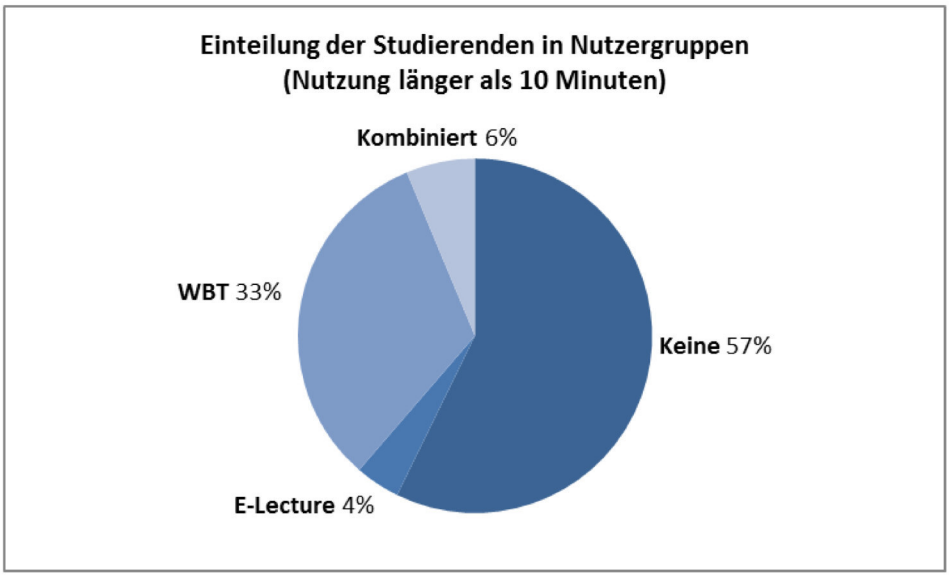

Abbildung 21. Einteilung der Studierenden in die vier Nutzergruppen „Keine Nutzung“, „E-Lecture-Nutzer“, „WBT-Nutzer“ und „kombinierte Nutzung“ (in Prozent)

[36] Aufgrund eines technischen Problems existieren leider keine Nutzungsaufzeichnungen für das WBT zum vierten Thema. Dieses Thema kann folglich bei den Berechnungen zur langfristigen Lernwirksamkeit nicht berücksichtigt werden.

In einem Chi-Quadrat-Test, mit dem die Häufigkeitsverteilung auf die vier Gruppen auf Signifikanz überprüft wurde, ergab sich ein signifikanter Unterschied (*Chi-Quadrat*(3) = 1023.48, $p < .01$). Signifikanzen ergaben sich auch in weiteren Tests zwischen den drei Nutzergruppen „E-Lecture“, „WBT“ und „Kombinierte Nutzung“ oder im Vergleich der beiden Nutzergruppen „E-Lecture“ und „WBT“. Die Häufigkeitsverteilung in den vier Gruppen ist folglich signifikant unterschiedlich voneinander. Die WBT-Nutzung ist signifikant höher als die E-Lecture-Nutzung oder die kombinierte Nutzung beider E-Learning-Anwendungen.

In der längerfristigen Akzeptanzbefragung am Ende des Wintersemesters erzielten die E-Lecture und das WBT vergleichbar gute Resultate. In keiner der Akzeptanzfragen schnitten die beiden Lehrmaterialien signifikant unterschiedlich voneinander ab. Auch hinsichtlich der Nutzungswünsche ergab sich lediglich ein signifikanter Unterschied dementsprechend, dass sich die Studierenden zum Nachholen verpasster Vorlesungen eher die E-Lectures wünschten, weniger die WBTs. Dieser Unterschied ist für die Betrachtung der Nutzung zur Klausurvorbereitung allerdings weniger bedeutsam.

Da nun die Einstellungsakzeptanz am Ende des Wintersemesters hinsichtlich beider Lehrmaterialien vergleichbar war, wäre Hypothese 9 demnach folgendermaßen anzupassen: „Die E-Lectures und die WBTs werden im Sommersemester gleich häufig freiwillig genutzt.“ Da die Vorbedingung (Unterschied in der Einstellungsakzeptanz) für die Hypothese nicht gegeben ist, ist Hypothese 9 nicht überprüfbar. Es kann allerdings zusammengefasst werden, dass die WBTs von ca. acht Mal so vielen Studierenden zur Klausurvorbereitung genutzt werden als die E-Lectures.

Dies spiegelt sich auch in den Ergebnissen zur abschließenden Akzeptanzbefragung wieder (siehe Abschnitt „3.2.3 Abschließende Einschätzungen“). Hier gaben die meisten Studierenden an, zur Vorbereitung auf die MAP das Skript und die WBTs genutzt zu haben.

3.2.4.1 Nutzungshäufigkeit und -zeit

Summe der Nutzungen

Vergleicht man die Nutzungshäufigkeit der E-Lecture und der WBTs, zeigt sich in der Summe der Nutzungen ganz deutlich, dass die Nutzer der WBTs im Schnitt fast doppelt so häufig eines der WBTs aufgerufen haben (M = 8.5 Mal) als die Nutzer der E-Lectures eine der E-Lectures (M = 5.3 Mal; siehe Tabelle 42).

Tabelle 42
Durchschnittliche Nutzungshäufigkeit der E-Lectures und WBTs für alle zwölf Themen sowie insgesamt (Summe und durchschnittliche Nutzung)

Nutzungen	E-Lecture			WBT		
	n	***M***	***SD***	***n***	***M***	***SD***
Thema 1	120	1.82	1.05	387	1.49	0.82
Thema 2	36	1.61	1.02	377	1.45	0.76
Thema 3	44	1.55	1.30	176	1.36	0.67
Thema 4	45	1.40	0.62	-	-	-[37]
Thema 5	48	1.79	1.35	295	1.34	0.68
Thema 6	43	1.40	0.76	246	1.22	0.49
Thema 7	62	1.58	0.82	362	1.55	0.85
Thema 8	39	1.62	1.04	363	1.42	0.77
Thema 9	39	1.31	0.61	379	1.37	0.73
Thema 10	22	1.18	0.50	258	1.17	0.40
Thema 11	24	1.38	0.65	213	1.19	0.48
Thema 12	30	1.50	0.68	198	1.22	0.64
Summe (>10 Min.)	140	5.48	4.77	525	8.40	5.52
Summe (>10 Min.) ohne Kombi	**55**	**5.25**	**5.16**	**440**	**8.45**	**5.62**
Durchschnitt (> 10 Min.)	140	1.56	0.70	525	1.30	0.45
Durchschnitt (>10 Min.) ohne Kombi	**55**	**1.56**	**0.91**	**440**	**1.30**	**0.46**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung; Für die Summe und den Durchschnitt sind jeweils die Nutzungen länger als 10 Minuten mit und ohne die Nutzergruppe „kombinierte Nutzung“ angegeben; dick umrahmt: Ergebnise der WBTs ohne Übungsaufgaben

Dieser Unterschied stellt sich in einem *t*-Test für unabhängige Stichproben[38] als signifikant heraus ($t(493) = -4.01$, $p < .01$; ohne Einbezug der „Kombi-Nutzer“). Die WBT-Nutzer haben folglich mehr Themen bearbeitet als die E-Lecture-Nutzer.

Mittlere Zugriffshäufigkeit

Betrachtet man jedoch die mittlere Zugriffshäufigkeit über die zwölf Themen, zeigt sich, dass einzelne E-Lectures von den E-Lecture-Nutzern im Mittel signifikant häufiger aufgerufen wurden ($M = 1.6$ Mal) als einzelne WBTs von den WBT-Nutzern ($M = 1.3$ Mal). Dabei scheint sowohl die Anzahl der Nutzer als auch die mittlere Zugriffshäufigkeit bei den WBTs ohne Übungsaufgaben (siehe Umrandung in Tabelle 42) niedriger zu sein als bei den WBTs mit Übungsaufgaben.

Für einen Mittelwertvergleich zwischen der E-Lecture- und der WBT-Gruppe (siehe Tabelle 43) mussten aufgrund unterschiedlicher Stichprobengrößen (und in diesem Fall auch Varianzheterogenität) Zufallsstichproben aus der WBT-Nutzergruppe gezogen werden, damit in etwa vergleichbar große Gruppen entstehen, die mittels *t*-Test verglichen werden können. Um sicherzugehen, dass durch die Ziehung einer Zufallsstichprobe kein falsches Ergebnis zustande kommt, wurde dieses Verfahren dreimal durchgeführt.

[37] Aufgrund technischer Schwierigkeiten wurden zu WBT 4 keine Nutzungszeiten erfasst.

[38] Trotz der hier sehr unterschiedlich großen Stichproben wurde dennoch ein *t*-Test gerechnet, da dieser bei Varianzhomogenität – welche in diesem Vergleich vorliegt – auch für ungleich große Stichproben als robust gilt (Bortz & Schuster, 2010, S. 122).

Wie in Tabelle 43 zu sehen, ergab sich in zwei der Auswertungen ein signifikanter Unterschied zugunsten der E-Lecture (mit jeweils etwa mittelgroßem Effekt). Einzelne E-Lectures wurden folglich signifikant häufiger aufgerufen als einzelne WBTs.

Tabelle 43
Ergebnisse der Mittelwertsanalysen zum Vergleich der mittleren Nutzungshäufigkeit der E-Lecture- und WBT-Nutzer (drei Zufallsziehungen)

	E-Lecture-Nutzer			WBT-Nutzer			Welch-Test			
	n	*M*	*SD*	*n*	*M*	*SD*	*t*	*df*	*p*	*d*
Zufallsgruppen 1	55	**1.56**	0.91	73	1.24	0.35	2.45	66.02	.02*	.47
Zufallsgruppen 2	55	**1.56**	0.91	73	1.23	0.39	2.52	69.26	.01*	.47
Zufallsgruppen 3	55	**1.56**	0.91	74	1.32	0.46	1.77	74.25	.08	-

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *t* = Teststatistik des Welch-Tests; *df* = Freiheitsgrade; *p* = Signifikanz; *d* = Effektstärke

Insgesamt haben die WBT-Nutzer folglich mehrere der zwölf WBTs aufgerufen, um sich für die Klausur vorzubereiten, die E-Lecture-Nutzer haben hingegen mit weniger E-Lectures gearbeitet, diese dafür häufiger aufgerufen.

Nutzungszeit

Dieser Unterschied findet sich auch in der Auswertung der Nutzungszeiten wieder (siehe Tabelle 44): Vergleicht man die **Summe der Nutzungszeiten** der E-Lecture- (19686 Sekunden = ca. 5.47 Stdn.) und WBT-Nutzer (16300 Sekunden = ca. 4.53 Stdn.), ergibt sich (mit $t(493) = 1.44$, $p = .15$) kein signifikanter Unterschied. Die einzelnen Nutzer haben sich folglich im Schnitt in etwa genauso lange mit den E-Lectures oder WBTs beschäftigt. Auch hier wurden die Kombi-Nutzer nicht in die Auswertungen einbezogen.

Tabelle 44
Durchschnittliche Nutzungszeit der E-Lectures und WBTs getrennt für alle zwölf Themen sowie insgesamt (Summe und durchschnittliche Nutzungszeit)

Nutzungszeit in Sekunden	E-Lecture			WBT		
	n	*M*	*SD*	*n*	*M*	*SD*
Thema 1	120	4151.30	4316.94	387	1965.49	2649.65
Thema 2	36	4540.64	4405.87	377	1627.65	2305.63
Thema 3	44	5820.66	4826.03	176	2027.15	2358.94
Thema 4	45	6181.73	4558.48	-	-	-
Thema 5	48	6329.54	4167.09	295	2951.40	3187.11
Thema 6	43	5147.70	5602.32	246	3065.70	3880.07
Thema 7	62	5814.79	5393.43	362	3458.37	4140.33
Thema 8	39	6141.67	4687.89	363	3243.52	3730.68
Thema 9	39	6165.69	5386.43	379	2641.98	3286.89
Thema 10	22	5364.50	4171.85	258	1950.53	2583.42
Thema 11	24	4635.83	3718.21	213	3262.01	4104.67
Thema 12	30	2482.93	2515.89	198	2503.85	3586.37
Summe (>10 Min.)	140	20382.74	20914.86	525	16138.13	15598.73
Summe (>10 Min.) ohne Kombi	**55**	**19685.89**	**20628.86**	**440**	**16299.68**	**15859.40**
Entspricht in Stunden	**55**	**5.47**	**5.73**	**440**	**4.53**	**4.41**
Durchschnitt (>10 Min.)	140	5189.35	3193.27	525	2691.01	2272.84
Durchschnitt (>10 Min.) ohne Kombi	**55**	**5380.59**	**2803.91**	**440**	**2734.29**	**111.36**
Entspricht in Stunden	**55**	**1.49**	**0.78**	**440**	**0.76**	**0.03**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung; Für die Summe und den Mittelwert sind jeweils die Nutzungen länger als 10 Minuten mit und ohne die Nutzergruppe „kombinierte Nutzung“ angegeben.

Betrachtet man jedoch auch hier die **mittlere Nutzungszeit** zu den zwölf Themen (dies wird der Übersicht halber auch noch einmal in Abbildung 22 veranschaulicht), zeigt sich eine signifikant höhere durchschnittliche Nutzungszeit für die Nutzungen einzelner E-Lectures. Studierende aus der E-Lecture-Nutzergruppe befassten sich ungefähr doppelt so lange mit einer E-Lecture als Studierende aus der WBT-Nutzergruppe mit einem WBT. Sie beschäftigen sich im Schnitt auch deutlich länger mit einer E-Lecture als deren Rohlaufzeit. Diese liegt über alle zwölf E-Lectures hinweg bei *M* = 3673.5 Sekunden = ca. 61.23 Minuten = ca. 1.02 Stunden.

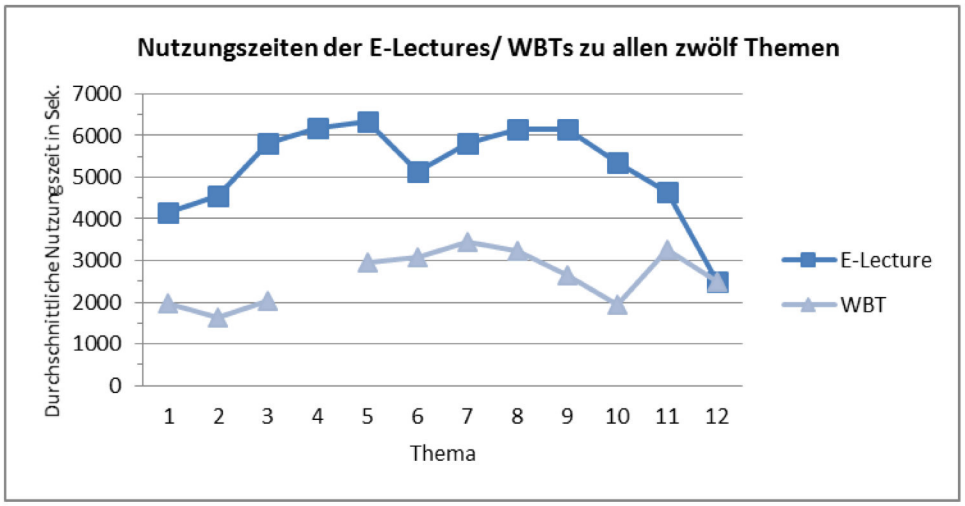

Abbildung 22. Durchschnittliche Nutzungszeit der E-Lectures und WBTs zu allen zwölf Themen

Auch dies wird durch dreimalige *t*-Tests für unabhängige Stichproben mit verschiedenen Zufallsstichproben aus der WBT-Nutzergruppe belegt (siehe Tabelle 45). Die Effektstärke ist hier jeweils als sehr stark zu beurteilen.

Tabelle 45
Ergebnisse der Mittelwertsanalysen zum Vergleich der mittleren Nutzungszeit der E-Lecture- und WBT-Nutzer

Nutzungszeit	E-Lecture-Nutzer			WBT-Nutzer			*t*-Test			
	n	*M*	*SD*	*n*	*M*	*SD*	*t*	*df*	*p*	*d*
Zufallsgruppen 1	55	**5380.59**	2803.91	73	2828.76	2617.05	5.30	126	<.01*	.94
Zufallsgruppen 2	55	**5380.59**	2803.91	73	2979.10	2299.27	5.99	126	<.01*	.94
Zufallsgruppen 3	55	**5380.59**	2803.91	74	2355.73	2251.46	6.79	127	<.01*	1.19

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *t* = Teststatistik des *t*-Tests; *df* = Freiheitsgrade; *p* = Signifikanz; *d* = Effektstärke

Das trotz der sehr viel längeren Nutzungsdauer der einzelnen E-Lectures eine im Schnitt gleichlange Nutzungszeit für die Gesamtbearbeitung der E-Lectures und WBTs resultiert, ist folglich auf die Anzahl der bearbeiteten Lerneinheiten zurückzuführen. Studierende, die mit den E-Lectures lernten, arbeiteten folglich häufiger und länger mit wenigen der E-Lectures, während Studierende, die sich mit den WBTs auf die MAP vorbereiteten, einzelne WBTs eher seltener und kürzer bearbeiteten, aber dafür mehrere verschiedene WBTs nutzten.

3.2.5 Zusammenfassung und Diskussion

Die Akzeptanz wurde sowohl kurzfristig einen Tag nach der Nutzung einer Anwendung bzw. dem Besuch der Vorlesung, längerfristig nach Nutzung aller drei Lehrmethoden am Ende des Wintersemesters als auch abschließend kurz vor der Modulabschlussprüfung nach dem Sommersemester erhoben. Zudem konnte durch die freiwillige Nutzung der E-Learning-Anwendungen im Sommersemester zu diesen beiden Lehrmethoden auch die Verhaltensakzeptanz erhoben werden, die ergänzend ein objektiveres Akzeptanzmaß darstellt als die Befragungen.

Interessant ist nun, inwieweit sich die Akzeptanzeinschätzungen zu den drei im Vordergrund betrachteten Lehrmethoden Vorlesung, E-Lecture und WBT – aber auch hinsichtlich der zusätzlich abgefragten Materialien, wie z.B. Übungsaufgaben oder Literatur – im Laufe der Projektzeit verändert haben. Dabei soll zunächst auf das grobe Akzeptanzmaß „Gesamtnote" eingegangen werden, welches in allen drei Befragungen von den Studierenden beurteilt werden sollte, danach sollen die feineren Akzeptanzurteile diskutiert werden. Zu guter Letzt wird der geäußerte Nutzungswunsch mit der tatsächlichen freiwilligen Nutzung der Materialien in Beziehung gesetzt.

3.2.5.1 Gesamtnote (1 - 6)

Betrachtet man die in Tabelle 46 angegebenen mittleren Beurteilungen der Lehrmaterialien/-methoden, zeigt sich, dass die im Vordergrund betrachteten Lehrmethoden „Vorlesung", „E-Lecture" und „Web-based Training" sowohl in der kurz- als auch in der längerfristigen Beurteilung in etwa gleich gut abschneiden. Zwischen diesen drei Methoden ergaben sich zu den beiden Evaluationszeitpunkten keine signifikanten Unterschiede.

Tabelle 46
Durchschnittliche Gesamtnote in der kurzfristigen, längerfristigen sowie abschließenden Evaluation zu den insgesamt sieben verschiedenen Lehrmaterialien/-methoden

	Kurzfristige Beurteilung	Längerfristige Beurteilung	Abschließende Beurteilung
Vorlesung	2.42	2.27	-
E-Lecture	2.44	2.10	2.56
WBT	**2.38**	2.21	**1.50**
Übungsaufgaben	-	**1.99**	-
Literatur	-	3.05	3.65
Vorlesungsskript	-	2.25	2.26
Mitschriften	-	-	2.68

Anmerkung. - = wurde in der jeweiligen Evaluation nicht erfragt

Interessant ist, dass die Beurteilung zu allen drei Methoden von der kurz- zur längerfristigen Beurteilung tendenziell etwas besser wird. In der kurzfristigen Beurteilung wurden die Lehrmethoden jeweils einzeln evaluiert, in der längerfristigen Beurteilung

mussten die Studierenden innerhalb einer Fragebogenseite für alle aufgeführten Lehrmethoden Noten vergeben. Evtl. hat hier das eher vergleichende Urteil sowie die Ergänzung durch die drei Lehrmaterialien „Übungsaufgaben“, „Literatur“ und „Vorlesungsskript“ zu einem etwas milderen Urteil beigetragen. Insbesondere die Literatur, die signifikant schlechter beurteilt wird als alle anderen Materialien, könnte hier einen Einfluss gehabt haben. Weitere Signifikanzen innerhalb der längerfristigen Beurteilung der Note ergaben sich zugunsten der Übungsaufgaben im Vergleich zur Vorlesung sowie dem Vorlesungsskript.

Besonders spannend ist im Vergleich zu den kurz- und längerfristigen Beurteilungen die abschließende Bewertung von E-Lecture und Web-based Training. Während die durchschnittliche Note für die E-Lectures nach der Möglichkeit zur freiwilligen Nutzung um eine halbe Note schlechter ausfällt (M = 2.1 vs. M = 2.6), hat sich die Beurteilung der Web-based Trainings deutlich verbessert (M = 2.2 vs. M = 1.5). In diesem abschließenden Urteil schneiden die beiden E-Learning-Anwendungen dann auch zum ersten Mal signifikant unterschiedlich zugunsten der WBTs ab.

Da die Vorlesung zu diesem Zeitpunkt schon ein halbes Jahr in der Vergangenheit lag und die abschließende Evaluation die für die Prüfungsvorbereitung nutzbaren Materialien fokussieren sollte, erschien es nicht sinnvoll, für die Vorlesung nun noch einmal eine weitere Beurteilung zu erheben. Dafür wurden hier sowohl das Vorlesungsskript als auch die selbst angefertigten Mitschriften in die Befragung integriert. Diese „vorlesungsnahen“ Lernmaterialien erhalten – genauso wie die E-Lecture – ebenfalls eine signifikant schlechtere Beurteilung als die WBTs. Zudem werden WBT, Skript, Mitschrift und E-Lecture auch hier signifikant besser beurteilt als die Literatur, deren Beurteilung noch einmal deutlich schlechter ausfällt als bereits in der längerfristigen Evaluation.

Die Beurteilung des Vorlesungsskripts verändert sich von der längerfristigen zur abschließenden Evaluation nicht. Dies könnte daran liegen, dass die Studierenden das Vorlesungsskript bereits während des Semesters häufig genutzt und gut kennengelernt haben und sich so bereits am Ende des Semesters eine klare Meinung zur Qualität des Skriptes bilden konnten. Die E-Learning-Anwendungen konnten die Studierenden im Gegensatz dazu bis zur längerfristigen Beurteilung nur einmal nutzen und auch die Literatur wurde vermutlich – wenn überhaupt – von den Studierenden erst im Rahmen ihrer Klausurvorbereitung zu Rate gezogen. In allen drei Fällen sind folglich Veränderungen der Beurteilung durch die zum Teil umfangreichere Nutzung während des Sommersemesters nicht überraschend. Dies kann jedoch nicht für alle Studierenden stimmen, da viele die Anwendungen nicht zur Prüfungsvorbereitung genutzt haben. Bedeutsamer erscheint deshalb, dass sich die Studierenden während der drei Evaluationszeitpunkte in unterschiedlichen Situationen befanden: Während die kurz- und längerfristige

Akzeptanzerhebung noch in der Phase der „Wissensaufnahme“ stattfand, in der die Klausur noch in weiter Ferne lag, erfolgte die abschließende Akzeptanzerhebung in der Phase der „Prüfungsvorbereitung“. Die Studierenden werden deshalb ihre Zufriedenheit mit der Lehrmethode in der gerade vorliegenden Phase wiedergegeben haben. So ist es nicht überraschend, dass die Lehrmethoden, die vornehmlich zur Wissensvermittlung geeignet sind (bzw. den Studierenden geeignet erschienen), in der kurz- und längerfristigen Beurteilung besser abschnitten als in der abschließenden Evaluation, während das WBT insbesondere in der abschließenden Evaluation hervorstach – vermutlich weil es (auch wegen der z.T. vorhandenen Übungsaufgaben) zur Prüfungsvorbereitung als geeigneter eingestuft wurde. Auf diesen Punkt wird noch einmal ausführlicher in Abschnitt „3.2.5.3 Nutzungswunsch und tatsächliche Nutzung“ eingegangen.

3.2.5.2 Feinere Akzeptanzbeurteilungen

Der Vergleich weiterer Akzeptanzbeurteilungen ist nicht ebenso geradlinig möglich wie bei der Gesamtnote, da in den Evaluationen jeweils etwas andere Facetten abgefragt wurden. Hier ist zumeist lediglich ein Vergleich der kurz- und längerfristigen Urteile möglich, da aus Effizienzgründen im Rahmen der abschließenden Evaluation auf eine tiefergehende Abfrage verzichtet werden musste.

Generell zeigen die kurz- und längerfristigen Urteile zur Lernförderlichkeit, Qualität der Lernumgebung, Flexibilität, Gesamtakzeptanz sowie zur Möglichkeit, sich konzentrieren zu können Vorteile für die E-Learning-Anwendungen gegenüber der Vorlesung. Ähnlich der kurz- und längerfristigen Beurteilung der Gesamtnote für die Lehrmethoden ergibt sich auch hier folglich ein stabiles Evaluationsurteil. Im Gegensatz zur Gesamtbenotung schneidet die Vorlesung bezogen auf die feineren Akzeptanzurteile jedoch häufig schlechter ab als die E-Learning-Anwendungen. Die einzelnen Akzeptanzmaße sollen aus diesem Grunde an dieser Stelle noch einmal getrennt voneinander betrachtet werden.

Lernförderlichkeit

Die Lernförderlichkeit wurde nur kurzfristig von den Studierenden für die im Vordergrund betrachteten Lehrmethoden Vorlesung, E-Lecture und WBT beurteilt. Alle drei werden von den Studierenden relativ gut eingeschätzt. Dabei werden das WBT und die E-Lecture signifikant besser bewertet als die Vorlesung.

Die vorliegenden Ergebnisse unterstützen zum Beispiel die der Studien von Glowalla (2003), Glowalla et al. (2004) sowie Görlich (2006), nach denen die Lernenden die Effektivität und Konzentrationserhaltung der untersuchten E-Learning-Anwendungen ebenfalls besser einschätzten als die der Vorlesung. Dieses Ergebnis ist auch nicht überraschend, hält man sich vor Augen, dass die Inhalte in der Vorlesung sehr flüchtig durch

gesprochene Sprache dargeboten werden. Auch wenn immer die Möglichkeit zum Nachfragen besteht, werden einzelne Studierende Vieles, was sie nicht verstanden haben, später mit einem Lehrbuch, dem Internet oder in Lerngruppen noch einmal tiefergehend bearbeiten müssen. Ein Eingehen auf Verständnisschwierigkeiten von Einzelnen ist häufig nicht möglich – insbesondere da nach Glowalla et al. (2001) lediglich 4% aller Studierenden überhaupt in einer vollbesetzten Vorlesung eine Frage stellen.

Nach McLeish (1968) konnten Lernende nach einem 45-minütigen Vortrag lediglich 20% der Lerninhalte erinnern. Es ist folglich ausgeschlossen, dass Studierende alle Informationen einer 90-minütigen Vorlesung vollständig aufnehmen können. Die E-Learning-Anwendungen bringen im Gegensatz dazu den klaren Vorteil mit sich, dass komplexe Inhalte im eigenen Tempo bearbeitet und Unverstandenes beliebig oft wiederholt werden kann. Zudem können die E-Learning-Anwendungen zu der Zeit und an dem Ort bearbeitet werden, der für jeden einzelnen Studierenden persönlich am sinnvollsten zum Lernen ist, während die Vorlesung zeitlich und örtlich klar festgelegt ist. So kann der Zeitpunkt der Vorlesung für den einzelnen Studierenden nicht ganz optimal sein. Dies bestätigt sich auch, betrachtet man die Bewertung der Lehrmethoden bezüglich der Skala „Flexibilität“ (siehe folgender Abschnitt).

Hinzu kommt ein klarer Nachteil gerade dieser Vorlesung: Es handelt sich hier um eine sehr große Veranstaltung, an der insgesamt 1300 Studierende teilnehmen. Aus kapazitären Gründen fand die Vorlesung deshalb schon in zwei parallelen Veranstaltungen innerhalb eines Semesters statt. Dennoch saßen mindestens 500 Studierende gleichzeitig in einem Hörsaal, was ganz nachvollziehbar immer wieder zur Störungen in der Konzentration der einzelnen Lernenden führen muss. Auch wenn bereits oben gesagt wurde, dass die Größe des Auditoriums nicht direkt mit der Lernwirksamkeit eines Vortrages korreliert, beeinflusst sie dennoch die Möglichkeit Einzelner für persönliche Kontakte mit dem Dozenten (Bligh, 2000; McKeachie et al., 1986). Auch der Vortragende kann unter diesen Umständen nicht auf individuelle Unterschiede der Lernenden eingehen. Dieses Problem ist sicherlich in Vorlesungen mit weniger Teilnehmern nicht so dramatisch. Je mehr Lernende jedoch anwesend sind, desto mehr störende Unruhe entsteht und desto mehr fallen einzelne in eine rezeptive Haltung. Dass durch diese Faktoren das Lernförderlichkeitsempfinden der Studierenden beeinflusst wird, ist nachvollziehbar.

Gesamtakzeptanz (Qualität der Lernumgebung, Flexibilität)

Ebenfalls nur kurzfristig erfasst wurden die Skalen „Qualität der Lernumgebung“ sowie „Flexibilität“. Es ergibt sich für beide Unterskalen sowie die Gesamtskala ein signifikanter Unterschied zwischen den Lehrmethoden. Hinsichtlich der Qualität der Lernumgebung schneidet die E-Lecture signifikant besser ab als das WBT und die Vorlesung. Das WBT

wiederum wird signifikant besser beurteilt als die Vorlesung. Hinsichtlich der Flexibilität und der Gesamtakzeptanz schneiden beide E-Learning-Anwendungen bedeutsam besser ab als die Vorlesung.

Die Ergebnisse hinsichtlich der feineren Akzeptanzskalen stützen ebenfalls Studien, wie beispielsweise die von Glowalla (2003), welche einen Vorteil für die E-Learning-Anwendungen im Vergleich zu der Präsenzlehre ergaben. Sicherlich kann man an dieser Stelle argumentieren, dass es sich hierbei um einen Neuigkeitseffekt (Kerres, 2009) handeln könnte, welcher besagt, dass E-Learning aufgrund der Neuheit zwar zunächst gut ankommt – technische Probleme und zeitlicher Aufwand dann aber negativen Einfluss auf die Akzeptanz von E-Learning-Anwendungen haben können (Döring, 2002). Die heutigen Studierenden haben allerdings sicherlich schon Erfahrungen mit verschiedenen Formen von E-Learning gemacht, so dass es hier wie oben erläutert sinnvoller erscheint, von einer Art „Abwechslungseffekt" zu sprechen – ein Effekt, der dadurch zustande kommt, dass die Lehre „endlich" mal anders abläuft, als nur innerhalb der gewohnten Vorlesungsstruktur. Die längerfristigen Akzeptanzergebnisse sprechen jedoch gegen einen solchen Effekt. Wie bereits im Rahmen der Gesamtnote gesehen, werden die durch die Studierenden vergebenen Beurteilungen sogar tendenziell besser als schlechter.

Motivation und Spaß

Inwieweit die Arbeit mit den Lehrmethoden Vorlesung (bzw. Vorlesungsskript/Mitschriften), E-Lecture und WBT motivierend war und Spaß bereitet hat, wurde in allen drei Evaluationen abgefragt. Dabei schnitt die Vorlesung in der kurzfristigen Evaluation bezogen auf beide Variablen signifikant besser ab als die E-Learning-Anwendungen. Die E-Lecture machte den Studierenden damals zudem ebenfalls mehr Spaß als das WBT.

In der längerfristigen Evaluation zeigte sich ein ähnliches Bild – auch hier wurde die Vorlesung hinsichtlich beider Variablen signifikant besser beurteilt als die E-Lecture und das WBT. Zu diesem Messzeitpunkt ergab sich allerdings kein Unterschied zwischen den beiden E-Learning-Anwendungen.

Nachdem die Studierenden die E-Learning-Anwendungen dann allerdings für ihre Klausurvorbereitung genutzt hatten, war es nun das WBT, welches bezogen auf seine motivierende Wirkung und den Spaß beim Lernen am besten abschnitt. Signifikant schlechter schnitten das Skript, die Mitschriften sowie die E-Lecture ab. Am wenigsten motivierend und Spaß bringend wurde die Arbeit mit der Literatur eingestuft.

Auch hier zeigt sich folglich, dass die WBTs nach der freiwilligen Nutzung im Sommersemester deutlich besser beurteilt werden. Die Beurteilung der E-Lectures fällt im

Gegensatz dazu schlechter aus und die „vorlesungsnahen“ Lernmaterialien (Skript, Mitschriften) schneiden nun genauso wie die E-Lecture schlechter ab als das WBT.
Auch dies kann daran liegen, dass die WBTs für die Prüfungsvorbereitung geeigneter erschienen, während die anderen Materialien mehr für das Nachholen verpasster Vorlesungssitzungen eingesetzt wurden. Dementsprechend wurden die WBTs in der Phase der Prüfungsvorbereitung häufiger genutzt und – weil gerade in der vorliegenden Situation hilfreicher – auch besser beurteilt.
Die Tatsache, dass die E-Learning-Anwendungen – insbesondere die Web-based Trainings – hinsichtlich der Variablen „Motivation“ und „Spaß“ im Rahmen des Wintersemesters/der Wissensvermittlungsphase schlechter abschneiden als die Vorlesung, spricht gegen die Annahme von u.a. Döring (2002), dass die Lernmotivation von Lernenden durch die größeren Gestaltungsmöglichkeiten, die höhere Interaktivität und ggf. Übungsmöglichkeiten bei E-Learning-Anwendungen per se erhöht ist. Dabei muss man im Auge behalten, dass sich eine gute Vorlesung auch dadurch auszeichnet, dass der Vortragende die Lernenden durch seinen Enthusiasmus anstecken und für seine Inhalte motivieren kann (Glowalla, 2004). Dies scheint in der vorliegenden Vorlesung gelungen zu sein. Hier spielt sicherlich auch die Möglichkeit zum sozialen Austausch mit anderen Studierenden eine bedeutende Rolle.
Damit konnte zumindest ein Ziel von E-Learning – Anwendungen zu erstellen, die die Lernenden besser motivieren können als traditionelle Lernumgebungen (Wache, 2003) – zumindest für die kurzfristige Nutzung der Anwendungen nicht erreicht werden.

3.2.5.3 Nutzungswunsch und tatsächliche Nutzung

Die Mehrheit der Studierenden wünschen sich im Rahmen der kurz- und längerfristigen Evaluation Übungsaufgaben für die Vorbereitung auf die MAP. Die E-Learning-Anwendungen und das Vorlesungsskript werden ebenfalls häufig für die Klausurvorbereitung gewünscht.
Verpasste Sitzungen möchten die Studierenden allerdings lieber mit den E-Lectures nachholen und nicht verstandene Inhalte mit WBTs oder E-Lectures erarbeiten. Für die Wissensaneignung und den Austausch über den Lernstoff wünschen sich die meisten Studierenden die Vorlesung.
Die Studierenden möchten die Lehr-/Lernmaterialien folglich für unterschiedliche Einsatzzwecke nutzen. Dabei schätzen sie die Vor- und Nachteile der einzelnen Lehrmethoden und -materialien gut ein und wissen, welches Ziel sie mit welcher Methode am besten erreichen können. Die Literatur wird dabei durchweg am wenigsten gewünscht. Wenn überhaupt, dann würden die Studierenden sie zur Wissensaneignung und zum Erarbeiten nicht verstandener Inhalte einsetzen.

In der abschließenden Evaluation gaben die Studierenden dann allerdings an, die WBTs und das Skript am häufigsten für die Klausurvorbereitung genutzt zu haben – weniger häufig ihre eigenen Mitschriften, noch seltener die E-Lectures und am seltensten die Literatur.
Auch die Nutzeraufzeichnungen im Sommersemester zeigen, dass 57% der Studierenden die E-Learning-Anwendungen tatsächlich gar nicht nutzten. Etwa 40% von ihnen lernten allerdings mit den WBTs, weitere 10% mit den E-Lectures. Darin enthalten sind 6%, die für ihre Klausurvorbereitung sowohl die E-Lectures als auch die WBTs nutzten.
Die meisten Studierenden bereiteten sich demnach vermutlich ausschließlich mit dem Vorlesungsskript auf die abschließende Klausur vor. Dies lässt sich eindeutig dadurch begründen, dass das Vorlesungsskript sehr umfangreich und damit für die Klausurvorbereitung sehr hilfreich ist. Hierüber wurden die Studierenden sicherlich auch durch vorangegangene Kohorten informiert.
Obwohl die WBTs folglich im Rahmen der kurz- und längerfristigen Evaluation genauso häufig für die Klausurvorbereitung gewünscht wurden wie die E-Lectures, wurden sie tatsächlich deutlich häufiger hierfür genutzt. Der Wunsch, beide E-Learning-Anwendungen in gleicher Weise für die Klausurvorbereitung nutzen zu wollen, könnte aus dem Bedürfnis entstanden sein, die zu lernenden Inhalte jederzeit und so häufig wie nötig wiederholen und nachschauen zu können. Das jedenfalls bieten beide E-Learning-Anwendungen in ähnlicher Weise und könnte die vergleichbar guten Einschätzungen beider Anwendungen erklären. Dass die WBTs dann allerdings deutlich häufiger tatsächlich genutzt wurden, liegt vermutlich zum Teil daran, dass die Hälfte der WBTs Übungsaufgaben enthielten, die von den Studierenden sowohl in der kurz- als auch in der längerfristigen Evaluation am meisten für die Klausurvorbereitung gewünscht wurden. Vergleicht man die Nutzerzahlen für die WBTs mit und ohne Übungsaufgaben (siehe Tabelle 42), zeigt sich tatsächlich auch ein Unterschied in der Hinsicht, dass die WBTs mit Übungsaufgaben von mehr Studierenden genutzt wurden als die WBTs ohne Übungsaufgaben. Dieser Unterschied ist allerdings bei Weitem nicht so deutlich wie der Unterschied zwischen den WBTs insgesamt und den E-Lectures, so dass die Übungsaufgaben nicht alleine für die vermehrte Nutzung der Web-based Trainings verantwortlich sein können. Bedeutsam ist sicherlich auch, dass es die WBTs genauso wie das Vorlesungsskript ermöglichen, schnell und gezielt ein bestimmtes Thema nachzuschlagen und zu bearbeiten. Darüber hinaus ist auch denkbar, dass sich die Studierenden beim Lernen mit geschriebenem Text wohler fühlten, da sie hiermit naturgemäß mehr Erfahrung haben.
Bis auf die unterschiedliche Nutzungshäufigkeit von E-Lecture und WBT ergibt sich folglich ein sehr kongruentes Bild zwischen der Einstellungs- und Verhaltensakzeptanz der Studierenden. Es zeigt sich allerdings auch, wie wichtig die Erhebung der Verhaltensakzeptanz war: Durch die Aufzeichnung der tatsächlichen freiwilligen Nutzung

im Sommersemester können die Akzeptanzunterschiede zwischen der längerfristigen und der abschließenden Evaluation hinsichtlich der E-Lecture und des WBTs gut erklärt und untermauert werden. Dies wäre durch eine rein auf Fragebogendaten basierende Erhebung nicht möglich gewesen. Daneben ergibt die Aufzeichnung der Verhaltensakzeptanz natürlich auch noch viele weitere Daten, die Aufschluss über die mögliche Zielsetzung der Studierenden bei der Nutzung der E-Learning-Anwendungen geben. Dies wird im nächsten Abschnitt erläutert.

Nutzungsverhalten im Rahmen der Klausurvorbereitung

Die einzelnen WBT- und E-Lecture-Nutzer nutzten die jeweiligen Anwendungen in der Summe in etwa gleich lang. Die E-Lecture-Nutzer betrachteten jedoch wenige der E-Lectures für längere Zeit, während die WBT-Nutzer mehrere der WBTs für kürzere Zeit bearbeiteten. Wie oben bereits erwähnt, zeigt sich darüber hinaus, dass insbesondere die Anzahl der Nutzer als auch die mittlere Zugriffshäufigkeit für die WBTs mit Übungsaufgaben höher zu sein scheint als für die WBTs ohne Übungsaufgaben.

Erklärbar sind solche Ergebnisse dadurch, dass die E-Lecture- und WBT-Nutzer sicherlich unterschiedliche Ziele verfolgten: Vermutlich wollten die E-Lecture-Nutzer einzelne Vorlesungssitzungen vollständig nachholen oder wiederholen, während die WBT-Nutzer die WBTs aufriefen, um gezielt bestimmte Lerninhalte noch einmal nachzulesen oder lediglich die Übungsaufgaben zu bearbeiten. Insbesondere die Tatsache, dass die mittlere Zugriffshäufigkeit sowie die Anzahl der Nutzer für die WBTs mit Übungsaufgaben höher sind als für die WBTs ohne Übungsaufgaben, unterstützt diese Erklärung. Betrachtet man zudem die durchschnittliche Zeit, die sich im Gegensatz dazu die E-Lecture-Nutzer mit einer E-Lecture beschäftigt haben, wird diese Interpretation noch gestärkt. Die Nutzungszeiten liegen im Schnitt deutlich über der Rohlaufzeit der E-Lectures. Dies zeigt, dass sich die E-Lecture-Nutzer einzelne E-Lectures zumeist vollständig angesehen, zwischendurch auch mal gestoppt haben oder sogar noch einmal zurückgegangen sind, um sich einzelne Passagen mehrfach anzuschauen.

Die Studierenden nutzten die Anwendungen folglich scheinbar tatsächlich so, wie in der längerfristigen Evaluation gewünscht: die E-Lectures wurden hauptsächlich zum Nachholen verpasster Sitzungen und die WBTs (sowie die Übungsaufgaben innerhalb dieser) zur konkreten Prüfungsvorbereitung (selektives Nachlesen und Üben) genutzt.

Hieraus lässt sich auch die „schlechte Einstufung“ von Vorlesung und Literatur für die Klausurvorbereitung ableiten. Die weitaus schlechtesten Beurteilungen für die angegebene Literatur erklären sich dadurch, dass diese nicht speziell nur die Inhalte der Vorlesung wiedergibt, sondern eine Sammlung von Büchern etc. darstellt, die weit über diese Inhalte hinausgehen. Die Studierenden haben zumeist das Ziel, so zu lernen, dass sie die Klausur

(als extern vorgegebenes Lernziel) hinreichend gut absolvieren können. Nur wenige möchten die Inhalte aus eigenem Interesse vertiefend bearbeiten. Aus der angegebenen Literatur das herauszufiltern, was speziell für das Erreichen einer guten Note in einer Klausur benötigt wird, ist eine Aufgabe, die viele Lernende abschreckt.

Die Vorlesung ist hingegen am besten geeignet für die Einführung in oder den Überblick über ein Thema (Glowalla, 2004). Dass sie über dies allerdings nicht unbedingt für die konkrete Klausurvorbereitung das Mittel der Wahl ist, kommt schon dadurch zustande, dass den Lernenden häufig nicht klar wird, was von dem Vorgetragenen nun konkret für die Prüfung gelernt werden soll (Soliman, 1999). Zudem erhalten die Lernenden innerhalb einer Vorlesung keine Rückmeldung über ihren eigenen Lernfortschritt, was wichtig für den Lernprozess wäre. Um Wissen langfristig zu lernen, ist es außerdem unentbehrlich, sich aktiv mit dem Lernstoff auseinanderzusetzen – und auch dies kommt im Rahmen einer Vorlesung meist zu kurz (Glowalla et al., 2004).

Die Vorlesung wird allerdings im Gegensatz dazu am seltensten von allen Lehrmedien/-materialien „gar nicht" gewünscht. Dies zeigt, dass sie den Studierenden trotz des umfangreichen Angebotes an E-Learning-Unterstützung immer noch sehr wichtig ist. Die Befürchtung mancher Dozenten, dass Studierende ihre Präsenzveranstaltungen nicht mehr besuchen würden, wenn inhaltsgleiche E-Learning-Angebote zur Verfügung stünden (Bell et al., 2001), ist nach diesen Ergebnissen (vergleichbar zu den Ergebnissen von Glowalla, 2004) folglich völlig unbegründet. Studierende fühlen sich durch das reine Lernen mit E-Learning-Anwendungen isoliert oder vernachlässigt („I am neglected"-Effekt), wünschen sich soziale Kontakte sowie die Möglichkeit, direkt Fragen stellen zu können. Dies kann nur innerhalb von Präsenzveranstaltungen geboten werden. E-Learning-Anwendungen werden in der Folge hauptsächlich zur Ergänzung gewünscht, z.B. in Phasen der Prüfungsvorbereitung.

Nutzung außerhalb des vorliegenden Settings

Die Frage, ob sie sich die Methoden/Materialien auch für andere studienrelevante Module wünschen, bejahen die Studierenden am häufigsten in Bezug auf die Übungsaufgaben, das Vorlesungsskript sowie die E-Lecture. Für interessante Themen außerhalb der Universität möchten sie allerdings am liebsten WBTs und E-Lectures nutzen – am seltensten die Übungsaufgaben. Dies untermauert, welches Ziel Lernende in der Hochschule im Gegensatz zu Themen außerhalb der Hochschule erreichen möchten und bestätigt die Erklärungen für das schlechte Abschneiden von Literatur und Vorlesung für die Prüfungsvorbereitung: Innerhalb der Hochschule möchten die Studierenden sich möglichst effizient auf die Prüfung vorbereiten. Lernen geschieht leider hauptsächlich extrinsisch gesteuert und das Lernziel – die Kenntnis der spezifischen Prüfungsinhalte – wird von außen vorgegeben.

Dieses Ziel erreicht man am einfachsten, indem man wieder und wieder überprüft, ob man das Wissen, das der Lehrende als relevant vorgibt, beherrscht und nicht Gewusstes gezielt wiederholt.
Außerhalb der Hochschule beschäftigt man sich jedoch zumeist nur mit einem Thema, wenn man auch ein klares intrinsisches Interesse daran hat. Hier möchte man sich interessehalber tiefer mit einem Thema auseinandersetzen, ohne zu dem Gelernten geprüft zu werden. Es geht mehr darum, das individuell spannende Wissen herauszufiltern und zu lernen. Dabei sind insbesondere E-Lectures und WBTs attraktiv, da sie ein zeit- und ortsunabhängiges Studieren sowie ein selbständiges Lernen ohne Lehrperson ermöglichen.

Im folgenden Abschnitt soll nun geschaut werden, inwieweit sich die Studierenden unterschiedlich vor/während und nach der Bearbeitung der Lehrmethoden (im Wintersemester) verhalten haben. Dabei wird gezielt auf die subjektiv eingeschätzte Vor- und Nachbereitungszeit, das Anfertigen von Notizen sowie die Pausenhäufigkeit und -zeit eingegangen.

3.3 Verhalten vor, während und nach Bearbeitung/Besuch des Lehrangebotes

Hypothese 7a:	In der Vorlesung machen die Studierenden mehr Notizen als bei Bearbeitung der E-Lecture.
Hypothese 7b:	In der Vorlesung machen die Studierenden mehr Notizen als bei Bearbeitung des Web-based Trainings.
Hypothese 7c:	Bei Bearbeitung der E-Lecture machen die Studierenden mehr Notizen als bei Bearbeitung des Web-based Trainings.
Hypothese 8a:	Während der Nutzung der E-Learning-Anwendungen legen die Studierenden häufiger eine Pause ein als während des Vorlesungsbesuchs.
Hypothese 8b:	Während der Nutzung der E-Learning-Anwendungen legen die Studierenden längere Pausen ein als während des Vorlesungsbesuchs.

In Bezug auf das Verhalten vor/während und nach der Bearbeitung des Lehrmediums bzw. dem Besuch der Vorlesung interessierte insbesondere, ob und wie viel Vor- und Nachbereitungszeit die Studierenden investiert, inwieweit und wie viele Pausen sie eingelegt und ob sie währenddessen oder danach Notizen angefertigt hatten. Tabelle 47 zeigt die deskriptiven Ergebnisse zu diesen vier Variablen.

Tabelle 47
Deskriptive Statistik zur Vor- und Nachbereitungszeit, Pausenhäufigkeit und -zeit sowie dem Machen von Notizen während der Vorlesung, der Bearbeitung der E-Lecture und dem WBT

	Vorlesung			E-Lecture			WBT		
	n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
Vor-/Nachbereitung	972	29.70	40.66	1017	32.55	42.47	996	**34.07**	32.47
Pausen	983	0.09	0.28	1033	**0.54**	0.50	1009	0.47	0.50
Pausenzeit	219	7.86	67.56	554	23.48	47.51	556	**24.95**	48.30
Notizen	982	**0.88**	0.32	1032	0.82	0.38	1003	0.80	0.40

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Auf den ersten Blick scheinen die Studierenden im Rahmen der Nutzung der E-Learning-Anwendungen mehr Vor- und Nachbereitungszeit investiert zu haben. Auch die angegebene Pausenhäufigkeit und die Pausenlänge scheinen bei Bearbeitung der E-Learning-Anwendungen höher ausgefallen zu sein als beim Besuch der Vorlesung. Das Anfertigen der Notizen scheint hingegen im Rahmen der Vorlesung etwas stärker ins Gewicht gefallen zu sein als bei der Bearbeitung der E-Learning-Anwendungen.

Um diese augenscheinlichen Unterschiede auf Signifikanz zu überprüfen, wurden für alle vier Variablen erneut abhängige Varianzanalysen bzw. Cochran Q-Tests berechnet. Die Ergebnisse dieser Analysen sind in Tabelle 48 abgetragen.

Tabelle 48
Ergebnisse der Mittelwertsanalysen zum Vergleich der Vor- und Nachbereitungszeit, Pausenhäufigkeit und -zeit sowie Notizenhäufigkeit während der Vorlesung, der Bearbeitung der E-Lecture und dem WBT

	n	Vorlesung		E-Lecture		WBT		ANOVA / Cochran-Test				Posthoc
		M	*SD*	*M*	*SD*	*M*	*SD*	*F/Q*	*df*	*p*	*Eta²/w*	
Vor-/Nach-bereitung	524	28.24	25.74	30.35	28.60	**32.16**	30.14	6.12	1.97	<.01*	.01	WBT>VL
Pausen	535	0.07	0.26	**0.53**	0.50	0.44	0.50	265.56	2	<.01*	.70	EL>WBT>VL
Pausenzeit	41	3.63	6.44	21.02	20.66	**22.38**	40.68	6.48	1.33	<.01*	.14	EL,WBT>VL
Notizen	532	**0.90**	0.31	0.84	0.37	0.80	0.40	24.51	2	<.01*	.21	VL>EL,WBT

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *Q* = Teststatistik des Cochran-Tests, *df* = Freiheitsgrade; *p* = Signifikanz, *Eta²* bzw. w = Effektstärke

Die Studierenden gaben folglich nach Bearbeitung der WBTs eine signifikant höhere Vor- und Nachbereitungszeit, Pausenzeit sowie -häufigkeit an als nach dem Besuch der Vorlesung. Während der Bearbeitung der E-Lectures machen sie ebenfalls häufiger und länger Pause als während des Besuchs der Vorlesung. Die Pausenhäufigkeit ist bei der Bearbeitung der E-Lecture ebenfalls höher als bei der Bearbeitung des WBTs. Die Effektstärke des Unterschieds bezüglich der Vor- und Nachbereitungszeit ist allerdings lediglich als gering einzustufen. Hinsichtlich der Unterschiede in Bezug auf die Pausenhäufigkeit und -zeit ergeben sich hingegen jeweils große Effektstärken. Die Hypothesen 8a und 8b sind demnach beizubehalten.

Im Gegensatz dazu werden während der Vorlesung signifikant häufiger Notizen gemacht als bei der Bearbeitung der E-Learning-Anwendungen. Dieser Unterschied ist als gering bis mittelstark einzustufen.
Die Hypothesen 7a und 7b sind folglich beizubehalten, Hypothese 7c ist zu verwerfen.

3.3.1 Zusammenfassung und Diskussion

3.3.1.1 Vor- und Nachbereitungszeit

Die Tatsache, dass Studierende, die die WBTs bearbeiteten, eine signifikant höhere Vor- und Nachbereitungszeit angaben, als Studierende, die die Vorlesung besucht hatten, ist sicherlich zum Teil darauf zurückzuführen, dass die Studierenden tatsächlich längere Zeit mit den WBTs verbrachten als der Besuch der Vorlesung in Anspruch genommen hätte (siehe „Lerneffizienz“). Diese Zeit entspricht zwar nicht der Vor- und Nachbereitungszeit, es ist aber vorstellbar, dass die Studierenden hier die Bearbeitungszeit zur Vor- und Nachbereitungszeit gezählt haben, da sie bereits während der Bearbeitung des WBTs zuhause am Schreibtisch saßen, um die Inhalte zu bearbeiten. Für die Vorlesung ist der konkrete Besuch der Vorlesung für die Studierenden im Gegensatz dazu klar trennbar von der Vor- und Nachbereitungsphase.
Darüber hinaus könnte die Arbeit mit den WBTs als anstrengender wahrgenommen worden sein als das Zuhören bei Vorlesung oder E-Lecture. Dadurch könnte ebenfalls die subjektive Einschätzung der Vor- und Nachbereitungszeit der Lehrmethoden beeinflusst sein. Die subjektive Erfassung der Vor- und Nachbereitungszeit erscheint demnach im Nachhinein wenig verlässlich.

3.3.1.2 Pausenhäufigkeit und -zeit

Hinsichtlich der Pausenhäufigkeit und -zeit gaben beide E-Learning-Gruppen höhere Werte an als die Vorlesungsgruppe. Dabei machen die Studierenden, die die E-Lecture bearbeiteten, signifikant am häufigsten Pause. Dies ist gar nicht überraschend, da es schlichtweg wenig attraktiv und sozial unerwünscht ist, aus einer Vorlesung herauszugehen als zuhause im Rahmen der Bearbeitung einer E-Learning-Anwendung eine Pause einzulegen. Die Möglichkeiten zur Nutzung einer solchen Pause sind im heimischen Rahmen viel attraktiver. Zudem kommt hier ein wichtiger Vorteil von E-Learning-Anwendungen zum Tragen: Wer in der Vorlesung eine Pause macht, verpasst Inhalte. Wer während der Bearbeitung von E-Learning-Anwendungen eine Pause einlegt, kann danach einfach dort weitermachen, wo er aufgehört hat.

3.3.1.3 Anfertigen von Notizen

Während einer Vorlesung werden signifikant häufiger Notizen angefertigt als während der Bearbeitung einer E-Learning-Anwendung. Die Tatsache, dass die Inhalte in der Vorlesung nur einmal dargeboten werden, scheint demnach dazu zu führen, dass die Studierenden in der Vorlesung vermehrt Notizen anfertigen. Die flüchtiger erscheinenden (weil gesprochenen) Inhalte im Rahmen der E-Lecture führten allerdings nicht dazu, dass die Studierenden bei der Bearbeitung der E-Lecture mehr Notizen anfertigten als bei Bearbeitung des Web-based Trainings mit geschriebenem Text. Bei der Entscheidung, Notizen zu machen oder nicht, spielt folglich nicht so sehr die Darstellung der Inhalte an sich eine Rolle, sondern eher die Frage, ob es möglich ist, die Inhalte beliebig häufig wiederholen zu können.

3.4 Übungsaufgaben

3.4.1 Beurteilung der Übungsaufgaben durch die Studierenden

Die Studierenden konnten in einem Teil der WBTs Übungsaufgaben bearbeiten. Darüber hinaus ist es wahrscheinlich, dass sie auch die kurzfristigen Lernleistungstests am Tag nach der Bearbeitung der E-Learning-Anwendungen bzw. dem Besuch der Vorlesung als Übungstests wahrnahmen. Deshalb sollte im Rahmen der kurzfristigen Evaluation eine feinere Beurteilung der Übungsaufgaben innerhalb der WBTs und der kurzfristigen Nachtests durch die Studierenden erfasst werden, um eine genauere Aussage über die Güte der Aufgaben aus Studierendensicht zu erhalten. Tabelle 49 gibt die durchschnittliche Bewertung der Aufgaben hinsichtlich verschiedener Aspekte wieder.

Tabelle 49
Deskriptive Statistik der Variablen zur Beurteilung der Übungsaufgaben

Beurteilung der Übungsaufgaben	***n***	***M***	***SD***
Fanden Sie die Übungsaufgaben sinnvoll, um das Thema des Vortages zu wiederholen?	939	3.45	0.69
Fanden Sie die Übungsaufgaben sinnvoll, um das Thema des Vortages besser zu verstehen?	941	2.98	0.89
Wie bewerten Sie den Schwierigkeitsgrad der Übungsaufgaben?	942	3.09	0.51
Wie bewerten Sie den inhaltlichen Umfang der Übungsaufgaben?	939	2.09	0.41
Fanden Sie das unmittelbare Feedback nach Bearbeitung jeder Frage hilfreich?	942	3.70	0.59
Fanden Sie das Feedback am Ende des Tests hilfreich?	940	3.28	0.88

Anmerkung. n = Stichprobengröße, M = Mittelwert, SD = Standardabweichung; Skalierung der Items 1, 2, 5 und 6: 1 = Nein, 2 = Eher nein, 3 = Eher ja, 4 = Ja; Skalierung Item 3: 1 = sehr leicht, 2 = eher leicht, 3 = eher schwer, 4 = sehr schwer, Skalierung Item 4: 1 = zu oberflächlich, 2 = angemessen, 3 = zu detailliert

Wie zu sehen ist, werden die Fragen von den Studierenden sehr positiv wahrgenommen. Im Schnitt finden die Studierenden die Aufgaben für das Wiederholen und das bessere Verstehen des Lernstoffes sinnvoll. Ihr Schwierigkeitsgrad wird dabei als schwer bewertet.

Inhaltlich werden die Aufgaben allerdings als angemessen wahrgenommen. Auch die Feedbacks nach jeder Frage und am Ende des Tests werden als hilfreich eingeschätzt.
Wie in Abschnitt „3.2.5 Zusammenfassung und Diskussion“ bereits erläutert, wurden von den Studierenden im Rahmen der kurz-, längerfristigen und langfristigen Evaluation weitere vergleichende Urteile zu den Aufgaben abgegeben. Auch diese sprechen für eine sehr hohe Akzeptanz der Aufgaben und bestätigen die hier dargestellten Ergebnisse.

3.4.2 Vergleich der WBTs mit und ohne Übungsaufgaben

Hypothese 3a: Kurzfristig erreichen WBTs mit Übungsaufgaben eine höhere Lerneffektivität als WBTs ohne Übungsaufgaben.
Hypothese 3b: Längerfristig erreichen WBTs mit Übungsaufgaben eine höhere Lerneffektivität als WBTs ohne Übungsaufgaben.
Hypothese 3c: Kurzfristig erreichen WBTs mit Übungsaufgaben eine höhere Lerneffizienz als WBTs ohne Übungsaufgaben.
Hypothese 3d: Längerfristig erreichen WBTs mit Übungsaufgaben eine höhere Lerneffizienz als WBTs ohne Übungsaufgaben.

Hinsichtlich aller relevanten Variablen der kurzfristigen Erhebung sowie der Probeklausur wurde zudem betrachtet, inwieweit sich die Beurteilung von WBTs mit und ohne Übungsaufgaben unterscheiden. Es zeigten sich lediglich bei zwei Variablen aus dem Akzeptanzbereich signifikante Unterschiede (siehe Tabelle 50, die Ergebnisse des vollständigen Vergleiches befinden sich in Anhang B, Tabelle B14).

Tabelle 50
Deskriptive Statistik und Ergebnisse der Mittelwertsanalysen zu den Lernleistungsvariablen sowie den signifikanten Akzeptanzvariablen zum Vergleich der WBTs mit und ohne Übungsaufgaben

		n	*M*	*SD*	*t*-Test			
					t	*df*	*p*	*d*
Gesamtakzeptanz	0	433	29.51	4.95	-2.00	790	.05*	.14
	1	359	**30.20**	4.63				
Eignung zur Wissensvermittlung	0	418	3.18	0.90	-2.31	766	.02*	.16
	1	350	**3.32**	0.84				
Kurzfristige Lernleistung	0	443	**68.43**	22.77	0.90	804	.37	-
	1	363	67.02	21.45				
Längerfristige Lernleistung	0	188	**49.04**	19.01	1.19	350	.24	-
	1	164	46.54	20.47				

Anmerkung. 0 = WBTs ohne Übungsaufgaben, *1* = WBTs mit Übungsaufgaben, *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *t* = Teststatistik des *t*-Tests, *df* = Freiheitsgrade; *p* = Signifikanz, *d* = Effektstärke

Die WBTs mit Übungsaufgaben erreichten eine höhere Gesamtakzeptanz bei den Studierenden und die Studierenden schätzten nach Bearbeitung der WBTs mit Übungsaufgaben die WBTs insgesamt besser zur Wissensvermittlung ein (jeweils begleitet von einer geringen Effektstärke). Die wenigen zu beobachtenden signifikanten Unterschiede liegen folglich im Akzeptanzbereich. Hinsichtlich der kurz- und

längerfristigen Lernwirksamkeit unterscheiden sich die beiden WBTs mit und ohne Übungsaufgaben nicht signifikant voneinander.

Betrachtet man den wöchentlichen Vergleich der kurzfristigen Lernwirksamkeit zwischen den drei Lehrmethoden WBT, E-Lecture und Vorlesung differenzierter (siehe „3.1.1 Kurzfristige Lerneffektivität und -effizienz"), ergeben sich auch hier zweimal signifikante Vorteile für ein WBT mit Übungsaufgaben und zweimal für ein WBT ohne Übungsaufgaben. Im längerfristigen Vergleich ist lediglich ein Unterschied für ein WBT mit Übungsaufgaben gegenüber der inhaltlich gleichen E-Lecture signifikant. Diese Ergebnisse sprechen ebenfalls gegen einen generellen Vorteil der WBTs mit Übungsaufgaben im Hinblick auf die Lernwirksamkeit. Die Hypothesen 3a und 3b müssen dementsprechend verworfen werden.

3.4.2.1 Lerneffizienz

Wie bereits im Rahmen der allgemeinen Auswertung zur Lerneffizienz berichtet, kann die Lerneffizienz lediglich für die Themen 7 bis 12 ausgewertet werden. So gehen in die folgende Auswertung drei WBTs ohne Übungsaufgaben (Themen 10 bis 12) und drei WBTs mit Übungsaufgaben (Themen 7 bis 9) ein.

Tabelle 51
Deskriptive Statistik sowie Ergebnisse der Mittelwertsanalysen zum Vergleich der Lernzeit und der kurz- und längerfristigen Lerneffizienz von WBTs mit und ohne Übungsaufgaben

		n	*M*	*SD*	Welch-Test			
					t	*df*	*p*	*d*
Lernzeit	0	172	3569.58	3334.40	-5.15	384.72	<.01*	.52
	1	215	**5503.74**	4058.72				
Kurzfristige Lerneffizienz	0	170	**31.61**	29.53	4.18	319.24	<.01*	.44
	1	212	19.98	23.59				
Längerfristige Lerneffizienz	0	82	**1.28**	1.54	3.41	131.41	<.01*	.53
	1	91	0.61	0.94				

Anmerkung. 0 = WBTs ohne Übungsaufgaben, *1* = WBTs mit Übungsaufgaben, *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *t* = Teststatistik des Welch-Tests, *df* = Freiheitsgrade; *p* = Signifikanz, *d* = Effektstärke

Nach diesen Ergebnissen (siehe Tabelle 51) sind die WBTs ohne Übungsaufgaben sowohl kurz- als auch längerfristig signifikant lerneffizienter (mittlere Effektstärke). Dies ist vor allem auf die deutlich längere Bearbeitungszeit der WBTs mit Übungsaufgaben zurückzuführen.

Dieses Ergebnis ist allerdings mit Vorsicht zu interpretieren, da ein starker inhaltlicher Einfluss anzunehmen ist: Gerade die Themen 7 („Denken und Problemlösen, Intelligenz" – 40 Inhaltsseiten), 8 („Entwicklungspsychologie" – 27 Inhaltsseiten) und 9 („Unterrichtsmethoden und -organisation" – 31 Inhaltsseiten), die alle Übungsaufgaben enthielten, waren inhaltlich eher umfangreich, so dass die Studierenden sie über alle drei Projektgruppen tendenziell länger bearbeiten mussten. Gleichzeitig gehörten die Themen 10 („Lernen mit Medien" – 21 Inhaltsseiten), 11 („Pädagogische Diagnostik" – 21

Inhaltsseiten) und 12 („Rückblick“ – 31 Inhaltsseiten, die aber mit vielen Abbildungen und wenig Text gefüllt sind) – für welche WBTs ohne Übungsaufgaben erstellt worden waren – zu den inhaltlich eher kürzeren WBTs. Abbildung B1 (siehe Anhang B) zeigt, dass auch die Vorlesungsgruppe insbesondere für die Themen 7 und 9 längere, für die Themen 10 und 12 eher kürzere Lernzeiten aufweist. Da leider für die WBTs zu den Themen 1 bis 6 – welche eine eher vergleichbare Länge aufweisen – zum großen Teil keine Zeitdaten vorliegen, verbietet sich aufgrund der vorliegenden Auswertung eine endgültige Entscheidung über die Hypothesen 3c und 3d.

Aus diesem Grunde wurden die Lernzeiten für die sechs betrachteten WBTs anhand der Anzahl ihrer Inhaltsseiten relativiert und hieraus sowohl für die kurz- als auch die längerfristige Lerneffizienz eine neue „inhaltsneutrale Lerneffizienz“ berechnet. Die WBTs mit und ohne Übungsaufgaben wurden anhand dieser Variablen erneut miteinander verglichen. Tabelle 52 gibt die Ergebnisse dieser Analysen wieder.

Tabelle 52
Deskriptive Statistik sowie Ergebnisse der Mittelwertsanalysen zum Vergleich der anhand der Inhaltsseiten relativierten Lernzeit und Lerneffizienz von WBTs mit und ohne Übungsaufgaben

		n	*M*	*SD*	*t*-/Welch-Test			
					t	*df*	*p*	*d*
Inhaltsneutrale Lernzeit	0	172	148.43	128.37	-1.54	385	.12	-
	1	215	**168.12**	122.09				
Inhaltsneutrale kurzfristige Lerneffizienz	0	170	**751.10**	717.14	1.84	380	.07	-
	1	212	620.63	667.06				
Inhaltsneutrale längerfristige Lerneffizienz	0	82	**31.94**	42.87	2.33	133.84	.02*	.36
	1	91	19.11	26.99				

Anmerkung. 0 = WBTs ohne Übungsaufgaben, *1* = WBTs mit Übungsaufgaben, *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *t* = Teststatistik des t-/Welch-Tests, *df* = Freiheitsgrade; *p* = Signifikanz, *d* = Effektstärke

Es zeigt sich, dass sich die WBTs mit und ohne Übungsaufgaben nach dieser Korrektur tatsächlich nicht mehr signifikant hinsichtlich ihrer Lernzeit unterscheiden. Auch in Bezug auf die „inhaltsneutrale kurzfristige Lerneffizienz“ ergibt sich nun keine Signifikanz mehr zwischen beiden WBT-Varianten. Längerfristig sind die WBTs ohne Übungsaufgaben allerdings immer noch signifikant lerneffizienter als die WBTs mit Übungsaufgaben. Hierbei handelt es sich allerdings nur noch um einen kleinen bis mittelstarken Effekt. Dennoch sind die Hypothesen 3c und 3d nach diesen Ergebnissen zu verwerfen.

3.4.3 Zusammenfassung und Diskussion

3.4.3.1 Beurteilung der Übungsaufgaben

Bereits Abschnitt „3.2 Akzeptanz“ zeigt, dass die Übungsaufgaben insgesamt sehr gut von den Studierenden beurteilt werden. Sie werden für die Klausurvorbereitung als am sinnvollsten eingeschätzt, sind nach Meinung der Studierenden am meisten motivierend

und ihre Bearbeitung macht (nach dem Besuch der Vorlesung!) am meisten Spaß. Es ist daher nicht verwunderlich, dass die meisten Studierenden zur Vorbereitung auf die MAP Übungsaufgaben nutzen möchten.

Auch Abschnitt „3.4 Übungsaufgaben“ zeigt, dass die Studierenden die Übungsaufgaben für die Wiederholung und das bessere Verstehen des Lernstoffes als sinnvoll erachten. Darüber hinaus schätzen sie die Übungsaufgaben als inhaltlich angemessen ein. Auch das Feedback, das nach jeder Frage und am Ende eines jeden Tests gegeben wurde, wird als hilfreich bewertet.

Die besonders guten Urteile zu den Übungsaufgaben sind wenig verwunderlich und bestätigen die Studien von zum Beispiel Glowalla et al. (2004), Görlich (2006) oder Klauser et al. (2002), welche ebenfalls eine hohe Akzeptanz für Übungsaufgaben bei den Lernenden feststellten. Schon von Glowalla et al. (2002) wird berichtet, dass Studierende sich zur Überprüfung ihres Wissens hauptsächlich an die Bearbeitung von Übungsaufgaben halten. Sie freuen sich über alle Angebote, mit denen sie ihr Wissen überprüfen können, ohne gleichzeitig benotet zu werden (Glowalla et al., 2004).

Der Schwierigkeitsgrad der Übungsaufgaben wird als schwer angegeben, was allerdings nicht unbedingt ein schlechtes Urteil darstellt, da die Studierenden auf diese Weise herausgefordert werden. Dies lässt das Abschneiden der Studierenden in den Wissenstests (ca. 45-65% richtige Antworten) in noch positiverem Licht erscheinen.

3.4.3.2 Vergleich der WBTs mit und ohne Übungsaufgaben

Im Vergleich zu den WBTs ohne Übungsaufgaben erreichen die WBTs mit Übungsaufgaben tatsächlich eine höhere Gesamtakzeptanz. Zudem schätzten die Studierenden diese WBTs insgesamt besser zur Wissensvermittlung ein als die WBTs ohne Übungsaufgaben. Zu allen anderen Variablen – auch zur Lernwirksamkeit – ergaben sich keinerlei weitere Signifikanzen zwischen den WBTs mit und ohne Übungsaufgaben. Hier ist erstaunlich, dass sich nicht auch bezogen auf die Lernförderlichkeitsskala, die Qualität der Lernumgebung oder den Nutzungswunsch für die Klausurvorbereitung weitere Signifikanzen ergeben haben. Unterschiede ergaben sich folglich nur wenige und diese lediglich im Rahmen einzelner Aspekte der Akzeptanzbefragung.

Alle Ergebnisse sprechen folglich nur für einen leichten Akzeptanzvorteil und nicht für einen Lernwirksamkeitsvorteil von WBTs mit Übungsaufgaben im Vergleich zu WBTs ohne Übungsaufgaben.

Dabei müssen natürlich auch inhaltliche Aspekte berücksichtigt werden. Es ist durchaus möglich, dass für die (nicht vorhandenen) Unterschiede die behandelten Themen verantwortlich sein könnten. So waren insbesondere die Themen 7 und 9, die inhaltlich sehr umfangreich waren und vermutlich deshalb (über alle drei Experimentalgruppen) zu

tendenziell schlechteren Testergebnissen führten, beide mit Übungsaufgaben ausgestattet. Alle anderen Themen waren wesentlich weniger umfangreich und verlangten den Studierenden folglich nicht so viel ab.

Auch beim Blick auf die Lernzeit sowie die kurz- und längerfristige Lerneffizienz ergab sich (nach einer Relativierung der WBTs in Bezug auf ihre Inhaltsseitenanzahl) nur noch hinsichtlich der längerfristigen Lerneffizienz ein signifikanter Vorteil der WBTs ohne gegenüber den WBTs mit Übungsaufgaben. Dieser war zudem begleitet von einer geringen Effektstärke.

Die Ergebnisse zur Lerneffizienz sprechen gegen die von Siegert et al. (2003), da die Autoren hinsichtlich der Lerneffizienz von WBTs mit und ohne Übungsaufgaben keine Unterschiede beobachten konnten. Sie sprechen auch gegen die Ergebnisse von Hamaker (1986), der in seiner Metaanalyse Lernwirksamkeitsvorteile von Übungsaufgaben auch dann beobachten konnte, wenn die Lernzeit für das Lernen eines Inhaltes durch die Versuchsleitung klar festgelegt war – es folglich für die Lernenden keine Möglichkeit gab, sich länger oder kürzer mit dem Lerninhalt auseinander zu setzen. An dieser Stelle wäre es sinnvoll, die Effekte erneut mit inhaltsgleichen WBTs mit und ohne Übungsaufgaben zu überprüfen, um Inhaltseffekte gänzlich ausschließen zu können. Schließlich kann eine Korrektur ausschließlich über die Anzahl der Inhaltsseiten Aspekte wie Komplexität oder Schwierigkeit der fachlichen Inhalte nicht berücksichtigen.

4 Studie I: Fazit

Ziel der ersten Studie war es, einen Vergleich zwischen zwei E-Learning-Anwendungen und der traditionellen Präsenzlehre in der Hochschule anzustellen. Als E-Learning-Anwendungen wurden Lehrmethoden gewählt, die häufig in der Praxis genutzt werden. Dies war zum einen die E-Lecture, die in der Darstellung der Inhalte der Vorlesung sehr ähnlich ist und häufig in der Hochschullehre zum Einsatz kommt. Zum anderen wurde das Web-based Training ausgewählt, da es die Lerninhalte deutlich anders als die E-Lecture und die Vorlesung darbietet und vornehmlich in der beruflichen Weiterbildung – vermehrt aber auch in der schulischen und hochschulischen Ausbildung – eingesetzt wird. Da die Studie als Felduntersuchung begleitend zu einer Großveranstaltung angelegt war, war es möglich, auf eine sehr umfangreiche Stichprobe von über 1.300 Studierenden zurückzugreifen. Dieses Studiendesign brachte einige Vorteile mit sich: So konnten innerhalb desselben Semesters mehrere Replikationen der Studie durchgeführt und der Versuchsplan so angelegt werden, dass sowohl quer- als auch längsschnittliche inhalts- und materialinvariante Vergleiche realisierbar waren. Zudem konnte die Akzeptanz und die Lernleistung der Studierenden in ihrer realen Lernsituation untersucht werden. Dies dürfte eine hohe ökologische Validität der Ergebnisse sicherstellen.
An dieser Stelle soll nun ein abschließendes Fazit aus den Ergebnissen der ersten Studie gezogen werden. Dabei werden zunächst abschließend die Ergebnisse zur Lernwirksamkeit und Akzeptanz beleuchtet. Im Anschluss werden Einschränkungen und Implikationen der Studie berichtet.

4.1 Lernwirksamkeit

Das Hauptaugenmerk der Studie lag auf der Lerneffektivität und -effizienz der drei Lehrmethoden Vorlesung, E-Lecture und Web-based Training. Unabhängig von der Lehrmethode erreichten die Studierenden innerhalb der kurz- und längerfristigen Erhebung etwa 50% bis 60% der möglichen Punktzahl. Die Lernwirksamkeit aller drei Lehrmethoden ist damit insgesamt als positiv zu beurteilen – und das, obwohl die Klausur erst ein halbes Jahr später anstand und die Studierenden dementsprechend noch nicht besonders stark motiviert gewesen sein dürften, die Inhalte schon zu diesem frühen Zeitpunkt vollständig zu verinnerlichen. Dies zeigt, dass alle drei Lehrmethoden die Inhalte sehr lernförderlich vermitteln konnten.
Ein eindeutiger Lerneffektivitätsvorteil einer Lehrmethode konnte nicht nachgewiesen werden. Hinsichtlich der Lerneffizienz zeigte sich jedoch ein klarer Vorteil der WBTs

gegenüber der Vorlesung. Dabei wird deutlich, dass es immer die E-Learning-Anwendungen (insbesondere die Web-based Trainings) sind, die bei signifikanten Lernwirksamkeits-unterschieden besser abschneiden. Die E-Learning-Anwendungen konnten folglich mindestens gleich gute – zum Teil auch bessere – Lernergebnisse bewirken. Alle drei Lehrmethoden können so mit gutem Gewissen eingesetzt werden.

Die kurzfristig besseren Leistungen der Studierenden, die mit den WBTs gelernt haben, könnten daher rühren, dass sie mehr als die anderen beiden Lehrmethoden erlauben, die Inhalte im eigenen Tempo zu bearbeiten. Auch wenn es im Rahmen der E-Lecture möglich ist zu pausieren und zurückzuspulen, wird dies nur selten gemacht (Görlich, 2006). Während der Nutzung der E-Lecture (oder auch dem Besuch der Vorlesung) entsteht womöglich zudem – ähnlich dem Schauen eines Films – die Gefahr des „Berieseln Lassens“ (siehe auch Salomon, 1984). Beim Lernen mit dem WBT *muss* allerdings Anstrengung investiert werden, um die Inhalte zu konsumieren. Dieses unterschiedliche Lernverhalten wird sicherlich Auswirkungen auf die Lernwirksamkeit haben.
Diese Erklärungsversuche stellen allerdings lediglich plausible Annahmen dar, die im Rahmen weiterer Studien belegt werden müssten. Um hierzu endgültige Aussagen zu machen, müsste mittels Videoanalyse oder/und Eye-Tracking das Verhalten der Nutzer bei der Bearbeitung der E-Learning-Anwendungen systematisch beobachtet werden. Da es aber auch durch solche Untersuchungssettings nicht möglich ist, Lernenden „in den Kopf zu schauen“, muss die Erhebung der Lernwirksamkeit auch in einer solchen Studie berücksichtigt werden.
Dass die Lernwirksamkeit längerfristig vergleichbar gut ausfällt, spricht vermutlich dafür, dass im längerfristigen Rahmen durch die Studierenden ggf. Kompensationen vorgenommen wurden. Evtl. haben die Studierenden insbesondere die Themen, die sie in der Vorlesung oder mittels E-Lecture zunächst entspannt konsumiert haben, dann vor Bearbeitung der Probeklausur noch einmal ausführlicher mit dem Vorlesungsskript nachbearbeitet als die Themen, für deren „Erstlernen“ sie bereits beim Lesen des WBTs Anstrengung investiert hatten. Solche Kompensationen durch mehr oder weniger fleißiges Arbeiten oder die Nutzung anderer Materialien können insbesondere vor der letzten Probeklausur und der MAP – hier wurde es ja erst richtig „ernst“ für die Studierenden – nicht ausgeschlossen werden. Diese Kompensationen hätte man zwar innerhalb einer kontrollierten Laboruntersuchung weitgehend ausschließen können. Dies wäre aber wiederum zulasten anderer Faktoren gegangen – allen voran der ökologischen Validität der Studie.

Und auch die Rolle der Übungsaufgaben innerhalb der WBTs kann nicht abschließend beurteilt werden. Nach vielen anderen Studien (z.B. Glowalla et al., 2004; Görlich, 2006 oder Klauser et al., 2002) konnte auch die vorliegende Studie zeigen, dass Übungsaufgaben bei Lernenden eine hohe Akzeptanz besitzen. Es konnten allerdings keine Signifikanzen zwischen den WBTs mit und ohne Übungsaufgaben hinsichtlich der Lernwirksamkeit festgestellt werden und bezogen auf die Lerneffizienz stellten sich sogar die WBTs ohne Übungsaufgaben längerfristig als vorteilhaft heraus. Wie bereits oben diskutiert (siehe Abschnitt „3.4 Übungsaufgaben"), muss in diesem Zusammenhang eine inhaltliche Konfundierung in Betracht gezogen werden. Deshalb wäre es an dieser Stelle sinnvoll, die Effekte erneut mit inhaltsgleichen WBTs mit und ohne Übungsaufgaben zu überprüfen, um Inhaltseffekte gänzlich ausschließen zu können.

4.2 Akzeptanz

Die Ergebnisse belegen eine gute Akzeptanz aller drei Lehrmaterialien/-methoden (Vorlesung, E-Lecture und WBT). Dabei zeigen die verschiedenen Materialien/Methoden jeweils unterschiedliche Stärken und Schwächen. So „brillieren" die E-Learning-Anwendungen zum Beispiel hinsichtlich ihrer Lernförderlichkeit und Flexibilität, während die Vorlesung von den Studierenden bezüglich ihres Spaßfaktors und ihrer motivierenden Wirkung besser beurteilt wird. Grundsätzlich kann man hieraus den Schluss ziehen, dass die Entwicklung von Blended Learning-Szenarien, innerhalb denen Vorteile aller Lehrmethoden genutzt und Nachteile ausgeglichen werden können, durchaus ihre Berechtigung hat. Die Befürchtung mancher Lehrender, dass Studierende Präsenzveranstaltungen nicht mehr besuchen würden, wenn inhaltsgleiche E-Learning-Angebote zur Verfügung stünden (Bell et al., 2001), ist eher unbegründet.

Darüber hinaus zeigen die Ergebnisse ganz deutlich, dass es eine wichtige Rolle spielt, in welcher Phase ihres Lernprozesses sich die Studierenden befinden. Die Studierenden möchten die Lehr-/Lernmaterialien folglich für unterschiedliche Einsatzzwecke nutzen und schätzen dabei die Vor- und Nachteile der einzelnen Lehrmethoden und -materialien sehr gut ein. Sie wissen intuitiv bzw. aus der Erfahrung heraus, welches Ziel sie mit welcher Methode am besten erreichen können. So fallen auch die Akzeptanzurteile sehr unterschiedlich aus, je nachdem, ob die Studierenden gerade „nur" das Ziel haben, Wissen „neu kennenzulernen" (kurz-/längerfristige Evaluation) oder ob sie sich schon konkret in der Phase der Prüfungsvorbereitung (abschließende Evaluation) befinden. Bei der Wahl der Lehrmethode muss folglich klar festgestellt werden, an welcher Stelle man die Lernenden

abholt und mit welchem Ziel sie die Lehrmethode nutzen werden. Die Lehrmethode muss also jeweils zum Lernziel passen.
Unterstützt wird dieser Schluss auch dadurch, dass die Studierenden sich für andere Module Ihres Studiums ebenfalls Übungsaufgaben, für freiwilliges Lernen außerhalb der Universität jedoch eher E-Lectures und WBTs wünschen. Darüber hinaus lassen auch die Nutzungszeiten und -häufigkeiten der E-Lecture- und WBT-Nutzer auf unterschiedliche Lernstrategien bzw. Lernziele schließen: Die E-Lecture-Nutzer haben womöglich vermehrt (verpasste) Sitzungen vollständig nachgeholt, während die WBT-Nutzer selektiv nachgelesen und Übungsaufgaben genutzt haben, um die Lerninhalte zu verinnerlichen. Um hierzu eine abschließende Aussage zu machen, wäre es hilfreich gewesen, die Studierenden zu ihrem Lernverhalten zu befragen oder sie im Rahmen einer Laboruntersuchung explizit bei der Nutzung zu beobachten. So könnte auch der Vermutung nachgegangen werden, ob die Studierenden die WBTs im Sommersemester nur deshalb so intensiv nutzten, weil sie zum Teil Übungsaufgaben boten. Dem könnte allerdings auch im Rahmen einer Felduntersuchung nachgegangen werden, wenn neben E-Lectures und WBTs (ohne/mit Übungsaufgaben) auch noch einmal gesondert Übungsaufgaben angeboten werden.

4.3 Limitationen von Studie I

Zu der vorliegenden Studie ist zunächst kritisch anzumerken, dass das Vorwissen der Studierenden nicht geprüft wurde. Erst durch die Differenz zwischen Nachtest und Vortest kann eine Aussage über den tatsächlichen Wissenszuwachs angestellt werden. Dies schränkt die Aussagekraft der Ergebnisse zur Lernwirksamkeit leider ein.
Auch wurde angestrebt, die Inhalte über die drei Versuchsbedingungen konstant zu halten. Dies kann allerdings nicht vollständig garantiert werden. So musste zum Beispiel im Rahmen des Themas „Lernen mit Medien“ in der E-Lecture-Bedingung auf die Darstellung verschiedener Lernsoftware verzichtet werden, während in der Vorlesung und im WBT unterschiedliche Lernsoftware vorgestellt werden konnte. Bei der Erstellung der Wissensdiagnostik wurde diese Abweichung allerdings berücksichtigt. Zudem wurde die Erhebung über zwölf verschiedene Themen vorgenommen und die Ergebnisse über alle Themen ergeben ein recht konsistentes Bild. Deshalb kann eine solche Divergenz sicherlich vernachlässigt werden.
Darüber hinaus muss sich die vorliegende Studie – wie auch andere Studien, die im Rahmen einer Felduntersuchung E-Learning-Anwendungen mit traditioneller Lehre vergleichen – die Kritik gefallen lassen, dass sicherlich eine Reihe von nicht-kontrollierbaren Störfaktoren existieren, die die Aussagekraft der Vergleiche einschränken. So ist es möglich, dass die

Studierenden die E-Learning-Anwendungen in einer ganz anderen Lernumgebung bearbeiteten als von den Entwicklern beabsichtigt (Glowalla et al., 2011). Die Lernsituation kann darüber hinaus für die Bearbeitung der E-Lectures ganz anders ausgesehen haben als für die Nutzung der WBTs. Innerhalb der Studie von Salomon (1984, „Television is easy and print is tough“) zeigte sich, dass Lernende der Meinung sind, mit manchen Medien, wie z.B. dem Fernsehen, relativ leicht lernen zu können, mit anderen Medien, wie z.B. Texten, jedoch mehr Anstrengung für das Lernen aufbringen zu müssen. Lernerfolge nach dem Lernen mit Text schreiben sie eher sich, nach dem Lernen mit einem Film eher dem Medium zu. Die Selbstwirksamkeit der Lernenden unterscheidet sich folglich, je nachdem mit welchem Medium sie lernen. Salomon geht demzufolge davon aus, dass Lernende sich weniger anstrengen, je geringer sie die Anforderungen eines Mediums einschätzen. Gegebenenfalls wurden die WBTs folglich konzentriert am Schreibtisch bearbeitet, da die Studierenden hier hauptsächlich einen Text bearbeiten mussten. Dies wurde von ihnen höchstwahrscheinlich als anstrengender wahrgenommen. Die E-Lectures hingegen könnten sich die Studierenden eher informell auf der Couch angesehen haben, die Vorlesung könnte auch als Treffpunkt zum Austausch mit Kommilitonen wahrgenommen worden sein. Die Studierenden könnten die beide letzteren Lernumgebungen folglich als „einfacher“ eingeschätzt haben. Wenn man hieraus schließt, dass sich die Lernenden innerhalb der WBT-Gruppe mehr anstrengten als diejenigen in der E-Lecture- und Vorlesungsgruppe, ist es verständlich, dass sich für die WBTs kurzfristig eine bessere Lernwirksamkeit zeigte. Längerfristig spielen vermutlich Kompensationen eine bedeutende Rolle – wie oben bereits erläutert.

Darüber hinaus wurde den Studierenden – um den Nachteil geringer Rücklaufquoten bei Felduntersuchungen zu verhindern – ein zusätzlicher Anreiz zum Mitarbeiten geboten, indem sie sich über die Mitarbeit Bonuspunkte für die abschließende Klausur verdienen konnten. Dies mag dazu geführt haben, dass manche Studierende nicht der Sache wegen teilnahmen und dementsprechend die E-Learning-Anwendungen und Evaluationen nur halbherzig bearbeiteten. Eine völlig freiwillige Teilnahme hätte allerdings den Nachteil haben können, dass weniger Studierende vollständig an den Evaluationen teilgenommen hätten und das es sich dann bei den teilnehmenden Studierenden um eine ausgewählte Gruppe gehandelt hätte. Auch dies hätte die Aussagekraft der Studie folglich eingeschränkt. Mit oder ohne Belohnung bringt eine Felduntersuchung jedoch immer den Nachteil mit sich, dass die Arbeit der Studierenden mit den Lehrmethoden und den nachfolgenden Befragungen – wie oben bereits erwähnt – nicht kontrolliert werden kann. Dadurch ist es nicht möglich, eine Aussage darüber zu treffen, inwieweit die Studierenden tatsächlich konzentriert mit den E-Learning-Anwendungen gearbeitet bzw. in der Vorlesung aufgepasst haben. Die der Lerneffizienz zugrunde gelegte Lernzeit ist so zudem nicht ganz verlässlich, da nicht protokolliert werden konnte, wie lange einzelne Studierende während

der Bearbeitung der E-Learning-Anwendungen bzw. dem Besuch der Vorlesung physisch oder zumindest kognitiv abwesend waren. Dieser Nachteil ist auch durch die Abfrage der Pausenhäufigkeiten und -zeiten nicht auszugleichen, da die Angaben der Studierenden hier sicherlich nicht objektiv und vollständig waren.

Auch für die Bearbeitung der Übungstests muss immer unterstellt werden, dass die Studierenden – trotz Aufforderung, dies nicht zu tun – in ihren Unterlagen oder ggf. im Internet nachgeschaut haben und somit die Ergebnisse der Lernleistungstests evtl. die wirkliche Lernleistung etwas überschätzen. Da allerdings davon ausgegangen werden kann, dass diese Tendenzen über alle Experimentalgruppen in gleicher Weise vorhanden sind und Effekte zudem in mehreren Replikationen mit den gleichen Studierenden in allen drei Experimentalbedingungen überprüft wurden, dürften die Ergebnisse der vorliegenden Studie nicht bedeutsam durch solche Fehlerquellen beeinflusst sein.

Neben diesen Nachteilen durch die Durchführung als Felduntersuchung ist allerdings in den Vordergrund zu rücken, dass diese auch klare Vorteile aufweist: Innerhalb einer Laboruntersuchung hätten die Studierenden nicht in ihrer gewohnten Lernumgebung lernen können. Man hätte Gruppen von Studierenden gemeinsam zu einem festen Zeitpunkt in ein Lernlabor bestellen müssen, was die Art und Weise ihres Lernens deutlich beeinflusst – und die Anzahl an Teilnehmern immens eingeschränkt – hätte. Dies hätte die ökologische Validität der Ergebnisse deutlich eingeschränkt.

Durch die sehr große Stichprobe und die Durchführung in mehreren Replikationen sowie die Möglichkeit, dass jeder Studierende in jeder Experimentalgruppe teilnehmen konnte, konnte hingegen der Einfluss einiger nicht-kontrollierbarer Faktoren verringert werden. Ganz auszuschließen sind solche Effekte in einer so lebensnahen Felduntersuchung allerdings nie. Durch die Messwiederholung war es allerdings möglich, inhalts- und personinvariante Vergleiche anzustellen.

Zu guter Letzt muss die Erstellung sowie die Qualität der Akzeptanz- und Lernleistungstests kritisch reflektiert werden: Sowohl die Akzeptanz- als auch die Lernleistungstests wurden begleitend zur Studie aufgrund von inhaltlichen Überlegungen erstellt. Auf eine Vorstudie zur Überprüfung dieser Fragebögen und Übungstests musste verzichtet werden, da diese aufgrund der großen Menge an Evaluationen sehr umfangreich hätte sein müssen. Dementsprechend wurden erst im Nachgang Faktorenanalysen bzw. Itemanalysen durchgeführt. Für einige Items der kurzfristigen Akzeptanzerhebung ergaben sich Skalen mit zufriedenstellenden Reliabilitäten von zumeist über $\alpha = .7$, häufig sogar über $\alpha = .8$. Leider weist die Skala „Einschätzung von E-Learning“ nur ein Cronbachs $\alpha = .60$ auf. Diese Skala wurde dementsprechend lediglich zur Beschreibung der Stichprobe genutzt und geht in keine weiteren Auswertungen ein. Eine Überarbeitung der Akzeptanzbefragungen ist folglich für weitere Studien essentiell.

4.4 Ausblick

Studie I konnte zeigen, dass alle drei untersuchten Lehrmethoden über eine gute Lernwirksamkeit und Akzeptanz verfügen. Die E-Learning-Anwendungen schneiden dabei häufig sogar etwas besser ab. Dementsprechend können Teile der Präsenzlehre durch solche Anwendungen ergänzt und auch ersetzt werden, ohne Nachteile für die Lernenden zu befürchten. Nach Blick auf die Ergebnisse stellt sich allerdings die Frage, inwieweit sich die Kosten für die Erstellung der E-Learning-Anwendungen auszahlen. Hier kann man argumentieren, dass sich eine Erstellung der E-Learning-Anwendungen bei vielen Nutzern und mehrfacher Nutzung immer lohnt. Die Vorlesung scheint dennoch ökonomischer, da nur eine Person über einen kurzen Zeitrahmen vonnöten ist, um alle Studierende gleichzeitig zu erreichen. Eine Vorlesung ist sicherlich im Vergleich zu den E-Learning-Anwendungen auch am einfachsten zu aktualisieren. Dieser Argumentation steht allerdings die erhöhte Flexibilität der E-Learning-Anwendungen entgegen, die sich auch in den Akzeptanzbeurteilungen niederschlägt. Die Orts- und Zeitflexibilität ermöglicht es auch den Studierenden die Lerninhalte zu verfolgen, die aufgrund anderer Verpflichtungen oder Krankheit nicht an der Vorlesung teilnehmen können. Dieser Aspekt der Chancengleichheit ist ein Nutzen von E-Learning, der bei der Abwägung der Kosten nicht außer Acht gelassen werden darf.

Dennoch wäre es erfreulich, wenn E-Learning-Anwendungen darüber hinaus einen Mehrwert gegenüber der traditionellen Präsenzlehre bieten und dieser sich insbesondere in der Lernwirksamkeit niederschlagen würde. Bei der Nutzung von E-Learning-Anwendungen spielt hier sicherlich auch eine Rolle, inwieweit Lernende die Anwendung akzeptieren und inwieweit sie diese motivierend und zielführend finden, da die Lernenden die E-Learning-Anwendungen hauptsächlich selbstgesteuert nutzen müssen.

Zwar konnten im Bereich der Lernwirksamkeit und Lerneffizienz bereits Vorteile der WBTs gegenüber der Vorlesung festgestellt werden. Die bisher gefundenen Vorteile sind allerdings nicht durchgängig zu finden und zum Teil nur in geringem Maße vorhanden.

Es stellt sich die Frage, wie gut die WBT-Nutzer in der vorliegenden Studie in den Lernleistungstests abgeschnitten hätten, wenn die WBTs hinsichtlich verschiedener lernfördernder Merkmale optimiert worden wären. Zum Beispiel wurde kurz- und längerfristig die motivierende Wirkung der Web-based Trainings schlechter beurteilt als für die Vorlesung und z.T. die E-Lecture. Motivierendere WBTs sollten jedoch beispielsweise zu einer längeren Beschäftigung, einer besseren subjektiven Einschätzung sowie in der Folge zu einer besseren Lernleistung führen. So sollte zum Beispiel eine Verbesserung der Motivationsstruktur dazu beitragen können, die Lernleistung der Lernenden zu verbessern.

Aus diesem Grunde soll die zweite Studie überprüfen, inwieweit die Akzeptanz und Lernwirksamkeit von Web-based Trainings durch die Variation verschiedener Gestaltungsmerkmale verbessert werden kann und welche Gestaltungsmerkmale sich hinsichtlich der Steigerung von Lernwirksamkeit und Akzeptanz deutlich auszahlen. Dies ist insbesondere vor dem Hintergrund entscheidend, dass die Produktion von E-Learning-Anwendungen mit lernförderlichen Aspekten, wie Übungsaufgaben, Lernspielen oder Animationen etc. sicherlich wesentlich kostenintensiver ist als die von vornehmlich einfacher gestalteten Anwendungen mit z.B. nur Text und Bild. Diese Kosten müssen durch einen Mehrwert gerechtfertigt sein, was es zu überprüfen gilt.

Für die Erstellung der E-Learning-Variationen fiel die Wahl auf die Web-based Trainings. So ist zum einen ein Vergleich zur ersten Studie möglich, zum anderen können die WBTs im Vergleich zu den E-Lectures wesentlich flexibler gestaltet werden. Es können relativ einfach verschiedene Variationen erstellt und diese vergleichend untersucht werden.

Nach Tergan (2004) ist der Einsatz eines strukturierten Instruktions-[39] oder/und Motivationsdesigns[40] essentiell, möchte man E-Learning-Anwendungen lernförderlich gestalten. Seiner Meinung nach stellt sich Lernerfolg dann ein (Tergan, 2004, S. 16), „…wenn ein aktuelles Lernangebot entsprechend den Präskriptionen eines Instruktionsdesign-Modells gestaltet wurde und es den Kriterien und Standards eines Evaluationsansatzes gerecht wird." Dementsprechend gilt es, für die zweite Studie ein geeignetes Modell auszuwählen, das strukturiert praktikable Hinweise für die lernförderliche Gestaltung der Anwendungen bietet und empirisch gut belegt ist.

Im Folgenden werden verschiedene Modelle kurz vorgestellt (ausführlichere Ausführungen finden sich jeweils unter der angegebenen Literatur). Daran anschließend wird erläutert, warum die Wahl letztendlich auf das „Macro Model of Motivation and Performance" von John Keller (2010) gefallen ist.

[39] Ein Instruktionsdesign dient dazu, Lernszenarien so zu gestalten, dass das intendierte Lernziel möglichst effizient erreicht werden kann.

[40] Das Ziel eines Motivationsdesigns ist nach Niegemann et al. (2008, S. 370) „…die Aktivierung von Lernhandlungen […], die Unterstützung der Lernenden bei realistischen Zielsetzungen und die Aufrechterhaltung dieser Lernaktivitäten (Ausdauer und Intensität)."

5 Studie II: Theoretische Grundlagen

In der Literatur finden sich mittlerweile einige Instruktions- und Motivationsmodelle für die Gestaltung von Lehre im Allgemeinen oder die Erstellung von E-Learning-Anwendungen im Speziellen. Von diesen soll im Folgenden eine Auswahl vorgestellt werden, bevor vergleichend auf das „Macro Model of Motivation and Performance" von John Keller (2010) eingegangen wird.

5.1 Überblick über verschiedene Instruktions- und Motivationsdesignmodelle

Nach dem Instruktionsmodell von **Klauer (1985)** sollten bei der Planung effizienter Lerneinheiten sechs Lehrfunktionen im Auge behalten werden: Motivation, Information, Informationsverarbeitung, Speichern und Abrufen, Transfer sowie Steuerung und Kontrolle. Damit Lernen optimal funktioniert, müssen die Lernenden folglich für die Lerneinheit motiviert werden, ihnen müssen die notwendigen Informationen gegeben und das Verständnis, die Speicherung sowie der Transfer der Information gesichert werden. Um diese Vorgänge sicherzustellen, muss überprüft werden, ob die Lernenden fähig sind, sich selbst zu kontrollieren und ihren Lernprozess zu steuern. Ist dies nicht der Fall, muss die Kontrolle und Steuerung von außen – durch die Lehrkraft – geschehen.

Gagné (2011) nennt acht Unterrichtsvorgänge, durch die effektives Lehren erreicht werden kann: Motive anregen, über Lehrziele informieren, Aufmerksamkeit gewinnen, Erinnern anregen, Lernhilfen bieten, Behalten und Transfer fördern sowie Leistung kontrollieren und Rückmeldungen bieten.

Das Modell von **Malone (1981)** möchte über Herausforderung, Fantasie und Neugier – in einer Erweiterung zusätzlich über Kontrolle, Kooperation, Wettbewerb und Wiedererkennung (Malone & Lepper, 1987) – die intrinsische Lernmotivation und damit den Erfolg eines Lernszenarios erhöhen.

Ein weiterer Ansatz, der in der Literatur Erwähnung findet, ist der Ansatz der „Supermotivation" nach **Spitzer (1995)**, welcher über die Förderung von Motivatoren (wie z.B. Aktion, Spaß, Abwechslung, Beteiligung, Wahlmöglichkeiten, Verantwortlichkeit, soziale Interaktion, Nutzung von Stärken, Fehlertoleranz, Herausforderung, Ermutigung) und den Abbau von Demotivatoren Arbeiter – bzw. im konkreteren Fall auch Lernende – effizienter zu machen versucht.

Wlodkowski und Ginsberg (Wlodkowski, 2004, Ginsberg & Wlodkowski, 2009 sowie Wlodkowski, 2003) haben ein Modell entwickelt, das auf kulturell unterschiedliche erwachsene Lernende fokussiert. Damit sich Lernende mit unterschiedlicher kultureller

Abstammung zum Lernen angeregt fühlen, müssen sie sich demnach respektiert und einbezogen fühlen (Establishing inclusion), ihnen müssen Relevanz und Wahlmöglichkeiten vermittelt werden (Developing attitude), sie sollten sich herausgefordert und engagiert fühlen (Enhancing meaning) sowie Effektivität und Authentizität erfahren (Engendering competence). Diese vier Bedingungen sind laut den Autoren zu unterschiedlichen Zeitpunkten des Lernprozesses bedeutsam.

Deci und Ryan (1993) formulieren drei psychologische Bedürfnisse, die für motiviertes, erfolgreiches Lernen bedeutsam sind: Streben nach Autonomie/Selbstbestimmung, Wunsch nach Kompetenz/Wirksamkeit sowie soziale Eingebundenheit. Das Modell wurde von Prenzel (1996) in Rahmen eines Modellversuchs an der Universität München erprobt. Nach den Ergebnissen lohnt es sich, diese Aspekte bei der Gestaltung von Lehre zu beachten. Prenzel (1996) ergänzte die drei Bedingungen allerdings um drei weitere: Interesse, inhaltliche Relevanz sowie Instruktionsqualität. Darüber hinaus gibt er zu jedem der sechs Bereiche verschiedene Maßnahmen an, die zur Optimierung beitragen sollen. Prenzel, Drechsel und Kramer (1998) erweiterten dieses Modell um eine weitere hinderliche Bedingung – das Erleben von Überforderung.

Astleitner (2000) zollt mit seinem Modell dem Aspekt Tribut, dass neben Kognition und Motivation auch die Emotion eine wichtige Rolle im Lernprozess spielt. Er nennt fünf instruktionspsychologisch relevante Dimensionen: Angst, Neid, Ärger, Sympathie und Freude (FEASP für fear, envy, anger, sympathy und pleasure). Durch zwanzig Instruktionsstrategien (je vier pro Emotion) sollen die negativen Gefühle Angst, Neid und Ärger verringert, die positiven Gefühle Sympathie und Freude verstärkt und in der Folge der Lernprozess erfolgreicher werden.

Schlag (2013) nennt 17 Strategien zur Förderung des Lernprozesses. Dieses sind die Herstellung von Zieltransparenz, die Modellierung von lernförderlichem Verhalten, der Aufbau einer vertrauensvollen Beziehung zur Lehrkraft, die Erzeugung von Wettstreit und Kooperation, das Wecken von Interesse, entdeckenlassendes und handlungsorientiertes Lernen, das Geben von Zeit, die Ansprache verschiedener Sinne, die individuelle Leistungsbeurteilung und Binnendifferenzierung, Gespräche zwischen Lehrkraft und Lernendem, das Verzichten auf Strafen, die Ermöglichung von Erfolgen und Verstärkung, die Förderung lernförderlicher Attributionen sowie die Unterstützung eines Leistungsklimas.

Astleitner, Pasuchin und Wiesner (2006) nennen weitere Modelle, die hier nicht mehr vorgestellt werden sollen. Im Folgenden sollen die eingeführten Modelle mit dem „Macro Model of Motivation and Performance“ von John Keller (2010) verglichen werden.

5.2 Vergleich zu Kellers „Macro Model of Motivation and Performance“ (2010)

Es gibt einige Gründe, warum die Wahl gerade auf das „Macro Model of Motivation and Performance“ von Keller (2010) fiel. Um einen Vergleich zu den oben aufgeführten Modellen zu ermöglichen, soll es im Folgenden kurz skizziert werden. Ausführlicher wird das Modell in Abschnitt „5.3 Das „Macro Model of Motivation and Performance“ nach Keller (2010)“ vorgestellt.

5.2.1 Einführung zum „Macro Model of Motivation and Performance“ (Keller, 2010)

Für Keller (2010) ist Motivationsdesign als Problemlösungsprozess zu verstehen, der Lösungen für Motivationsprobleme bieten soll. „You can stimulate your students to learn, or you can kill their motivation.“ (Keller, 2010, S. 1). Lernende zu motivieren bedeutet für ihn demnach „sie zum Lernen zu bringen“. Und insbesondere im E-Learning-Szenario, in dem – wie oben erläutert – Lernende selbstgesteuert lernen, spielt die Motivationserhaltung eine besondere Rolle. Keller betont dies bereits im ersten Absatz seines Buches (2010, S. vii): „[…] motivation […] is a powerful influence on performance […].“ Und so steht im Mittelpunkt seines „Macro Model of Motivation and Performance“ (wie bereits anhand des Titels des Motivationsmodells festzustellen) die Lernleistung, die durch die Anstrengung des Lernenden beeinflussbar ist (siehe Abbildung 23). Die Anstrengung wird nach Keller (2010) wiederum durch das motivationale Design beeinflusst. Das Modell zeigt folglich, wie eine motivierende Lernumgebung die Anstrengung einer Person verbessern und wie diese Anstrengung in Kombination mit den Fähigkeiten, Fertigkeiten und dem Wissen der Person zur Lernleistung führen kann. Um die Lernenden zu mehr Anstrengung anzuspornen, sollte eine Lernumgebung nach Keller so gestaltet sein, dass sie die Aufmerksamkeit (A, attention), das Relevanzgefühl (R, relevance), die Erfolgszuversicht (C, confidence) und die Zufriedenheit (S, satisfaction) der Lernenden erhöht. Aus diesem Grunde wird im Folgenden häufig auch vom „ARCS-Modell“ die Rede sein.

Dreht man folglich an den „Keller’schen Rädchen“, sollte man auch die Lernförderlichkeit einer Lehrmethode – und nicht nur die Motivation der Lernenden – verbessern. Wie in Abbildung 23 (siehe Abschnitt „5.3 Das „Macro Model of Motivation and Performance“ nach Keller (2010)“) zu sehen, ist Motivation selbst gar nicht konkret als Variable im Modell enthalten. Keller bildet die Motivation durch die Anstrengung ab, die die Lernenden zu erbringen bereit sind.

5.2.2 Vergleich zu den anderen Instruktions- und Motivationsdesignmodellen

Betrachtet man die Strategien, die Keller einsetzt, um die Anstrengungsbereitschaft der Lernenden zu erhöhen (Aufmerksamkeit, Relevanzgefühl, Erfolgszuversicht und Zufriedenheit) zeigen sich klare Überschneidungen zu den oben kurz vorgestellten Modellen. So finden sich die vier Komponenten des ARCS-Modells mit ihren Unterkategorien in vielen der Modelle wieder – auch wenn sie zum Teil anders benannt wurden. Aspekte, die Keller beispielsweise unter „Relevanz" fasst (Lehrzielorientierung, Motivanpassung und Vertrautheit), finden sich bei Gagné (2011) unter „Motive anregen, über Lehrziele informieren", bei Spitzer (1995) unter „Verantwortlichkeit, soziale Interaktion oder Nutzung von Stärken", bei Ginsberg und Wlodkowski (2009) unter „Relevanz und Wahlmöglichkeiten (Developing attitude)", bei Prenzel (1996) unter „inhaltliche Relevanz", bei Astleitner (2000) unter „Sympathie" und bei Schlag (2013) unter „Herstellung von Zieltransparenz". Ganz analog könnte man diesen Vergleich für die anderen drei Keller'schen Komponenten vornehmen.

Die Wahl fiel deshalb auf das Modell von Keller (2010), da die oben kurz vorgestellten Modelle zur „Erfolgssteigerung" von Lehre zum Teil sehr abstrakt oder grob gehalten sind, nicht oder nicht vollständig stringent aus der Literatur hergeleitet wurden, keine konkreten Maßnahmen bzw. keinen Designprozess für die Praxis bieten oder/und bislang nicht ausreichend empirisch untersucht wurden. Auf der Suche nach einem Modell, das sehr konkrete Methoden zur Erfolgssteigerung von E-Learning-Anwendungen zur Verfügung stellt, stach das Modell von Keller (2110) deshalb hervor. Es verfügt über eine sehr umfangreiche aber gleichzeitig auch sehr systematische Methodensammlung und ist gut auf ein E-Learning-Szenario anwendbar. Außerdem wurde das Modell bereits sehr häufig empirisch überprüft. Nach Means, Jonassen und Dwyer (1997, S. 5) ist das ARCS-Modell „[…] the only coherent and comprehensive instructional design model accommodating motivation […]". Deimann (2007, S. 49) bezeichnet es als das „[…] fundierteste und am weitesten verbreitete […]" Modell. Darüber hinaus sei es das bekannteste Modell, das zudem weltweit am häufigsten eingesetzt würde (Astleitner, 2005; Holzinger, 2000).

Nach Keller (2010) kann Motivationsdesign eine Brücke zwischen der Motivationsforschung und der Praxis bilden, in der die Motivation von Lernenden erhöht und verändert werden soll. Zum einen sollten hier Konzepte, Theorien und Prinzipien aus der Motivationsforschung einbezogen werden, zum anderen aber auch aus der Arbeit von Praktikern stammende Prozeduren, erfolgreiche Strategien etc., die das Ziel verfolgen, die Motivation von Lernenden zu erhöhen. Keller selbst nutzte deshalb die Erkenntnisse zu verschiedenen Motivationstheorien und -praktiken, um sein Modell zu erstellen (siehe u.a. Keller, 2010). Darüber hinaus entwickelte er einen passenden Evaluationsfragebogen, der

auch im Rahmen der vorliegenden Studie (zumindest als Grundlage) genutzt werden konnte.

Alle diese Vorteile führten zu der Entscheidung, das Modell im Rahmen der vorliegenden Studie als Grundlage für die Erstellung der WBT-Varianten zu nutzen:

- Die stringente Ableitung aus Forschung und Praxis,
- die strukturierte Darstellung der möglichen Maßnahmen,
- die starken Überschneidungen zu Instruktionsdesignmodellen,
- der häufige Einsatz und die umfangreiche empirische Überprüfung des Modells sowie
- das Vorhandensein eines Designprozesses und eines Evaluationsfragebogens.

Es wird aus diesem Grunde im Folgenden ausführlich erläutert.

5.3 Das „Macro Model of Motivation and Performance“ nach Keller (2010)

Wie oben bereits dargestellt, leitete Keller (2010) sein „Macro Model of Motivation and Performance“ stringent aus der Forschung ab und setzte es in der Folge mehrfach und in verschiedenen Kontexten ein, um es zu validieren und zu überarbeiten. In Abschnitt „5.3.5 Wirksamkeit des Modells“ wird deutlich, dass das Modell darüber hinaus von einigen weiteren Autoren evaluiert wurde. Nach Keller und Kopp (1987, S. 289) beinhaltet das Modell Komponenten aus „social learning theory, environmental theories, humanistic theories, and aspects of attitude theory, decision theory, attribution theory, cognitive evaluation theory, equity theory, cognitive dissonance theory, locus of control, and learned helplessness.“ So bemüht sich das Modell, viele Strategien einzubeziehen, die für Instruktionsdesigner nützlich sein können.

Aus dem ursprünglichen Modell, das 1983 für Lehrveranstaltungen und Lehrmaterial entworfen wurde, entwickelte Keller bereits Ende der 80er Jahre Empfehlungen für die didaktische Konzeption von multimedialen Lernangeboten (Domagk & Niegemann, 2008; siehe Keller & Kopp, 1987). Keller passte das Modell darüber hinaus laut Deimann (2007) für eine ganze Reihe von weiteren Anwendungsmöglichkeiten an – wie z.B. für Fernstudiengänge oder auch adaptive Instruktionseinheiten.

Um den erfolgreichen Einsatz des Modells zu überprüfen, entwickelte Keller 2006 einen Fragebogen zur Erfassung von ARCS-relevanten Motivationskomponenten – den *Instructional Material Motivational Survey (IMMS)*. Dieser wurde in der englischen Version validiert und bietet die Möglichkeit zu überprüfen, ob die einzelnen ARCS-Komponenten gut umgesetzt wurden.

Wie bereits erwähnt, beinhaltet das Modell vier Komponenten: A (attention bzw. Aufmerksamkeit), R (relevance bzw. Relevanz), C (confidence bzw. Erfolgszuversicht)

sowie S (satisfaction bzw. Zufriedenheit). Für diese vier Komponenten gibt Keller (2010) ganz konkrete fördernde Maßnahmen an, die die Anstrengungsbereitschaft der Lernenden und damit ihre Lernleistung erhöhen sollen. Abbildung 23 veranschaulicht das „Macro Model of Motivation and Performance“.

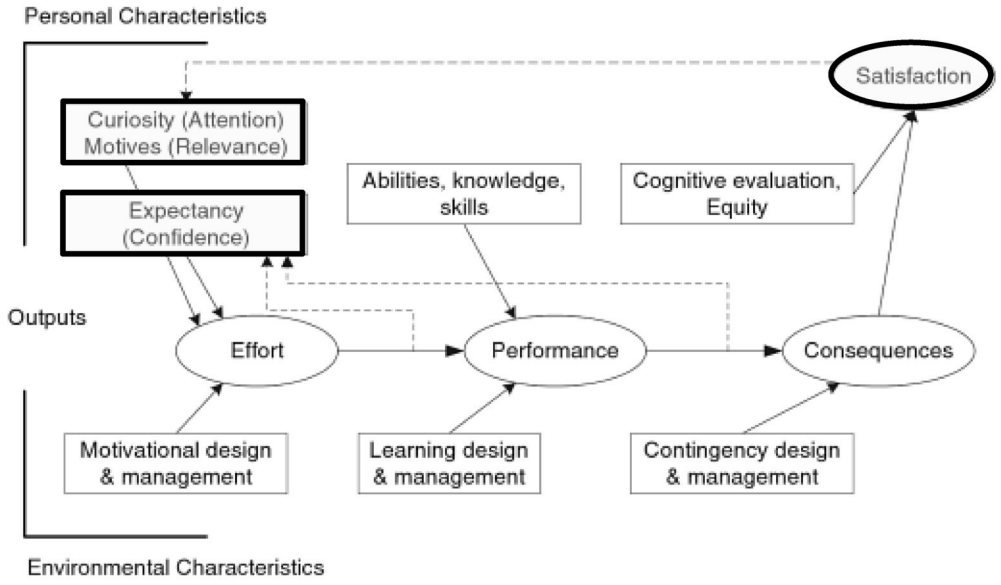

Abbildung 23. Kellers Macro Model of Motivation and Performance (nach Keller, 2010)

Diesem Modell folgend wird die Anstrengungsbereitschaft des Lernenden zuallererst durch seine Aufmerksamkeit (attention) beeinflusst. Der Lernende muss dem Lernmaterial zunächst einmal seine Aufmerksamkeit schenken. Zudem ist es von Bedeutung, ob er das Lernmaterial für sich selbst als relevant erachtet (relevance) und dass er die Erwartung hat, die Lernziele erreichen zu können (confidence). Wenn diese Aspekte gegeben sind, sollte der Lernende ausreichend Anstrengung (effort) investieren, um das Lernmaterial zu bearbeiten. Dies kann durch das Motivationsdesign des Lernmaterials unterstützt werden. Zusammen mit seinen Fähigkeiten, seinem Wissen und seinen Fertigkeiten (abilities, knowledge, skills) mündet diese Anstrengung in eine Lernleistung (performance). An dieser Stelle spielt auch das Instruktionsdesign des Lernmaterials eine wichtige Rolle. Die Leistung wiederum ist begleitet von bestimmten Konsequenzen (consequences), die wiederum durch das Kontingenzdesign beeinflusst werden können. Führen diese Konsequenzen nach der Beurteilung durch den Lernenden hinsichtlich seiner Erwartungen und nach Überprüfung der Fairness zu positiven Empfindungen, ist der Lernende zufrieden (satisfaction). Diese Zufriedenheit wirkt sich wiederum auf die Aufmerksamkeit und das Relevanzgefühl des Lernenden im weiteren Lernprozess aus. Durch das Erreichen der Leistungen (performances) wird zudem die Erfolgszuversicht für zukünftige vergleichbare Aufgaben erhöht.

Betrachtet man das Modell genauer, findet man zum Beispiel klare Überschneidungen zu den Modellen von Vollmeyer und Rheinberg (1998) sowie dem Erweiterten Kognitiven Motivationsmodell von Heckhausen und Rheinberg (1980). Auch Vollmeyer und Rheinberg (1998) betrachten personelle sowie situative Variablen, die in Interaktion zu Motivation, damit zu höherer Anstrengung etc. und wiederum zu verbesserter Lernleistung führen. Auch nach Heckhausen und Rheinberg (1980) führen u.a. die Lernleistungen zu bestimmten Konsequenzen für den Lernenden. Diese Beispiele zeigen eindrücklich die stringente Ableitung des Modells aus der Forschung.

Die vier von Keller formulierten Komponenten Aufmerksamkeit, Relevanz, Erfolgszuversicht und Zufriedenheit werden im Folgenden ausführlich erläutert. Die praktische Anwendung der hier theoretisch dargestellten Aspekte wird im Methodenteil anhand von Beispielen illustriert.

5.3.1 Aufmerksamkeit

Wie oben bereits dargelegt muss der Lernende laut Keller (2010) zunächst einmal seine Aufmerksamkeit auf das zu Lernende richten, damit überhaupt ein Lernprozess beginnen kann. Um diese Aufmerksamkeit zu erregen und zu erhalten, ist es notwendig, aktivierende Elemente in die Lernanwendung einzubauen. Eine Überfrachtung sollte allerdings vermieden werden, da diese zu einer Ablenkung vom eigentlichen Lehrstoff führen kann. Keller nennt drei Möglichkeiten, durch die dies erreicht werden kann:

1 **Wahrnehmungsbezogene Aufmerksamkeit**

   Sie wird dadurch erreicht, dass die Neugier des Lernenden angeregt oder der Lernende durch ungewöhnliche Ereignisse überrascht wird. Auch persönliche oder emotionale Inhalte wecken das Interesse des Lernenden.

2 **Fragehaltung schaffen**

   Dies entspricht weitestgehend dem Wecken von Neugier. Neugier kann dadurch erreicht werden, dass durch Fragen zum Nachdenken angeregt wird oder Widersprüche innerhalb des Lerninhaltes hervorgehoben werden.

3 **Variabilität/Abwechslung**

   Durch Abwechslung erreicht man, dass die Aufmerksamkeit des Lernenden aufrechterhalten bleibt. Abwechslung wird erzeugt, indem das Darbietungsformat variiert oder unerwartete Inhalte eingebunden werden. Auch konkrete Analogien zu bekannten Wissensbereichen oder Beispiele aus der Alltagswelt können Variabilität erzeugen.

5.3.2 Relevanz

Nach Keller (2010) ist es neben der Aufmerksamkeitsförderung zudem notwendig, Lernenden zu vermitteln, dass die Lerninhalte für sie persönlich relevant sind. Erst dann werden sie motiviert sein, sich mit dem Lernstoff auseinanderzusetzen. Es ist deshalb sinnvoll, die Lerninhalte so darzustellen, dass sie die Interessen und Ziele der Lernenden treffen.

Keller (2010) empfiehlt auch hier drei Methoden, mit denen dies bewirkt werden kann:

1 **Lehrzielorientierung**

Man wird den Bedürfnissen der Lernenden am besten gerecht, indem man Anwendungsbeispiele anbietet, Lernziele vorgibt oder die Lernenden anregt, diese selbst zu definieren. Auch ist es sinnvoll, immer wieder auf den Nutzen und die Wichtigkeit dieser Lernziele hinzuweisen.

2 **Motivanpassung**

Indem man Erfolgsmöglichkeiten (z.B. durch Rückmeldung von erbrachten Leistungen) bietet, kooperative Aktivitäten anregt (z.B. durch Chats oder Foren) oder Verantwortlichkeiten an Lernende vergibt (z.B. Leitung von Diskussionen im Forum), kann das Lehrmaterial auf die Lernstile und das persönliche Interesse von Lernenden angepasst werden.

3 **Vertrautheit**

Durch konkrete Beispiele oder Analogien zur Lebenswelt der Lernenden, persönliche Ansprache der Lernenden oder den Einsatz eines unterstützenden pädagogischen Agenten kann das Lehrmaterial mit den Erfahrungen der Lernenden verbunden und Vertrautheit geschaffen werden.

5.3.3 Erfolgszuversicht

Der dritte Aspekt, der nach Keller (2010) den Lernprozess positiv beeinflusst, ist die Erfolgszuversicht. Hierfür ist es insbesondere notwendig, dass der Lernende ein Gefühl der Kontrolle empfindet. Dieses Gefühl ist verbunden mit der eigenen Fähigkeitseinschätzung und der Erwartung, das gewünschte Lernergebnis durch Anstrengung erreichen zu können. Hat der Lernende nicht das Gefühl, die Fähigkeiten zu besitzen, die er benötigt, um das Lernziel zu erreichen, wirkt sich dies in Stress oder Angst aus. Dies beeinträchtigt verständlicherweise den Lernerfolg. Ist ein Lernender jedoch optimistisch, das Lernziel mit seinen Fähigkeiten erreichen zu können – seine Erfolgszuversicht ist demnach hoch – ist er gleichzeitig sehr motiviert, seine Lernaufgabe erfolgreich abzuschließen. Die Arbeit mit einer Lernanwendung sollte demnach zwar herausfordernd, nicht jedoch überfordernd sein. Lernende sollten also gefordert werden. Sie sollten allerdings möglichst nicht

scheitern (Niegemann et al., 2008). Diese Forderungen finden sich zum Beispiel auch im Rahmen des Flow-Konzeptes von Csikszentmihalyi (1975) wieder.

Für die Förderung der Erfolgszuversicht in Lernanwendungen gibt es nach Keller (2010) ebenfalls drei Maßnahmen, die bei der Erstellung der Anwendungen umgesetzt werden sollten:

1 **Lernanforderungen**

Die Lernerfahrung selbst sollte den Glauben des Lernenden an seine Kompetenz und an den Lernerfolg unterstützen oder sogar erhöhen. Dies kann gelingen, indem viele abwechslungsreiche und herausfordernde Aufgaben mit unterschiedlichem Leistungsniveau zur Bearbeitung zur Verfügung gestellt werden.

2 **Erfolgsmöglichkeiten**

Zudem sollte der Aufbau einer positiven Erfolgserwartung unterstützt werden, indem die Struktur der Lernanwendung sowie die Lernanforderungen und das Bewertungssystem transparent dargestellt werden. Dies schafft beim Lernenden Vertrauen und positive Erwartungen.

3 **Selbstkontrolle**

Ebenso sollte der Lernende den Grund für seinen Lernerfolg in seiner eigenen Anstrengung und Fähigkeit verankern. Dies kann unterstützt werden, indem viele Möglichkeiten zur Kontrolle geboten werden – zum Beispiel durch die Möglichkeit, den Lernweg durch die Anwendung selbst zu wählen. Auch Rückmeldungen, die den Erfolg auf die Anstrengung attribuieren, können hier unterstützend wirken.

5.3.4 Zufriedenheit

Wenn ein Lernender eine Lernanwendung erfolgreich bearbeitet hat, muss ihn dies laut Keller (2010) nicht automatisch zufriedenstellen. Nach Teo (2010) stellt die Zufriedenheit des Lernenden allerdings die wichtigste Bedingung einer erfolgreichen E-Learning-Anwendung dar, da diese einen positiven Effekt auf die Motivation des Lernenden hat. Wenn Lernende zufrieden mit einer E-Learning-Anwendung sind, sind sie motivierter und beschäftigen sich intensiver mit dem Lernmaterial. Dementsprechend sind sie auch erfolgreicher. Auch Holzinger (2000, S. 256) schreibt: „Schließlich sollte als Konsequenz eine Befriedigung (satisfaction) der Lernenden (Lernspaß, enjoyment) über die eigenen erbrachten Leistungen eintreten,…“

Sieht der Lernende selbst die Zeit, die er für die Bearbeitung aufgebracht hat, als Zeitverschwendung an, wird er eher demotiviert und unzufrieden sein. Dem Lernenden muss folglich das Gefühl vermittelt werden, dass das Lernergebnis in einem guten Verhältnis zum Aufwand steht und die Bewertung aller Lernenden im Vergleich gerecht ist.

Weitere Aspekte, die die Zufriedenheit von Lernenden beeinflussen können, sind extrinsische und intrinsische Verstärkung. Hier empfiehlt Keller (2010), mehr Wert auf intrinsische Motivation zu legen, da extrinsische Belohnungen alleine die Zufriedenheit negativ beeinflussen können, wenn der Lernende die Belohnung – ggf. aufgrund schlechter Leistungen – nicht erhält. Im Rahmen der intrinsischen Motivation liegt jedoch die Belohnung im Lernprozess selbst und wird dadurch so gut wie immer erreicht.
Zur Förderung der Zufriedenheit der Lernenden empfiehlt Keller (2010) wiederum drei Maßnahmen:

1 **Intrinsische Verstärkung**
   Die intrinsische Freude an der Arbeit mit der Lernanwendung kann man fördern bzw. unterstützen, indem man dem Lernenden regelmäßig Rückmeldungen und Informationen gibt, die dazu führen, dass er seine eigene Anstrengung positiv beurteilt. Auch Lernspiele oder Übungen, bei denen neues Wissen angewandt werden kann oder der Rückgriff bzw. der Verweis auf bereits Gelerntes können Lernende intrinsisch verstärken.
2 **Extrinsische Verstärkung**
   Verbales Lob, echte oder symbolische Belohnungen bzw. Prämien, selbst wählbare Belohnungen (z.B. Spielmöglichkeiten etc.) oder die Möglichkeit, die Ergebnisse der eigenen Anstrengung zeigen zu können, sind für Lernende belohnend und können sie extrinsisch verstärken.
3 **Gleichheit**
   Lernende erhalten ein Gefühl einer fairen Behandlung bzw. Bewertung, indem die Leistungsanforderungen und Erwartungen an alle Lernenden gleichgehalten werden und für alle Aufgaben und Lernende gleiche Beurteilungsstandards gelten.

5.3.5 Empirische Evidenz für das ARCS-Modell

Das Modell von Keller regte einige empirische Forschung an. So konnte wiederholt gezeigt werden, dass die Strategien die Anstrengungsbereitschaft und damit die Lernleistung in verschiedenen Lernumgebungen beeinflussen können (Astleitner & Lintner, 2004). Einige der Studien sollen im Folgenden erläutert werden.
Means, Jonassen und Dwyer (1997) zeigten, dass sich die Anreicherung von Texten mit Relevanzstrategien vor allem bei den Studierenden positiv auf Motivation und Lernleistung auswirkte, für deren Studium die Lerninhalte selbst irrelevant waren. Studierende, für die die Inhalte studienrelevant waren, zeigten eine höhere Motivation und eine bessere Lernleistung als Studierende, für die dies nicht galt. Dieser Unterschied konnte allerdings durch die Anreicherung mit Relevanzstrategien aufgefangen werden. Studierende, für die

die Inhalte irrelevant waren, die dann aber das angereicherte Lehrmedium erhielten, erreichten vergleichbare Motivations- und Lernleistungswerte wie Studierende, für die die Inhalte studienrelevant waren. Diese wiederum zogen geringere Vorteile aus der angereicherten Lerneinheit. Die Studie zeigt demnach, dass der Einsatz von Relevanzstrategien besonders dann sinnvoll und notwendig ist, wenn den Lernenden die Relevanz des Lerninhaltes nicht per se bewusst ist.

Chyung, Winiecki und Fenner (1999) nutzten das Modell von Keller im Rahmen eines computergestützten Fernunterrichts, da sie hier in den Vorjahren sehr hohe Abbruchquoten von 44% beobachtet hatten. Nachdem die ARCS-Methoden umgesetzt worden waren, wurde der Fernunterricht von den Lernenden als ansprechender und seine Inhalte als relevanter eingeschätzt. Auch die Erfolgszuversicht und die Zufriedenheit der Lernenden mit der Lernumgebung stiegen merkbar an und die Autoren konnten einen besseren Lernerfolg durch den Fernunterricht nachweisen. Am erfreulichsten war jedoch die deutliche Abnahme der Abbruchquote auf nur noch 22%. Von den Abbrechern wurden nun auch andere Gründe für den Abbruch angegeben: Während vorher unter anderem auf Aspekte in der Lernumgebung verwiesen wurde, gaben die Abbrecher nun eher persönliche oder berufliche Probleme für ihren Abbruch an.

Astleitner und Hufnagl (2003) überprüften die ARCS-Komponenten hinsichtlich ihrer Auswirkungen auf Motivation und Lernerfolg, indem sie sie im Rahmen einer E-Lecture einsetzten. Sie konnten feststellen, dass sich die ARCS-Anreicherung insbesondere für Lernende mit lernförderlichen Erwartungen (z.B. niedriger Situations-Ergebnis-Erwartung) positiv auswirkte. Diese Lernenden zeichnen sich dadurch aus, dass sie ihren Lernerfolg auf ihre eigene Anstrengung zurückführen. Sie sind dementsprechend der Überzeugung, ihr Erfolg sowie ihr Misserfolg seien von ihrem eigenen Handeln abhängig. Deshalb sind sie leichter zur Anstrengung zu motivieren. Für diese Lernenden ergaben sich höhere Motivationswerte, höhere Erfolgserwartungen sowie eine bessere Wissensaneignung. Im Gegensatz dazu machen Lernende mit lernhinderlichen Erwartungen (z.B. hoher Situations-Ergebnis-Erwartung) ihren Lernerfolg von für sie nicht beeinflussbaren Variablen abhängig. Diese Lernenden lassen sich nicht so einfach motivieren. Dementsprechend konnten diese Lernenden durch die ARCS-Strategien nicht profitieren.

Visser (1998) entwickelte ein Unterstützungssystem für Lernende (*Motivational Message Support System, MMSS*), indem sie Strategien nach dem ARCS-Modell für die Kommunikation mit Lernenden anwendete. Mit motivierenden Nachrichten innerhalb eines E-Learning-Lernprozesses sollten die Lernenden (je nach vorab erhobenen Eigenschaften, innerhalb des Lernprozesses auftretenden Reaktionen bzw. dem aktuellen Stand im Lernprozess) zum Weitermachen motiviert werden. Dadurch sollte in der Folge auch die

Abbruchquote gesenkt werden. Beispielsweise wurden zur Weckung der Aufmerksamkeit überraschende Nachrichten zugesandt, eine Betonung des Zusammenhanges zwischen konzentriertem Lernen und Lernerfolg sollte die Erfolgszuversicht erhöhen etc. Visser (1998) erreichte hierdurch eine niedrigere Abbruchquote und eine höhere Zufriedenheit der Lernenden. Eine motivationsfördernde Kommunikation mit den Lernenden wird von der Autorin demnach als effektive und zudem sehr kostengünstige Intervention beurteilt.

In der Studie von Astleitner und Hufnagl (2003) wurde ein Lehrtext, der mit ARCS-Strategien angereichert war, mit einem Lehrtext verglichen, der ohne ARCS-Merkmale auskam. Hierbei wurden von allen vier ARCS-Komponenten Aspekte umgesetzt. Die Anreicherung durch ARCS-Strategien verbesserte die Situations-Ergebnis-Erwartungen sowie die Bewertungen des Lernprozesses, verschlechterte jedoch die Wahrnehmung der Anreizwerte, der Handlungskontrolle und die Leistung im ersten Wissenstest. Im zweiten Wissenstest sowie im abschließenden längerfristigen Wissenstest schnitten beide Lerngruppen vergleichbar gut ab. Die Autoren erklären die negativen Effekte der ARCS-Strategien auf manche der Variablen durch die Verlängerung des Lehrtextes z.B. durch die zusätzlichen Instruktionen und die damit verbundene Erhöhung der kognitiven Last.

Farahat und Astleitner (2004) reicherten einen E-Learning-Kurs (web-based instruction) mit ARCS-Strategien an und verglichen ihn mit einem traditionellen Lernsetting. Studierende, die den E-Learning-Kurs nutzten, zeigten in der Folge ein höheres theoretisches Wissen, Studierende, die traditionell unterrichtet wurden, ein höheres praktisches Wissen. Die Motivation der Studierenden nach dem Lernen mit dem E-Learning-Kurs war zwar etwas höher als die Motivation vor dem Lernen mit dem Kurs, dieser Unterschied war allerdings nicht statistisch signifikant. Die Motivation vor dem Kurs sowie die Einstellung zum Kurs korrelierten signifikant positiv mit der Lernleistung.

Song und Keller (2001) überprüften den adaptiven Einsatz von Motivationsstrategien nach dem ARCS-Modell. Dafür verglichen sie eine adaptive Version einer E-Learning-Anwendung (computer-assisted instruction) mit einer motivational gesättigten und einer motivational minimierten Version hinsichtlich ihrer Wirkung auf Motivation, Lernwirksamkeit und Lerneffizienz. Die adaptive E-Learning-Version erreichte hinsichtlich der Lernleistung, Motivation sowie der Skala „Aufmerksamkeit“ bessere Werte als die anderen beiden Versionen und hinsichtlich der Skala „Relevanz“ bessere Werte als die gesättigte E-Learning-Version. Die Autoren erklären dieses Ergebnis damit, dass die Menge an Inhalten innerhalb der gesättigten E-Learning-Version die Lernenden von den relevanzförderlichen Strategien ablenkte. Die Autoren geben allerdings zu Bedenken, dass sich die Lernleistungen der beiden Kontroll-Versionen (gesättigt vs. minimiert) nicht signifikant unterschieden. Die Tatsache, dass sich die adaptive und die minimierte Version der E-Learning-Anwendung als lerneffizienter herausstellten als die gesättigte Version,

zeigt nach den Autoren, dass der adaptive Einsatz motivationaler Strategien je nach Bedarf der Lernenden sinnvoller ist als die „Überfrachtung“ einer Anwendung mit (annähernd) allen Strategien nach dem ARCS-Modell.
Auch Bellon und Oates (2002) konnten bestätigen, dass die Effekte verschiedener motivierender Strategien in Online-Kursen unterschiedlich ausfallen, wenn man Persönlichkeitsmerkmale von Lernenden beachtet. Über alle Lernenden hinweg betrachtet zeigte sich zudem, dass die verschiedenen motivierenden Strategien als unterschiedlich stark motivierend wahrgenommen wurden. Während z.B. Zusammenfassungen als sehr motivierend angenommen wurden, gaben die Lernenden an, dass Cartoons, Bilder oder Visualisierungen sie weniger motivierten.

Nach Deimann (2007) ist die Effektivität der ARCS-Strategien wissenschaftlich gut belegt – und das für verschiedene Lehr- und Lernangebote, wie Texte, E-Lectures, Web-based Trainings etc. Diese unterschiedliche Anwendbarkeit zeigt sich auch in den oben aufgeführten Studien und untermauert nach Deimann (2007) die Validität des Modells. Auch nach Astleitner (2005) zeigt die wissenschaftliche Befundlage, dass durch den Einsatz von ARCS-Komponenten hauptsächlich positive Auswirkungen auf die Motivation und den Lernerfolg erreicht werden können. Dabei ist allerdings eine sorgfältige Gestaltung unter Berücksichtigung der jeweiligen Bedingungen und Zielgruppe sowie eine sinnvolle Evaluation über den Strategieeinsatz unabdingbar.
In den vorgestellten Untersuchungen wurden bislang zumeist entweder alle vier oder lediglich eine bis zwei der Komponenten des ARCS-Modells integriert und die resultierenden Versionen miteinander verglichen. In der vorliegenden Studie werden allerdings sechs WBT-Variationen erstellt und ihr Einsatz evaluiert. Eine solche Evaluation entlang des gesamten Modells von Keller (2010) wurde nach gründlicher Sichtung der Literatur bislang noch nicht durchgeführt. Welche sechs Varianten miteinander verglichen werden und welche Fragestellungen zugrundeliegen, wird im Folgenden erläutert.

5.4 Fragestellungen der zweiten Studie

Im Rahmen der zweiten Studie dieser Arbeit werden insgesamt sechs WBT-Variationen miteinander verglichen:

- Ein WBT, das möglichst keine fördernden Aspekte beinhaltet („0-WBT“),
- jeweils ein WBT, in dem je eine der vier Komponenten optimiert wurde: Aufmerksamkeit („A-WBT“), Relevanz („R-WBT“), Erfolgszuversicht („C-WBT“) oder Zufriedenheit („S-WBT“) – im Folgenden „Einzelkomponenten-WBTs“ genannt – sowie

- ein WBT, bei dessen Erstellung möglichst alle von Keller (2010) vorgeschlagenen Komponenten berücksichtigt wurden („ARCS-WBT“).

Durch den Vergleich dieser Variationen soll eine Aussage darüber getroffen werden, welche der Komponenten „Aufmerksamkeit“, „Relevanz“, „Erfolgszuversicht“ und „Zufriedenheit“ tatsächlich umgesetzt werden sollten, um die Lernwirksamkeit und die Akzeptanz der Lerneinheit zu optimieren. Ist es notwendig, alle vier Komponenten zu integrieren oder spielen einzelne der Komponenten eine besonders wichtige Rolle? Überfrachtet der Einsatz aller Komponenten evtl. sogar die Lernanwendung? Wirkt sich der Einsatz einer Komponente nur auf den Bereich aus, den sie im Speziellen fördern soll (z.B. Aufmerksamkeitsstrategien nur zur Förderung der Aufmerksamkeit), oder können durch den Einsatz von Maßnahmen einer Komponente auch andere Bereiche beeinflusst werden (z.B. Aufmerksamkeitsstrategien fördern auch die Relevanzwahrnehmung)? Diese Fragen sind auch im Rahmen des Kosten-Nutzen-Denkens relevant, verlangt die Integration der ARCS-Komponenten doch zum Teil erheblichen Mehraufwand im Gegensatz zur Produktion einer Lerneinheit ohne diese Aspekte.

Vergleichend zu diesen sechs Varianten wird eine Gruppe von Studierenden die Lerninhalte mit einer E-Lecture bearbeiten. Dies stellt eine Vergleichsmöglichkeit zur Studie I her, da es sich um die gleichen E-Lectures handelt, die bereits hier zum Einsatz kamen. Dementsprechend resultieren sieben Experimentalgruppen, die hinsichtlich ihrer Lernmotivation, Akzeptanz, Nutzungszeit und Lernleistung (Lerneffektivität und -effizienz) miteinander verglichen werden können. Nur durch die Erhebung aller dieser Variablen kann eine differenzierte Aussage zu den Unterschieden zwischen den verschiedenen Lernanwendungen gemacht werden (siehe u.a. Glowalla et al., 2011 bzw. Abschnitt „1.5 Evaluation von E-Learning-Anwendungen“).

Darüber hinaus werden die Variationen hinsichtlich drei inhaltlich unterschiedlicher Themen untersucht. Dadurch sind vergleichbar zu Studie I erneut inhaltsinvariante Vergleiche möglich. Eine durchgängige person-/materialinvariante Untersuchung ist leider nicht möglich, da jeder einzelne Studierende hierzu sieben Mal an einer Erhebung teilnehmen müsste. Dies würde den Rahmen der Studie sprengen, da hierzu folglich 7 x 6 WBT-Varianten erstellt werden müssten. Die Teilnahme an so vielen Evaluations-zeitpunkten dürfte die Studierenden vermutlich überfordern, zumindest aber verärgern. Deshalb werden die Studierenden so über die drei Erhebungen rotiert, dass jeder Studierende einmal eine E-Lecture und zweimal ein WBT (unterschiedliche Varianten) nutzen und evaluieren kann.

Im darauffolgenden Sommersemester werden den Studierenden erneut Anwendungen zur Vorbereitung auf die Modulabschlussprüfung zur Verfügung gestellt. Sie erhalten zu allen

elf Vorlesungsthemen[41] ein 0-WBT (dieses wurde den Studierenden als „Kurzversion (K-WBT)“ vorgestellt), ein ARCS-WBT (welches in der Kommunikation mit den Studierenden als „elaboriertes WBT (E-WBT)“ bezeichnet wurde) sowie eine E-Lecture. Welche der insgesamt 33 Anwendungen sie wie intensiv für ihre Prüfungsvorbereitung nutzen, können die Studierenden selbst bestimmen.

Im Folgenden werden die Fragestellungen der Studie ausführlich erläutert. Dabei wird zunächst auf die Fragestellungen zur Lernleistung, Nutzungszeit und Lerneffizienz eingegangen. Danach folgen Fragestellungen zur Erhebung der ARCS-Skala und Akzeptanz. In einem separaten Abschnitt werden abschließend Fragestellungen zur freiwilligen Nutzung der Anwendungen im Sommersemester vorgestellt.

5.4.1 Lernleistung, Nutzungszeit und Lerneffizienz

Im Rahmen der Lernwirksamkeit wird erneut zwischen Lernleistung und Lerneffizienz unterschieden. Hierzu wird das Wissen sowohl vor als auch nach der Nutzung der Lernanwendungen durch hierfür erstellte Tests erhoben und die Differenz beider Erhebungen gebildet. Wie in Abschnitt „1.5.2 Evaluationskriterien“ bereits erwähnt, ist die zusätzliche Auswertung der Lerneffizienz sinnvoll, um insbesondere bei Lernanwendungen, die hinsichtlich der Lernleistung vergleichbar gut abschneiden, eine Aussage über die für die gleiche Lernleistung notwendigen Anstrengungen machen zu können.

Um über die kurz- und längerfristige Lernleistung und -effizienz berichten zu können, wird sowohl ein Tag nach der Nutzung der E-Learning-Maßnahme als auch am Ende des Wintersemesters die Lernleistung erhoben.

5.4.1.1 Lernleistung

Die Studien von Astleitner und Hufnagl (2003), Chyung, Winiecki und Fenner (1999), Farahat und Astleitner (2004) sowie Means, Jonassen und Dwyer (1997) konnten Lernwirksamkeitsvorteile von mit ARCS-Strategien angereicherten Anwendungen nachweisen. Gleichzeitig zeigten Astleitner und Lintner (2004) sowie Song und Keller (2001), dass durch die vollständige Umsetzung aller Komponenten die Gefahr einer Verlängerung bzw. Überfrachtung der Lerneinheit besteht, was sich in den Studien negativ auf die Lernwirksamkeit auswirkte.

Dennoch wird für die vorliegende Studie angenommen, dass die vollständige Umsetzung aller ARCS-Komponenten die Lernleistung deutlicher verbessert als eine nur teilweise Anreicherung mit nur einer der Komponenten. Sicherlich ist eine Verlängerung der

[41] Auf das Thema 12 „Rückblick“ wurde diesmal verzichtet, da es lediglich eine Zusammenfassung der anderen elf Themen darstellt.

Lerneinheit durch die vollständige Umsetzung des ARCS-Modells nicht zu vermeiden. Dies ist im Rahmen der Diskussion der Ergebnisse zu berücksichtigen. Eine Überfrachtung soll jedoch durch einen ausgewogenen Einsatz der Strategien möglichst vermieden werden.
Darüber hinaus sollte auch durch die teilweise Anreicherung mit nur einer der ARCS-Komponenten (Einzelkomponenten-WBTs) die Lernwirksamkeit höher ausfallen als die des 0-WBTs. Da die Strategien innerhalb der E-Lecture ebenfalls teilweise umgesetzt sind, wird diese vergleichbar eingestuft wie die Einzelkomponenten-WBTs.
Es resultieren hieraus die folgenden Hypothesen zur kurz- und längerfristigen Lernleistung (zum Verständnis werden diese jeweils kurz durch eine Gleichung zusammengefasst):

- *Hypothese 1a: Die kurzfristige Lernleistung (Differenz zwischen Vortest und kurzfristigem Nachtest) ist nach Bearbeitung des ARCS-WBTs signifikant höher als nach Bearbeitung aller anderen E-Learning-Varianten.*
  ➔ *ARCS > A, R, C, S, 0, EL*
- *Hypothese 1b: Die kurzfristige Lernleistung (Differenz zwischen Vortest und kurzfristigem Nachtest) ist nach Bearbeitung der Einzelkomponenten-WBTs sowie der E-Lecture signifikant höher als nach Bearbeitung des 0-WBTs.*
  ➔ *A, R, C, S, EL > 0*
- *Hypothese 2a: Die längerfristige Lernleistung (Differenz zwischen Vortest und Probeklausur) ist nach Bearbeitung des ARCS-WBTs signifikant höher als nach Bearbeitung aller anderen E-Learning-Varianten.*
  ➔ *ARCS > A, R, C, S, 0, EL*
- *Hypothese 2b: Die längerfristige Lernleistung (Differenz zwischen Vortest und Probeklausur) ist nach Bearbeitung der Einzelkomponenten-WBTs sowie der E-Lecture signifikant höher als nach Bearbeitung des 0-WBTs.*
  ➔ *A, R, C, S, EL > 0*

Zwar findet die Vorlesung in den Wochen der E-Learning-Bearbeitung nicht statt, die Lernleistung wird längerfristig (im Rahmen der Probeklausur am Ende des Wintersemesters) jedoch auch für die Themen erhoben, zu denen Vorlesungssitzungen stattfinden. Vergleichbar zur Studie I wird angenommen, dass die Bearbeitung der E-Learning-Anwendungen zu einer höheren Lernleistung führt als der Besuch der Vorlesung (Begründung siehe Abschnitt „1.6.2 Fragestellungen“).
Hieraus ergibt sich für den Vergleich der E-Learning-Anwendungen und der Vorlesung folgende Hypothese:

- *Hypothese 3: Die längerfristige Lernleistung (Differenz zwischen Vortest und Probeklausur) ist nach Nutzung der E-Learning-Anwendungen signifikant höher als nach dem Besuch der Vorlesung.*
  ➔ *EL, ARCS, A, R, C, S, 0 > VL*

5.4.1.2 Nutzungszeit

Mit Nutzungszeit ist die Zeit gemeint, die die Studierenden benötigen, um die jeweilige E-Learning-Anwendung zu bearbeiten. Diese wird über das Learning-Management-System *Moodle* erfasst.

Wie oben bereits angeklungen, ist davon auszugehen, dass für die „ARCS-Version" des WBTs eine längere Bearbeitungszeit resultiert als für die E-Lecture, die Einzelkomponenten-WBTs sowie das 0-WBT. Das 0-WBT sollte im Gegensatz dazu eine kürzere Bearbeitungszeit aufweisen als alle anderen E-Learning-Versionen, da hier keinerlei ausschmückende Details enthalten sind. Dementsprechend werden folgende Hypothesen zur Nutzungszeit aufgestellt:

- *Hypothese 4a: Die Nutzungszeit des ARCS-WBTs ist signifikant länger als die Nutzungszeit aller anderen E-Learning-Anwendungen.*
  ➔ *ARCS > A, R, C, S, 0, EL*
- *Hypothese 4b: Die Nutzungszeit der Einzelkomponenten-WBTs und der E-Lecture sind signifikant länger als die Nutzungszeit des 0-WBTs.*
  ➔ *A, R, C, S, EL > 0*

5.4.1.3 Lerneffizienz

Aufgrund der vorangegangenen Hypothesen zur Lernleistung und Nutzungszeit sollten alle E-Learning-Anwendungen in der Summe über eine in etwa gleich gute Lerneffizienz verfügen. Zwar wird angenommen, dass das ARCS-WBT die höchste Lerneffektivität erreicht. Gleichzeitig wird ihm aber auch die höchste Bearbeitungszeit zugeordnet. Für das 0-WBT werden im Gegensatz eine geringere Lernwirksamkeit und eine niedrigere Bearbeitungszeit hypothetisiert als für die anderen E-Learning-Anwendungen.

Wie oben dargestellt, zeigen die Ergebnisse von Astleitner und Lintner (2004) bzw. Song und Keller (2001) allerdings, dass sich die vollständige Umsetzung aller Komponenten gegenüber einer teilweisen oder sogar keinen Umsetzung der Komponenten negativ auf die Lerneffizienz der Lernanwendung auswirken kann. Dies erklären die Autoren damit, dass die ARCS-WBTs aus deutlich mehr Seiten bestehen als WBT-Varianten mit nur teilweise oder gar nicht umgesetzten ARCS-Komponenten. Dies zeigt sich auch in der vorliegenden Studie. Für das Thema Konditionierung enthält das ARCS-WBT zum Beispiel 60 Seiten während das 0-WBT mit 30 Seiten auskommt. Die Einzelkomponenten-WBTs enthalten 37 (A- und S-WBT), 48 (R-WBT) bzw. 49 Seiten (C-WBT).

Deshalb wird angenommen, dass das ARCS-WBT über eine geringere Lerneffizienz verfügt als die anderen E-Learning-Anwendungen, sich die höhere Nutzungszeit also nicht vollständig durch die bessere Lernleistung amortisiert. Es ergeben sich dementsprechend folgende Hypothesen zur kurz- und längerfristigen Lerneffizienz:

- *Hypothese 5: Die kurzfristige Lerneffizienz (Differenz zwischen Vortest und kurzfristigem Nachtest * 1000 / Nutzungszeit) des ARCS-WBTs ist signifikant geringer als die der anderen E-Learning-Anwendungen.*
  ➔ *ARCS < A, R, C, S, 0, EL*
- *Hypothese 6: Die längerfristige Lerneffizienz (Differenz zwischen Vortest und Probeklausur * 1000 / Nutzungszeit) des ARCS-WBTs ist signifikant geringer als die der anderen E-Learning-Anwendungen.*
  ➔ *ARCS < A, R, C, S, 0, EL*

Zudem wurde angenommen, dass die längerfristige Lernwirksamkeit der Vorlesungssitzungen geringer ausfällt als die der E-Learning-Anwendungen. Studie I zeigte schon, dass die Zeit, die die Studierenden innerhalb der Vorlesung verbrachten, kürzer war als die Zeit, die sie mit der Bearbeitung der WBTs verbrachten. Die Lerneffizienz der Vorlesung fiel in Studie I dennoch sowohl kurz- als auch längerfristig schlechter aus als die Lerneffizienz der WBTs. Dementsprechend ergibt sich für Studie II folgende Hypothese:

- *Hypothese 7: Die längerfristige Lerneffizienz (Differenz zwischen Vortest und Probeklausur * 1000 / Nutzungszeit bzw. Vorlesungszeit) der E-Learning-Anwendungen ist signifikant höher als die der Vorlesung.*
  ➔ *EL, ARCS, A, R, C, S, 0 > VL*

5.4.2 ARCS-Skala

Neben der Lernwirksamkeit soll untersucht werden, ob durch die Integration der einzelnen ARCS-Komponenten die von Keller postulierten Aspekte Aufmerksamkeitserhaltung, Relevanzgefühl, Erfolgszuversicht sowie Zufriedenheit erhöht werden können. Darüber hinaus ist es interessant zu betrachten, inwieweit die Umsetzung einer Komponente nur jeweils Auswirkungen auf ihren jeweiligen Zielaspekt hat (also inwieweit sich z.B. die Aufmerksamkeitsstrategien nur auf die Aufmerksamkeitserhaltung auswirken) oder ob sie sich auch auf andere Aspekte (z.B. Relevanzgefühl, Erfolgszuversicht oder auch die Zufriedenheit) auswirkt. Es wird davon ausgegangen, dass sich die Studierenden durch…

- …das ARCS- und A-WBT hinsichtlich ihrer Aufmerksamkeitshaltung deutlicher unterstützt fühlen als durch die anderen E-Learning-Anwendungen.
- …das ARCS- und R-WBT hinsichtlich der Vermittlung der Relevanz des Lerninhaltes besser beraten fühlen als durch die anderen E-Learning-Anwendungen.
- …das ARCS- und C-WBT deutlicher in ihrer Erfolgszuversicht bestärkt fühlen als durch die anderen E-Learning-Anwendungen.

- …das ARCS- und S-WBT stärker zufriedengestellt fühlen als durch die anderen E-Learning-Anwendungen.

Das ARCS-WBT sollte dementsprechend in der Summe hinsichtlich der Gesamt-ARCS-Skala eine signifikant bessere Einstufung erhalten als alle anderen WBTs, die Einzelkomponenten-WBTs sollten eine bessere Einschätzung erreichen als das 0-WBT.

Die E-Lecture wird an dieser Stelle aus den Hypothesen ausgeklammert, da in ihr unstrukturiert verschiedene der einzelnen Komponenten berücksichtigt sind und zudem an dieser Stelle der strukturierte Vergleich der sechs WBT-Varianten im Vordergrund stehen soll. Die Ergebnisse zur E-Lecture werden dennoch mit ausgewertet und berichtet. Es resultieren konkret folgende Hypothesen:

- *Hypothese 8a: Die Studierenden fühlen sich durch das ARCS- und das A-WBT hinsichtlich ihrer Aufmerksamkeitshaltung signifikant besser unterstützt als durch die anderen WBT-Varianten.*

  ➔ *ARCS, A > R, C, S, 0*
- *Hypothese 8b: Die Studierenden fühlen sich durch das ARCS- und das R-WBT hinsichtlich ihres Relevanzgefühls signifikant besser unterstützt als durch die anderen WBT-Varianten.*

  ➔ *ARCS, R > A, C, S, 0*
- *Hypothese 8c: Die Studierenden fühlen sich durch das ARCS- und das C-WBT hinsichtlich ihrer Erfolgszuversicht signifikant besser unterstützt als durch die anderen WBT-Varianten.*

  ➔ *ARCS, C > A, R, S, 0*
- *Hypothese 8d: Die Studierenden fühlen sich durch das ARCS- und das S-WBT signifikant besser zufriedengestellt als durch die anderen WBT-Varianten.*

  ➔ *ARCS, S > A, R, C, 0*
- *Hypothese 8e: Die Studierenden fühlen sich durch das ARCS-WBT hinsichtlich aller vier ARCS-Aspekte (Gesamt-ARCS-Skala) signifikant besser unterstützt als durch die anderen WBT-Varianten.*

  ➔ *ARCS > A, R, C, S, 0*
- *Hypothese 8f: Die Studierenden fühlen sich durch die Einzelkomponenten-WBTs hinsichtlich aller vier ARCS-Aspekte (Gesamt-ARCS-Skala) signifikant besser unterstützt als durch das 0-WBT.*

  ➔ *A, R, C, S > 0*

5.4.3 Akzeptanz

Die Akzeptanz einer Lernanwendung ist insbesondere deshalb von Bedeutung, da sie in der Regel mit einer verstärkten Nutzung dieser Anwendung einhergeht (Glowalla et al., 2011). Das kann sich wiederum gewinnbringend auf die Lernleistung der Lernenden auswirken. Dementsprechend ist es sinnvoll, den subjektiven Eindruck der Studierenden zu den Lernanwendungen zu erheben und zu vergleichen. Es ist davon auszugehen, dass die Lernanwendungen, durch die sich die Studierenden insgesamt stärker in ihrem Lernprozess unterstützt fühlen (Einschätzung hinsichtlich der Gesamt-ARCS-Skala), auch eine höhere Akzeptanz erreichen. Dementsprechend sollte das ARCS-WBT die höchsten Akzeptanzwerte erreichen. Danach folgend sollte die Akzeptanz für die Einzelkomponenten-WBTs sowie die E-Lecture einzuordnen sein. Die geringste Akzeptanz ist für das 0-WBT zu erwarten. Daraus ergeben sich sowohl für die feinere Akzeptanzerhebung (Akzeptanzskala) als auch für die gröbere Akzeptanzeinstufung der Anwendungen auf einer Notenskala von 1 bis 6 folgende Hypothesen:

- *Hypothese 9a: Für das ARCS-WBT ergibt sich eine signifikant höhere Lernerakzeptanz (Akzeptanzskala) als für alle anderen E-Learning-Anwendungen.*
  ➔ *ARCS > A, R, C, S, 0, EL*
- *Hypothese 9b: Für die Einzelkomponenten-WBTs sowie die E-Lecture ergibt sich eine signifikant höhere Lernerakzeptanz (Akzeptanzskala) als für das 0-WBT.*
  ➔ *A, R, C, S, EL > 0*
- *Hypothese 10a: Für das ARCS-WBT ergibt sich eine signifikant höhere Lernerakzeptanz (Noteneinstufung) als für alle anderen E-Learning- Anwendungen.*
  ➔ *ARCS > A, R, C, S, 0, EL*
- *Hypothese 10b: Für die Einzelkomponenten-WBTs sowie die E-Lecture ergibt sich eine signifikant höhere Lernerakzeptanz (Noteneinstufung) als für das 0-WBT.*
  ➔ *A, R, C, S, EL > 0*

5.4.4 Freiwillige Nutzung im Folgesemester

Im Sommersemester wurden den Studierenden mehrere Monate lang ARCS-WBTs (elaborierte Version = E-WBT[42]), 0-WBTs (Kurze Version = K-WBT) sowie E-Lectures zu elf Vorlesungsthemen zum Lernen für die abschließende Prüfung (MAP) zur Verfügung

[42] Die neue Benennung von ARCS- und 0-WBT erfolgte, um den Studierenden zu erläutern, auf welche Art die beiden WBT-Versionen differieren. Den Studierenden sollte klar dargestellt werden, dass beide WBTs inhaltlich gleich sind und sich lediglich durch Elaborationen und Sonderfunktionen unterscheiden. Ihnen sollte das Gefühl vermittelt werden, dass sie sich mit beiden WBT-Versionen sowie den E-Lectures vergleichbar gut auf die Klausur vorbereiten können.

gestellt. Die Studierenden konnten frei wählen, welche Anwendungen sie für ihre Klausurvorbereitung nutzen oder ob sie sich gar nicht mit den Anwendungen vorbereiten. Die freiwillige Nutzung soll eine Aussage über die Verhaltensakzeptanz dieser drei E-Learning-Versionen ermöglichen (Müller-Böling & Müller, 1986). Dieses objektivere Maß für die Akzeptanz soll vergleichbar zur Studie I die bislang erhobene subjektive Einstellungsakzeptanz ergänzen.

Durch das Abschneiden der Studierenden innerhalb einer erneuten Probeklausur sowie in der Modulabschlussprüfung soll darüber hinaus untersucht werden, welche der aus dieser freiwilligen Wahl resultierenden Nutzergruppen den besten Lernerfolg erreicht. Eine abschließende Befragung zu einzelnen Komponenten des ARCS-Modells soll darüber hinaus zeigen, welche der eingesetzten Methoden die Studierenden selbst als besonders wichtig für ihren Lernprozess empfinden.

Im Folgenden werden die Fragestellungen zu diesem Teil der Studie erläutert. Dabei werden zunächst die Fragestellungen zur Nutzungshäufigkeit und -zeit sowie zur Lernleistung der Nutzergruppen vorgestellt. Abschließend werden die Auswertungen hinsichtlich der wichtigen Eigenschaften von E-Learning-Anwendungen eingeführt.

5.4.4.1 Nutzungshäufigkeit und -zeit

Da die Einstellungsakzeptanz nach Bürg und Mandl (2005) Hauptprädiktor für die Verhaltensakzeptanz ist, sollte die E-Learning-Anwendung mit der höchsten Einstellungsakzeptanz am Ende des Wintersemesters im Sommersemester am häufigsten und längsten genutzt werden (Verhaltensakzeptanz). Es ergeben sich für die drei E-Learning-Anwendungen (ARCS- bzw. E-WBT, 0- bzw. K-WBT und E-Lecture) folgende Hypothesen:

- *Hypothese 11a: Die E-Learning-Anwendung, die im Wintersemester durch die Studierenden die höchsten Akzeptanzwerte erreicht, wird im Sommersemester signifikant häufiger genutzt als die anderen E-Learning-Anwendungen.*
- *Hypothese 11b: Die E-Learning-Anwendung, die im Wintersemester durch die Studierenden die höchsten Akzeptanzwerte erreicht, wird im Sommersemester signifikant länger genutzt als die anderen E-Learning-Anwendungen.*

5.4.4.2 Lernleistung der Nutzergruppen

Vergleichbar zu den Ergebnissen der Studie von Glowalla (2003), in der die Nutzer der E-Learning-Anwendung in der abschließenden Klausur besser abschnitten als die Nicht-Nutzer, ergeben sich für die Lernleistung der Studierenden (Probeklausur und MAP), die im Sommersemester die E-Learning-Anwendungen für ihre Klausurvorbereitung nutzen bzw. nicht nutzen, folgende Hypothesen:

- *Hypothese 12a: Studierende, die angeben, sich hauptsächlich mit den E-Learning-Anwendungen auf den Teil A der abschließenden Klausur vorbereitet zu haben, schneiden signifikant besser in der Probeklausur ab als Studierende, die angeben, sich hauptsächlich mit Offline-Material (z.B. dem Vorlesungsskript) vorbereitet zu haben.*
  ➔ *E-Learning > Offline-Material*
- *Hypothese 13a: Studierende, die angeben, sich hauptsächlich mit den E-Learning-Anwendungen auf den Teil A der abschließenden Klausur vorbereitet zu haben, schneiden hier signifikant besser ab als Studierende, die angeben, sich hauptsächlich mit Offline-Material (z.B. dem Vorlesungsskript) vorbereitet zu haben.*
  ➔ *E-Learning > Offline-Material*

Darüber hinaus ist davon auszugehen, dass eine längere Auseinandersetzung mit dem Lernmaterial zu einer höheren Lernleistung führt. Ausgehend von Hypothese 11b sollte folglich die Nutzergruppe signifikant besser abschneiden, die sich insgesamt am längsten mit einer der E-Learning-Anwendungen auf den Teil A der Klausur vorbereitet hat. Daraus resultieren folgende Hypothesen:

- *Hypothese 12b: Die Nutzergruppe, die sich am längsten mit den E-Learning-Anwendungen auf den Teil A der abschließenden Klausur vorbereitet hat, schneidet signifikant besser in der Probeklausur ab.*
- *Hypothese 13b: Die Nutzergruppe, die sich am längsten mit den E-Learning-Anwendungen auf den Teil A der abschließenden Klausur vorbereitet hat, schneidet hier signifikant besser ab.*

Die Art der Vorbereitung auf den Teil A der Klausur sollte allerdings keine Unterschiede hinsichtlich der Leistungen der Studierenden in Teil B nach sich ziehen. Sind die Studierenden, die nach den Hypothesen 13a und 13b in Teil A besser abschneiden, auch in Teil B besser, ist anzunehmen, dass es sich um diejenigen handelt, die aufgrund ihrer Leistungsfähigkeit und ihrer Lernmotivation generell besser abschneiden. Der Einfluss der Lehrmethode ist damit in Frage zu stellen. Deshalb soll im Vergleich zum Klausurergebnis in Teil A auch das Ergebnis zu Teil B der Klausur betrachtet werden, um die Interpretation der Ergebnisse zu Teil A abzusichern.

5.4.4.3 Wichtige Eigenschaften von E-Learning-Anwendungen

In der letzten Befragung vor der Modulabschlussprüfung sollen die Studierenden auf einer Skala von 1 (nicht wichtig) bis 10 (sehr wichtig) einschätzen, wie wichtig sie im Rahmen von E-Learning-Szenarien welche von 20 (im Falle einer vorangegangenen hauptsächlichen WBT-Nutzung) bzw. sechs Eigenschaften (im Falle einer hauptsächlichen E-Lecture-

Nutzung) finden. Die abgefragten Eigenschaften beinhalten wichtige Aspekte der vier Komponenten des ARCS-Modells von John Keller (2010), wie z.B. den Einsatz eines Lernbegleiters, die Vorgabe von Lernzielen, die Möglichkeit für einen flexiblen Lernweg etc. An dieser Stelle soll lediglich deskriptiv untersucht werden, welche Eigenschaften die Studierenden als besonders wichtig bewerten. Dabei soll auch darauf geachtet werden, ob es insbesondere Maßnahmen einer ARCS-Komponente sind, die als besonders wichtig eingestuft werden. Aufgrund der Menge an Eigenschaften erscheint hier ein hypothesengesteuerter Vergleich wenig sinnvoll.

Im folgenden Abschnitt wird die genutzte Methodik der zweiten Studie erläutert. Dabei wird zunächst die zugrundeliegende Stichprobe beschrieben, bevor die Durchführung und das Versuchsmaterial ausführlich dargestellt werden.

6 Studie II: Methode

Die Datenerhebung zur zweiten Studie fand im Wintersemester 2010/2011 erneut im Rahmen der Vorlesung „Psychologie für Lehramtsstudierende A“ von Professor Glowalla statt. Wie bereits im Zusammenhang mit Studie I erläutert, ist die Vorlesung Teil des Grundmoduls „Psychologie für Lehramtsstudierende“ an der Justus-Liebig-Universität Gießen. Zum Modul gehört eine weitere Vorlesung, die jeweils im Sommersemester gehalten wird, sowie ein Seminar, in welchem einzelne Vorlesungsthemen vertieft behandelt werden. Die Abschlussprüfung zu dem Modul wird nach dem Sommersemester abgelegt und behandelt die Inhalte beider Vorlesungen. Alle Lehramtsstudierenden müssen dieses Modul durchlaufen. Empfohlen werden hierfür die ersten beiden Studiensemester.
Aufgrund der erneut hohen Studierendenanzahl von insgesamt 1232 Studierenden wurde die Vorlesung auch im Wintersemester 2010/2011 zu zwei Terminen angeboten, für die sich die Studierenden erneut über das universitätsinterne Lehrverwaltungssystem *Stud.IP* (siehe Abschnitt „2.3 Material“) anmeldeten. Die Vorlesungen fanden mittwochs von 12 bis 14 Uhr sowie 18 bis 20 Uhr statt. 837 Studierende (67.9%) meldeten sich für die Mittagsvorlesung und 392 Studierende (31.8%) für die Abendvorlesung an. Drei weitere Studierende (0.2%) nehmen zwar an der Studie teil, sind aber offiziell in keiner der beiden Veranstaltungen in *Stud.IP* angemeldet.

6.1 Stichprobe

Die Gesamtstichprobe besteht aus den oben genannten 1232 Lehramtsstudierenden, die im Wintersemester 2010/2011 die Vorlesung von Professor Glowalla besuchten. Für diese wurde bereits zu Beginn des Semesters eine Vorbefragung zur Erhebung der demographischen Angaben und dem Vorwissen freigeschaltet (siehe Anhang C). Lediglich 69 (5.6%) der Studierenden nahmen an dieser ersten Befragung nicht teil, 1149 Studierende (93.3%) füllten den Fragebogen vollständig aus, 14 weitere (1.1%) beantworteten ihn unvollständig. Im Folgenden werden die Fehlendwerte nur dann angegeben, wenn sie von den 69 Studierenden abweichen. Die prozentualen Angaben beziehen sich auf die jeweilige Stichprobe.

6.1.1 Allgemeine Angaben

1163 (94.4%) der Studierenden gaben im Rahmen der Vorbefragung ihr Geschlecht an. 64.6% von ihnen sind weiblich, 35.4% männlich. Diese Geschlechtsverteilung ist für die

Gesamtpopulation an Lehramtsstudierenden in Deutschland nicht ungewöhnlich (siehe u.a. Streitenberger, nicht datiert oder Kühn, 2009).

Dabei variierte das Alter der Studierenden zwischen 18 und 52 Jahren (siehe Abbildung 24). Im Mittel waren die Studierenden zum Zeitpunkt der Datenerhebung M = 21.9 Jahre alt.

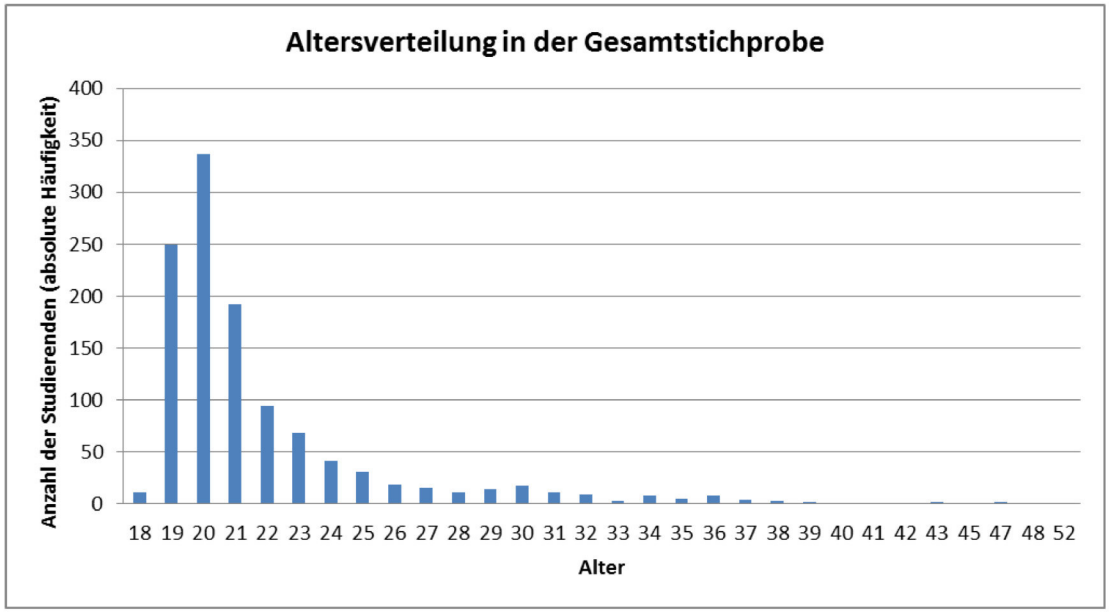

Abbildung 24. Altersverteilung der Stichprobe (n = 1163)

Die Studierenden wurden zudem dazu befragt, welche Art von Lehramtsstudium sie studieren. Unter den Studierenden, die hierzu eine Angabe machten, befanden sich 9.8% L1-Studierende (Lehramt an Grundschulen), 30.5% L2-Studierende (Lehramt an Haupt- und Realschulen), 40.8% L3-Studierende (Lehramt an Gymnasien), 7.6% L4- bzw. BaBBB-Studierende (Lehramt an beruflichen Schulen) sowie 10.8% L5-Studierende (Lehramt an Förderschulen). Abbildung 25 zeigt die Verteilung der Studierenden auf die Studiengänge.

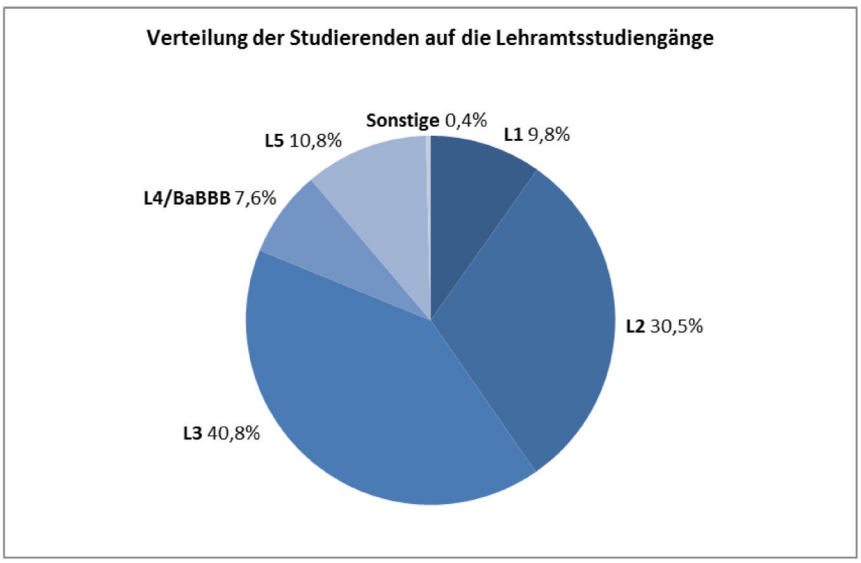

Abbildung 25. Prozentuale Verteilung der Studierenden auf die Lehramtsstudiengänge

Anmerkung. L1 (Grundschule), L2 (Haupt-/Realschule), L3 (Gymnasium), L4/BaBBB (Berufsschule) und L5 (Sonderschule) sowie die Kategorie „Sonstige“

0.4% (5 Studierende) studierten ein anderes Studienfach. Dabei gab ein Studierender an, B.Sc. Chemie zu studieren, eine Person studiert BBB-Nahrungsgewerbe und eine die Studienfachkombination L1 und Bildung und Förderung der Kindheit. Zwei der Studierenden mit anderem Studienfach machten keine Angabe zu ihrer Fachrichtung.

Die Studierenden befanden sich zum Zeitpunkt der Datenerhebung insgesamt im ersten bis 15. Semester (1162 Angaben). Dabei waren die meisten (95.1%) im ersten, 3.6% im zweiten bis fünften und 1.3% bereits in einem höheren Semester.

98.5% derjenigen, die uns hierzu eine Angabe machten, besuchten die Vorlesung zum ersten Mal, 1.5% (18 Studierende) haben sie bereits früher schon einmal besucht. Elf von diesen Studierenden waren bisher in einem früheren Semester in der Vorlesung, zwei schon zweimal und eine Person sogar dreimal. Die vier restlichen machten keine Angabe, wie oft sie die Vorlesung vorher bereits besucht hatten.

Die Studierenden sollten zudem angeben, wo sie ihr Abitur absolviert haben. Dabei gaben 98.7% (von 1162 Studierenden) an, ihr Abitur in Deutschland abgelegt zu haben. Abbildung 26 zeigt die Angaben dieser Studierenden auf die Frage hin, in welchem Bundesland sie es erworben haben. Die allermeisten haben ihr Abitur in Hessen abgelegt (74.3%). Häufiger genannt werden auch die Nachbarbundesländer Nordrhein-Westfalen (8.5%), Rheinland-Pfalz (3.3%), Niedersachsen (3.0%), Thüringen (2.2%) und Baden-Württemberg (1.6%).

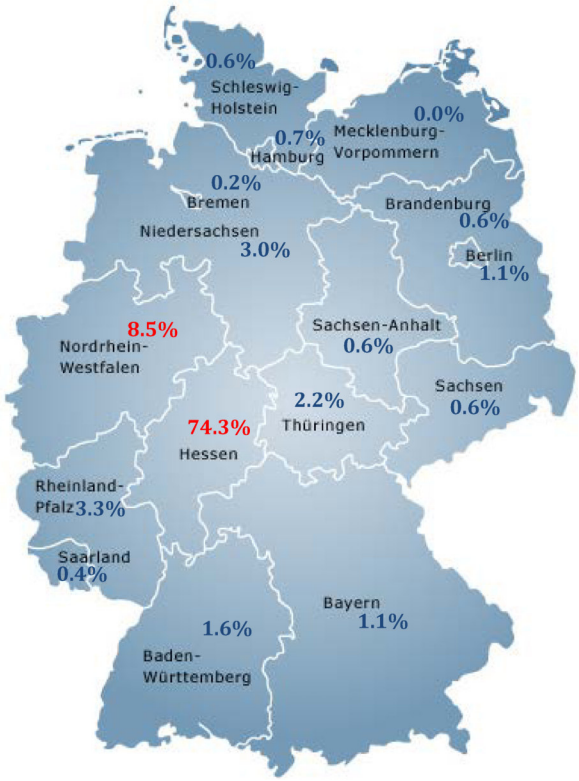

Abbildung 26. Gültige prozentuale Angaben der Studierenden auf die Frage: "Wo haben Sie Ihr Abitur erworben?[43]

[43] Karte der Bundesländer: Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz (2016)

13 Studierende (1.1%) erlangten ihre Hochschulzugangsberechtigung in einem anderen Staat als Deutschland: jeweils zwei Personen in Griechenland und der Türkei sowie jeweils eine Person in Weißrussland, Frankreich, Mazedonien, Polen, Russland und den USA. Zwei Personen gaben nicht an, in welchem Staat sie ihre Hochschulzugangsberechtigung erreicht hatten. Eine Person teilte mit, sie habe kein Abitur.

Darüber hinaus sollten die Studierenden Angaben darüber machen, wie gut ihr Abitur ausgefallen war. Die Abiturnoten der Gesamtstichprobe variierten zwischen 1.0 und 3.8. Im Mittel erreichte die Stichprobe einen Abiturschnitt von $M = 2.56$ ($SD = 0.54$, $n = 1158$).

6.1.2 Interesse und Vorkenntnisse hinsichtlich der Vorlesung und des Projekts

Bereits im Rahmen der Vorbefragung wurden die Studierenden gebeten anzugeben, wie sie ihre Teilnahme an dem Projekt empfanden. 96.8% waren der Teilnahme gegenüber positiv eingestellt, 3.2% sahen ihre Beteiligung an dem Projekt kritisch[44].

Auf die Frage, ob sie sich insgesamt für die Inhalte der Vorlesung „Psychologie für Lehramtsstudierende“ interessieren würden, gaben 65.6% der Studierenden an, sich sehr für die Vorlesungsinhalte zu interessieren, 32.3% interessieren sich etwas, 2.1% wenig und eine Person (0.1%) gar nicht für die Inhalte ($n = 1162$).

Diese Ergebnisse zeigen, dass die Studierenden zum größten Teil interessiert an die Vorlesung und das Projekt herangingen und dementsprechend grundlegend motiviert waren, die Inhalte der Veranstaltung zu lernen.

Darüber hinaus sollten die Studierenden ihre Vorkenntnisse zu den Inhalten der Vorlesung subjektiv einschätzen. 1163 Studierende machten uns hierzu eine Angabe. 33.2% von ihnen gaben an, keinerlei Vorkenntnisse zu besitzen, 62.7% waren der Meinung, über geringe Vorkenntnisse zu den Vorlesungsinhalten und 4.1% meinten sogar, über gute oder sehr gute Vorkenntnisse zu verfügen.

Der Vortest, der auf die Befragung folgte und zu Beginn des Wintersemesters die Vorkenntnisse zu den Inhalten der Vorlesung zudem objektiv erfassen sollte (siehe Abschnitt „6.3.2“), wurde von 1149 Studierenden (93.3%) vollständig bearbeitet. Diese erreichten zwischen 2.1 und 26.0 der maximal 30 möglichen Punkte (Verteilung siehe Abbildung 27).

---

[44] Dies wurden die Studierenden erneut zum Schluss des Projektes gefragt. Von den 502 Studierenden, die zu dieser Frage erneut eine Auskunft gaben, schätzten 89.5% die Teilnahme an dem Projekt immer noch positiv ein. 8.2% waren unentschieden und weitere 2.4% waren unzufrieden. Aufgrund des hohen Drop-outs spricht dies lediglich dafür, dass diejenigen, die bis zum Ende teilnahmen, zufrieden waren. Über die Gründe derjenigen, die diesen Fragebogen nicht mehr ausgefüllt haben, ist leider keine Aussage möglich.

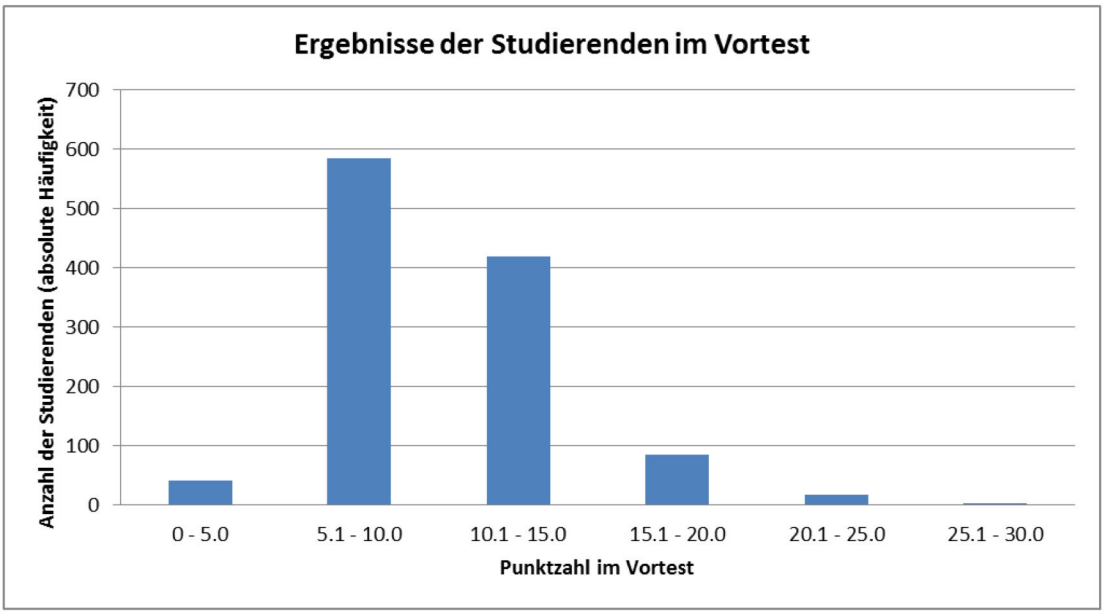

Abbildung 27. Kategorisierte Vortestergebnisse der Studierenden

Im Durchschnitt erlangten die Studierenden M = 10.23 Punkte (SD = 3.56). Sie konnten demnach gut ein Drittel der Fragen bereits zu Beginn der Vorlesung korrekt beantworten. Die subjektiven Einschätzungen der Studierenden decken sich folglich zufriedenstellend mit den objektiv erhobenen Vortestergebnissen.

Es gab allerdings sicherlich ein paar Studierende, die den Vortest vorab anonym geübt hatten (ohne Angabe ihrer Matrikelnummer). Es ist zudem davon auszugehen, dass weitere Studierende zum Bearbeiten das Skript zu Hilfe nahmen, obwohl darum gebeten wurde, dies nicht zu tun. Durch die Wahl der Aufgabenformate (u.a. Mehrfach- und Einfachauswahl-Aufgaben) kommt darüber hinaus eine gewisse Ratewahrscheinlichkeit zustande, so dass das wahre Vorwissen der Studierenden vermutlich unterhalb der Vortestergebnisse einzustufen ist.

6.1.3 Computerbezogene Fragen

Auf die Frage, wie oft sie ihren Computer nutzen, machten 1162 Studierende eine Aussage. 67.1% von ihnen gaben an, täglich vor ihrem Computer zu sitzen, 24.2% nutzen ihn vier bis sechs Tage und 7.7% immerhin ein bis drei Tage in der Woche. Nur 0.3% verwenden ihren Computer lediglich ein bis drei Mal im Monat, weitere 0.6% sogar noch seltener und ein Studierender (0.1%) gab sogar an, nie einen Computer zu benötigen. Abbildung 28 gibt diese Verteilung noch einmal bildlich wieder.

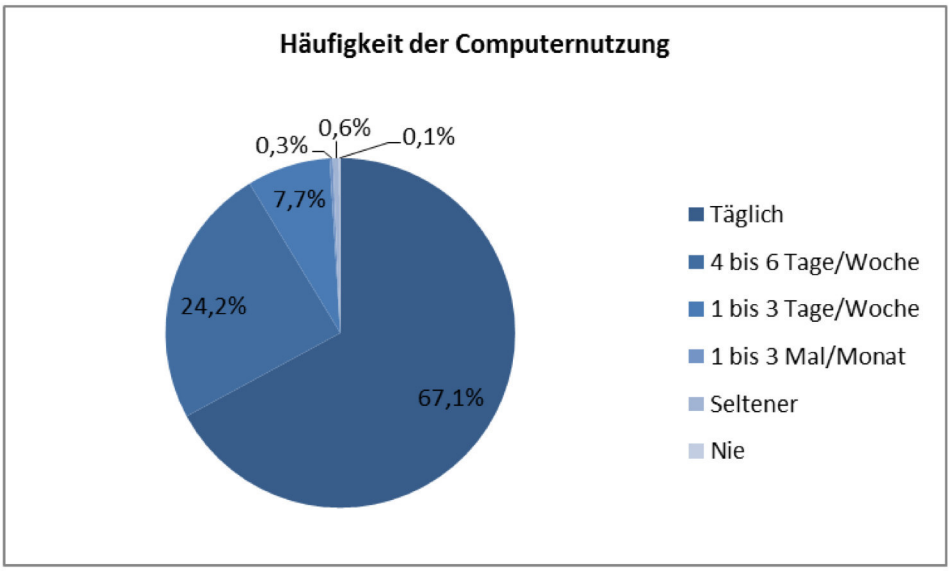

Abbildung 28. Prozentuale Häufigkeit der studentischen Angaben zu der Frage „Wie oft nutzen Sie Ihren Computer?“

Dabei nutzen die allermeisten (95.4% von insgesamt 1162 Studierenden) das Betriebssystem *Windows* von *Microsoft*, 3.8% nutzen *MacOS*, weitere 0.6% *Linux* und 0.2% ein anderes Betriebssystem. Eine Person war sich hinsichtlich ihres Betriebssystems unsicher.

Die meisten der Studierenden (n = 1163) schätzten ihre Fähigkeiten im Umgang mit dem Computer sehr positiv ein. Nur 21.5% waren der Meinung, im Umgang mit dem Computer Anfänger zu sein, 72.6% schätzten sich als fortgeschritten ein und 5.8% bezeichneten sich selbst als Experte.

Dies zeigen auch die Ergebnisse der Erhebung der Kurzversion des „*Fragebogens zur computerbezogenen Selbstwirksamkeit*“ (*CUSE-D-r*, nach Bescherer und Spannagel, 2011). Hiermit wurde die computerbezogene Selbstwirksamkeit durch ein objektives Maß erhoben. Die zwölf Fragebogenitems wurden von insgesamt 1162 Studierenden (94.3%) vollständig ausgefüllt. Maximal konnten in dem Fragebogen 72 Punkte erreicht werden, minimal 12 Punkte. Die Punktzahl der Studierenden lag im gesamten möglichen Bereich zwischen 12 und 72 Punkten.

Betrachtet man die Verteilung der studentischen Punktzahlen in Abbildung 29, zeigt sich eine deutliche Rechtsschiefe. Die meisten der Studierenden erreichten demnach ein hohes Ergebnis im *CUSE-D-r* – im Mittel erlangten sie einen Testwert von M = 54.04 (SD = 9.90). Dies entspricht einer relativ guten Computerselbstwirksamkeit und spricht unter anderem dafür, dass die Beteiligung am Projekt bei den allermeisten Studierenden weder zu Problemen noch zu Ängsten führen sollte.

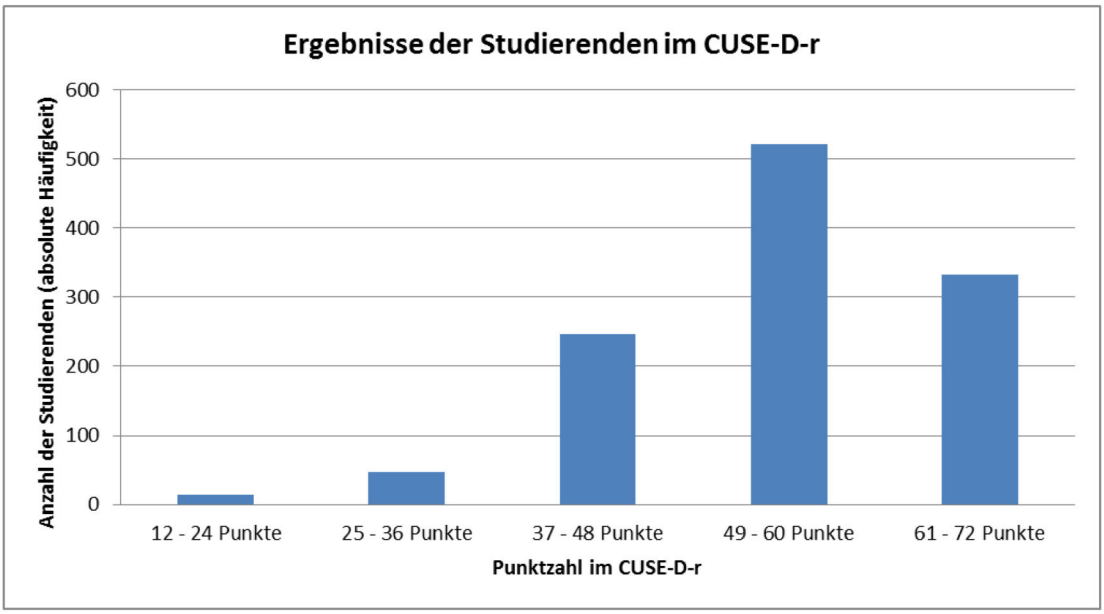

Abbildung 29. Kategorisierte Ergebnisse der Studierenden im CUSE-D-r

6.1.4 Personenvariablen

In zwei Nachbefragungen wurden zusätzlich verschiedene Personenvariablen erhoben, die einen Einfluss auf die Auswirkung der ARCS-Komponenten und ihre Einschätzung durch die Studierenden haben könnten. Nach Astleitner (2005) ist die Analyse der Zielgruppe wichtig für die Planung des ARCS-Einsatzes und auch die Ergebnisse von Astleitner und Hufnagl (2003), Means, Jonassen und Dwyer (1997), Song und Keller (2001) sowie Visser (1998) sprechen dafür, den Einsatz der ARCS-Strategien gezielt für die Lernendengruppe auszuwählen und nicht per se alle Komponenten zu integrieren. Ein solcher zielgruppenangemessener Einsatz der Komponenten wird in der Studie nicht umgesetzt, da alle sechs WBT-Varianten miteinander verglichen werden sollen. Dennoch wurden zur Ergänzung verschiedene Personenvariablen erhoben, um durch deren Kenntnis die Interpretation der Lernwirksamkeits- und Akzeptanzergebnisse zu unterstützen. Explizit wurden der „*Fragebogen zur Selbstbeschreibung der Aufmerksamkeit*“ von Bühner, Schmidt-Atzert, Richter und Grieshaber (2002), die „*Skalen zur Erfassung der Lern- und Leistungsmotivation von Studierenden*“ (*SELLMO-ST*) von Spinath, Stiensmeier-Pelster, Schöne und Dickhäuser (2002), die Skala „*Erfolgszuversicht*“ des „*Leistungsmotivationsinventar (LMI)*“ von Schuler, Prochaska und Frintrup (2001) sowie die „*Satisfaction With Life Scale*“ von Diener, Emmons, Larsen und Griffin (1985) erhoben. Die nähere Beschreibung der Fragebögen befindet sich in Abschnitt „6.3.2 Erhobene Variablen“. An dieser Stelle sollen die Mittelwerte und Standardabweichungen der Stichprobe zu den sieben Skalen vorgestellt und hinsichtlich der Normwerte, Vergleichsgruppenmittelwerte

oder dem möglichen Wertebereich der Skala eingeordnet werden. Tabelle 53 zeigt die Mittelwerte und Standardabweichungen der Studierenden für die Skalen.

Tabelle 53
Deskriptive Statistik der Personen-Skalen für die Gesamtstichprobe

	n	*M*	*SD*
Fragebogen zur Selbstbeschreibung der Aufmerksamkeit	771	14.44	7.66
SELLMO-ST - Lernzielorientierung	863	34.52	4.47
SELLMO-ST - Annäherungsleistungszielorientierung	863	22.36	4.43
SELLMO-ST - Vermeidungsleistungszielorientierung	863	18.84	6.12
SELLMO-ST - Arbeitsvermeidung	863	18.49	5.96
Skala Erfolgszuversicht aus dem LMI	771	43.83	9.36
Satisfaction With Life Scale	863	24.97	5.66

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Hinsichtlich der Skala „*Fragebogen zur Selbstbeschreibung der Aufmerksamkeit*“ von Bühner et al. (2002) konnte leider weder auf eine Normierung noch auf Mittelwerte einer Vergleichsstichprobe zurückgegriffen werden. Die Skala hat allerdings einen Wertebereich von 0 bis 54, was für einen sehr niedrigen Mittelwert (*M* = 14.44) der Studierendengruppe spricht. Dies spricht für eine relativ gute Selbstbeschreibung der Studierenden hinsichtlich ihrer Aufmerksamkeitsleistung, da sie im Mittel die meisten der Aufmerksamkeits-schwierigkeiten, die durch den Fragebogen abgefragt werden, als selten bis gar nicht vorhanden einstuften.

Da ebenfalls auf keine Normierung für den *SELLMO-ST* zurückgegriffen werden kann, wurden die mittleren Ergebnisse der Studierenden hinsichtlich der vier Skalen mit den bei Spinath et al. (2002) angegebenen Mittelwerten aus einer Vergleichsstichprobe verglichen. Die Autoren ermittelten für ihre Vergleichsstichprobe hinsichtlich der Skala „*Lernzielorientierung*“ einen Mittelwert von *M* = 33.67 (*SD* = 3.93), hinsichtlich der Skala „*Annäherung-Leistungszielorientierung*“ einen Mittelwert von *M* = 19.50 (*SD* = 5.18), für die Skala „*Vermeidungs-Leistungszielorientierung*“ einen Mittelwert von *M* = 16.31 (*SD* = 6.10) und für die Skala „*Arbeitsvermeidung*“ einen Mittelwert von *M* = 15.95 (*SD* = 5.43). Die vorliegende Stichprobe erreichte folglich bezüglich aller vier Skalen einen etwas höheren Mittelwert (siehe Tabelle 53). Die Studierenden waren insgesamt etwas höher lern- und deutlich höher leistungszielorientiert als die Vergleichsstichprobe, jedoch auch deutlich stärker arbeitsvermeidend.

Im Vergleich mit einer Normgruppe entspricht der Wert von *M* = 43.8 Punkten hinsichtlich der Skala „*Erfolgszuversicht*“ im Leistungsmotivationsinventar von Schuler et al. (2001) einem Prozentrang von 34. Die Studierendengruppe, die an der zweiten Studie teilnahm, erreichte folglich hinsichtlich der Erfolgszuversicht einen eher niedrigen mittleren Wert.

Auch für die *SWLS* von Diener et al. (1985) sind leider keine Normen verfügbar. Nach Schumacher (2003) spricht ein Mittelwert von *M* = 25.0 für eine eher zufriedene Lebenseinstellung der Studierenden. Der Wertebereich der Skala bewegt sich zwischen 5

und 31 Punkten, ein Wert von 20 würde nach Schumacher (2003) einer neutralen Lebenszufriedenheit entsprechen, ab 19 und niedriger spricht der Autor von Unzufriedenheit.
Nach diesen Ergebnissen sollten die Studierenden durch die E-Learning-Anwendungen insbesondere hinsichtlich ihrer Erfolgszuversicht unterstützt werden. Dies wird im Rahmen der Diskussion zu besprechen sein.

6.2 Durchführung

Die Untersuchung zur zweiten Studie wurde von Oktober 2010 bis Juli 2011 durchgeführt. Im Folgenden wird zunächst die vorbereitende Planungsphase dargestellt, bevor ausführlich auf die Datenerhebungsphase eingegangen wird.

6.2.1 Planung und Vorbereitung

Wie oben bereits erwähnt, fand die Durchführung der Studie erneut begleitend zur Vorlesung von Professor Glowalla statt. Dies ermöglicht einen Vergleich zur ersten Studie. Zudem waren die notwendigen Materialien durch die erste Studie schon zum Teil vorhanden. Da allerdings in der zweiten Studie Vergleiche zwischen sechs verschiedenen WBT-Versionen angestellt werden sollten, war es nicht möglich, die Studie erneut über alle zwölf Vorlesungsthemen durchzuführen. Aus diesem Grunde wurden drei Vorlesungsthemen ausgewählt. Dabei wurde unter anderem darauf geachtet, dass die Themen ausreichend Möglichkeiten für die WBT-Variationen boten und es möglich war, viele der ARCS-Komponenten sinnvoll einzubauen. Es musste zum Beispiel möglich sein, die Themen in kurze unabhängige Einheiten einzuteilen, damit zwischendurch Übungsaufgaben integriert oder einzelne Teilbereiche des Themas durch die Lernenden in freier Reihenfolge bearbeitet werden können. Zudem sollten die Inhalte Basis für anschauliche Anwendungsbeispiele und -aufgaben sowie Lernspiele bieten.
Am geeignetsten erschienen nach diesen Kriterien die Themen „Klassische und operante Konditionierung“ (im Folgenden kurz „Konditionierung“ genannt), „Unterrichtsmethoden, Unterrichtsplanung und -organisation“ („Unterrichten“) sowie „Pädagogische Diagnostik“ („Diagnostik“).
Einigen Schwierigkeiten, die im Verlauf der Durchführung der ersten Studie aufgetreten waren, konnte nun vorgebeugt werden. Die Studierenden wurden diesmal langsamer an die Durchführung herangeführt. Dies war durch die Durchführung der Studie in nur drei (anstatt in zwölf) Vorlesungswochen leichter möglich.

Den Studierenden wurde deshalb bereits in der ersten Vorlesungssitzung mitgeteilt, wie die Nutzung der E-Learning-Anwendungen ablaufen würde. Um sie zu motivieren, an der vorlesungsbegleitenden Studie teilzunehmen, wurden ihnen aussagekräftige Ergebnisse aus Studie I präsentiert, die ihnen zeigten, dass den Studierenden die Teilnahme an der Studie Spaß gemacht hatte und sie mit den E-Learning-Anwendungen mehr gelernt hatten als in der Vorlesung. Außerdem wurde den Studierenden mitgeteilt, dass diejenigen, die in ausreichendem Maße an der Studie teilnehmen, die E-Learning-Anwendungen auch zu ihrer Vorbereitung auf die MAP nutzen dürfen.
Da in Studie I einige technische Probleme insbesondere mit den E-Lectures auftraten, erhielten sie darüber hinaus die Möglichkeit, bereits eine „Probe-E-Lecture“ und ein „Probe-WBT“ zu nutzen (jeweils zum Thema „Gedächtnis I“), um zu testen, ob die Anwendungen auf ihrem Rechner problemlos funktionieren. So konnten die Tutoren den Studierenden bei Problemen noch im Vorfeld helfen und die Studierenden selbst konnten sich schon einmal mit den Anwendungen vertraut machen.

6.2.2 Datenerhebung

Abbildung 30 zeigt im Überblick die verschiedenen Phasen der Datenerhebung zu Studie II. Die einzelnen Phasen werden im Folgenden ausführlich erläutert.



Abbildung 30. Überblick über die Datenerhebung zu Studie II

6.2.2.1 Vorbefragung und Vortest

Mittels des Online-Befragungstools *Unipark 7.0* (siehe Studie I) wurde den Studierenden zwischen dem 28. Oktober und dem 05. November 2010 der erste Fragebogen freigeschaltet (siehe Abschnitt „6.3.2 Abhängige Variablen“).
Wie bereits im Rahmen der Stichprobenbeschreibung erläutert, nahmen 1149 (93.6%) der Studierenden vollständig an dieser Vorerhebung teil. Um die Motivation der Studierenden für die weitere Teilnahme an der Studie zu erhöhen, wurden den Studierenden die Ergebnisse des Vortests in einer zeitnahen Vorlesungssitzung präsentiert.

6.2.2.2 Dreifache E-Learning-Nutzung und -Evaluation

Insgesamt wurde die Vorlesung dreimal durch eine E-Learning-Anwendung ersetzt. Dies geschah – wie oben erläutert – zu den Themen „Konditionierung“, „Unterrichten“ und

„Diagnostik“. An diesen drei Erhebungszeitpunkten bearbeiteten die Studierenden eine von sechs möglichen WBT-Varianten oder eine E-Lecture. Alle E-Learning-Anwendungen waren möglichst inhaltsgleich und variierten lediglich wie in Abschnitt „6.3.1 Lehrmaterial: E-Lectures und Web-based Trainings“ beschrieben.

Die Phase der Nutzung der E-Learning-Anwendungen begann im Dezember 2010. Hier erhielten die Studierenden per Email einen Überblick über den Ablauf der ersten E-Learning-Nutzung und -Evaluation. Diese Email enthielt zudem den Zugangsschlüssel für die Online-Veranstaltung innerhalb der Lernplattform.

Tabelle 54
Experimentalgruppen der Studie II mit Variationen über die drei Erhebungen

Gruppe	1. Erhebung	2. Erhebung	3. Erhebung
1	E-Lecture	0-WBT	A-WBT
2	E-Lecture	0-WBT	R-WBT
3	E-Lecture	0-WBT	C-WBT
4	E-Lecture	0-WBT	S-WBT
5	E-Lecture	0-WBT	ARCS-WBT
6	E-Lecture	A-WBT	0-WBT
7	E-Lecture	A-WBT	R-WBT
8	E-Lecture	A-WBT	C-WBT
9	E-Lecture	A-WBT	S-WBT
10	E-Lecture	A-WBT	ARCS-WBT
11	E-Lecture	R-WBT	0-WBT
12	E-Lecture	R-WBT	A-WBT
13	E-Lecture	R-WBT	C-WBT
14	E-Lecture	R-WBT	S-WBT
15	E-Lecture	R-WBT	ARCS-WBT
16	E-Lecture	C-WBT	0-WBT
17	E-Lecture	C-WBT	A-WBT
18	E-Lecture	C-WBT	R-WBT
19	E-Lecture	C-WBT	S-WBT
20	E-Lecture	C-WBT	ARCS-WBT
21	E-Lecture	S-WBT	0-WBT
22	E-Lecture	S-WBT	A-WBT
23	E-Lecture	S-WBT	R-WBT
24	E-Lecture	S-WBT	C-WBT
25	E-Lecture	S-WBT	ARCS-WBT
26	E-Lecture	ARCS-WBT	0-WBT
27	E-Lecture	ARCS-WBT	A-WBT
28	E-Lecture	ARCS-WBT	R-WBT
29	E-Lecture	ARCS-WBT	C-WBT
30	E-Lecture	ARCS-WBT	S-WBT
31	0-WBT	E-Lecture	A-WBT
32	0-WBT	E-Lecture	R-WBT
33	0-WBT	E-Lecture	C-WBT
34	0-WBT	E-Lecture	S-WBT
35	0-WBT	E-Lecture	ARCS-WBT
36	A-WBT	E-Lecture	0-WBT
37	A-WBT	E-Lecture	R-WBT
38	A-WBT	E-Lecture	C-WBT
39	A-WBT	E-Lecture	S-WBT
40	A-WBT	E-Lecture	ARCS-WBT
41	R-WBT	E-Lecture	0-WBT
42	R-WBT	E-Lecture	A-WBT
43	R-WBT	E-Lecture	C-WBT
44	R-WBT	E-Lecture	S-WBT
45	R-WBT	E-Lecture	ARCS-WBT
46	C-WBT	E-Lecture	0-WBT
47	C-WBT	E-Lecture	A-WBT
48	C-WBT	E-Lecture	R-WBT
49	C-WBT	E-Lecture	S-WBT
50	C-WBT	E-Lecture	ARCS-WBT
51	S-WBT	E-Lecture	0-WBT
52	S-WBT	E-Lecture	A-WBT
53	S-WBT	E-Lecture	R-WBT
54	S-WBT	E-Lecture	C-WBT
55	S-WBT	E-Lecture	ARCS-WBT
56	ARCS-WBT	E-Lecture	0-WBT
57	ARCS-WBT	E-Lecture	A-WBT
58	ARCS-WBT	E-Lecture	R-WBT
59	ARCS-WBT	E-Lecture	C-WBT
60	ARCS-WBT	E-Lecture	S-WBT
61	0-WBT	A-WBT	E-Lecture
62	0-WBT	R-WBT	E-Lecture
63	0-WBT	C-WBT	E-Lecture
64	0-WBT	S-WBT	E-Lecture
65	0-WBT	ARCS-WBT	E-Lecture
66	A-WBT	0-WBT	E-Lecture
67	A-WBT	R-WBT	E-Lecture
68	A-WBT	C-WBT	E-Lecture
69	A-WBT	S-WBT	E-Lecture
70	A-WBT	ARCS-WBT	E-Lecture
71	R-WBT	0-WBT	E-Lecture
72	R-WBT	A-WBT	E-Lecture
73	R-WBT	C-WBT	E-Lecture
74	R-WBT	S-WBT	E-Lecture
75	R-WBT	ARCS-WBT	E-Lecture
76	C-WBT	0-WBT	E-Lecture
77	C-WBT	A-WBT	E-Lecture
78	C-WBT	R-WBT	E-Lecture
79	C-WBT	S-WBT	E-Lecture
80	C-WBT	ARCS-WBT	E-Lecture
81	S-WBT	0-WBT	E-Lecture
82	S-WBT	A-WBT	E-Lecture
83	S-WBT	R-WBT	E-Lecture
84	S-WBT	C-WBT	E-Lecture
85	S-WBT	ARCS-WBT	E-Lecture
86	ARCS-WBT	0-WBT	E-Lecture
87	ARCS-WBT	A-WBT	E-Lecture
88	ARCS-WBT	R-WBT	E-Lecture
89	ARCS-WBT	C-WBT	E-Lecture
90	ARCS-WBT	S-WBT	E-Lecture

Sobald die Studierenden sich in diese Veranstaltung eingeschrieben hatten, wurden sie zufällig in eine der insgesamt 90 Experimentalgruppen (siehe Tabelle 54) eingeteilt, die alle

möglichen Gruppierungsvariationen abdeckten und es möglich machten, Reihenfolgeeffekte über die drei Evaluationen hinweg auszuschließen.

89 der 90 Gruppen umfassten jeweils 12 oder 13 Studierende, eine Gruppe bestand aufgrund einer Umgruppierung aus 14 Studierenden.

Für jede Evaluationswelle wurden zehn Gruppen den sechs verschiedenen WBT-Variationen zugeordnet, 30 Gruppen nutzten jeweils die inhaltlich vergleichbare E-Lecture. Damit wurde sichergestellt, dass jeder Studierende über alle drei Termine hinweg zwei WBT-Varianten sowie einmal die E-Lecture kennenlernen konnte.

Die Studierenden konnten die E-Learning-Anwendungen in der Folge jeweils von Mittwoch (Tag der zu ersetzenden Vorlesungssitzung) bis Sonntag so häufig wie gewünscht nutzen. Nachdem die Anwendungen gesperrt worden waren, wurde die zugehörige Online-Nachbefragung freigeschaltet. Diese konnte bis zum darauffolgenden Freitag von den Studierenden bearbeitet werden. Die dritte Evaluationswelle zum Thema „Diagnostik" musste zeitlich etwas gekürzt werden, da aus organisatorischen Gründen die Erhebung der Probeklausur begonnen werden musste. Die Nachbefragung stand hier deshalb lediglich bis zum Mittwoch zur Verfügung.

Tabelle 55 zeigt noch einmal im Überblick die Erhebungszeiträume sowie die Beteiligung der Studierenden an den drei Evaluationswellen.

Tabelle 55
Zeitraum der Evaluationswellen sowie Anzahl der Studierenden, die jeweils teilgenommen haben

Thema	Zeitraum	Nutzung der Anwendung		Evaluation	Nutzung der Anwendung und Evaluation	
		Gesamt	> 5 Min. < 5 Std.		Gesamt	> 5 Min. < 5 Std.
Konditionieren	08. – 17.12.2010	779	620	949	722	589
Unterrichten	19. – 28.01.2011	887	712	924	823	665
Diagnostik	02. – 09.02.2011	803	622	860	724	574

Anmerkung. Angegeben ist nur die Nutzung der Anwendung, nur die Teilnahme an der Evaluation sowie die vollständige Teilnahme an Nutzung und Evaluation; Dabei stellen die Angaben in der Spalte „> 5 Min./< 5 Std." die Nutzung mit realistischer Zeitdauer dar; Die letzte Spalte entspricht der Stichprobe, die in die Auswertungen eingeht.

Bei Betrachtung der Nutzungszeiten wurde leider festgestellt, dass es Studierende gab, für die zwar eine Nutzung der Web-based Trainings, jedoch keine Nutzungszeit aufgezeichnet worden war. Dies war vermutlich dadurch begründet, dass diese Studierenden die ihnen zugeordnete Anwendung in *Moodle* zwar aufgerufen hatten, dann aber am Popupblocker gescheitert waren und das WBT doch nicht nutzen konnten. Diese Studierenden wurden aus der Auswertung herausgenommen.

Des Weiteren wurden vergleichbar zu Studie I alle diejenigen nicht für die Auswertungen berücksichtigt, die laut den Aufzeichnungen die E-Learning-Anwendungen kürzer als fünf Minuten oder länger als fünf Stunden bearbeitet hatten.

Insgesamt gingen in die Auswertungen demnach nur diejenigen Studierenden ein, die die jeweilige E-Learning-Anwendung im sinnvollen Zeitrahmen genutzt und die zugehörige Evaluation durchgehend bearbeitet hatten. In den einzelnen WBT-Nutzergruppen verblieben trotz dieser strengen Einschränkungen jeweils zwischen 47 und 80, in der E-Lecture-Nutzergruppe waren es generell über 200 Studierende. Über die drei Evaluationswellen hinweg nahmen unter Berücksichtigung dieser Einschränkungen von den 1232 Studierenden 27.6% an keiner der drei Evaluationswellen, 21.6% an einer, 25.6% an zweien und 25.2% an allen drei Erhebungen vollständig teil.

6.2.2.3 Probeklausur

Um die längerfristige Lernleistung sowie eine Einschätzung des Projektes zu erheben, wurde den Studierenden vom 10. bis zum 15. Februar 2011 mittels dem Online-Befragungstool *Unipark 7.0* eine Probeklausur (siehe Abschnitt 6.3.2) mit zwei abschließenden Evaluationsfragen zur Verfügung gestellt. In der letzten Vorlesungssitzung wurde die Auswertung dieser Probeklausur über alle Teilnehmer hinweg vorgestellt.
682 Studierende (55.4% der Gesamtstichprobe) bearbeiteten die Probeklausur vollständig, wobei ein Teil von ihnen die beiden Evaluationsfragen am Ende leider nicht mehr beantwortete.

6.2.2.4 Freiwillige Nutzung und abschließende Datenerhebung

Alle Studierenden, die an zwei Evaluationswellen vollständig teilgenommen hatten, durften die 33 E-Learning-Anwendungen ab dem Mai 2011 für ihre Klausurvorbereitung nutzen. Die Einschränkung der Nutzungszeit von mehr als fünf Minuten und weniger als fünf Stunden wurde hier nicht vorgenommen – diese war nur für die Daten-Auswertung von Bedeutung. Dementsprechend wurden die Anwendungen für 784 Studierende freigeschaltet. 639 (81.5%) von ihnen nahmen sie bis zur Modulabschlussprüfung in Anspruch.
Für alle elf Vorlesungsthemen wurden je drei Anwendungen zur Verfügung gestellt: eine E-Lecture, ein WBT ohne (0-WBT) sowie ein WBT mit Umsetzung der ARCS-Komponenten (ARCS-WBT). Dies sollte einen Blick darauf ermöglichen, welche E-Learning-Anwendung von den Studierenden freiwillig zur Klausurvorbereitung genutzt wird. Den Studierenden wurden die drei E-Learning-Varianten erläutert. Insbesondere wurden das 0-WBT als kurzgefasstes WBT (K-WBT) und das ARCS-WBT als elaboriertes WBT (E-WBT) umschrieben. Es wurde ihnen darüber hinaus deutlich gemacht, dass die Anwendungen für jedes Thema inhaltsgleich sind. Sie wurden darauf hingewiesen, dass sie keine Inhalte übersehen, wenn sie zu jedem Thema nur jeweils eine der drei Anwendungen bearbeiten

und ihnen wurde empfohlen, dass sie sich hierfür diejenige Variante auswählen sollten, mit der sie am besten lernen können.
Etwa zwei Wochen vor der Modulabschlussprüfung wurde zudem noch einmal die Probeklausur zuzüglich einer kurzen adaptiven Nachbefragung erhoben. Am 25. und 26. Juli 2011 fand in mehreren Studierendengruppen die Modulabschlussprüfung statt (siehe Abschnitt „6.3.2“).

6.3 Material

6.3.1 Lehrmaterial: E-Lectures und Web-based Trainings

Bei den genutzten E-Lectures handelte es sich um die bereits in Studie I genutzten Anwendungen. Diese wurden ausführlich in den Abschnitten „1.4.2 E-Lecture“ sowie „2.3.1 Lehrmaterialien: Vorlesung, E-Lecture und Web-based Trainings“ beschrieben.
Auch die eingesetzten Web-based Trainings sind auf Grundlage derjenigen entstanden, die in Studie I zum Einsatz kamen. Diese wurden ebenfalls in den Abschnitten „1.4.3 Web-based Training (WBT)“ sowie „2.3.1 Lehrmaterialien: Vorlesung, E-Lecture und Web-based Trainings“ ausführlich vorgestellt. Für Studie II wurden allerdings – wie oben bereits erläutert – zu den drei Themen „Konditionierung“, „Unterrichten“ sowie „Diagnostik“ sechs neue Variationen erstellt:

- Das „0“-WBT, für dessen Erstellung keine der ARCS-Komponenten berücksichtigt wurde,
- die „A (Aufmerksamkeit)“, „R (Relevanz)“, „C (Erfolgszuversicht)“ und „S (Zufriedenheit)“-WBTs, innerhalb derer jeweils eine der von Keller (2010) vorgeschlagenen Komponenten umgesetzt wurde sowie
- das „ARCS“-WBT, das insgesamt nach dem Keller’schen Modell elaboriert wurde.

Die sechs WBT-Variationen werden im Folgenden ausführlich anhand von Beispielseiten vorgestellt.

0-WBT

Die erste WBT-Variation wurde so erstellt, dass möglichst keine der von Keller (2010) genannten Aspekte umgesetzt waren. Es sollte lediglich das reine Faktenwissen sowie die grundlegend notwendigen Steuerungsmöglichkeiten enthalten. Dazu wurden die WBTs aus Studie I als Grundlage genommen und die darin integrierten Aspekte, wie z.B. eine direkte

Anrede oder Alltagsbeispiele, möglichst vollständig entfernt. Zum Einsatz kam hier erneut das Autorentool *Lectora*[45].

Die resultierenden 0-WBTs enthalten auf jeder Seite lediglich einen Text und ein einheitliches Bild – wenn nicht sowieso ein notwendiges informierendes Bild auf der Seite enthalten war (siehe Abbildung 31, rechts oben sowie links unten).



Abbildung 31. Beispielseiten aus dem 0-WBT zum Thema "Konditionierung"
Anmerkung. Links oben: Inhaltsverzeichnis; rechts oben: typische Seite ohne Bild; links unten: typische Seite mit Bild; rechts unten: Ende-Seite

Jede Inhaltsseite weist das gleiche dreispaltige Format auf. Auch enthalten die 0-WBTs weder Videos noch Übungsaufgaben, wodurch sie möglichst abwechslungsarm wirken sollten. Die Gestaltung selbst sollte nicht die Aufmerksamkeitserhaltung der Lernenden unterstützen. Deshalb – und auch um das Relevanzgefühl der Studierenden nicht zu fördern – ist der Ausdruck im 0-WBT insgesamt sehr kurz gehalten. Ausschmückende oder humoristische Ausführungen wurden aus dem Text entfernt und die Lernenden wurden nicht persönlich angesprochen. Die Texte bieten folglich lediglich Fakten – Beispiele, Anwendungsbezüge und Fragen an die Lernenden wurden eliminiert, da die Textbeschaffenheit selbst keine Neugier oder Interesse erregen sollte.

Zudem sollte auch die Transparenz des 0-WBTs nicht unterstützt werden. Deshalb wurde auf Hervorhebungen von wichtigen Inhalten, Lernziele, eine Einleitung sowie eine Angabe

[45] Autorenwerkzeug der Firma *Trivantis* zur Erstellung von Web-based Trainings (siehe Trivantis, 2016)

der Gesamtseitenzahl verzichtet. Lediglich die aktuelle Seitenzahl wird dem Nutzer angegeben. Auch wird zu Beginn des WBTs ein nur sehr grob gehaltenes nicht-interaktives Inhaltsverzeichnis dargeboten (siehe Abbildung 31, links oben). Innerhalb des 0-WBTs wurde auf Überleitungen und Abschnittsfolien verzichtet. Die Studierenden haben dementsprechend lediglich die Möglichkeit, über Vor- und Zurückbuttons zu navigieren und sich über den Text inhaltlich zu orientieren. Auch auf Übungstests wurde vollständig verzichtet.

Um auch die Zufriedenheitsaspekte des ARCS-Modells nicht zu berücksichtigen, wurden alle Textstellen entfernt, an denen das ursprüngliche WBT an das Vorwissen oder die Erfahrungen der Studierenden angeknüpft hatte. Zudem werden keine weiterführenden Quellen oder Lernspiele angeboten.

Das 0-WBT endet dementsprechend nach den faktisch und einheitlich gehaltenen Inhaltsseiten mit einer schmucklosen „Ende"-Seite (siehe Abbildung 31, rechts unten), dem Impressum und einer Seite, die das Schließen des WBTs erläutert.

Aufmerksamkeits(A)-WBT

Für die Erstellung des „A (Aufmerksamkeits)"-WBTs wurde das 0-WBT um solche Komponenten ergänzt, die nach Keller (2010) die Aufmerksamkeit der Nutzer erhöhen sollen.

Um **Wahrnehmungsvielfalt** zu erzeugen, sollte die Darstellung der Lerninhalte abwechslungsreich erfolgen, aber nicht so abwechslungsreich, dass die Studierenden vom Lerninhalt abgelenkt werden. Neben Texten und informierenden Bildern wurden dementsprechend auch dekorative Bilder (z.B. thematisch passende Comics, siehe z.B. Abbildung 32, rechts oben), inhaltlich passende Videos, Audios oder Animationen (siehe Abbildung 32, rechts Mitte) sowie unerwartete Ereignisse (z.B. Begrüßungsmelodie auf der ersten Seite des WBTs, siehe Abbildung 32, rechts oben) ergänzt. Damit die Animationen nicht vom Lerninhalt ablenken, stoppen sie automatisch nach kurzer Zeit. Die Videos und Audios können von den Nutzern selbst gesteuert werden.

Auch der Wechsel zwischen Theorie und Übung wurde angestrebt. Da Übungstests mit Rückmeldungen nicht integriert werden durften (dies ist Bestandteil des Erfolgszuversichts(C)-WBTs), erfolgte dies über Denkfragen, die die Studierenden zwischendurch zum Nachdenken anregen sollen (siehe Abbildung 32, links unten). Um den Effekt des „Wachrüttelns" durch die Denkfragen zu verstärken, wurden sie zusätzlich durch die „*Jeopardy*-Melodie" unterlegt, die durch die Nutzer aber jederzeit beendet werden konnte. Zudem wurde abstraktes Wissen in den Texten konkreter dargestellt und ggf. visualisiert.

Im Aufmerksamkeits(A)-WBT sollten überdies **Neugier und Eigeninitiative** der Studierenden gefördert werden. Die Lernenden werden innerhalb des WBTs deshalb vor Probleme gestellt, die durch die folgenden Lerninhalte gelöst werden können. Dies geschieht zu Beginn der WBTs durch einen provokativen Einstieg (als Beispiel siehe Abbildung 32, links Mitte) und zwischendurch durch die bereits erwähnten Denkfragen.
Auch in Bezug auf Format und Textstil sollte angemessen **Variabilität und Abwechslung** erzeugt werden. Wichtiges wurde durch Markierungen, Umrahmungen oder durch direkte Hinweise („Ganz wichtig:…“) hervorgehoben (siehe Abbildung 32, rechts Mitte), um die Aufmerksamkeit klar zu lenken. Des Weiteren wurden die Lerninhalte durch Abschnittsfolien in kurze Einheiten unterteilt (siehe Abbildung 32, rechts unten).

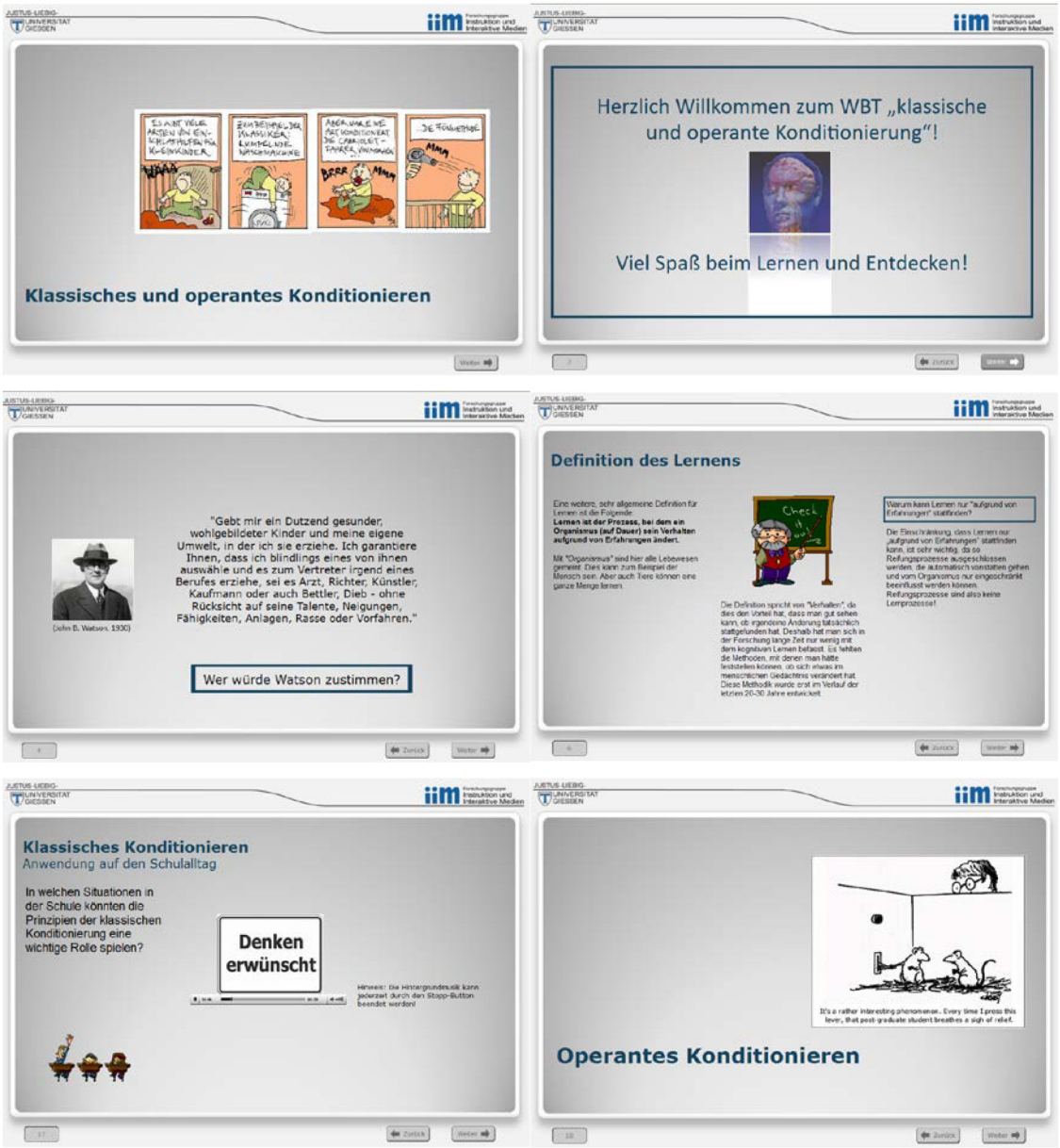

Abbildung 32. Beispielseiten aus dem A-WBT zum Thema "Konditionierung"
Anmerkung. Links oben: Titelseite mit Comic; rechts oben: Begrüßungsseite; links Mitte: Provokativer Einstieg; rechts Mitte: Beispielseite mit Animation, Hervorhebungen und Variabilität in der Gestaltung; links unten: Denkfrage; rechts unten: Abschnittsfolie mit Comic

Relevanz(R)-WBT

Für das „R (Relevanz)"-WBT wurde das 0-WBT hinsichtlich der Komponenten optimiert, die die Ziel- und Prozessrelevanz verbessern sollen. Dabei beinhaltet der Zielaspekt, dass die Lernenden den Lernstoff als nützlich für das Erreichen bestimmter Ziele wahrnehmen. Der Prozessaspekt bezieht sich auf die Relevanz, die durch den Lernprozess selbst entsteht (Niegemann et al., 2008).

Um die **Lehrzielorientierung** („Zielrelevanz") zu verbessern, werden den Studierenden innerhalb der R-WBTs sowohl kurz- als auch langfristige Lernziele sowie der Nutzen der Lerneinheit angegeben. Kurzfristig beziehen sich die Lernziele auf das Lehramtsstudium, langfristig wird auch der Nutzen für den späteren Unterrichtsalltag dargestellt. Zudem bieten die R-WBTs Anwendungsbeispiele für das theoretische Wissen (siehe Abbildung 33).

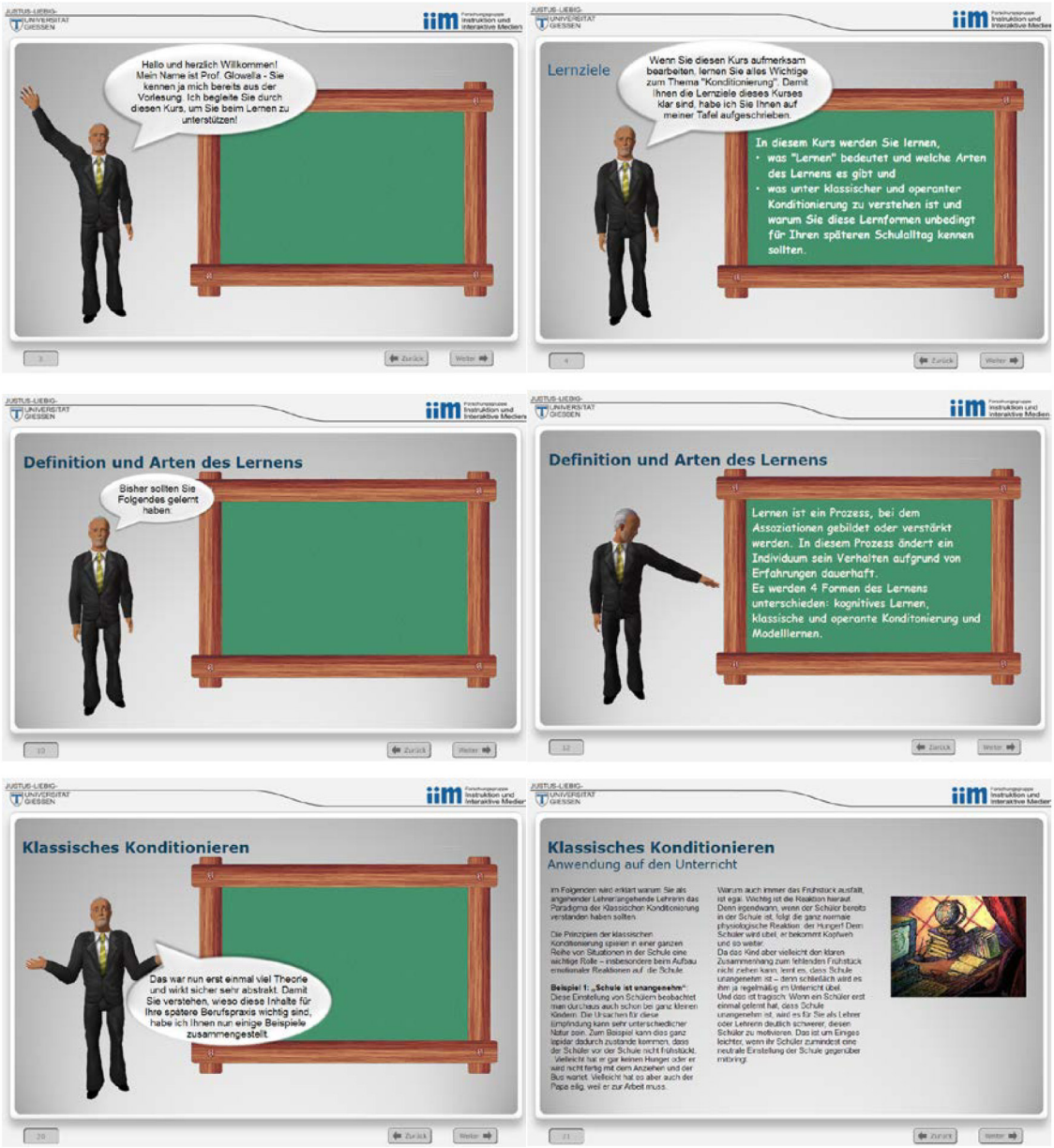

Abbildung 33. Beispielseiten aus dem R-WBT zum Thema "Konditionierung"
Anmerkung. Links oben: Begrüßung durch den pädagogischen Agenten; rechts oben: Vorstellung der Lernziele; Mitte links und rechts: Erinnerung an die Lernziele; unten links und rechts: Hinweis auf Anwendungsbeispiele

Die Prozessrelevanz soll laut Keller (2010) unter anderem dadurch gestärkt werden, dass eine **Vertrautheit** mit dem Lernmaterial hergestellt wird. Dies wird im WBT durch wiederkehrende Anknüpfung des neuen Wissens an das Vorwissen und an die Erfahrungen der Studierenden umgesetzt. Das theoretisch zu Lernende wird immer wieder mit praktischen Anwendungsmöglichkeiten untermalt, die Bezüge zu den Erfahrungen und dem späteren Berufsalltag der Studierenden bieten.

Darüber hinaus werden die Studierenden innerhalb der Texte der R-WBTs persönlich angesprochen (z.B. „Wie Sie sehen…"). Die Texte im WBT wurden so angepasst, dass sie weniger formell klingen – es wurde allerdings darauf geachtet, dass der Sprachstil dennoch seriös wirkte, damit die E-Learning-Anwendung weiterhin von den Studierenden ernstgenommen wird. Zudem wurde das WBT durch einen pädagogischen Agenten ergänzt. Dabei wurde eine Avatarvisualisierung von Professor Glowalla genutzt, um das Vertrautheitsgefühl der Studierenden zu steigern. Dieser „kleine Professor Glowalla" spricht die Studierenden persönlich an und erfüllt verschiedene Aufgaben: Zu Beginn des WBTs begrüßt er die Studierenden, am Ende verabschiedet er sie. Außerdem obliegt ihm die Aufgabe, den Lernenden zu Beginn und ggf. auch zwischendurch die Lernziele und den Nutzen der Inhalte zu vermitteln sowie am Ende noch einmal auf diese Lernziele hinzuweisen und den Studierenden zu empfehlen, den Lernstoff wiederholt zu bearbeiten (siehe Abbildung 33).

Der dritte von Keller genannte Punkt „**Motivanpassung**" konnte leider im Rahmen dieser Untersuchung nicht umgesetzt werden. Eine individuelle Anpassung der Lernanwendung an die Motive der einzelnen Lernenden (Leistungsmotiv, Anschlussmotiv, Machtmotiv; siehe Niegemann et al., 2008) wäre an vielen Stellen auf Schwierigkeiten gestoßen. Zum Beispiel hätten nur für die Studierenden, die diese WBT-Variation nutzen, zur Anregung des Anschlussmotivs weitere Kommunikationsmöglichkeiten (Chats, Foren) angeboten werden müssen. Dies wäre in dem vorliegenden Setting nicht kontrollierbar gewesen. Daneben wären auch zusätzliche Präsenz-Maßnahmen, wie z.B. Gruppenarbeiten, unerlässlich gewesen, was aufgrund der Menge an Studierenden nicht möglich war und zudem die Vergleichbarkeit zu den anderen WBT-Variationen deutlich eingeschränkt hätte. Die nicht vollständige Umsetzung der Relevanzstrategien muss allerdings im Rahmen der Ergebnisinterpretation bedacht werden.

Erfolgszuversichts(C)-WBT

Zur Erstellung des „C (Erfolgszuversicht)"-WBTs wurden diejenigen Komponenten umgesetzt, die dazu dienen sollen, die Erfolgserwartungen der Studierenden zu erhöhen.

Um den Lernenden den Glauben an die eigene Kompetenz und an den Lernerfolg zu vermitteln, sollten den Studierenden die **Lernanforderungen** transparent dargestellt

werden. Das Erfolgzuversichts(C)-WBT verfügt aus diesem Grunde als einziges über Übungstests, die zudem unterschiedliche Leistungsniveaus ansprechen. Die drei unterschiedlich schwierigen Tests (leicht, mittel und anspruchsvoll) bestehen aus jeweils zehn bis zwölf Fragen mit unterschiedlichen Aufgabenformaten. Nach Bearbeitung jeder Aufgabe erhält der Lernende ein informatives Feedback. Am Ende jedes Tests wird ein Überblick über die erreichte Punktzahl sowie eine Interpretationshilfe geboten. Dadurch soll die Transparenz der Bewertungskriterien im Test verdeutlicht werden (siehe Abbildung 34, Mitte und unten).

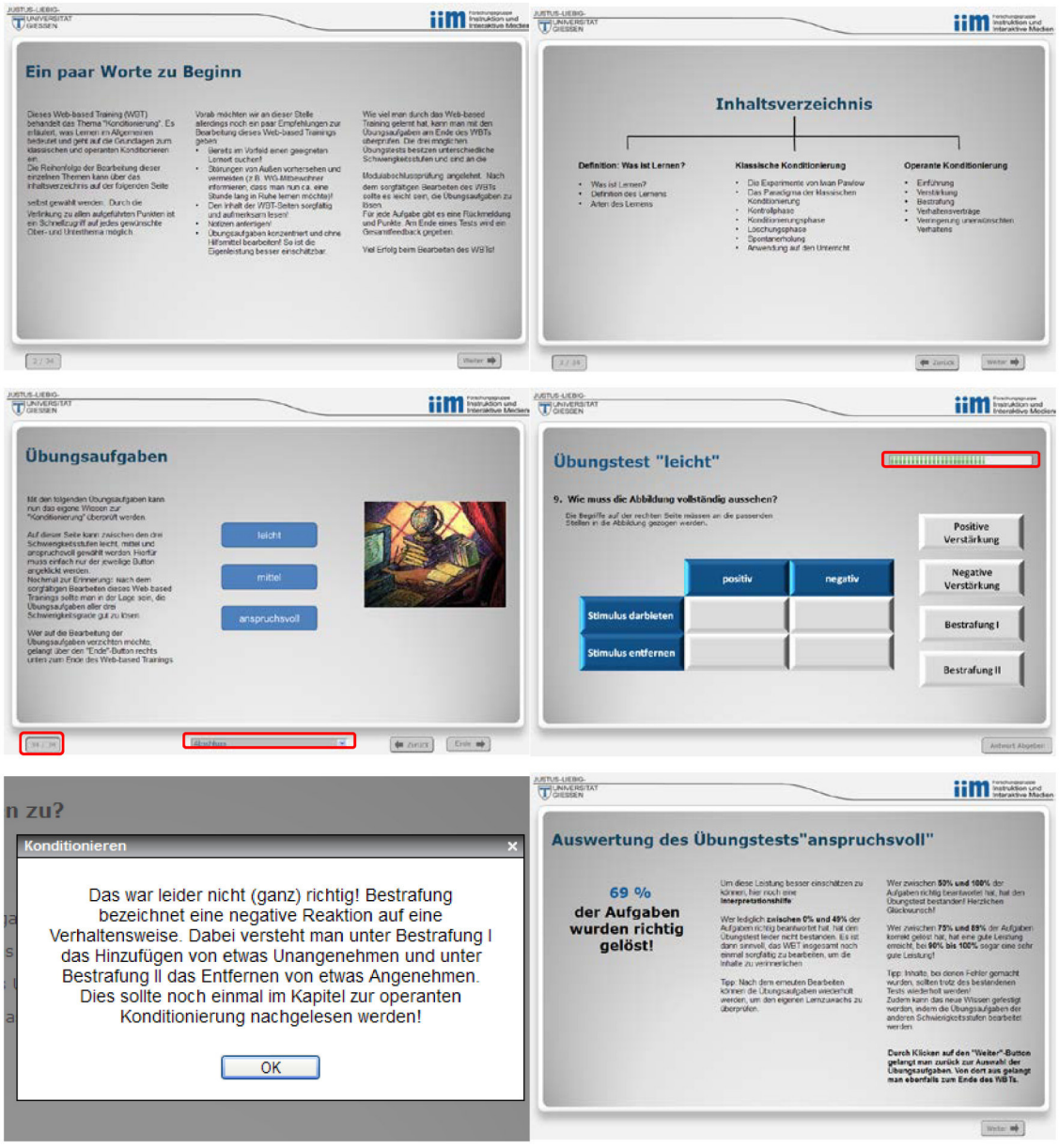

Abbildung 34. Beispielseiten aus dem C-WBT zum Thema "Konditionierung"
Anmerkung. Links oben: Instruktionsseiten zu Beginn des C-WBTs; rechts oben: Interaktives Inhaltsverzeichnis; links Mitte: Einführung in die Übungstests – man sieht hier auch die Seitenzahlen und das Inhaltsverzeichnis (jeweils rot markiert); rechts Mitte: Beispiel für eine Übungsaufgabe mit Fortschrittsbalken (rot markiert); unten links: Aufgabenfeedback und unten rechts: Testfeedback

Um die **Erfolgsmöglichkeiten** der Lernenden zu erhöhen, wird zu Beginn des WBTs knapp erläutert, was mit der Lernanwendung gelernt werden soll[46]. Darüber hinaus wird ein Überblick über die Struktur des WBTs gegeben. Zudem wird eine Instruktionsseite dargeboten, die neben dem Aufbau des WBTs die Navigation erläutert und Empfehlungen gibt, wie man effizient mit einem WBT lernt (Abbildung 34, oben links). Um die Transparenz über das WBT hinweg zu erhalten, werden die Studierenden neben der Seitenzahl auch über die Gesamtseitenzahl informiert (Abbildung 34, Mitte links) und jedes Kapitel beginnt mit einer eigenen Titelseite. Wiederkehrende Überschriften wurden durchnummeriert, begleitend zu den Übungsaufgaben wurde ein Fortschrittsbalken eingeblendet (Abbildung 34, Mitte rechts). All dies sollte den Studierenden Vertrauen in ihren Lernfortschritt und -erfolg vermitteln.

Um den Studierenden darüber hinaus die Möglichkeit zu bieten, ihren Lernprozess selbst zu kontrollieren, wird der Inhalt zu Beginn des WBTs im Rahmen eines interaktiven Organigramms dargeboten (Abbildung 34, oben rechts) – so können die Studierenden selbst wählen, welche Inhalte sie in welcher Reihenfolge bearbeiten möchten. Außerdem existiert auf jeder Seite die Möglichkeit, über ein Inhaltsverzeichnis in ein anderes Kapitel zu springen (Abbildung 34, Mitte links). Dieses Gefühl der **Selbstkontrolle** sollte auch durch die Rückmeldungen in den Übungstests verstärkt werden. Innerhalb der Rückmeldungen wird der Erfolg/Misserfolg hinsichtlich jeder Frage auf die Anstrengung des Lernenden attribuiert (Abbildung 34, unten links). Dies sollte die Lernenden zu weiteren Lernhandlungen anregen.

Zufriedenheits(S)-WBT

Um das „S (Zufriedenheits)"-WBT zu erstellen, wurde das 0-WBT um Komponenten ergänzt, die nach Keller (2010) die Zufriedenheit der Lernenden durch intrinsische und extrinsische Verstärkung sowie die Vermittlung von Gleichheit bzw. Gerechtigkeit erhöhen sollen.

Die **intrinsische Verstärkung** erfolgt innerhalb der S-WBTs unter anderem durch die praktische Anwendung des theoretisch Gelernten in Lernspielen und Anwendungsaufgaben. Die Spielregeln zu jedem Lernspiel werden klar erläutert (siehe Abbildung 35, unten links) und für die Lernenden besteht kein Zwang, die Spiele zu bearbeiten – sie können sie einfach überspringen. Darüber hinaus wird wann immer möglich auf bereits Gelerntes zurückgegriffen bzw. hingewiesen, um den Lernenden das positive Gefühl zu vermitteln, schon etwas gelernt zu haben. Außerdem werden zusätzliche

[46] Es handelt sich hier nicht um eine konkrete Nennung der Lernziele, wie dies im Rahmen der Relevanzförderung der Fall ist, sondern lediglich um eine kurze Beschreibung der Inhalte des WBTs. Beispiel für das WBT zum Thema „Konditionierung": „Dieses Web-based Training (WBT) behandelt das Thema „Konditionierung". Es erläutert, was Lernen im Allgemeinen bedeutet und geht auf die Grundlagen zum klassischen und operanten Konditionieren ein."

Informationen und Internetquellen für Interessierte zur Verfügung gestellt, welche durch einen Klick auf eine kleine Glühbirne angesteuert werden können (siehe Abbildung 35 oben).



Abbildung 35. Beispielseiten aus dem S-WBT zum Thema "Konditionierung"
Anmerkung. Oben: Erläuterung der Glühbirne und Beispiel für eine Glühbirne; links unten: Lernspiel mit Spielregeln; rechts unten: Belohnungsspiel am Ende des WBTs

Extrinsisch wurden die Studierenden vornehmlich dadurch verstärkt, dass sie am Ende des WBTs eine Belohnung erhielten – ein inhaltlich passendes sowie unterhaltendes Video oder ein interaktives Lernspiel (Abbildung 35, unten rechts). Auch wurden die Rückmeldungen innerhalb der Lernspiele sehr enthusiastisch gehalten. Dies sollte die Studierenden verbal für die erbrachten Leistungen loben.
Darüber hinaus werden die Anforderungen innerhalb des WBTs für alle Studierenden gleich gehalten, so dass die Studierenden sich gerecht behandelt fühlen und so ein Gefühl der **Gleichheit** bzw. Gerechtigkeit empfinden sollten.

ARCS-WBT

Ganz zum Schluss wurde das ARCS-WBT erstellt, das alle oben genannten Komponenten der Aufmerksamkeits(A)-, Relevanz(R)-, Erfolgszuversichts(C)- und Zufriedenheits(S)-WBTs in sich vereint. Zur Erstellung dieser WBT-Variante wurden zur Erhaltung der Vergleichbarkeit die Aspekte aus den Einzelkomponenten-WBTs sinnvoll zusammengefügt. Dabei wurden nicht alle Inhalte der vier Einzelkomponenten-WBTs übernommen, da dies

zu einer Überfrachtung geführt hätte und dementsprechend eher lernhinderlich gewesen wäre. Zum Beispiel gibt es innerhalb des ARCS-WBTs nicht die Möglichkeit, zwischen drei Übungstests mit unterschiedlichen Schwierigkeitsgraden auszuwählen. Hier stehen lediglich einmal zehn unterschiedlich anspruchsvolle Übungsfragen zur Verfügung, die den Lerninhalt möglichst repräsentativ abdecken. Zum anderen konnten selbstverständlich nicht alle Lernspiele, Bilder oder Videos integriert werden. Aus diesen Medien wurde eine sinnvolle Auswahl getroffen, um eine möglichst abwechslungsreiche Darbietung des Lernstoffs zu erreichen, ohne die Studierenden durch zu viel Abwechslung zu überfordern und damit vom Lerninhalt abzulenken.

6.3.2 Abhängige Variablen

Zu insgesamt sieben Messzeitpunkten wurden im Laufe der Studie mit Hilfe eines Vortests, dreier kurzfristiger Nachbefragungen, zweier Probeklausuren sowie der Modulabschlussprüfung Daten erhoben. Die hierfür genutzten Fragebögen (siehe Überblick in Abbildung 37 am Ende des Abschnitts) sollen im Folgenden kurz erläutert werden. Die vollständigen Fragebögen befinden sich in Anhang C.
Alle Datenerhebungen fanden – außer der Modulabschlussprüfung – online statt und wurden mit dem Programm *Unipark 7.0*[47] erstellt und durchgeführt. Zu Beginn jeder Befragung erfolgte eine ausführliche Instruktion der Studierenden. Es wurde explizit auf die vertrauliche und anonyme Handhabung der Daten hingewiesen.

Vorbefragung und Vortest

Der Vortest erhob einige demographische Variablen, die Einstellung zum Projekt, die Anzahl der vorhergegangenen Vorlesungsbesuche, die inhaltlichen Vorkenntnisse, das Interesse an der Vorlesung sowie die Computernutzung der Studierenden. Zudem wurde der „*Fragebogen zur computerbezogenen Selbstwirksamkeit*“ in seiner reduzierten Fassung mit zwölf Items (*CUSE-D-r*; Bescherer und Spannagel, 2011) erfragt. Hierbei handelt es sich um die Kurzversion des Fragebogens, die mittels einer sechsstufigen Skala von „trifft überhaupt nicht zu“ bis „trifft völlig zu“ Aussagen wie z.B. die Folgende abfragt: „Ich bin sehr unsicher über meine Fähigkeiten im Umgang mit Computern.“ Der Fragebogen wurde bereits durch Bescherer und Spannagel (2011) überprüft. Für die Stichprobe der vorliegenden Dissertation erreicht er ein Cronbachs $\alpha = .87$.
Im Anschluss sollte zudem das Vorwissen zu den Vorlesungsthemen erfragt werden, damit im Unterschied zu Studie I eine Aussage über den Lernzuwachs möglich ist. Der

[47] Online-Befragungssoftware der Firma *Questback* (siehe Questback, 2016)

Vorwissenstest beinhaltete zwei Fragen zu jedem Thema, das während des Semesters innerhalb der Präsenzvorlesung behandelt wurde (insgesamt 9 Themen) sowie vier Fragen zu den drei Themen, zu denen die Studierenden E-Learning-Anwendungen bearbeiteten. Dementsprechend resultierte ein Test mit insgesamt 30 Fragen. Hierin enthalten waren Aufgaben mit kurzen offenen, Einfach- sowie Mehrfachauswahl-Fragen. Für jede korrekt beantwortete Frage erhielten die Studierenden jeweils einen Punkt, so dass maximal 30 Punkte erreicht werden konnten. Nach jeder Frage erhielten sie zudem eine Rückmeldung, ob ihre Antwort richtig oder falsch war. Die richtige Antwort wurde bei falscher Beantwortung genannt.

Tabelle 56 gibt einen Überblick über die im Vortest erhobenen Aspekte. Zwischen den Fragebogenblöcken wurde den Studierenden jeweils eine kurze Instruktion dargeboten.

Tabelle 56
Überblick über die Vortestitems

Abschnitt	Inhalte
Fragebogen (26 Items)	• Demographische Angaben (Geschlecht, Studiengang, Alter, Fachsemester, Abiturnote und -land) • Einstellung zum Projekt • Früherer Vorlesungsbesuch und Anzahl der Vorlesungsbesuche • Fachliche Vorkenntnisse und Interesse an den Vorlesungsinhalten • Computernutzung (Häufigkeit, genutztes Betriebssystem, Computerkenntnisse) • CUSE-D-r (Kurzform)
Wissenstest (30 Fragen)	• Je 4 Fragen zu den Themen der drei E-Learning-Nutzungen • Je 2 Fragen zu den Themen der neun Vorlesungssitzungen

Nachbefragungen

Nach jeder E-Learning-Nutzung erfolgte eine Nachbefragung (siehe Überblick in Tabelle 57). Grundsätzlich wurde ein Fragebogen zur Erhebung der ARCS-Komponenten erhoben („ARCS-Fragebogen"), desweiteren Validierungsskalen für diesen neu entwickelten Fragebogen, eine Akzeptanzskala sowie ein Wissenstest mit zehn Aufgaben.

Tabelle 57
Grundlegender Aufbau der drei kurzfristigen Nacherhebungen

Abschnitt	Anzahl Items
ARCS-Fragebogen	• 6 Items zur Aufmerksamkeit • 7 Items zur Relevanz • 8 Items zur Erfolgszuversicht • 7 Items zur Zufriedenheit
Validierungsskalen	18 Items aus verschiedenen Fragebögen (s.u.)
Akzeptanz	• Skala „Akzeptanz der Lernumgebung" (9 Items) • Einstufung auf einer Notenskala (1 Item)
Wissenstest	10 Fragen zu den Inhalten der E-Learning-Anwendung

ARCS-Fragebogen

Der ARCS-Fragebogen wurde in Anlehnung an den *IMMS* („*Instructional Material Motivation Survey*") von Keller (2006) entwickelt. Für jede der vier Unterkategorien

(„Aufmerksamkeit“, „Relevanz“, „Erfolgszuversicht“ und „Zufriedenheit“) beinhaltet er eine eigene Skala, die in der neuen Version zunächst jeweils acht Fragen enthielt. Die genutzte fünfstufige Antwortskala besteht dabei aus den Antwortmöglichkeiten „trifft nicht zu“, „trifft eher nicht zu“, „weder/noch“, „trifft eher zu“ sowie „trifft voll zu“.

Huang, Huang, Diefes-Dux und Imbrie (2006) überprüften den *IMMS* hinsichtlich seiner Güte. Ihre Ergebnisse zeigten, dass instruktionale Eigenschaften die Validität des Fragebogens beeinflussen und der *IMMS* dementsprechend auf die jeweilige Lernsituation angepasst werden sollte. Dies stach bereits bei der Vorbereitung des Fragebogens für den Einsatz innerhalb der vorliegenden Studie hervor: Der Fragebogen wurde zunächst ins Deutsche übersetzt. Dabei fiel auf, dass manche Aspekte im *IMMS* für die vorliegende Lernumgebung nicht geeignet waren. Andere Aspekte waren gar nicht enthalten. Weiterhin fragten mehrere Items des Original-Fragebogens gleichzeitig zwei Aspekte ab. Es konnte deshalb nicht bei der reinen Übersetzung des Original-Fragebogens bleiben. Dementsprechend wurden einzelne Items umformuliert oder sogar eliminiert bzw. neue Items hinzugefügt.

Nach der ersten Nachbefragung erfolgte eine Validierung (diese wird im folgenden Abschnitt „Validierungsskalen“ erläutert) sowie eine Faktorenanalyse mit Varimax-Rotation. Aufgrund dieser Analyse wurde der Fragebogen von 32 auf 28 Items reduziert. Tabelle 58 gibt für jede der vier Skalen sowie für den Gesamtfragebogen jeweils ein Beispielitem, das Minimum und Maximum der erreichbaren Punktzahl, die Reliabilität, den Mittelwert sowie die Standardabweichung wieder.

Tabelle 58
Beispielitems und deskriptive Statistiken zu den vier Skalen des ARCS-Fragebogens sowie dem Fragebogen insgesamt

Skala (Anzahl Items)	**Beispielitem**	***Min/Max***	**Reliabilität**	***M***	***SD***
Aufmerksamkeit (6 Items)	Ich konnte mich gut auf die Inhalte der E-Learning-Anwendung konzentrieren.	6/30	.85	23.35	4.42
Relevanz (7 Items)	Die E-Learning-Anwendung vermittelte mir den Eindruck, dass der Inhalt wissenswert ist.	7/35	.80	27.21	4.07
Erfolgszuversicht (8 Items)	Der Inhalt war so gut strukturiert, dass ich sicher war, ihn lernen zu können.	8/40	.77	28.32	4.98
Zufriedenheit (7 Items)	Es hat mir richtig Spaß gemacht, mich mit dem Thema zu befassen.	7/35	.87	22.91	5.53
Insgesamt (28 Items)	-	28/140	.92	100.09	15.08

Anmerkung. Min = Minimale Punktzahl, *Max* = Maximale Punktzahl, *M* = Mittelwert, *SD* = Standardabweichung

Die einzelnen Skalen weisen eine akzeptable bis gute, die Gesamtskala sogar eine exzellente Reliabilität auf.

Im Rahmen der Faktorenanalyse zeigte sich, dass manche Items auf mehr als einem Faktor luden. Dies ist allerdings dem Konzept des ARCS-Modells zuzuschreiben, weshalb hier bei

der Eliminierung der Items nicht allzu streng vorgegangen wurde. Schon das ARCS-Modell selbst widerspricht der Unabhängigkeit der vier Dimensionen, da diese hier zum Teil als voneinander abhängig dargestellt werden (siehe Abbildung 23). Auch Means et al. (1997) stellten die Unabhängigkeit der vier Dimensionen in Frage. Sie belegten dies sogar durch ihre Studie, bei der die Anreicherung einer Lerneinheit mit Relevanzstrategien zwar am meisten zur Erhöhung der durch die Lernenden empfundenen Relevanz führte, daneben aber auch zur Erhöhung der wahrgenommenen Aufmerksamkeitsförderung, Erfolgszuversicht und Zufriedenheit. Die vier Dimensionen scheinen demnach zwar eine gewisse Spezifität aufzuweisen, als völlig unabhängig voneinander können sie allerdings nicht angesehen werden.

Validierungsskalen

Zur Validierung des ARCS-Fragebogens wurden zusätzlich Vergleichsskalen mit insgesamt 18 Items erhoben. Für die Skala „Aufmerksamkeit“ wurden die drei Items der Skala „Interessantheit“ von Rindermann und Amelang (1994) erfasst, die über ein Cronbachs α = .85 verfügt. Die Skala „Relevanz“ wurde durch die beiden Skalen „Thema“ (vier Items, Cronbachs α = .75) und „Auseinandersetzung mit dem Thema – Breite“ (drei Items, Cronbachs α = .84) von Rindermann und Amelang (1994) überprüft. Für die Skala „Zufriedenheit“ wurde ebenfalls aus Rindermann und Amelang (1994) die Skala „Allgemeinbeurteilung“ (drei Items, Cronbachs α = .84) erhoben. Für alle drei Skalen mussten die Studierenden Aussagen auf einer siebenstufigen Skala hinsichtlich ihres Zutreffens beurteilen. Zur Validierung der Skala „Erfolgszuversicht“ wurden die Skalen „Themenspezifisches Selbstkonzept“ (vier Items, Cronbachs α = .85, 6stufige Skala von „stimmt überhaupt nicht“ bis „stimmt genau“) sowie „Themenspezifische Kompetenzeinschätzung“ (kein Cronbachs Alpha, da nur ein Item mit einer Notenskala von 1 bis 6) von Krause (2007) genutzt. Tabelle 59 gibt einen Überblick über die Ergebnisse der Validitätsberechnungen.

Tabelle 59
Korrelationen der ARCS-Skalen mit den Validierungsskalen

ARCS-Skala	Interessantheit	Thema	Thema - Breite	Selbstkonzept	Allgemeinbeurteilung
Aufmerksamkeit	**.74**	.35	.43	.33	.70
Relevanz	.42	**.59**	**.61**	.34	.57
Erfolgszuversicht	.45	.33	.40	**.34**	.55
Zufriedenheit	.32	.21	.26	.11	**.45**

Anmerkung. n = 648; alle Korrelationen sind signifikant mit $p < .01$

Hier zeigt sich, dass jede Skala zumindest mittelstark mit ihrer Validierungsskala korreliert, die Skalen Aufmerksamkeit und Relevanz sogar stark. Allerdings korrelieren die Skalen ebenfalls mittelstark (bis selten auch stark) zu den Validierungsskalen der anderen ARCS-Skalen. Dies ist aber nicht überraschend, da die ARCS-Skalen selbst nicht unabhängig

voneinander sind. Kritisch ist allerdings, dass die Skala „Erfolgszuversicht“ mit den anderen Validierungsskalen zum größten Teil höher korreliert als mit der eigenen. Allerdings zeigte sich auch in der Faktorenanalyse, dass die Items dieser Skala häufig auch auf den anderen Skalen luden. Auch hinsichtlich der Reliabilität ist dies die schwächste Skala.

Akzeptanz

Um auch die Nutzerakzeptanz zu erheben, wurden die neun Items der Skala „Akzeptanz der Lernumgebung“ von Krause (2007) erhoben. Diese Skala weist ein Cronbachs α = .81 auf und musste von den Studierenden auf einer sechsstufigen Skala von „stimmt überhaupt nicht“ bis „stimmt genau“ beantwortet werden. Ein Beispielitem lautet „Ich habe das Gefühl, durch die Arbeit mit der E-Learning-Anwendung viel gelernt zu haben.“

Darüber hinaus sollten die Studierenden die genutzte Anwendung zusammenfassend auf einer Notenskala von 1 bis 6 bewerten.

Wissenstest

Zu guter Letzt erfolgte nach jeder der drei Nachbefragungen die Erhebung der kurzfristigen Lernleistung mittels zehn Aufgaben, die formell denen des Vortests ähnlich waren. In der ersten Nachbefragung stellten drei der Aufgaben dabei Anwendungsaufgaben (Beispiel siehe Abbildung 36) dar, sieben Fragen waren reine Wissensfragen. Nach jeder Frage erhielten die Studierenden eine Rückmeldung über die Korrektheit ihrer Antwort. Am Ende des Tests wurde ihnen zudem ihr Gesamtergebnis mitgeteilt.

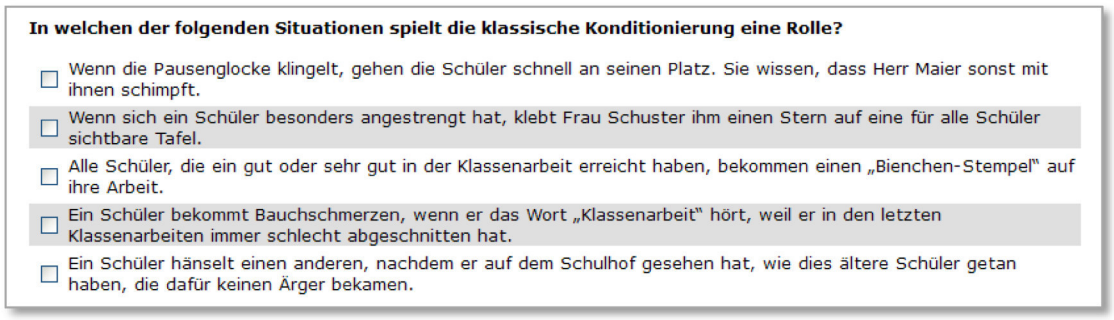

Abbildung 36. Beispielanwendungsaufgabe aus der ersten Nachbefragung („Konditionierung“)

Die Nachbefragungen 2 und 3 unterschieden sich darüber hinaus hinsichtlich verschiedener Aspekte von der Nachbefragung 1. Diese sollen im Folgenden erläutert werden.

Nachbefragung 2

In der Nachbefragung 2 wurde zum einen eine der Validierungsskalen zur Relevanz nicht mehr erhoben, da dies nicht mehr sinnvoll erschien. Zum anderen wurden Persönlichkeitseigenschaften erhoben, die einen möglichen Einfluss auf die beiden ARCS-Komponenten „Relevanz“ und „Zufriedenheit“ haben könnten. Hierbei handelte es sich

zunächst um den Fragebogen zur allgemeinen Lebenszufriedenheit (*„Satisfaction with Life Scale“, SWLS*) von Diener et al. (1985), der aus fünf Items besteht, die auf einer siebenstufigen Skala bewertet werden müssen (stimme überhaupt nicht zu, stimme nicht zu, stimme eher nicht zu, weder/noch, stimme eher zu, stimme zu und stimme völlig zu). Die Skala weist ein Cronbachs α = .88 auf.

Desweiteren wurden die *„Skalen zur Erfassung der Lern- und Leistungsmotivation von Studierenden“ (SELLMO-ST)* von Spinath et al. (2002) erhoben, die über 31 Items die allgemeine Lernmotivation erfassen und auf einer fünfstufigen Skala (stimmt gar nicht, stimmt eher nicht, weder/noch, stimmt eher, stimmt genau) beurteilt werden müssen. Die vier Skalen verfügen über ein Cronbachs Alpha zwischen α = .75 und α = .90.

Der Test zur Lernleistung bestand wie bereits bei der ersten Evaluation aus zehn Fragen – diesmal setzte er sich allerdings aus fünf Anwendungs- und fünf Wissensfragen zusammen, um den Anwendungsfragen einen höheren Stellenwert zu geben.

Nach oberflächlicher Auswertung der Daten zur ersten Nachbefragung fiel auf, dass auch Studierende, die die E-Learning-Anwendung nicht genutzt hatten, die Nachbefragung beantworteten. Grund hierfür war vermutlich der Umstand, dass auch diese Studierenden den Lernleistungstest für sich nutzen wollten. Um diesem Wunsch gerecht zu werden, wurde nun zu Beginn des Fragebogens abgeklärt, ob die E-Learning-Anwendung genutzt worden war. Studierende, die dies verneinten, wurden nicht zum Evaluationsfragebogen weitergeleitet, sondern zu einer Seite, auf der sie angeben sollten, warum sie die Anwendung nicht genutzt hatten und welche Verbesserungsvorschläge sie hätten. Danach hatten auch sie die Möglichkeit, den Wissenstest zu bearbeiten.

Nachbefragung 3

In der dritten Nachbefragung wurden Persönlichkeitseigenschaften erhoben, die ggf. mit den ARCS-Komponenten „Aufmerksamkeit“ und „Erfolgszuversicht“ in Verbindung stehen. Dies war zum einen der *„Fragebogen zur Selbstbeschreibung der Aufmerksamkeit“* von Bühner et al. (2002), der 18 Items enthält, die jeweils auf einer vierstufigen Skala (nie oder selten, manchmal, ziemlich häufig, sehr häufig sowie eine Kategorie betrifft mich nicht) hinsichtlich ihrer Auftretenshäufigkeit beurteilt werden müssen. Dieser Fragebogen weist ein Cronbachs Alpha von .75 für eine studentische Stichprobe auf.

Zum Zweiten wurde die Skala *„Erfolgszuversicht“* aus dem *„Leistungsmotivationsinventar“* (*LMI*; Cronbachs α = .84) von Schuler, Prochaska und Frintrup (2001) erhoben, welche über zehn Items verfügt, die auf einer siebenstufigen Skala hinsichtlich ihres Zutreffens („trifft gar nicht zu“ bis „trifft vollständig zu“) bewertet werden müssen.

Der Wissenstests bestand wiederum aus zehn Aufgaben. Erneut waren dies fünf Anwendungs- und fünf Wissensfragen.

Erhebung der längerfristigen Lernleistung nach dem Wintersemester

Die Probeklausur enthielt 60 Aufgaben zu den zwölf Themen der Veranstaltung. Dabei wurden zu jedem Thema, zu dem eine Vorlesungssitzung stattgefunden hatte, vier Fragen gestellt, zu den drei Themen, zu denen die Studierenden eine E-Learning-Anwendung bearbeitet hatten, sollten sie jeweils acht Fragen beantworten. Um Vergleichbarkeit zu erzielen, waren alle Fragen des Vortests in der Probeklausur enthalten. Durch die hierdurch möglichen Erinnerungsleistungen werden die Lernleistungen der Studierenden sicherlich etwas überschätzt – dies sollte sich aber über alle Gruppen in etwa gleich auswirken, da der Vortest für alle Studierenden zur gleichen Zeit stattfand. Die Auswirkungen sollten zudem eher gering sein, da zwischen Vortest und Probeklausur fast vier Monate lagen. Die ursprünglich 30 Vortestfragen wurden folglich durch weitere 30 neue Fragen ergänzt. Auch hier erhielten die Studierenden nach Beantwortung jeder Frage eine Rückmeldung zur Korrektheit ihrer Antwort sowie ein abschließendes Feedback zu ihrem Abschneiden in der Probeklausur.

Die Probeklausur schloss mit zwei Evaluationsfragen: Zum einen sollten die Studierenden angeben, wie sie es fanden, an dem Projekt teilzunehmen (mittels siebenstufiger Kounin-Skala[48]), zum anderen hatten sie hier Gelegenheit, innerhalb einer offenen Frage abschließend Lob oder Kritik zu äußern.

Probeklausur nach dem Sommersemester/vor der MAP

Bevor erneut die Probeklausur bearbeitet werden konnte, wurde eine kurze Abschlussevaluation vorgeschaltet. Damit die Studierenden die Möglichkeit hatten, die Probeklausur mehrfach zu beantworten, ohne jedes Mal auch die Evaluationsfragen beantworten zu müssen, wurde zunächst gefragt, ob sie diese schon beantwortet hatten. Bei einem Klick auf „Ja“ wurden sie dann direkt zur Probeklausur weitergeleitet. Sicher führte dies dazu, dass einige Studierende die Evaluation gar nicht bearbeiteten und lediglich die Probeklausur zur Klausurvorbereitung nutzten. Da den Studierenden im Laufe des Projektes allerdings schon Einiges abverlangt worden war, wurde dieses Risiko in Kauf genommen.

Die Studierenden, die den Fragebogen bearbeiteten, wurden zunächst nach den Materialien gefragt, die sie hauptsächlich zum Lernen für die MAP genutzt hatten. Je nachdem, was sie hier angaben, wurden sie dann zu einer Frage weitergeleitet, warum sie die anderen Materialien nicht oder nicht häufiger genutzt hatten. Danach wurden die Studierenden, die E-Learning-Anwendungen für ihre MAP-Vorbereitung genutzt hatten, zu ihrer Einschätzung der Wichtigkeit von einzelnen Aspekten von Kellers ARCS-Modell (z.B. Einsatz

[48] Sieben Gesichtsausdrücke von traurig bis fröhlich, von denen die Studierenden den zutreffenden Gesichtsausdruck ankreuzen sollten.

eines Lernbegleiters, Vorgabe von Lernzielen, Möglichkeit für einen flexiblen Lernweg etc.) befragt. Alle Studierenden konnten zu guter Letzt noch einmal äußern, wie sie die Teilnahme an dem Projekt fanden (Kounin-Skala). Zudem konnten sie in einem offenen Textfeld Lob oder Kritik äußern. Die abschließende Probeklausur war vergleichbar zur Studie I identisch zu der, die bereits nach dem Wintersemester zur Verfügung gestellt worden war. Da fast ein halbes Jahr zwischen den beiden Erhebungen lag, wurde erneut davon ausgegangen, dass die wiederholte Nutzung der Probeklausur sich nur in sehr geringem Maße auf die Lernleistungen der Studierenden auswirken sollte bzw. die Auswirkungen für alle Evaluationsgruppen in etwa gleich ausfallen sollten.

MAP

Die Modulabschlussprüfung wurde aufgrund der Menge an Studierenden in mehreren Gruppen geschrieben. Deshalb kamen erneut verschiedene Versionen der Klausur zum Einsatz. Diese sind vergleichbar zu den Klausurversionen, die ausführlich im Rahmen von Studie I beschrieben und ausgewertet wurden (siehe Abschnitt „2.3.2 Erhobene Variablen“). Aus diesem Grunde wird an dieser Stelle auf eine erneute ausführliche Darstellung der Modulabschlussprüfung verzichtet.

In die Auswertung der vorliegenden Arbeit geht im Vordergrund der Teil der Modulabschlussprüfung ein, der die Inhalte der Vorlesung „Psychologie für Lehramtsstudierende A“ abprüft. Im Gegensatz zur ersten Studie werden auch die Ergebnisse von Teil B der Klausur zu Rate gezogen, um eine Aussage über die Hintergründe von etwaigen Lernleistungsunterschieden in Teil A machen zu können. Dies wurde bereits in Abschnitt „5.4.4 Freiwillige Nutzung im Sommersemester“ erläutert.

User-Tracking

Erneut wurde im Lehr-Management-System *Moodle* direkt erfasst, welcher Studierende wann und wie lange mit der E-Lecture oder einer WBT-Version gearbeitet hat. Diese Daten gehen in die Berechnung der Lerneffizienz der Lehrmethoden ein. Im Nachgang wurden – wie oben bereits erläutert und im Rahmen von Studie I begründet – für die kurzfristigen Lernzeiten diejenigen Nutzungen im Wintersemester eliminiert, die kürzer als fünf Minuten oder länger als fünf Stunden andauerten.

Für die Auswertungen im Sommersemester wurden diesmal keine Nutzungen ausgeschlossen, da die Gruppierung der Studierenden nicht aufgrund der Nutzungszeiten erfolgte, sondern aufgrund der subjektiven Angabe der Studierenden, mit welchem Lernangebot sie am **häufigsten** gelernt hatten. Die Nutzungszeiten dienten hier lediglich als Kontrollmöglichkeit zu den subjektiven Angaben der Studierenden (siehe Abschnitt „7.4 Freiwillige Nutzung im Sommersemester“). Da aufgrund der online erhobenen

Nutzungsdaten nicht beurteilt werden kann, wie viel Zeit die Studierenden tatsächlich (online *sowie* offline) in die Klausurvorbereitung investierten, wurde auf die Berechnung der Lerneffizienz für die Nutzung im Sommersemester erneut vollständig verzichtet und auch die Ergebnisse zur Lerneffektivität werden zu diskutieren sein.

Abbildung 37 gibt abschließend einen Überblick über die Evaluationsfragebögen der zweiten Studie.

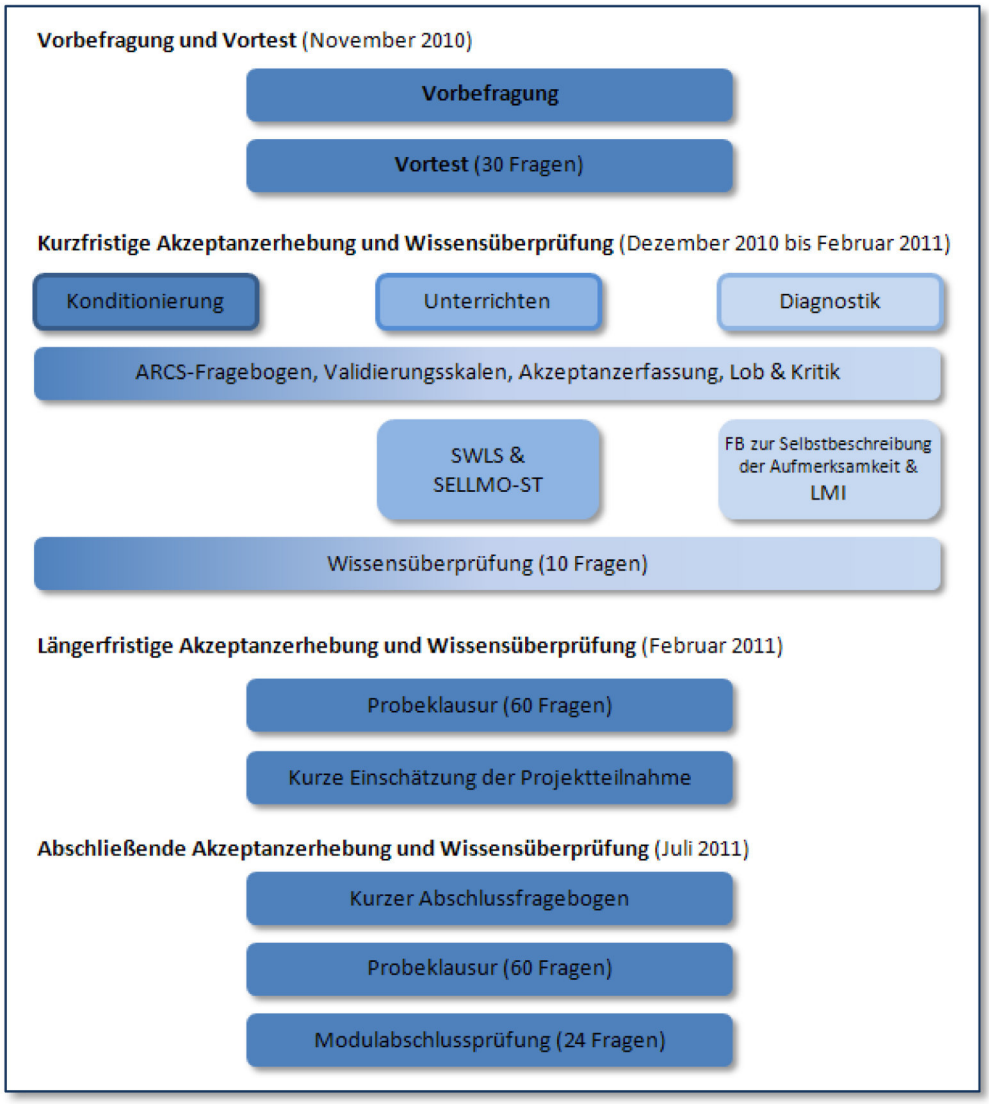

Abbildung 37. Überblick über die genutzten Fragebögen in Studie II

7 Studie II: Ergebnisse und Diskussion

Die Überprüfung der Hypothesen erfolgte mit dem Statistikprogramm *SPSS* in den Versionen 21 und 22. Für die Prüfung der Hypothesen anhand von Mittelwertsvergleichen wurde erneut das von Bortz und Schuster (2010) empfohlene Signifikanzniveau von $\alpha = .05$ zugrundegelegt.

Zur Hypothesenprüfung wurden hauptsächlich einfaktorielle Varianzanalysen (ANOVAs) berechnet. Die einfaktorielle Varianzanalyse setzt voraus, dass die zugrundeliegenden Daten normalverteilt und varianzhomogen sind. Die Überprüfung der Normalverteilung erfolgte deshalb mittels Kolmogorov-Smirnov-Test, die Überprüfung der Varianzhomogenität mittels Levene-Test. Ist Varianzhomogenität gegeben, Normalverteilung allerdings nicht, werden dennoch der *F*-Wert und die Ergebnisse der Post-Hoc-Tests (paarweise *t*-Tests mit Bonferroni-Adjustierung) angegeben. Nach Bortz und Schuster (2010) verlieren die Annahmen der Varianzanalyse mit steigender Stichprobengröße an Bedeutung. Zudem ist die ANOVA bei vergleichbar großen Stichproben relativ robust gegenüber Verletzungen ihrer Annahmen. Da in der vorliegenden Studie die geringste Stichprobengröße $n = 27$ beträgt[49] und die Stichproben vergleichbar groß sind (oder dies durch Ziehung von Zufallsstichproben aus sehr großen Gruppen[50] erzeugt wurde), kann die ANOVA hier berechnet werden. In Fällen, in denen jedoch auch die Annahme der Varianzhomogenität verletzt ist, werden Ergebnisse des Brown-Forsythe-Tests und Post-Hoc-Ergebnisse aus paarweisen Games-Howell-Tests berichtet.

In manchen Fällen werden nur zwei Gruppen miteinander verglichen. Hierfür werden *t*-Tests durchgeführt. Bei nicht gegebener Varianzhomogenität werden die Ergebnisse des Welch-Tests ausgegeben. Die zugehörige Effektstärke wurde mittels des Programms *GPower 3.1* berechnet.

Anhand der angegebenen Freiheitsgrade ist im konkreten Fall zu erkennen, welcher Test genutzt wurde.

Für die einfaktoriellen Varianzanalysen mit Messwiederholung besteht die Annahme der Sphärizität. Diese wurde mittels Mauchly-Test überprüft. Es wird allerdings generell das

[49] Zumeist sind die Stichprobenumfänge deutlich höher.

[50] Im Wintersemester sind aufgrund des Versuchsplans die E-Lecture-Nutzergruppen etwa dreimal so groß wie die WBT-Nutzergruppen. Im Sommersemester sind aufgrund der freiwilligen Wahlmöglichkeit der Studierenden hingegen die Gruppen größer, die sich hauptsächlich mit dem Vorlesungsskript oder mit den elaborierten WBTs auf die MAP vorbereiteten. Deshalb wurden mehrere Zufallsstichproben gezogen und mit diesen einfaktorielle Varianzanalysen durchgeführt. Die Ergebnisse werden nur dann als signifikant bewertet, wenn die Mehrzahl dieser Auswertungen signifikante Unterschiede ergab.

Verfahren von Box (Greenhouse-Geisser-Korrektur) angewandt, da die Messwiederholungs-Varianzanalyse anfällig für große Strichprobenumfänge und für die Verletzung der Normalverteilungsannahme ist. Zu diesem Vorgehen raten Rasch et al. (2008).

Da insbesondere für große Stichproben schon sehr kleine Effekte signifikant werden können, werden neben der Signifikanz stets auch die Effektstärken angegeben. Dabei wird vergleichbar zu Studie I nach Rasch, Friese und Hofmann (2009) ein Eta^2 von .01 als schwacher, ein Eta^2 von .06 als mittlerer und ein Eta^2 von .14 als starker Effekt interpretiert. Wenn auf die Effektstärke *d* zurückgegriffen wird, wird nach Cohen (1988) $d = .2$ als kleiner, $d = .5$ als mittlerer und $d = .8$ als starker Effekt interpretiert.

In der Folge werden zunächst die Ergebnisse zu Nutzungszeit, Lerneffektivität und Lerneffizienz berichtet. Es folgt ein Blick auf die Ergebnisse zur ARCS-Skala sowie zur Akzeptanz der E-Learning-Varianten. Im Anschluss werden gesondert hiervon die Ergebnisse dargestellt, die sich im Rahmen der freiwilligen Nutzung im Sommersemester ergaben. Dabei werden für jeden Abschnitt zunächst die Ergebnisse präsentiert. Diese werden jeweils in einem abschließenden Unterkapitel zusammengefasst und diskutiert.

7.1 Nutzungszeit, Lerneffektivität und -effizienz

Zunächst wird die Lernleistung der Studierenden sowie ihre Nutzungszeit/-häufigkeit und die daraus resultierende Lerneffizienz der sieben verschiedenen E-Learning-Varianten betrachtet. Dabei werden in jedem Unterkapitel erst die kurzfristigen Ergebnisse betrachtet. Danach folgen die Ergebnisse zur längerfristigen Lerneffektivität bzw. -effizienz. Die beschriebenen Nutzungszeiten und -häufigkeiten liegen naturgemäß sowohl der kurz- als auch der längerfristigen Lerneffizienzberechnung zugrunde.

7.1.1 Kurz- und längerfristige Lerneffektivität

7.1.1.1 Kurzfristige Lerneffektivität

Hypothese 1a: Die kurzfristige Lernleistung (Differenz zwischen Vortest und kurzfristigem Nachtest) ist nach Bearbeitung des ARCS-WBTs signifikant höher als nach Bearbeitung aller anderen E-Learning-Variationen.
➔ **ARCS > A, R, C, S, 0, EL**

Hypothese 1b: Die kurzfristige Lernleistung (Differenz zwischen Vortest und kurzfristigem Nachtest) ist nach Bearbeitung der Einzelkomponenten-WBTs sowie der E-Lecture signifikant höher als nach Bearbeitung des 0-WBTs.
➔ **A, R, C, S, EL > 0**

Die kurzfristige Lerneffektivität wurde durch Nachtests kurz nach der Nutzung der E-Learning-Anwendungen erhoben. Da in diesem Rahmen die Vortestergebnisse ebenfalls zu Rate gezogen werden müssen, um etwaige Gruppenunterschiede zu kontrollieren, werden die Vor- und Nachtestergebnisse jeweils in Prozent wiedergegeben, da im Vortest lediglich vier Punkte erreicht werden konnten, im Nachtest maximal zehn Punkte.

Zur Anschaulichkeit werden im Folgenden zunächst die Vor- und Nachtestergebnisse sowie der Lernzuwachs getrennt für die drei Themenbereiche sowie insgesamt über die Gruppen verglichen. Danach erfolgt eine Kovarianzanalyse zum Vergleich der Nachtestergebnisse mit der Kovariate „Vortestergebnis".

Tabelle 60 zeigt, dass die Vortestleistungen sich deutlich zwischen den Themenbereichen unterscheiden: Während in den Bereichen „Konditionierung" und „Diagnostik" von den Studierenden bereits etwa 40% der Vortestfragen korrekt beantwortet werden konnten, sind es für das Thema „Unterrichten" lediglich rund 27%. Über die drei Themen hinweg können die Studierenden etwa 35% der Fragen im Vortest korrekt beantworten.

Tabelle 60

Deskriptive Statistik zu den Vor-, Nachtestergebnissen und dem Lernzuwachs für die E-Learning-Gruppen getrennt für die drei Themenbereiche sowie insgesamt

Vortest	**Konditionierung**			**Unterrichten**			**Diagnostik**			**Gesamt**		
	n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
0-WBT	63	41.67	25.43	74	29.56	22.81	51	**45.59**	21.33	188	**37.97**	24.24
A-WBT	63	42.22	23.24	79	24.84	20.07	77	35.88	22.15	219	33.72	22.80
R-WBT	68	**45.29**	25.30	69	**30.25**	21.90	59	35.59	19.26	196	37.08	23.20
C-WBT	56	42.95	24.32	65	20.58	20.06	54	40.74	22.67	175	33.96	24.47
S-WBT	80	37.69	22.30	60	24.79	18.77	54	37.04	25.23	194	33.52	22.81
ARCS-WBT	56	37.95	18.26	68	23.71	21.18	71	39.79	24.49	195	33.65	22.79
E-Lecture	197	41.73	22.98	242	27.74	22.10	202	39.85	21.65	641	35.85	23.10
Gesamt	**583**	**41.39**	**23.20**	**657**	**26.46**	**21.44**	**568**	**39.19**	**22.34**	**1808**	**35.28**	**23.28**
Nachtest	**Konditionierung**			**Unterrichten**			**Diagnostik**			**Gesamt**		
	n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
0-WBT	64	58.75	16.52	74	**59.66**	13.72	52	58.84	22.55	190	59.13	17.35
A-WBT	64	58.75	16.60	82	59.50	15.64	78	56.02	21.34	224	58.07	18.06
R-WBT	68	61.40	16.23	69	59.30	15.92	59	**61.42**	22.03	196	**60.67**	18.01
C-WBT	56	62.77	17.84	66	56.67	16.69	54	56.50	21.93	176	58.56	18.90
S-WBT	80	62.44	19.11	61	56.38	18.10	54	51.73	21.93	195	57.58	20.03
ARCS-WBT	57	**64.21**	20.91	69	57.91	13.73	73	55.16	23.83	199	58.71	20.21
E-Lecture	200	61.80	15.55	244	56.33	17.39	204	57.23	21.13	648	58.30	18.27
Gesamt	**589**	**61.50**	**17.14**	**665**	**57.60**	**16.29**	**574**	**56.79**	**21.90**	**1828**	**58.61**	**18.59**
Lernzuwachs	**Konditionierung**			**Unterrichten**			**Diagnostik**			**Gesamt**		
	n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
0-WBT	63	16.98	29.93	74	30.09	25.24	51	13.23	31.60	188	21.13	29.45
A-WBT	63	16.67	30.09	79	35.14	24.72	77	20.65	28.21	219	24.73	28.59
R-WBT	68	16.10	27.82	69	29.04	23.11	59	**25.83**	26.95	196	23.59	26.45
C-WBT	56	19.82	26.29	65	**36.04**	21.56	54	15.76	31.52	175	24.59	27.79
S-WBT	80	24.75	26.93	60	31.84	20.67	54	14.69	34.76	194	24.14	28.33
ARCS-WBT	56	**26.96**	27.41	68	34.01	24.03	71	15.32	33.85	195	**25.18**	29.83
E-Lecture	197	20.51	25.41	242	28.62	26.64	202	17.46	30.00	641	22.61	27.77
Gesamt	**583**	**20.33**	**27.29**	**657**	**31.20**	**24.71**	**568**	**17.69**	**30.78**	**1808**	**23.45**	**28.18**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung; Bei den Werten handelt es sich jeweils um die prozentuale Angabe der richtigen Antworten

Im Vergleich dazu ist die Variation zwischen den Nachtestergebnissen deutlich geringer. Hier können die Studierenden über alle drei Themenbereiche zwischen 57% und 62% der Punkte erreichen – insgesamt resultiert ein durchschnittliches Ergebnis von M = 58.6%.
In Tabelle 60 sowie Abbildung 38 wird zudem der durchschnittliche Lernzuwachs in Prozent dargestellt. Hier zeigt sich, dass der Lernzuwachs über alle Gruppen zum Themenbereich „Unterrichten“ mit M = 31.2% (vs. M = 20.3% und M = 17.7% für die anderen Themen) am deutlichsten ausfällt. Über alle Themen hinweg erreichen die Studierenden einen Lernzuwachs von M = 23.5%. Es fällt auf, dass sich augenscheinlich keine E-Learning-Anwendung ausfindig machen lässt, deren Bearbeitung hinsichtlich aller drei Themen zu einem besonders hohen Lernzuwachs geführt hat. Hier ergibt sich ein sehr gemischtes Bild. Auf den ersten Blick scheinen das 0-WBT (M = 21.1%) und die E-Lecture (M = 22.6%) zu eher niedrigen Lernzuwachsraten geführt zu haben. Das ARCS-WBT erreicht mit M = 25.2% über alle drei Themen hinweg den durchschnittlich höchsten Lernzuwachs.

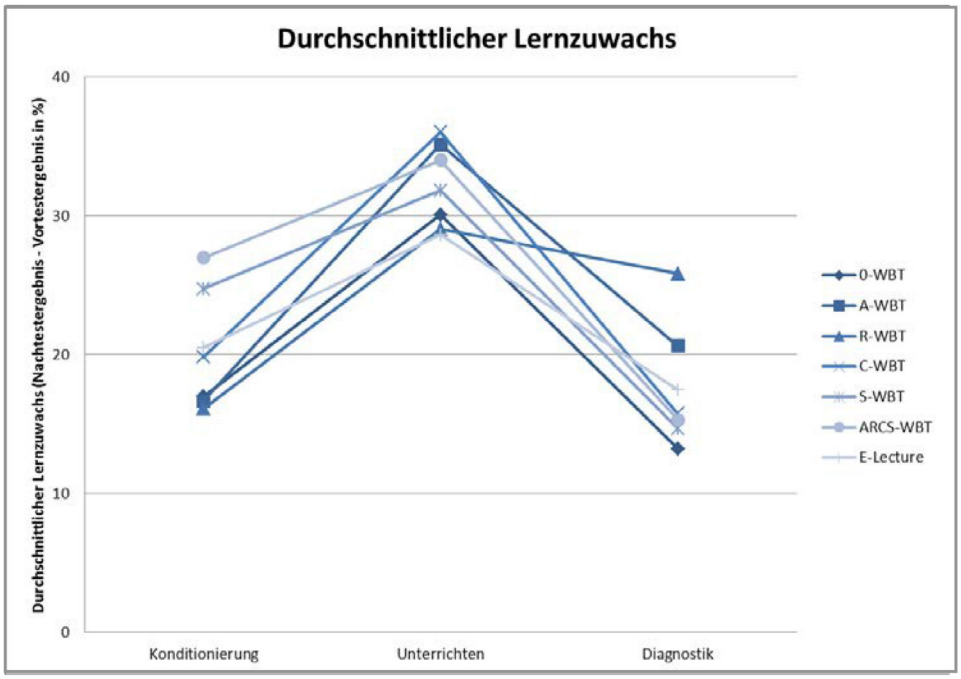

Abbildung 38. Durchschnittlicher Lernzuwachs durch die Bearbeitung der verschiedenen E-Learning-Varianten zu allen drei Themenbereichen

Diese Beobachtungen gilt es nun mittels einfaktoriellen Varianzanalysen zu überprüfen. Die Ergebnisse dieser Analyse zeigt Tabelle 61. Es ergeben sich weder zwischen den Vor- noch zwischen den Nachtestergebnissen der sieben Gruppen signifikante Unterschiede. Dies bestätigt sich ebenfalls in Varianzanalysen zum Lernzuwachs in % (siehe Anhang D, Tabelle D1).

Tabelle 61
Ergebnisse der Mittelwertsanalysen zum Vergleich der Vor- und Nachtestergebnisse der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Vortest	**ANOVA/Brown-Forsythe-Test**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	0.97	442,6	.45	-	
Zufallsauswahl 2	0.99	416.54,6	.43	-	-
Zufallsauswahl 3	0.90	448,6	.49	-	
Unterrichten					
Zufallsauswahl 1	1.86	486,6	.09	-	
Zufallsauswahl 2	1.76	488,6	.11	-	-
Zufallsauswahl 3	2.33	492,6	.03*	.03	
Diagnostik					
Zufallsauswahl 1	1.38	426,6	.22	-	
Zufallsauswahl 2	1.66	427,6	.13	-	-
Zufallsauswahl 3	1.45	426,6	.20	-	
Gesamt					
Zufallsauswahl 1	1.26	1368,6	.27	-	
Zufallsauswahl 2	1.17	1373,6	.32	-	-
Zufallsauswahl 3	1.24	1380,6	.28	-	
Nachtest	**ANOVA**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	0.87	446,6	.51	-	
Zufallsauswahl 2	0.95	448,6	.46	-	-
Zufallsauswahl 3	0.95	452,6	.46	-	
Unterrichten					
Zufallsauswahl 1	0.68	492,6	.67	-	
Zufallsauswahl 2	0.57	496,6	.75	-	-
Zufallsauswahl 3	0.75	498,6	.61	-	
Diagnostik					
Zufallsauswahl 1	1.04	431,6	.40	-	
Zufallsauswahl 2	1.29	432,6	.26	-	-
Zufallsauswahl 3	1.07	430,6	.38	-	
Gesamt					
Zufallsauswahl 1	0.58	1368,6	.75	-	
Zufallsauswahl 2	0.54	1390,6	.78	-	-
Zufallsauswahl 3	0.56	1394,6	.76	-	

Anmerkung. F = Teststatistik der Varianzanalyse bzw. des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Da die Vortestergebnisse signifikant mit den Nachtestergebnissen korrelieren, ist es methodisch sinnvoll, zusätzlich Kovarianzanalysen zwischen den sieben Gruppen hinsichtlich der Nachtestleistungen zu berechnen, bei denen das jeweilige Vortestergebnis auspartialisiert wird (Ergebnisse siehe Tabelle 62).

Hier zeigt sich, dass die Kovariate „Vortestergebnis“ für die Mehrzahl der Analysen zu den Themenbereichen „Konditionierung“ und „Unterrichten“ sowie insgesamt einen Einfluss auf das Nachtestergebnis hat. Dies gilt nicht für den Themenbereich „Diagnostik“ (siehe Tabelle 62, jeweils Zeile „Vortest“). Die Effekte der gefundenen Unterschiede sind als klein bis maximal mittelstark zu beurteilen.

Im Gegensatz dazu spielt es in allen drei Erhebungen für die Höhe des Nachtestergebnisses keine Rolle, welche E-Learning-Anwendung die Studierenden vorher bearbeitet hatten

(siehe Tabelle 62, jeweils Zeile „Gruppe“). Die Hypothesen 1a und 1b sind demnach eindeutig zu verwerfen.

Tabelle 62
Ergebnisse der Kovarianzanalysen zum Vergleich der Nachtestergebnisse der E-Learning-Gruppen unter Einbezug der Kovariate „Vortestergebnis“ getrennt für die drei Themen sowie insgesamt

Nachtest, Kovariate Vortest Konditionierung		**Kovarianzanalyse** *F*	*df*	*p*	*Eta²*
Zufallsauswahl 1	Vortest	5.04	1	.03*	.01
	Gruppe	1.09	6	.37	-
Zufallsauswahl 2	Vortest	4.51	1	.03*	.01
	Gruppe	1.13	6	.35	-
Zufallsauswahl 3	Vortest	3.13	1	.08	-
	Gruppe	1.19	6	.31	-
Unterrichten					
Zufallsauswahl 1	Vortest	15.50	1	<.01*	.03
	Gruppe	0.63	6	.71	-
Zufallsauswahl 2	Vortest	14.95	1	<.01*	.03
	Gruppe	0.43	6	.86	-
Zufallsauswahl 3	Vortest	23.09	1	<.01*	.05
	Gruppe	0.86	6	.52	-
Diagnostik					
Zufallsauswahl 1	Vortest	0.68	1	.41	-
	Gruppe	1.04	6	.40	-
Zufallsauswahl 2	Vortest	0.07	1	.80	-
	Gruppe	1.31	6	.25	-
Zufallsauswahl 3	Vortest	0.84	1	.36	-
	Gruppe	1.06	6	.38	-
Gesamt					
Zufallsauswahl 1	Vortest	15.24	1	<.01*	.01
	Gruppe	0.44	6	.85	-
Zufallsauswahl 2	Vortest	12.06	1	<.01*	.01
	Gruppe	0.38	6	.89	-
Zufallsauswahl 3	Vortest	17.44	1	<.01*	.01
	Gruppe	0.40	6	.88	-

Anmerkung. F = Teststatistik der Kovarianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

7.1.1.2 Längerfristige Lernwirksamkeit

Hypothese 2a: Die längerfristige Lernleistung (Differenz zwischen Vortest und Probeklausur) ist nach Bearbeitung des ARCS-WBTs signifikant höher als nach Bearbeitung aller anderen E-Learning-Variationen.
➔ **ARCS > A, R, C, S, 0, EL**

Hypothese 2b: Die längerfristige Lernleistung (Differenz zwischen Vortest und Probeklausur) ist nach Bearbeitung der Einzelkomponenten-WBTs sowie der E-Lecture signifikant höher als nach Bearbeitung des 0-WBTs.
➔ **A, R, C, S, EL > 0**

Die längerfristige Lernwirksamkeit wurde durch die Probeklausur am Ende des Wintersemesters erhoben. Da hier erneut mehr Punkte erreicht werden konnten als in der Vortesterhebung, werden in Tabelle 63 die Ergebnisse (Nachtestergebnisse und

Lernzuwachsraten) der sieben Gruppen prozentual wiedergegeben. Die Vortestergebnisse wurden oben bereits dargestellt.

Tabelle 63
Deskriptive Statistik zur längerfristigen Lernwirksamkeit der E-Learning-Gruppen für die drei Themen sowie insgesamt

Probeklausur	**Konditionierung**			**Unterrichten**			**Diagnostik**			**Gesamt**		
	n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
0-WBT	46	67.45	15.61	51	59.80	21.33	41	65.32	22.62	138	63.99	20.16
A-WBT	47	71.70	23.01	52	59.01	22.22	57	61.13	21.51	156	63.61	22.72
R-WBT	47	65.77	22.42	52	**59.98**	23.06	46	67.53	21.35	145	**64.25**	22.41
C-WBT	41	72.44	22.26	44	59.09	24.65	40	60.31	26.05	125	63.86	24.91
S-WBT	54	**72.66**	22.19	46	56.93	29.17	30	56.67	25.96	130	63.40	26.69
ARCS-WBT	38	62.86	26.00	60	56.46	22.99	53	**72.41**	20.34	151	63.67	23.79
E-Lecture	131	67.81	21.82	173	54.44	25.03	159	59.41	24.18	463	59.93	24.42
Gesamt	**404**	**68.64**	**22.00**	**478**	**57.03**	**24.27**	**426**	**62.60**	**23.58**	**1308**	**62.43**	**23.82**
Lernzuwachs	**Konditionierung**			**Unterrichten**			**Diagnostik**			**Gesamt**		
	n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
0-WBT	45	26.44	29.84	51	28.19	25.29	40	18.05	28.37	136	24.63	27.90
A-WBT	47	29.68	30.93	50	32.50	27.02	56	27.01	30.85	153	29.62	29.57
R-WBT	47	19.60	27.37	52	26.56	33.71	46	30.30	24.19	145	25.49	29.05
C-WBT	41	28.05	32.28	44	**35.51**	29.43	40	19.07	24.65	125	27.80	29.55
S-WBT	54	**34.98**	28.31	45	30.97	33.07	30	17.92	37.86	129	29.61	32.82
ARCS-WBT	38	26.15	26.93	60	32.71	29.77	52	**31.01**	26.20	150	**30.46**	27.80
E-Lecture	130	24.92	30.83	172	26.42	28.59	157	19.13	31.50	459	23.50	30.35
Gesamt	**402**	**26.81**	**29.89**	**474**	**29.34**	**29.37**	**421**	**22.67**	**29.94**	**1297**	**26.39**	**29.83**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung; Bei den Werten handelt es sich jeweils um die prozentuale Angabe der richtigen Antworten.

Die Probeklausurleistungen der Gesamtgruppe bewegen sich zwischen 57% und 69% und sind damit weniger homogen als die Ergebnisse der kurzfristigen Nachtests (siehe auch Abbildung 39).

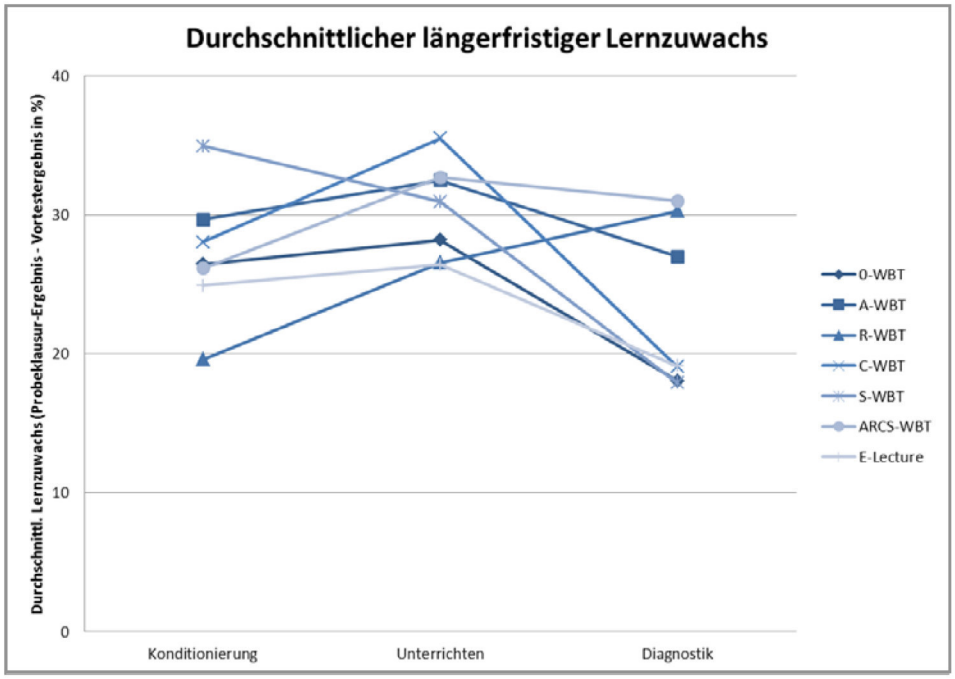

Abbildung 39. Längerfristiger Lernzuwachs der E-Learning-Versionen für alle drei Themen

Über alle Themen hinweg können die Studierenden *M* = 62.4% der Punkte in der Probeklausur erreichen. Innerhalb der Themen variieren die Probeklausurergebnisse ebenfalls mit maximal 12% zwischen den Gruppen. Der mittlere Lernzuwachs beträgt zwischen 23% und 29%. Über alle Themen hinweg resultiert damit ein Lernzuwachs von *M* = 26.4%.

Dabei ist erneut keine Anwendung auszumachen, die über alle Themen hinweg zu einem besonders hohen oder niedrigen Lernzuwachs geführt hat. Es ergibt sich – wie bereits bei den kurzfristigen Leistungsmessungen – ein sehr gemischtes Bild über die Gruppen. Über die drei Themen hinweg erreicht die E-Lecture mit *M* = 23.5% den niedrigsten, das ARCS-WBT mit *M* = 30.5% den höchsten längerfristigen Lernzuwachs.

Tabelle 64
Ergebnisse der Mittelwertsanalysen zum Vergleich der längerfristigen Lernwirksamkeit der E-Learning-Versionen getrennt für die drei Themen sowie insgesamt

Probeklausur	**ANOVA/Brown-Forsythe-Test**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	1.60	310,6	.15	-	
Zufallsauswahl 2	1.46	309,6	.19	-	-
Zufallsauswahl 3	1.56	310,6	.16	-	
Unterrichten					
Zufallsauswahl 1	0.48	354,6	.83	-	
Zufallsauswahl 2	0.63	358,6	.71	-	-
Zufallsauswahl 3	0.21	355,6	.97	-	
Diagnostik					
Zufallsauswahl 1	2.81	314,6	.01*	.05	ARCS > EL
Zufallsauswahl 2	2.38	316,6	.03*	.04	-
Zufallsauswahl 3	2.56	309,6	.02*	.05	
Gesamt					
Zufallsauswahl 1	0.32	946.74,6	.93	-	
Zufallsauswahl 2	0.95	951.39,6	.46	-	-
Zufallsauswahl 3	0.61	938.21,6	.73	-	

Anmerkung. F = Teststatistik der Varianzanalyse bzw. des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Die einfaktoriellen Varianzanalysen zum Vergleich der E-Learning-Gruppen hinsichtlich des Probeklausurergebnisses ergeben allerdings in der Mehrzahl keine signifikanten Unterschiede (siehe Tabelle 64). Dies wird auch durch die Auswertungen der Lernzuwachsraten bestätigt (siehe Anhang D, Tabelle D2)

Wie bereits im Rahmen der kurzfristigen Lernwirksamkeit sollen auch hier Kovarianzanalysen zu den Unterschieden zwischen den Probeklausurergebnissen berechnet werden. Als Kovariate dient erneut das themengleiche Vortestergebnis. Die Ergebnisse dieser Analysen sind in Tabelle 65 abgetragen.

Tabelle 65
Ergebnisse der Kovarianzanalysen zum Vergleich der Probeklausurergebnisse der E-Learning-Gruppen unter Einbezug der Kovariate „Vortestergebnis" getrennt für die drei Themen sowie insgesamt

Probeklausur, Kovariate Vortest Konditionierung	**Kovarianzanalyse**				***Eta²***
		F	***df***	***p***	
Zufallsauswahl 1	Vortest	9.27	1	<.01*	.03
	Gruppe	1.42	6	.21	-
Zufallsauswahl 2	Vortest	5.66	1	.02*	.02
	Gruppe	1.41	6	.21	-
Zufallsauswahl 3	Vortest	8.57	1	.<.01*	.02
	Gruppe	1.53	6	.17	-
Unterrichten					
Zufallsauswahl 1	Vortest	6.25	1	.01*	.02
	Gruppe	0.66	6	.68	-
Zufallsauswahl 2	Vortest	8.26	1	<.01*	.02
	Gruppe	0.37	6	.90	-
Zufallsauswahl 3	Vortest	10.33	1	<.01*	.02
	Gruppe	0.32	6	.93	-
Diagnostik					
Zufallsauswahl 1	Vortest	13.38	1	<.01*	.03
	Gruppe	1.69	6	.12	-
Zufallsauswahl 2	Vortest	22.10	1	<.01*	.06
	Gruppe	1.50	6	.18	-
Zufallsauswahl 3	Vortest	17.68	1	<.01*	.05
	Gruppe	1.81	6	.10	-
Gesamt[51]					
Zufallsauswahl 1	Vortest	35.95	1	<.01*	.04
	Gruppe	0.51	6	.80	-
Zufallsauswahl 2	Vortest	41.23	1	<.01*	.04
	Gruppe	0.75	6	.61	-
Zufallsauswahl 3	Vortest	45.73	1	<.01*	.05
	Gruppe	0.76	6	.61	-

Anmerkung. F = Teststatistik der Kovarianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Hier ergibt sich über alle drei Themenbereiche sowie insgesamt jeweils ein signifikanter Einfluss der Kovariate „Vortestergebnis" auf das Probeklausurergebnis. Im Gegensatz dazu spielt es vergleichbar zur kurzfristigen Lernwirksamkeit keine Rolle für die Höhe des Probeklausurergebnisses, welche E-Learning-Anwendung die Studierenden vorher bearbeitet hatten. Die Hypothesen 2a und 2b sind demnach ebenfalls zu verwerfen.

7.1.1.3 Längerfristige Lernwirksamkeit – Vergleich Vorlesung vs. E-Learning

Hypothese 3: Die längerfristige Lernleistung (Differenz zwischen Vortest und Probeklausur) ist nach Nutzung der E-Learning-Anwendungen signifikant höher als nach Besuch der Vorlesung.
➔ **EL, ARCS, A, R, C, S, 0 > VL**

[51] Trotz Verletzung der Annahme der Varianzhomogenität wurde hier dennoch eine Kovarianzanalyse berechnet, da die Auswertungen auf sehr großen Stichproben (insgesamt etwa 1000 Studierende) beruhen und nach Bortz und Schuster (2010) die Annahmen der Varianzanalyse mit steigender Stichprobengröße an Bedeutung verlieren.

Durch die umfassende Erhebung von Vortest und Probeklausur über alle zwölf Vorlesungsthemen ist es möglich, die Probeklausurergebnisse nach Bearbeitung der drei E-Learning-Anwendungen mit den Ergebnissen nach dem Besuch der neun planmäßig gehaltenen Vorlesungssitzungen zu vergleichen. Da zwischen den einzelnen E-Learning-Anwendungen kein signifikanter Unterschied besteht, werden hierfür die Ergebnisse der sieben E-Learning-Gruppen zusammengefasst. Es resultiert folglich für jeden Studierenden ein Vortest- und ein Probeklausurergebnis für die neun Vorlesungsthemen und ein Vortest- und Probeklausurergebnis für die drei Themen, die durch E-Learning-Anwendungen bearbeitet wurden.

Tabelle 66 gibt die Ergebnisse der abhängigen *t*-Tests wieder. Es wird deutlich, dass sich die Ergebnisse im Vortest nicht unterscheiden – hinsichtlich beider Präsentationsvarianten (E-Learning/Vorlesung) erreichen die Studierenden ein Vortestergebnis von ca. 34%. In der Probeklausur und für den Lernzuwachs ergeben sich dann jedoch deutliche Mittelwertsunterschiede zwischen den Präsentationsvarianten (*M* = 48.8% für die Vorlesungsthemen vs. *M* = 58.5% für die E-Learning-Themen). Der Lernzuwachs ist nach Bearbeitung der E-Learning-Anwendungen (*M* = 23.3%) um fast 10% höher als nach dem Besuch der Vorlesung (*M* = 13.9%). Diese Unterschiede sind begleitet durch mittelstarke Effekte.

Tabelle 66
Ergebnisse der Mittelwertsanalysen zum Vergleich des Vortestergebnisses, der längerfristigen Lernwirksamkeit sowie des längerfristigen Lernzuwachses zu den Themen, die durch die Vorlesung und den Themen, die durch E-Learning-Anwendungen gelehrt wurden

Vortest (*n* = 1149)	**Deskriptive Statistik**		***t*-Test**			
	M	***SD***	***t***	***df***	***p***	***d***
Vorlesungsthemen	34.03	12.51	-.38	1148	.71	-
E-Learning-Themen	**34.20**	16.01				
Probeklausur (*n* = 682)						
Vorlesungsthemen	48.84	18.17	-18.61	681	<.01*	.71
E-Learning-Themen	**58.54**	21.40				
Lernzuwachs (*n* = 674)						
Vorlesungsthemen	13.91	19.83	-12.94	673	<.01*	.50
E-Learning-Themen	**23.31**	22.84				

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *t* = Teststatistik des *t*-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *d* = Effektstärke

Hypothese 3 ist demnach beizubehalten: Nach Nutzung einer E-Learning-Anwendung ist die längerfristige Lernleistung höher als nach dem Besuch der Vorlesung.

7.1.2 Nutzungszeit und -häufigkeit

Hypothese 4a: Die Nutzungszeit des ARCS-WBTs ist signifikant länger als die Nutzungszeit aller anderen E-Learning-Anwendungen.
➔ **ARCS > A, R, C, S, 0, EL**
Hypothese 4b: Die Nutzungszeit der Einzelkomponenten-WBTs und der E-Lecture sind signifikant länger als die Nutzungszeit des 0-WBTs.
➔ **A, R, C, S, EL > 0**

7.1.2.1 Nutzungszeit

Die für die sieben verschiedenen E-Learning-Anwendungen aufgezeichneten durchschnittlichen Nutzungszeiten in Sekunden werden in Tabelle 67 wiedergegeben und in Abbildung 40 veranschaulicht.

Tabelle 67
Deskriptive Statistik zur Nutzungszeit (in Sek.) aller E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Nutzungs-zeit	Konditionierung			Unterrichten			Diagnostik			Gesamt		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
0-WBT	64	3101.36	2485.80	74	4423.34	3633.13	52	2844.58	2393.72	190	3545.96	3034.44
A-WBT	64	3913.88	2906.51	82	3207.54	2952.64	78	2907.33	2881.04	224	3304.81	2929.93
R-WBT	68	3998.99	2647.59	69	3781.48	2649.61	59	3407.68	2787.74	196	3744.42	2688.03
C-WBT	56	4793.18	4015.28	66	4901.11	3772.71	54	2985.04	2425.88	176	4278.88	3589.32
S-WBT	80	4871.84	3445.34	61	4009.69	3701.38	54	2358.22	2776.27	195	3906.06	3497.37
ARCS-WBT	57	4568.16	3854.27	69	5885.87	3276.67	73	3927.15	3071.80	199	4789.91	3468.03
E-Lecture	200	**6095.11**	3438.44	244	**6272.45**	3629.66	204	**4600.26**	2907.39	648	**5691.29**	3433.61
Gesamt	**589**	**4853.10**	**3454.30**	**665**	**5046.52**	**3624.83**	**574**	**3640.09**	**2918.86**	**1828**	**4542.57**	**3415.70**

Anmerkung. n = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Im Schnitt bearbeiteten die Studierenden die Anwendungen etwa 4500 Sekunden (entspricht etwa 75 Minuten). Es zeigt sich deutlich, dass über alle drei Themenbereiche sowie insgesamt die E-Lecture die längsten Nutzungszeiten aufweist – insgesamt durchschnittlich etwa 95 Minuten. Direkt danach folgt das ARCS-WBT mit rund 80 Minuten. Die C-WBTs werden über alle Themen hinweg ca. 70 Minuten genutzt, die S-WBTs etwa 65 Minuten und die R-WBTs rund 62.5 Minuten. Die relativ kürzesten Bearbeitungszeiten weisen das 0-WBT (ca. 60 Minuten) und das Aufmerksamkeits(A)-WBT (55 Minuten) auf. Augenscheinlich ist dabei die Bearbeitungsdauer für die meisten E-Learning-Varianten zum Thema „Unterrichten“ mit durchschnittlich etwa 84 Minuten am längsten, zum Thema „Diagnostik“ (60 Minuten) am kürzesten.

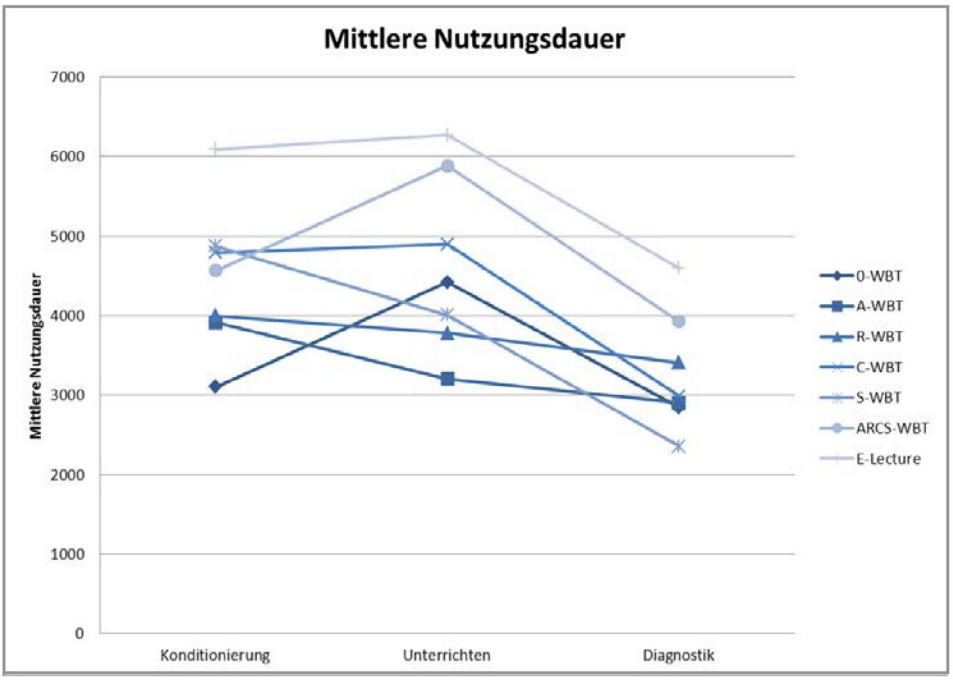

Abbildung 40. Mittlere Nutzungsdauer (in Sek.) der E-Learning-Gruppen zu allen drei Themen

Betrachtet man die Ergebnisse der einfaktoriellen Varianzanalysen zum Vergleich der Nutzungszeiten zwischen den E-Learning-Anwendungen (siehe Tabelle 68), zeigen sich für alle drei Themenbereiche sowie insgesamt signifikante Unterschiede mit mittleren Effektstärken zwischen den Anwendungen.

Tabelle 68
Ergebnisse der Mittelwertsanalysen zum Vergleich der Nutzungszeit der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Nutzungszeit	**ANOVA/Brown-Forsythe-Test**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	3.82	377.77,6	<.01*	.05	EL > 0, A, R; S > 0
Zufallsauswahl 2	5.47	381.11,6	<.01*	.07	EL > 0, A, R; S > 0
Zufallsauswahl 3	7.13	380.95,6	<.01*	.09	EL > 0, A, R, S, ARCS; S > 0
Unterrichten					
Zufallsauswahl 1	8.39	492,6	<.01*	.09	EL > A, R, S, 0; ARCS > A, R, S; C > A
Zufallsauswahl 2	9.45	459.79,6	<.01*	.10	EL > A, R, S, 0; ARCS > A, R, S
Zufallsauswahl 3	7.37	498,6	<.01*	.08	EL, ARCS > A, R, S
Diagnostik					
Zufallsauswahl 1	6.24	431,6	<.01*	.08	EL > R, C, A, 0, S; ARCS > S
Zufallsauswahl 2	3.94	432,6	<.01*	.05	EL > A, S; ARCS > S
Zufallsauswahl 3	4.22	430,6	<.01*	.06	EL > S, 0, A; ARCS > S
Gesamt					
Zufallsauswahl 1	13.01	1303.34,6	<.01*	.05	EL > 0, A, R, C, S; ARCS > 0, A, R
Zufallsauswahl 2	14.23	1311.71,6	<.01*	.06	EL > 0, A, R, C, S; ARCS > 0, A, R
Zufallsauswahl 3	13.84	1317.53,6	<.01*	.06	EL > 0, A, R, C, S; ARCS > 0, A, R

Anmerkung. F = Teststatistik der Varianzanalyse bzw. des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Die Posthoc-Analysen getrennt für die drei Themenbereiche zeigen, dass die E-Lectures signifikant längere Bearbeitungszeiten aufweisen als die 0-, Aufmerksamkeits(A)-, Relevanz (R)- sowie Zufriedenheits(S)-WBTs. Zum Thema „Konditionierung“ wird auch das

Zufriedenheits(S)-WBT signifikant länger bearbeitet als das 0-WBT, zum Thema „Unterrichten“ bearbeiten die ARCS-Nutzer ihr WBT zudem länger als die Nutzer des Aufmerksamkeits(A)-, Relevanz(R)- oder Zufriedenheits(S)-WBTs, zum Thema „Diagnostik“ wird ebenfalls das ARCS-WBT signifikant länger bearbeitet als das Zufriedenheits(S)-WBT. Über alle Themen hinweg wird die E-Lecture signifikant länger bearbeitet als die 0-, A-, R-, C- und S-WBTs, die ARCS-WBTs bedeutsam länger als die 0-, A- und R-WBTs.

Die Hypothesen 4a und 4b sind nach diesen Ergebnissen zu verwerfen. Es sind im Gegensatz eher die E-Lectures, die signifikant länger bearbeitet werden als die meisten Einzelkomponenten-WBTs. Graut man die Ergebnisse zu den E-Lectures einmal aus (wie zur Anschauung in Tabelle 68 geschehen), wird das ARCS-WBT in zwei Themenbereichen sowie in der Gesamtauswertung länger bearbeitet als manche der Einzelkomponenten-WBTs. Das 0-WBT wird lediglich für ein Thema signifikant kürzer bearbeitet als das S-WBT und erreicht in der Gesamtauswertung eine kürzere Bearbeitungszeit als das ARCS-WBT.

Tabelle 69
Ergebnisse der Mittelwertsanalysen zum Vergleich der Nutzungszeit aller Gruppen zwischen den Themen

Nutzungszeit	**Deskriptive Statistik**		**Abhängige Varianzanalyse**				**Posthoc**
(*n* = 310)	***M***	***SD***	***F***	***df***	***p***	**Partielles *Eta²***	
Konditionierung	5039.37	3334.35					
Unterrichten	**5342.80**	3372.93	27.84	1.99	<.01*	.08	Kon, Unt > Dia
Diagnostik	3752.15	2949.79					

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, Partielles *Eta²* = Effektstärke

Der Vergleich der Nutzungszeiten zwischen den drei Themenbereichen über alle E-Learning-Anwendungen hinweg mittels abhängiger Varianzanalyse mit Greenhouse-Geisser-Korrektur ergibt einen signifikanten Unterschied mit mittlerer Effektstärke (siehe Tabelle 69). Dabei werden die Anwendungen zu den Themen „Konditionierung“ und „Unterrichten“ signifikant länger bearbeitet als die Anwendungen zum Thema „Diagnostik“, was sicherlich auf den unterschiedlichen inhaltlichen Umfang der Themen zurückzuführen ist.

7.1.2.2 Nutzungshäufigkeit

Im Rahmen der Nutzungshäufigkeiten ergeben sich zwischen den E-Learning-Anwendungen augenscheinlich nur wenige Unterschiede (siehe Tabelle 70 sowie Abbildung 41). Die Anwendungen werden im Nutzungszeitraum im Schnitt zwischen 1.4 und 2.5 Mal von den Studierenden genutzt. Über alle Themen hinweg resultiert eine durchschnittliche Nutzungshäufigkeit von *M* = 1.8 Mal. Vor allem ist zu beobachten, dass keine der Anwendungen über alle Themenbereiche besonders häufig oder besonders selten genutzt

wird. Es ergibt sich erneut ein eher gemischtes Bild. Über alle drei Themen hinweg werden die ARCS-WBTs augenscheinlich mit *M* = 1.95 Nutzungen durchschnittlich am häufigsten, die E-Lectures mit *M* = 1.66 Nutzungen am seltensten genutzt.

Tabelle 70
Deskriptive Statistik zur Nutzungshäufigkeit der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Nutzungshäufigkeit	Konditionierung			Unterrichten			Diagnostik			Gesamt		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
0-WBT	64	1.91	1.41	74	**1.99**	1.64	52	1.50	0.87	190	1.83	1.40
A-WBT	64	**2.44**	2.00	82	1.74	1.08	78	1.63	1.28	224	1.90	1.49
R-WBT	68	2.09	1.97	69	1.78	1.26	59	1.42	0.84	196	1.78	1.47
C-WBT	56	2.39	1.53	66	1.80	1.19	54	1.52	1.06	176	1.90	1.32
S-WBT	80	2.01	1.63	61	1.84	1.31	54	1.54	0.91	195	1.83	1.37
ARCS-WBT	57	2.42	1.94	69	1.87	1.47	73	**1.66**	1.04	199	**1.95**	1.52
E-Lecture	200	1.73	1.13	244	1.70	1.10	204	1.53	1.07	648	1.66	1.10
Gesamt	**589**	**2.03**	**1.59**	**665**	**1.79**	**1.25**	**574**	**1.55**	**1.04**	**1828**	**1.79**	**1.33**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Über alle Anwendungen hinweg ist allerdings deutlich eine Abnahme der Nutzungshäufigkeit zu erkennen. Werden die Anwendungen zum Thema „Konditionierung" im Schnitt ca. 2.0 Mal genutzt, sind es zum Thema „Unterrichten" noch ca. 1.8 Nutzungen und zum Thema „Diagnostik" durchschnittlich 1.6 Nutzungen.

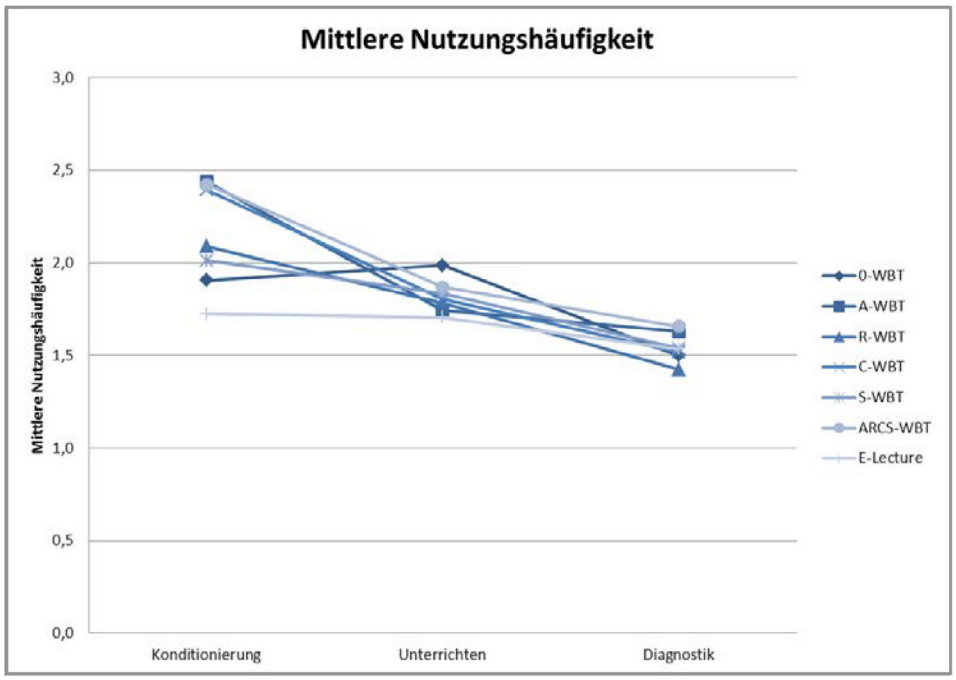

Abbildung 41. Mittlere Nutzungsdauer der E-Learning-Gruppen zu allen drei Themen

Die einfaktoriellen Varianzanalysen zum Vergleich zwischen den E-Learning-Anwendungen bestätigen diese Beobachtungen, indem sie lediglich jeweils für eine Zufallsauswahl für das Thema „Konditionierung" sowie insgesamt einen knapp signifikanten Unterschied

aufzeigen (siehe Tabelle 71). Die Mehrheit der Analysen spricht folglich für eine etwa gleichhäufige Nutzung der Anwendungen.

Tabelle 71
Ergebnisse der Mittelwertsanalysen zum Vergleich der Nutzungshäufigkeit der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Nutzungshäufigkeit	**ANOVA/Brown-Forsythe-Test**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	2.14	446,6	<.05*	.03	
Zufallsauswahl 2	1.77	448,6	.10	-	-
Zufallsauswahl 3	1.50	452,6	.18	-	
Unterrichten					
Zufallsauswahl 1	0.44	442.37,6	.85	-	
Zufallsauswahl 2	0.33	496,6	.92	-	-
Zufallsauswahl 3	0.39	498,6	.89	-	
Diagnostik					
Zufallsauswahl 1	1.13	392.25,6	.35	-	
Zufallsauswahl 2	0.65	432,6	.69	-	-
Zufallsauswahl 3	1.26	430,6	.27	-	
Gesamt					
Zufallsauswahl 1	2.11	1293.74,6	<.05*	.01	EL < A, C, ARCS
Zufallsauswahl 2	1.26	1390,6	.27	-	-
Zufallsauswahl 3	0.39	1394,6	.89	-	-

Anmerkung. F = Teststatistik der Varianzanalyse bzw. des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Der Vergleich zwischen den drei Themenbereichen über alle E-Learning-Varianten hinweg ergibt hingegen einen signifikanten Unterschied mit mittlerem Effekt (siehe Tabelle 72). Die Anwendungen zum Thema „Konditionierung" werden dabei signifikant häufiger genutzt als die Anwendungen zu den beiden anderen Themen. Zudem werden auch die E-Learning-Varianten zum Thema „Unterrichten" häufiger aufgerufen als die Anwendungen zum Thema „Diagnostik".

Tabelle 72
Ergebnisse der Mittelwertsanalysen zum Vergleich der Nutzungshäufigkeit aller Gruppen zwischen den Themen

Nutzungshäufigkeit	**Deskriptive Statistik**		**Abhängige Varianzanalyse**				**Posthoc**
(*n* = 310)	***M***	***SD***	***F***	***df***	***p***	**Partielles *Eta²***	
Konditionierung	**2.09**	1.75					
Unterrichten	1.80	1.27	19.98	1.73	<.01*	.06	Kon > Unt > Dia
Diagnostik	1.48	1.03					

Anmerkung. n = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, Partielles *Eta²* = Effektstärke

7.1.3 Kurz- und längerfristige Lerneffizienz

7.1.3.1 Kurzfristige Lerneffizienz

Hypothese 5: Die kurzfristige Lerneffizienz (Differenz zwischen Vortest und kurzfristigem Nachtest * 1000 / Nutzungszeit) des ARCS-WBTs ist signifikant geringer als die der anderen E-Learning-Anwendungen.
➔ **ARCS < A, R, C, S, 0, EL**

Für die Berechnung der Lerneffizienz wurde vergleichbar zur Studie I der Lernzuwachs durch die Nutzungszeit in Sekunden dividiert. Um die Darstellung der Ergebnisse dieser Formel zu vereinfachen, wurde das Ergebnis mit 1000 multipliziert. Es resultiert zusammengefasst folgende Formel:

*Lerneffizienz = (Lernzuwachs * 1000) / Nutzungszeit in Sekunden*

Je höher das Ergebnis dieser Berechnung, desto effizienter ist die jeweilige Lernanwendung.

Tabelle 73 gibt die kurzfristigen Lerneffizienzwerte für die E-Learning-Anwendungen über alle drei Themenbereiche sowie insgesamt wieder. Veranschaulicht werden die Werte zusätzlich in Abbildung 42.

Tabelle 73
Deskriptive Statistik zur kurzfristigen Lerneffizienz der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Kurzfristige Lerneffizienz	Konditionierung			Unterrichten			Diagnostik			Gesamt		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
0-WBT	63	8.49	29.70	74	16.72	30.27	51	6.08	38.07	188	11.08	32.55
A-WBT	63	3.28	33.24	79	**23.45**	30.09	77	14.62	33.74	219	**14.54**	33.17
R-WBT	68	9.05	26.18	69	14.27	20.82	59	14.02	26.68	196	12.38	24.58
C-WBT	56	9.92	41.74	65	15.87	24.28	54	**16.95**	37.61	175	14.30	34.75
S-WBT	80	9.00	17.93	60	20.80	31.85	54	12.61	60.39	194	13.65	38.32
ARCS-WBT	56	**18.85**	39.57	68	7.70	8.37	71	7.44	23.15	195	10.81	26.22
E-Lecture	197	5.93	18.57	242	8.77	22.17	202	7.59	22.05	641	7.53	21.09
Gesamt	**583**	**8.33**	**27.90**	**657**	**13.70**	**24.93**	**568**	**10.42**	**32.85**	**1808**	**10.94**	**28.63**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Hinsichtlich der kurzfristigen Lerneffizienz ist deutlich zu beobachten, dass sich erneut ein sehr gemischtes Bild ergibt: Während das ARCS-WBT zum Thema „Konditionierung" am effizientesten genutzt wurde, schneiden die ARCS-WBTs zu den beiden anderen Themenbereichen relativ schlecht hinsichtlich ihrer Lerneffizienz ab. Die Lerneffizienz der Einzelkomponenten-WBTs fällt im Gegensatz dazu insbesondere für das Thema „Konditionierung" schlecht, für die anderen beiden Themen deutlich besser aus (dies gilt insbesondere für die A-WBTs). Die Lerneffizienz der E-Lectures liegt für alle drei Themen im unteren Bereich, die der 0-WBTs fällt ebenfalls für zwei der Themen eher niedrig aus. Hinsichtlich des Themas „Unterrichten" erreicht das 0-WBT ein zu den Einzelkomponenten-WBTs vergleichbar gutes Ergebnis. Die Variation zwischen den drei Themenbereichen

(zwischen *M* = 8.3 und *M* = 13.7) ist als eher gering einzuschätzen. Über alle Themen hinweg erreichen die A-WBTs mit *M* = 14.5 den höchsten Effizienzwert, die E-Lectures erscheinen mit *M* = 7.5 als am wenigsten lerneffizient. Alle anderen Anwendungen erreichen Lerneffizienzwerte zwischen 10.8 und 14.3.

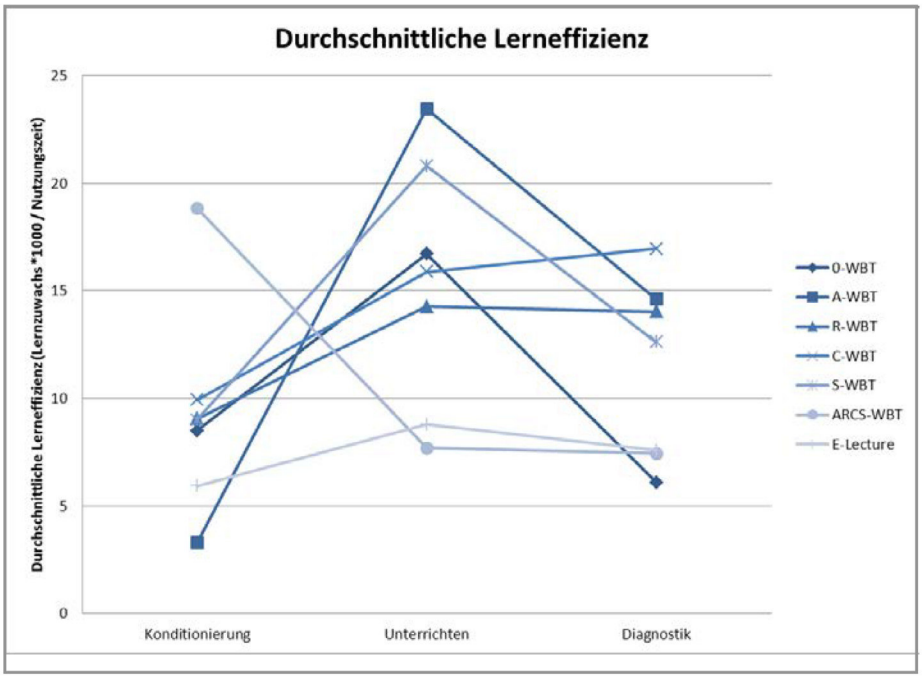

*Abbildung 42. K*urzfristige Lerneffizienz der E-Learning-Gruppen für alle drei Themen

Betrachtet man die Ergebnisse der Brown-Forsythe-Tests (siehe Tabelle 74), ergeben sich lediglich zum Thema „Unterrichten" signifikante Unterschiede mit geringem bis mittelstarkem Effekt.

Tabelle 74
Ergebnisse der Mittelwertsanalysen zum Vergleich der kurzfristigen Lerneffizienz der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Kurzfristige Lerneffizienz	**Brown-Forsythe-Test**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	1.32	310.23,6	.25	-	
Zufallsauswahl 2	1.54	290.40,6	.16	-	-
Zufallsauswahl 3	1.45	317.56,6	.20	-	
Unterrichten					
Zufallsauswahl 1	3.42	397.41,6	<.01*	.04	A > ARCS, EL; S > ARCS
Zufallsauswahl 2	4.23	356.76,6	<.01*	.05	A > ARCS, EL; S > ARCS
Zufallsauswahl 3	4.01	379.15,6	<.01*	.05	A > ARCS, EL; S > ARCS
Diagnostik					
Zufallsauswahl 1	0.90	238.52,6	.49	-	
Zufallsauswahl 2	0.92	231.88,6	.48	-	-
Zufallsauswahl 3	0.67	242.42,6	.68	-	
Gesamt					
Zufallsauswahl 1	1.21	1210.73,6	.30	-	
Zufallsauswahl 2	1.74	1132.18,6	.11	-	-
Zufallsauswahl 3	1.16	1199.41,6	.32	-	

Anmerkung. F = Teststatistik des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Für dieses Thema zeichnet sich das Aufmerksamkeits(A)-WBT durch eine höhere kurzfristige Lerneffizienz aus als das ARCS-WBT und die E-Lecture. Das Zufriedenheits(S)-WBT ist zudem signifikant lerneffizienter als das ARCS-WBT.
Insgesamt ist Hypothese 5 nach der Mehrzahl der Ergebnisse folglich zu verwerfen. Das ARCS-WBT erreicht nicht per se eine schlechtere kurzfristige Lerneffizienz als die anderen E-Learning-Varianten.

7.1.3.2 Längerfristige Lerneffizienz

Hypothese 6: Die längerfristige Lerneffizienz (Differenz zwischen Vortest und Probeklausur * 1000 / Nutzungszeit) des ARCS-WBTs ist signifikant geringer als die der anderen E-Learning-Anwendungen.
➔ **ARCS < A, R, C, S, 0, EL**

Die Berechnung der längerfristigen Lerneffizienz erfolgt vergleichbar zur kurzfristigen. Als Maß für die Lernwirksamkeit wird hier nun allerdings die Differenz zwischen dem Probeklausurergebnis und dem Vortestergebnis genutzt. Tabelle 75 gibt einen Überblick über die längerfristige Lerneffizienz der sieben E-Learning-Anwendungen. Veranschaulicht wird dies zusätzlich in Abbildung 43.

Tabelle 75
Deskriptive Statistik zur längerfristigen Lerneffizienz der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Längerfristige Lerneffizienz	Konditionierung			Unterrichten			Diagnostik			Gesamt		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
0-WBT	45	14.47	33.27	51	17.98	28.76	40	9.69	24.96	136	14.38	29.30
A-WBT	47	12.69	30.36	50	**20.65**	29.98	56	15.97	35.89	153	**16.49**	32.32
R-WBT	47	10.40	28.75	52	11.90	21.40	46	15.46	22.80	145	12.54	24.35
C-WBT	41	14.79	49.37	44	13.89	24.89	40	**20.04**	32.63	125	16.15	36.65
S-WBT	54	12.87	18.04	45	16.82	42.01	30	14.84	57.47	129	14.71	38.64
ARCS-WBT	38	**27.03**	54.72	60	7.80	10.87	52	12.03	26.08	150	14.14	32.91
E-Lecture	130	6.74	15.20	172	7.15	18.32	157	7.43	26.80	459	7.13	20.86
Gesamt	**402**	**12.29**	**31.48**	**474**	**11.89**	**24.73**	**421**	**11.95**	**31.25**	**1297**	**12.03**	**29.09**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Es zeigt sich, dass die längerfristige Lerneffizienz zu allen drei Themenbereichen mit einem Wert von etwa 12 Punkten vergleichbar ausfällt. Innerhalb der Themen variiert die Lerneffizienz jedoch zwischen den E-Learning-Anwendungen um bis zu 21 Punkte.
Dabei ist erneut keine E-Learning-Anwendung ausfindig zu machen, deren Lerneffizienz zu jedem Thema am besten ausfällt. Es ergibt sich analog zur kurzfristigen Lerneffizienz ein sehr heterogenes Bild. Noch deutlicher als bereits im Rahmen der kurzfristigen Lerneffizienz ist zu erkennen, dass die E-Lectures zu allen drei Themenbereichen hinsichtlich ihrer Lerneffizienz augenscheinlich am schlechtesten abschneiden. Die ARCS-WBT erreichen erneut für das erste Thema die beste Lerneffizienz, fallen dann für die beiden anderen Themen jedoch relativ weit ab. Die anderen WBT-Varianten bewegen sich

über alle Themen hinweg sehr nah beieinander im oberen Effizienzbereich. Über alle Themen hinweg erreichen alle WBT-Varianten damit eine sehr vergleichbare längerfristige Lerneffizienz (12.5 bis 16.5), die Lerneffizienz der E-Lecture liegt mit *M* = 7.1 deutlich darunter.

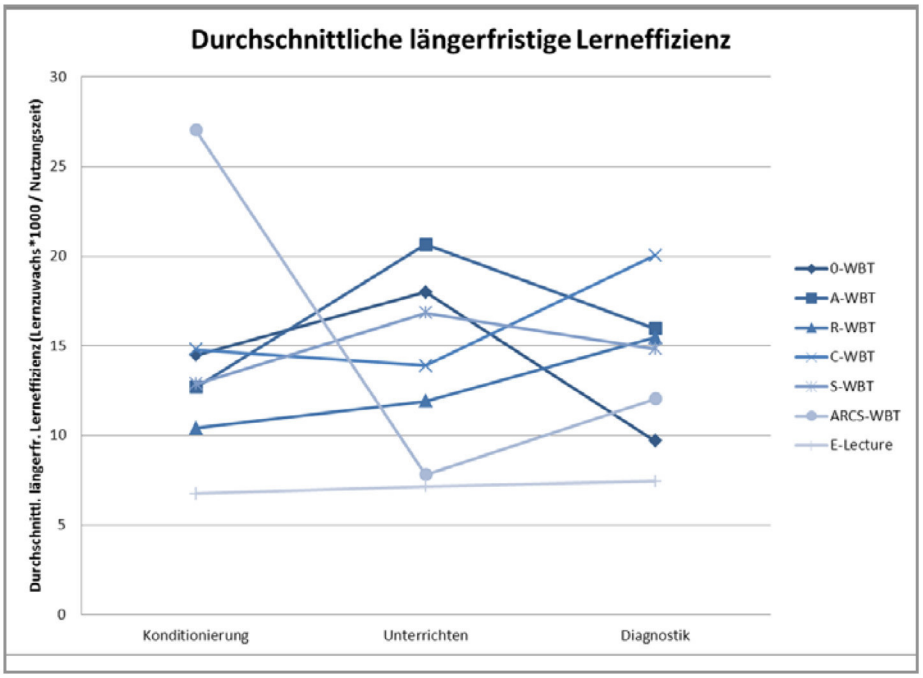

Abbildung 43. Längerfristige Lerneffizienz der E-Learning-Gruppe für alle drei Themen

Die Analyse dieser Beobachtungen mittels Brown-Forsythe-Tests (Ergebnisse siehe Tabelle 76) ergibt in der deutlichen Mehrzahl der Analysen keine signifikanten Unterschiede zwischen den sieben E-Learning-Anwendungen.

Tabelle 76
Ergebnisse der Mittelwertsanalysen zum Vergleich der längerfristigen Lerneffizienz der E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Längerfristige Lerneffizienz	**Brown-Forsythe-Test**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	1.3	164.02,6	.26	-	
Zufallsauswahl 2	1.23	173.21,6	.30	-	-
Zufallsauswahl 3	1.21	170.40,6	.31	-	
Unterrichten					
Zufallsauswahl 1	1.83	234.66,6	.10	-	-
Zufallsauswahl 2	2.47	194.44,6	.03*	.04	0, A > EL
Zufallsauswahl 3	1.67	208.03,6	.13	-	-
Diagnostik					
Zufallsauswahl 1	0.56	143.83,6	.77	-	
Zufallsauswahl 2	0.83	127.34,6	.55	-	-
Zufallsauswahl 3	0.70	134.79,6	.65	-	
Gesamt					
Zufallsauswahl 1	1.37	860.85,6	.22	-	
Zufallsauswahl 2	1.99	798.19,6	.07	-	-
Zufallsauswahl 3	1.25	822.29,6	.28	-	

Anmerkung. F = Teststatistik des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Lediglich für eine Zufallsauswahl zum Thema „Unterrichten“ erreicht die E-Lecture eine signifikant schlechtere Lerneffizienz als das 0- und das A-WBT. Insgesamt ist Hypothese 6 folglich zu verwerfen.

7.1.3.3 Längerfristige Lerneffizienz – Vergleich Vorlesung vs. E-Learning

Hypothese 7: Die längerfristige Lerneffizienz (Differenz zwischen Vortest und Probeklausur * 1000 / Nutzungszeit bzw. Vorlesungszeit) der E-Learning-Anwendungen ist signifikant höher als die der Vorlesung.
➔ **EL, ARCS, A, R, C, S, 0 > VL**

Vergleichbar zur längerfristigen Lernwirksamkeit wurde aus der Differenz zwischen Vortest- und Probeklausurergebnis dividiert durch die durchschnittliche Vorlesungszeit für die neun Vorlesungsthemen ebenfalls ein Lerneffizienzwert berechnet. Für die Berechnung der Lerneffizienz der Vorlesungsthemen wurde dabei vergleichbar zur Studie I als recht strenge Schätzung der Vorlesungszeit die Rohlaufzeit der E-Lecture zugrundegelegt. Diese beträgt für die neun Themen im Schnitt 3404 Sekunden (entspricht etwa 57 Minuten) pro Vorlesung, während die E-Learning-Anwendungen von den 358 in die Auswertung eingehenden Studierenden im Schnitt 4920 Sekunden (entspricht ca. 82 Minuten) genutzt wurden. Da sich die Stichprobe durch Einbezug der Nutzungszeiten deutlich reduziert, werden in Tabelle 77 erneut die Ergebnisse für die Vortestergebnisse, die Probeklausurergebnisse und den Lernzuwachs wiedergegeben, bevor die Lerneffizienzmaße der Vorlesungs- und E-Learning-Themen verglichen werden.

Tabelle 77
Ergebnisse der Mittelwertsanalysen zum Vergleich der längerfristigen Lernwirksamkeit und -effizienz zu den Themen, die durch die Vorlesung und den Themen, die durch E-Learning-Anwendungen gelehrt wurden

Vortest (n = 358)	**Deskriptive Statistik**		***t*-Test**			
	M	***SD***	***t***	***df***	***p***	***d***
Vorlesungsthemen	34.83	12.72	1.75	357	.08	-
E-Learning-Themen	**36.24**	15.35				
Probeklausur						
Vorlesungsthemen	50.83	17.54	18.08	357	<.01*	1.02
E-Learning-Themen	**63.22**	19.22				
Lernzuwachs						
Vorlesungsthemen	15.99	19.90	11.06	357	<.01*	.60
E-Learning-Themen	**26.98**	22.22				
Lerneffizienz						
Vorlesungsthemen	4.70	5.85	4.07	357	<.01*	.31
E-Learning-Themen	**8.91**	20.79				

Anmerkung. n = Stichprobengröße, M = Mittelwert, SD = Standardabweichung, t = Teststatistik des t-Tests, df = Freiheitsgrade, p = Signifikanz, d = Effektstärke

Wie bereits oben für die umfassendere Stichprobe berichtet, ergeben sich auch für die hier betrachtete Teilstichprobe keine Unterschiede zwischen den Vortestleistungen der

Studierenden zu den Themen, die sie in der Vorlesung gehört und den Themen, die sie mit E-Learning-Anwendungen bearbeiteten haben.

Die Leistungen in der Probeklausur – und daraus resultierend die Lernzuwachs-Unterschiede – ergeben hingegen signifikante Unterschiede. Auch diese Teilstichprobe erreicht in der Probeklausur ein deutlich besseres Ergebnis für die Themenbereiche, die mit einer E-Learning-Anwendung bearbeitet wurden als für die Themenbereiche, die in der Vorlesung behandelt wurden (starker Effekt). Der Lernzuwachs für die in der Vorlesung gehörten Themen liegt mit M = 16.0% um 11% niedriger als der Lernzuwachs für die Themen, die die Studierenden mit einer E-Learning-Anwendung bearbeitet haben (M = 27.0%, mittelstarker Effekt).

Obwohl die Studierenden deutlich länger mit den E-Learning-Anwendungen arbeiteten als sie Zeit mit der Darbietung der Lerninhalte in der Vorlesung verbrachten, resultiert auch für die Lerneffizienz ein signifikant besseres Ergebnis zugunsten der E-Learning-Anwendungen (siehe Tabelle 77). Dies ist begleitet von einem kleinen bis mittelstarken Effekt. Hypothese 7 ist daher beizubehalten.

7.1.4 Zusammenfassung und Diskussion

7.1.4.1 Lernleistung

Bereits zu Beginn des Semesters wurden die Studierenden zu allen Themen der kommenden Vorlesung befragt. Dabei zeigte sich, dass sie die Fragen zu den Themen „Konditionierung“ und „Diagnostik“ im Durchschnitt bereits zu 40% korrekt beantworten konnten, zum Thema „Unterrichten“ erreichten sie lediglich 27% der möglichen Gesamtpunktzahl. Die Werte dieses Vortests wurden im Folgenden dazu genutzt, um etwaige Vorwissensunterschiede zwischen den Experimentalgruppen auszuschließen. Dementsprechend wurde das Vortestergebnis entweder als Kovariate auspartialisiert oder ein Lernzuwachswert (Nachtest- bzw. Probeklausurergebnis minus Vortestergebnis) herangezogen.

Kurzfristige Lernleistung

Direkt nach Bearbeitung der E-Learning-Anwendungen wurde im Rahmen eines zehn Fragen umfassenden Nachtests die kurzfristige Lernleistung der Studierenden erfasst. Dabei zeigte sich, dass die Variation zwischen den Nachtestergebnissen wesentlich geringer ausfiel als zwischen den oben erwähnten Vortestergebnissen. Die Studierenden konnten in den Nachtests zwischen 57% und 62% der Fragen korrekt beantworten. Dementsprechend resultiert für das Thema „Unterrichten“ (M = 31.1%) ein deutlich höherer Lernzuwachs als für die beiden anderen Themen (M = 20.1% und M = 17.6%). Über alle Themen hinweg

ergibt sich ein Lernzuwachs von *M* = 23.5%. Verglichen mit den Ergebnissen von McLeish (1968, lediglich 20% nach 45-minütigem Vortrag) handelt es sich hierbei um sehr gute Lernresultate, da die Themen wesentlich längeren Vorträgen entsprachen. Interessant ist in diesem Zusammenhang, dass für das umfangreichste Thema „Unterrichten" (siehe hierzu auch „7.1.4.2 Nutzungszeit und -häufigkeit") der höchste Lernzuwachs beobachtet werden konnte. Dies hängt sicherlich mit dem relativ niedrigen Vortestergebnis der Studierenden hinsichtlich dieses Themas zusammen.

Vergleichbar zu Studie I muss erneut berücksichtigt werden, dass die Modulabschlussprüfung, für die die Studierenden die Inhalte lernen mussten, erst mehr als ein halbes Jahr später stattfand. Zum Zeitpunkt der kurz- und auch der längerfristigen Wissensüberprüfung war folglich noch viel Zeit, um die Inhalte zu verinnerlichen. Es ist erfreulich, dass die Lernzuwachsraten dennoch so gut ausgefallen sind. Dies spricht für die Lernförderlichkeit der eingesetzten Anwendungen und dafür, dass alle Variationen im Hinblick auf den Lernerfolg der Lernenden mit gutem Gewissen eingesetzt werden können. Bestärkt wird dies auch dadurch, dass sich beim Vergleich der sieben E-Learning-Varianten keine signifikanten Unterschiede ergaben. Die Kovarianzanalyse zum Vergleich der Nachtestergebnisse mit der Kovariate „Vortestergebnis" zeigt zwar einen signifikanten Einfluss des Vortestergebnisses auf das Nachtestergebnis für die Themenbereiche „Konditionierung" und „Unterrichten" sowie für die Gesamtauswertung über alle Themen. Es spielte jedoch für die Höhe des Nachtestergebnisses keine Rolle, welche der E-Learning-Varianten vorher bearbeitet worden war. Damit konnte der in Studie I gefundene kurzfristige Lernwirksamkeitsvorteil der WBTs gegenüber den E-Lectures nicht repliziert werden. Dabei muss bedacht werden, dass sich für die hier untersuchten Themen (Konditionierung, Unterrichten und Diagnostik) in Studie I ebenfalls keine Lernwirksamkeitsunterschiede zwischen den WBTs und E-Lectures ergeben hatten (siehe Anhang B). Eine Optimierung der Web-based Trainings durch die ARCS-Strategien konnte demnach nicht dazu beitragen, die kurzfristige Lernwirksamkeit gegenüber der „alten" WBT-Version bzw. der anderen in dieser Studie genutzten WBT-Varianten und der E-Lecture zu erhöhen. Auch in früheren Evaluationsstudien zum ARCS-Modell von Keller (siehe Abschnitt „5.3.5 Wirksamkeit des Modells") zeigte sich eine eher uneinheitliche Befundlage hinsichtlich der Auswirkungen der ARCS-Strategien auf die Lernwirksamkeit. Astleitner und Lintner (2004) erklären dabei zum Beispiel das schlechte Abschneiden ihres mit ARCS-Strategien angereicherten Lehrtextes dadurch, dass dieser sehr viel länger war als der nicht-angereicherte Text. Auch Song und Keller (2001) sprechen von einer möglichen Überfrachtung einer E-Learning-Anwendung mit allen ARCS-Strategien, welche sich negativ auf u.a. die Lernwirksamkeit auswirken kann. Dies könnte auch im Falle der zweiten Studie eine mögliche Erklärung für die nicht-signifikanten Lernwirksamkeits-

unterschiede zwischen den E-Learning-Varianten darstellen, da das ARCS-WBT deutlich länger war als die anderen WBT-Varianten. Ein Blick auf die Ergebnisse zur Skala „Aufmerksamkeit“ (siehe Abschnitt „5.4.2 ARCS-Skala“) spricht allerdings eher gegen eine solche Interpretation.

Längerfristige Lernleistung

Die längerfristige Lernleistung wurde am Ende des Wintersemesters durch eine – ähnlich dem Vortest alle Themen umfassende – Probeklausur erfasst. Die Probeklausurleistungen der Studierenden bewegen sich insgesamt zwischen 57% und 69%. Der längerfristige Lernzuwachs liegt damit für die drei Themen zwischen insgesamt 23% und 29% (über alle Themen bei $M = 26.4\%$) und fällt für die Themen „Konditionierung“ und „Diagnostik“ – und damit auch für die Gesamtauswertung – sogar etwas höher aus als der kurzfristige Lernzuwachs. Zum Thema „Unterrichten“ sinkt der Lernzuwachs um ca. 2% von der kurzfristigen zur längerfristigen Erhebung. Dies spricht dafür, dass die Studierenden nicht nur durch die E-Learning-Nutzung sondern vermutlich auch durch die Erhebung der kurzfristigen Lernleistung profitiert haben, was in Abschnitt „7.1.4.3 Vergleich der Vorlesung mit den E-Learning-Anwendungen“ eine bedeutsame Rolle spielen wird. Doch auch die Tatsache, dass die Probeklausur durch die Erhebung des Vortests bereits zur Hälfte bekannt war, wird die Leistungen sicherlich positiv beeinflusst haben.

Allerdings ergeben sich auch längerfristig zwischen den Lernleistungen der verschiedenen E-Learning-Gruppen keine signifikanten Unterschiede. Erneut ergibt die Kovarianzanalyse zum Vergleich der Probeklausurergebnisse mit der Kovariate „Vortestergebnis“ einen signifikanten Einfluss der Kovariate auf das Probeklausurergebnis. Jedoch spielt es auch hier keine Rolle für die Höhe des Probeklausurergebnisses, welche E-Learning-Anwendung im Vorhinein bearbeitet worden war. Dies bestätigt nun die Ergebnisse der ersten Studie, in welcher ebenfalls keine signifikanten längerfristigen Lernwirksamkeitsunterschiede zwischen WBTs und E-Lectures nachgewiesen werden konnten.

Die längerfristigen Lernleistungen fallen dabei deutlich höher aus innerhalb der ersten Studie. Hier konnten die Studierenden etwa 50% der Probeklausurfragen korrekt beantworten. Eine Aussage über den wahren Lernzuwachs kann für die erste Studie allerdings nicht gemacht werden, da keine Erhebung des Vorwissens stattgefunden hatte. Sicherlich kann ein Lernzuwachs von $M = 26.4\%$ nach einer so langen Zeit nach der E-Learning-Nutzung als positiv beurteilt werden. Im Vergleich zu den Ergebnissen von McLeish, der bereits eine Woche nach einer Vorlesung nur noch einen Lernzuwachs von 17% feststellen konnte, ist das Ergebnis in der Probeklausur insbesondere deshalb erfreulich, da die Daten für die drei Themen zwischen einer Woche und zwei Monaten nach Bearbeitung der E-Learning-Anwendungen erhoben wurden. Hierbei muss allerdings im

Auge behalten werden, dass – wie oben bereits angeklungen – die Hälfte der Fragen bereits im Vortest enthalten waren und die Studierenden auch im Vortest nach jeder Frage ein Feedback erhielten. Dies wird sich sicherlich auf das Ergebnis positiv ausgewirkt haben. Hierfür spricht auch die Beobachtung, dass die längerfristige Lernwirksamkeit für zwei Themen etwas besser ausfällt als die kurzfristige, in welcher keine Fragen der Vortests enthalten waren. Der Vergleich der Testleistungen der Studierenden hinsichtlich der „alten" und „neuen" Fragen in der Probeklausur[52] bestätigt dies. Hier ergibt sich (mit n = 682, $t(681) = 16.11$) ein signifikanter Unterschied mit $p < .01$ und $Eta^2 = .62$. Für die bereits bekannten Fragen erreichen die Studierenden einen Mittelwert von $M = 16.78$ ($SD = 5.72$), für die unbekannten Fragen einen Mittelwert von $M = 14.86$ ($SD = 5.72$). Dies verbietet zwar eine klare Einschätzung der Güte der längerfristigen Lernwirksamkeit, wird sich jedoch über alle Experimentalgruppen in etwa gleich ausgewirkt haben. Daher sollten die Vergleiche zwischen den Varianten hiervon unbeeinflusst sein.

7.1.4.2 Nutzungszeit, Nutzungshäufigkeit und Lerneffizienz

Die Nutzungszeit, -häufigkeit und Lerneffizienz sollen an dieser Stelle gemeinsam diskutiert werden, da sie naturgemäß sehr eng zusammenhängen.

Die Nutzungszeiten der E-Learning-Anwendungen wurden durch das Lehrmanagement-system *Moodle* aufgezeichnet. Für alle drei Themenbereiche ergaben sich im Vergleich der E-Learning-Anwendungen signifikante Unterschiede mit mittleren Effektstärken. Dabei werden die E-Lectures nach der Mehrzahl der Analysen signifikant länger bearbeitet als die 0-, Aufmerksamkeits(A)-, Zufriedenheits(S)- und Relevanz(R)-WBTs. Darüber hinaus bearbeiten die ARCS-Nutzer ihr WBT mehrheitlich signifikant länger als die Zufriedenheits(S)-WBT-Nutzer. Für die Auswertung über alle Themen hinweg zeigten sich zudem signifikant längere Bearbeitungszeiten der E-Lectures gegenüber den C-WBTs. Die ARCS-WBTs wurden über alle Themen hinweg länger bearbeitet als die 0-, A- und R-WBTs.

Die Lerneffizienz ergibt sich aus dem Quotienten des Lernzuwachses (*1000) und der Nutzungszeit der E-Learning-Anwendungen bzw. der Vorlesungsdauer. Dabei wurde ein Lernzuwachsmaß für die kurzfristige (kurzfristiges Nachtestergebnis minus Vortestergebnis) und für die längerfristige Lerneffizienz (Probeklausurergebnis minus Vortestergebnis) ermittelt. Für beide Effizienzmaße ist die gleiche Nutzungszeit bzw. Vorlesungszeit zugrundegelegt.

Im Rahmen der kurzfristigen Lerneffizienz ergeben sich nur geringe Variationen zwischen den drei Themenbereichen und auch die Lerneffizienz der sieben verschiedenen E-Learning-Anwendungen unterscheiden sich nur im Themenbereich „Unterrichten"

[52] In beiden Teilen der Probeklausur konnten insgesamt 30 Punkte erreicht werden.

signifikant voneinander. Dieser Unterschied weist zudem einen nur geringen bis mittelstarken Effekt auf. Das Aufmerksamkeits(A)-WBT ist hier lerneffizienter als das ARCS-WBT und die E-Lecture. Das Zufriedenheits(S)-WBT zeichnet sich ebenfalls durch eine höhere Lerneffizienz aus als das ARCS-WBT. In der Mehrzahl der Analysen ergeben sich folglich keine signifikanten Vorteile einer der E-Learning-Anwendungen im Hinblick auf die kurzfristige Lerneffizienz.
Die längerfristige Lerneffizienz liegt für alle drei Themenbereiche bei ca. 12 Punkten und ist dementsprechend im Themenvergleich noch homogener als bereits die kurzfristige Lerneffizienz. Im Vergleich zwischen den E-Learning-Varianten variiert die Lerneffizienz mit bis zu 21 Punkten stärker. Es ergeben sich zwischen den Varianten allerdings vergleichbar zur kurzfristigen Lerneffizienz keine signifikanten Unterschiede.

Die zumeist vergleichbaren Nutzungszeiten und die resultierende vergleichbare Lerneffizienz der WBT-Varianten sind erstaunlich, da insbesondere die ARCS- und C-Varianten unter anderem durch die Übungsaufgaben deutlich längere Bearbeitungszeiten erwarten ließen. Für die ARCS- und C-WBTs sind zwar tendenziell längere Nutzungszeiten zu erkennen, betrachtet man die Mittelwerte über alle drei Themen hinweg. Die Unterschiede werden jedoch nicht durchgängig signifikant. Aber auch die anderen Einzelkomponenten-WBTs wurden im Vergleich zu den 0-WBTs durch für die Lerner zeitaufwendige Aspekte ergänzt. Es ist sehr überraschend, dass das 0-WBT nicht deutlich kürzer bearbeitet wurde als ein Teil der Einzelkomponenten-WBTs. Dies könnte darin begründet sein, dass eine Darstellung der Inhalte ohne Ausschmückung und Abwechslung, Beispiele und Übungsmöglichkeiten etc. auf Dauer sehr ermüdend für die Studierenden war. Dementsprechend könnte die Bearbeitung dieser WBT-Version im Gegensatz zu den anderen WBT-Varianten vermehrt dazu geführt haben, dass die Studierenden „abdrifteten“ oder Pausen einlegen mussten, was dann in einer Verlängerung der Nutzungszeiten des 0-WBTs resultierte. Ein Hinweis hierauf ist die augenscheinlich um durchschnittlich fünf Minuten kürzere Bearbeitungszeit des A-WBTs, in welchem gerade diejenigen Maßnahmen umgesetzt wurden, welche die Aufmerksamkeitserhaltung erleichtern sollten. Es könnte allerdings auch dafür sprechen, dass die schmucklose, faktische Darstellung der Lerninhalte innerhalb des 0-WBTs die Studierenden veranlasste, die Inhalte vermehrt selbständig zu durchdenken. Gegebenenfalls suchten sie selbst nach Anwendungsmöglichkeiten und Anknüpfungspunkten zu bereits Gelerntem.
Um dem Grund für die nicht signifikanten Lernzeitunterschiede nachzugehen, bedarf es einer Beobachtung der Lernenden während der Nutzung der E-Learning-Varianten. Es wäre hier interessant zu beobachten, wie sie die einzelnen WBT-Varianten nutzen, ob es ihnen beispielweise im Rahmen der 0-WBTs deutlich schwerer fällt, bei der Sache zu bleiben als

bei der Bearbeitung von anderen WBT-Varianten oder ob sie während der Bearbeitung selbständig „Denkpausen“ einlegen. Die unten berichteten Ergebnisse zur ARCS-Skala zeigen bereits, dass das ARCS-WBT hinsichtlich der Aufmerksamkeitsskala signifikant besser beurteilt wird als das 0-WBT. Dies könnte bereits ein erster Hinweis auf Schwierigkeiten sein, die die Studierenden mit dem 0-WBT hatten. Es handelt sich hierbei allerdings lediglich um subjektive Fragebogendaten und es wäre interessant, diese durch objektive Beobachtungen zu ergänzen.

Im Vergleich der drei Themenbereiche zeigt sich, dass die Bearbeitungszeiten zu den Themen „Konditionierung“ und „Unterrichten“ signifikant länger ausfallen als zum Thema „Diagnostik“ (mittelstarker Effekt). Dies ist sicherlich ganz einfach dadurch zu erklären, dass die Themen „Konditionierung“ und „Unterrichten“ inhaltlich umfangreicher sind als das Thema „Diagnostik“. Beispielsweise besteht das 0-WBT zum Thema „Konditionierung“ aus 30 Seiten, das 0-WBT zum Thema „Unterrichten“ aus 29 Seiten und das 0-WBT zum Thema „Diagnostik“ aus 22 Seiten. Vergleichbar fällt eine Gegenüberstellung für die anderen WBT-Varianten aus. Hinsichtlich des Themas „Diagnostik“ ist zudem zu bedenken, dass dieses sehr spät im Semester bearbeitet wurde – zu einem Zeitpunkt, an dem die Studierenden bereits in der Vorbereitung zu anderen Prüfungen waren. Auch dieser Umstand kann zulasten der Bearbeitungszeit gegangen sein.

Für die Nutzungshäufigkeit der E-Learning-Anwendungen ergibt sich kein signifikanter Unterschied zwischen den Anwendungsvarianten. Allerdings zeigt sich auch in Bezug auf die Nutzungshäufigkeit eine häufigere Nutzung der E-Learning-Anwendungen zum Thema „Konditionierung“ (mittelstarker Effekt). Diese Anwendungen werden im Schnitt häufiger genutzt als die Anwendungen zu den beiden anderen Themen. Dabei werden die E-Learning-Anwendungen zum Thema „Unterrichten“ ebenfalls häufiger genutzt als diejenigen zum Thema „Diagnostik“. Auch dies könnte ggf. ein Zeichen für das abfallende Interesse an den E-Learning-Anwendungen sein – es ist jedoch auch möglich, dass die Studierenden mit den Anwendungen vertrauter wurden und dementsprechend weniger Neustarts vornehmen mussten.

7.1.4.3 Vergleich der Vorlesung mit den E-Learning-Anwendungen

Zudem wurden die längerfristigen Lernleistungen über die drei Themen, die die Studierenden mit E-Learning-Anwendungen bearbeitet hatten („E-Learning-Themen“) mit den Leistungen bezüglich der neun Themen verglichen, für die die Studierenden die Vorlesung besucht hatten („Vorlesungsthemen“).

Die Vortestergebnisse dieser beiden Themengruppen lagen jeweils bei 34% – unterschieden sich demnach nicht signifikant voneinander. Die Studierenden hatten folglich kein unterschiedlich gutes Vorwissen hinsichtlich der Vorlesungs- bzw. E-Learning-Themen. Allerdings ergaben sich Signifikanzen im Vergleich der Probeklausurergebnisse und dementsprechend für den Lernzuwachs: Die Studierenden erreichten durch die Bearbeitung der E-Learning-Anwendungen einen durchschnittlichen längerfristigen Lernzuwachs von M = 23.3%, durch den Besuch der Vorlesung einen Lernzuwachs von M = 13.9%.

Diese Ergebnisse sind allerdings mit Bedacht zu interpretieren. Zum einen ist anzumerken, dass es zwar für die Bearbeitung der E-Learning-Anwendung eine Kontrolle gab – hier zeichnete das Lehrmanagement-System auf, dass der betreffende Studierende die jeweilige E-Learning-Anwendung zumindest aufgerufen hatte (eine Kontrolle darüber, was der Studierende vor dem Computer wirklich getan hat, existiert im Rahmen einer solchen Erhebung allerdings auch nicht). Im Gegensatz dazu wurde die Teilnahme an der Vorlesung nicht kontrolliert. Es kann also nicht sicher gesagt werden, dass alle, die in die Auswertung der Vorlesungsleistungen eingingen, auch wirklich in allen betreffenden Vorlesungen anwesend waren. Klar ist aber, dass das aufmerksame Verfolgen der Ausführungen des Dozenten in einer Vorlesung (besonders in dieser Größenordnung) genauso wenig kontrollierbar ist wie die aufmerksame Bearbeitung einer E-Learning-Anwendung.

Zum anderen bearbeiteten die Studierenden nach Nutzung der E-Learning-Anwendungen die kurzfristigen Wissensüberprüfungen, durch die sie sich erneut mit den behandelten Inhalten auseinandersetzen mussten. Zwar handelt es sich hierbei nicht um die gleichen Fragen, die später in der Probeklausur gestellt wurden. Es ist allerdings wahrscheinlich, dass diese Wissensüberprüfungen ähnliche Effekte haben wie Übungsaufgaben („Testing effect“, siehe zum Beispiel Hamaker, 1986; Larsen, Butler & Roediger III, 2009; Nungester & Duchastel, 1982 oder Roedinger III & Karpicke, 2006). Dies wird insbesondere deshalb der Fall sein, da die Studierenden nach jeder Frage ein Feedback zur Richtigkeit ihrer Antwort erhielten. Eine solche Nacherhebung wurde nach dem Besuch der Vorlesung nicht durchgeführt. Dies wird die deutlichen Lernwirksamkeitsunterschiede zumindest mitverursacht haben.

Insgesamt müssen die Ergebnisse zur Lernwirksamkeit sowie zur Lerneffizienz im Vergleich der E-Learning-Varianten mit der Vorlesung folglich mit Vorsicht interpretiert werden. Um beide Lehrvarianten vergleichbar zu halten, wäre es notwendig gewesen, die Studierenden in jeder Vorlesungssitzung auf einer Teilnehmerliste unterschreiben und sie auch nach jeder Vorlesungssitzung eine kurze Nacherhebung bearbeiten zu lassen. Eine solche zusätzliche Erhebung in neun Vorlesungssitzungen mit über 1200 Studierenden hätte die Versuchsdurchführung allerdings viel aufwendiger und für die Studierenden

deutlich unbequemer gemacht. Insbesondere die Kontrolle der Teilnahme an der Vorlesung wäre zudem nie vollständig möglich gewesen, da auch die Teilnehmerliste keine Auskunft darüber gegeben hätte, welche Studierende wirklich bis zum Ende der Vorlesung anwesend (und aufmerksam) waren.

Trotz dieser Einschränkungen soll abschließend die längerfristige Lerneffizienz der E-Learning- und Vorlesungsthemen betrachtet werden. Es ergibt sich hier ein signifikanter Unterschied zwischen den Themen, die die Studierenden mit E-Learning-Anwendungen bearbeitet haben und den Themen, die sie in der Vorlesung gehört haben. Dieser fällt zugunsten der E-Learning-Themen aus, obwohl die Studierenden mehr Zeit mit den E-Learning-Anwendungen verbrachten als in der Vorlesung (etwa durchschnittlich 82 Minuten Bearbeitung der E-Learning-Anwendungen vs. ca. 57 Minuten Vorlesungszeit). Im Gegensatz zur Lernwirksamkeit zeigte sich dieses Ergebnis bereits in Studie I. Auch hier erreichten die WBTs eine signifikant bessere Lerneffizienz als die Vorlesung. Die E-Lectures konnten in Studie I leider nicht hinsichtlich ihrer Nutzungszeiten ausgewertet werden. Ähnlich zur Studie I kommt der Lerneffizienzvorteil aufgrund einer besseren Lerneffektivität zustande – nicht durch eine kürzere Bearbeitungszeit der E-Learning-Anwendungen. Dies gibt den oben erläuterten Einschränkungen neues Gewicht. Dagegen kann allerdings die Tatsache gestellt werden, dass die Schätzung der Vorlesungszeit durch die Rohlaufzeit der E-Lecture mit durchschnittlich 57 Minuten eine sehr strenge darstellt. Diese sollte einer Unterschätzung der Lerneffizienz der Vorlesung vorbeugen. Dies wäre sicherlich der Fall gewesen, wenn man für die Vorlesungszeit die regulären 90 Minuten zugrundegelegt hätte. Die Annahme von durchschnittlich 57 Minuten beinhaltet allerdings nur die Zeit, die im Rahmen der ursprünglichen Aufzeichnung der E-Lectures ausschließlich für die Vermittlung von Inhalten genutzt wurde. Nachfragen von Studierenden oder organisatorische Informationen – welche ebenfalls zu einer Vorlesung gehören – werden hier außer Acht gelassen. Dies erhöht sicherlich die Vergleichbarkeit. Die Effizienz der Vorlesung könnte dadurch nun allerdings überschätzt werden. Das würde das Ergebnis der Analysen noch untermauern, denn womöglich verbringen die Studierenden bei der Bearbeitung der E-Learning-Anwendungen auch Zeit mit inhaltsfernen Dingen. Dies ist im Rahmen einer Felduntersuchung nicht überprüfbar. Hier wäre zur Ergänzung der vorliegenden Ergebnisse eine weitere Untersuchung im Labor wünschenswert.

Durch die sehr strenge Schätzung der Lernzeit für die Vorlesung relativiert sich folglich zum Teil der zuvor diskutierte Vorteil der E-Learning-Anwendungen hinsichtlich der Lernwirksamkeitserhebung. Da sich die Ergebnisse zur Lerneffizienz mit denen in Studie I decken, kann trotz aller Einschränkungen von einer besseren Lerneffizienz der E-Learning-Anwendungen ausgegangen werden.

7.2 ARCS-Skala

Hypothese 8a: Die Studierenden fühlen sich durch das ARCS- und das A-WBT hinsichtlich ihrer Aufmerksamkeitshaltung signifikant besser unterstützt als durch die anderen WBT-Varianten.
➔ **ARCS, A > R, C, S, 0**

Hypothese 8b: Die Studierenden fühlen sich durch das ARCS- und das R-WBT hinsichtlich ihres Relevanzgefühls signifikant besser unterstützt als durch die anderen WBT-Varianten.
➔ **ARCS, R > A, C, S, 0**

Hypothese 8c: Die Studierenden fühlen sich durch das ARCS- und das C-WBT hinsichtlich ihrer Erfolgszuversicht signifikant besser unterstützt als durch die anderen WBT-Varianten.
➔ **ARCS, C > A, R, S, 0**

Hypothese 8d: Die Studierenden fühlen sich durch das ARCS- und das S-WBT signifikant besser zufriedengestellt als durch die anderen WBT-Varianten.
➔ **ARCS, S > A, R, C, 0**

Hypothese 8e: Die Studierenden fühlen sich durch das ARCS-WBT hinsichtlich aller vier ARCS-Aspekte (Gesamt-ARCS-Skala) signifikant besser unterstützt als durch die anderen WBT-Varianten.
➔ **ARCS > A, R, C, S, 0**

Hypothese 8f: Die Studierenden fühlen sich durch die Einzelkomponenten-WBTs hinsichtlich aller vier ARCS-Aspekte (Gesamt-ARCS-Skala) signifikant besser unterstützt als durch das 0-WBT.
➔ **A, R, C, S > 0**

Inwieweit sich die Studierenden durch die verschiedenen E-Learning-Varianten in ihrem Lernprozess unterstützt fühlten, wurde durch den ARCS-Fragebogen erhoben. Im Folgenden werden zunächst die Unterskalen „Aufmerksamkeit“, „Relevanz“, „Erfolgszuversicht“ und „Zufriedenheit“ betrachtet, um zu überprüfen, ob sich die Studierenden, die die Einzelkomponenten-WBTs nutzten, insbesondere hinsichtlich der jeweils geförderten Komponente unterstützt gefühlt haben (z.B. Aufmerksamkeits(A)-WBT führt zu hoher Bewertung auf der Skala „Aufmerksamkeit“).

Zu guter Letzt werden die Vergleiche zur Gesamtskala vorgestellt. Dabei wird jeweils auch ein Blick darauf geworfen, inwieweit sich die Beurteilungen der Studierenden zwischen den drei Themen unterscheiden.

Da nach der Analyse des Fragebogens für die vier Unterskalen unterschiedlich viele Items resultierten (siehe Abschnitt „6.3.2 Erhobene Variablen“), werden für jede Skala jeweils die mit der Anzahl ihrer Items gewichteten Mittelwerte berichtet. Für die Berechnung der Gesamtskala „ARCS“ wird auf diese gewichteten Mittelwerte zurückgegriffen. So resultiert

für die Unterskalen ein Wertebereich von 1 bis 5 Punkten, für die Gesamtskala ein Wertebereich von 4 bis 20 Punkten.

7.2.1 Skala „Aufmerksamkeit“

Betrachtet man die Mittelwerte in Tabelle 78 (zur Verdeutlichung auch noch mal abgebildet in Abbildung 44), zeigt sich, dass die sieben E-Learning-Varianten recht vergleichbar hinsichtlich der Skala „Aufmerksamkeit“ beurteilt wurden.

Tabelle 78
Deskriptive Statistik zur gewichteten Skala "Aufmerksamkeit" für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Aufmerksamkeit	Konditionierung			Unterrichten			Diagnostik			Gesamt		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
0-WBT	64	3.71	0.73	74	3.11	0.83	50	3.12	1.00	188	3.32	0.89
A-WBT	64	4.07	0.80	81	3.35	0.81	78	3.33	0.66	223	3.55	0.83
R-WBT	68	3.88	0.68	66	3.22	0.90	58	3.04	0.89	192	3.40	0.90
C-WBT	56	3.85	0.76	64	3.21	0.75	51	3.29	0.71	171	3.44	0.79
S-WBT	80	3.94	0.67	61	3.39	0.76	53	3.24	0.77	194	3.58	0.79
ARCS-WBT	57	**4.15**	0.69	68	**3.59**	0.75	73	**3.52**	0.76	198	**3.73**	0.79
E-Lecture	200	3.89	0.70	242	3.11	0.75	203	3.02	0.79	645	3.32	0.84
Gesamt	**589**	**3.92**	**0.72**	**656**	**3.24**	**0.80**	**566**	**3.18**	**0.81**	**1811**	**3.44**	**0.84**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

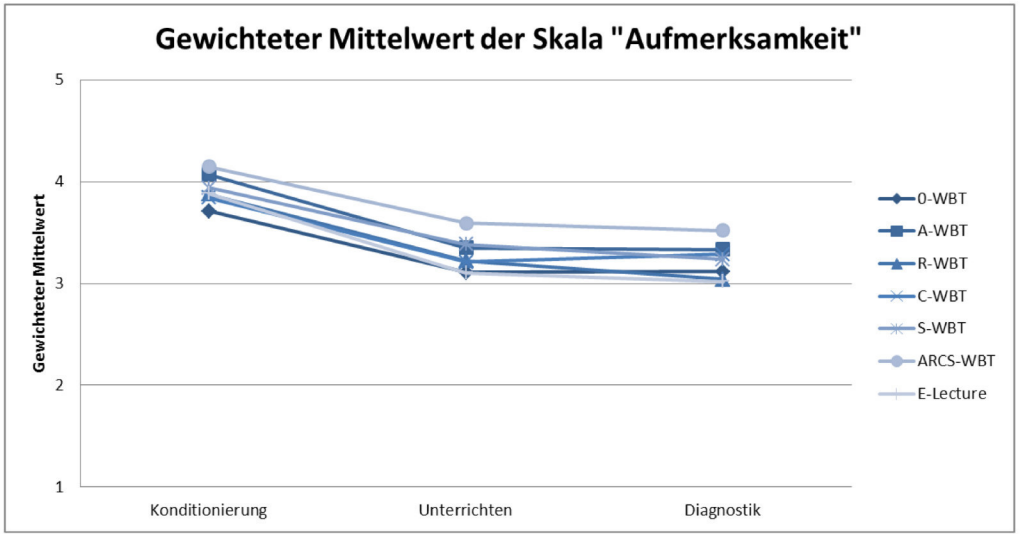

Abbildung 44. Gewichtete Mittelwerte der Skala "Aufmerksamkeit" für die E-Learning-Gruppen zu allen drei Themen

Über alle drei Erhebungen erreichen die E-Learning-Varianten einen Mittelwert zwischen 3.0 und 4.2 – es resultiert ein Gesamtmittelwert über alle Erhebungen und Varianten von $M = 3.4$.

Augenscheinlich erreichen das ARCS- sowie das Aufmerksamkeits(A)-WBT jedoch wie erwartet etwas höhere Werte als die anderen Varianten. Zudem fällt auf, dass die Mittelwerte über alle sieben Gruppen insbesondere von der ersten zur zweiten Erhebung deutlich sinken. Diese Beobachtungen werden im Folgenden inferenzstatistisch überprüft.

Tabelle 79
Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Aufmerksamkeit" zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Aufmerksamkeit	**ANOVA/Brown-Forsythe-Test**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	2.71	446,6	.01*	.04	ARCS > 0
Zufallsauswahl 2	2.55	448,6	.02*	.03	ARCS > 0
Zufallsauswahl 3	2.54	452,6	.02*	.03	ARCS > 0
Unterrichten					
Zufallsauswahl 1	4.00	485,6	<.01*	.05	ARCS > 0, EL
Zufallsauswahl 2	3.22	488,6	<.01*	.04	ARCS > 0, EL
Zufallsauswahl 3	3.09	490,6	<.01*	.04	ARCS > 0, EL
Diagnostik					
Zufallsauswahl 1	3.42	363.38,6	<.01*	.05	ARCS > EL, R
Zufallsauswahl 2	4.54	352.26,6	<.01*	.06	ARCS > EL, R; A > EL
Zufallsauswahl 3	2.56	353.30,6	.02*	.04	ARCS > R
Gesamt					
Zufallsauswahl 1	7.53	1368,6	<.01*	.03	ARCS > 0, R, C, EL; A, S > EL
Zufallsauswahl 2	6.82	1375,6	<.01*	.03	ARCS > 0, R, C, EL; A, S > EL
Zufallsauswahl 3	5.62	1379,6	<.01*	.02	ARCS > 0, R, C, EL; S > 0

Anmerkung. F = Teststatistik der Varianzanalyse bzw. des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Fasst man die Ergebnisse der einfaktoriellen Varianzanalysen/Brown-Forsythe-Tests zu den drei Erhebungen zusammen, ergeben sie (siehe Tabelle 79) einen Vorteil für die ARCS-WBTs. Es wird hinsichtlich seiner aufmerksamkeitsfördernden Wirkung in der Mehrheit der Analysen gegenüber den 0-WBTs und E-Lectures signifikant besser beurteilt. Signifikant bessere Ergebnisse zeigen sich darüber hinaus gegenüber den R-WBTs. Diese treten jedoch lediglich für das Thema „Diagnostik" auf. Die Effektstärken dieser Mittelwertsunterschiede sind als gering bis mittelstark einzustufen.

Innerhalb der Auswertungen über alle Erhebungen hinweg ergeben sich darüber hinaus mehrheitlich signifikante Vorteile der ARCS-WBTs gegenüber den R- und C-WBTs. Auch die A- und S-WBTs erreichen hier ein signifikant besseres Ergebnis als die E-Lectures.

Trotz dieser tendenziell hypothesenkonformen Ergebnisse ist die sehr streng formulierte Hypothese 8a demnach zu verwerfen.

Tabelle 80
Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Aufmerksamkeit" über alle Gruppen zwischen den Themen

Aufmerksamkeit	**Deskriptive Statistik**		**Abhängige Varianzanalyse**				**Posthoc**
n = 304	***M***	***SD***	***F***	***df***	***p***	**Partielles *Eta²***	
Konditionierung	**4.01**	0.68					
Unterrichten	3.26	0.85	123.55	1.97	<.01*	.29	Kon > Unt, Dia
Diagnostik	3.24	0.77					

Anmerkung. n = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, Partielles *Eta²* = Effektstärke

Über die sieben Gruppen hinweg wird die Aufmerksamkeitsförderung durch die E-Learning-Anwendungen zum Thema „Konditionierung" von den Studierenden deutlich

besser beurteilt als zu den Themen „Unterrichten“ und „Diagnostik“ (siehe Tabelle 80). Die Effektstärke dieses Unterschiedes ist als stark zu beurteilen.

7.2.2 Skala „Relevanz“

Abbildung 45 (siehe auch Tabelle 81) stellt die gewichteten Mittelwerte der Skala „Relevanz“ für alle sieben E-Learning-Varianten dar.

Tabelle 81
Deskriptive Statistik zur gewichteten Skala "Relevanz" für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Relevanz	Konditionierung			Unterrichten			Diagnostik			Gesamt		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
0-WBT	64	3.69	0.67	74	3.79	0.62	50	3.49	0.83	188	3.67	0.70
A-WBT	64	3.84	0.56	81	3.76	0.71	78	3.60	0.64	223	3.73	0.65
R-WBT	68	**4.13**	0.58	66	**3.81**	0.56	58	3.66	0.77	192	3.88	0.66
C-WBT	56	3.75	0.67	64	3.72	0.53	51	3.62	0.64	171	3.70	0.61
S-WBT	80	3.91	0.45	61	3.74	0.58	53	3.59	0.68	194	3.77	0.58
ARCS-WBT	57	4.08	0.50	68	**3.81**	0.64	73	**3.82**	0.62	198	**3.89**	0.60
E-Lecture	200	3.91	0.58	242	3.70	0.63	203	3.49	0.70	645	3.70	0.65
Gesamt	**589**	**3.90**	**0.58**	**656**	**3.74**	**0.62**	**566**	**3.59**	**0.70**	**1811**	**3.75**	**0.65**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

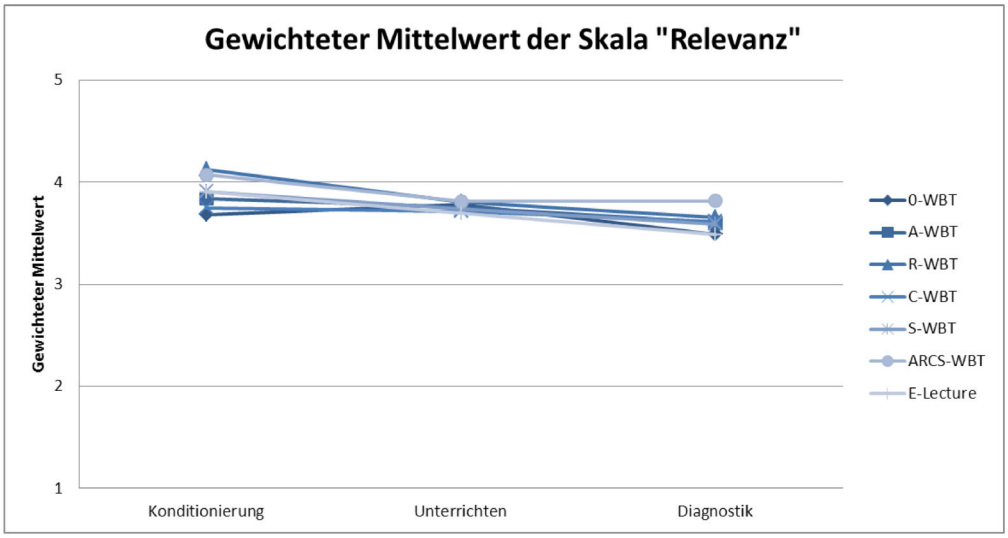

Abbildung 45. Gewichtete Mittelwerte der Skala "Relevanz" für die E-Learning-Gruppen zu allen drei Themen

Es ist deutlich zu sehen, dass sich die Bewertungen zwischen den E-Learning-Anwendungen gering unterscheiden. Sie liegen alle zwischen 3.5 und 4.2. Über alle Erhebungen und Varianten hinweg resultiert ein Mittelwert von *M* = 3.75. Eine E-Learning-Anwendung, die hinsichtlich der Relevanzförderung besonders gut oder schlecht abgeschnitten hat, ist hier augenscheinlich nicht auszumachen. Die minimal höchsten Mittelwerte finden sich allerdings hypothesenkonform für die R- und ARCS-WBTs.

Es ist erneut ein Abwärtstrend von der ersten zur dritten Erhebung festzustellen. Wenn dieser auch nicht so deutlich ausfällt wie im Rahmen der Skala „Aufmerksamkeit“.

Tabelle 82
Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Relevanz" zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Relevanz	**ANOVA**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	5.08	446,6	<.01*	.06	R, ARCS > 0; R > C
Zufallsauswahl 2	4.77	448,6	<.01*	.06	R, ARCS > 0; R > C
Zufallsauswahl 3	5.12	452,6	<.01*	.06	R, ARCS > 0; R > C
Unterrichten					
Zufallsauswahl 1	0.76	485,6	.60	-	
Zufallsauswahl 2	0.22	488,6	.97	-	-
Zufallsauswahl 3	0.33	490,6	.92	-	
Diagnostik					
Zufallsauswahl 1	1.78	423,6	.10	-	
Zufallsauswahl 2	2.06	425,6	.06	-	-
Zufallsauswahl 3	1.46	423,6	.19	-	
Gesamt					
Zufallsauswahl 1	3.97	1368,6	<.01*	.02	ARCS, R > EL, 0
Zufallsauswahl 2	3.86	1375,6	<.01*	.02	ARCS, R > EL, 0
Zufallsauswahl 3	3.48	1379,6	<.01*	.02	ARCS, R > 0

Anmerkung. F = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Die einfaktoriellen Varianzanalysen (siehe Tabelle 82) ergaben für die erste Erhebung sowie die Gesamtauswertung signifikante Effekte zwischen den E-Learning-Anwendungen. Für das Thema „Konditionierung" (mittelstarker Effekt) werden das Relevanz(R)- und das ARCS-WBT signifikant relevanzfördernder eingeschätzt als das thematisch gleiche 0-WBT. Das Relevanz(R)-WBT wird zudem als relevanzfördernder beurteilt als das Erfolgszuversichts(C)-WBT. Ähnliches ergab sich für die Gesamtauswertung (mit geringerem Effekt): Hier erreichten die ARCS- und R-WBTs signifikant bessere Ergebnisse als die 0-WBTs und die E-Lectures.

Die sehr eng gefasste Hypothese 8b muss dennoch verworfen werden. Durch den Einsatz der Relevanzstrategien konnte das Relevanzgefühl der Studierenden nicht signifikant besser unterstützt werden als durch die anderen Einzelkomponenten-WBTs. Es ist allerdings erfreulich, dass sich die ARCS- und R-WBTs zumindest gegenüber den 0-WBTs und E-Lectures abheben konnten.

Der Vergleich der drei Erhebungen über alle sieben Gruppen hinweg mittels abhängiger Varianzanalyse ergab erneut einen signifikanten mittelstarken Effekt (siehe Tabelle 83).

Tabelle 83
Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Relevanz" über alle Gruppen zwischen den Themen

Relevanz	**Deskriptive Statistik**		**Abhängige Varianzanalyse**				**Posthoc**
n = 304	***M***	***SD***	***F***	***df***	***p***	**Partielles *Eta²***	
Konditionierung	**3.92**	0.61					
Unterrichten	3.79	0.62	28.29	2.00	<.01*	.09	Kon > Unt > Dia
Diagnostik	3.62	0.64					

Anmerkung. n = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, Partielles *Eta²* = Effektstärke

Dabei werden die E-Learning-Anwendungen zum Thema „Konditionierung“ von den Studierenden als relevanzfördernder wahrgenommen als die Anwendungen zum Thema „Unterrichten“ und „Diagnostik“. Auch die Anwendungen zum Thema „Unterrichten“ fördern aus Sicht der Studierenden die Relevanzwahrnehmung mehr als die Anwendungen zum Thema „Diagnostik“.

7.2.3 Skala „Erfolgszuversicht“

Die Beurteilung der Erfolgszuversichtsförderung liegt für alle sieben Varianten insgesamt zwischen 3.0 und 4.0 (siehe Tabelle 84). Es resultiert ein Gesamtmittelwert über alle Erhebungen und Varianten von $M = 3.4$.

Tabelle 84
Deskriptive Statistik zur gewichteten Skala "Erfolgszuversicht" für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Erfolgs-zuversicht	Konditionierung			Unterrichten			Diagnostik			Gesamt		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
0-WBT	64	3.22	0.59	74	3.10	0.65	50	3.02	0.72	188	3.12	0.65
A-WBT	64	3.32	0.54	81	3.18	0.61	78	3.14	0.61	223	3.21	0.60
R-WBT	68	3.60	0.59	66	3.26	0.61	58	3.23	0.76	192	3.37	0.67
C-WBT	56	3.84	0.68	64	3.59	0.52	51	3.58	0.66	171	3.67	0.63
S-WBT	80	3.74	0.53	61	3.49	0.58	53	3.46	0.66	194	3.58	0.59
ARCS-WBT	57	**3.93**	0.60	68	**3.84**	0.64	73	**3.76**	0.62	198	**3.83**	0.62
E-Lecture	200	3.44	0.61	242	3.25	0.61	203	3.03	0.58	645	3.24	0.62
Gesamt	**589**	**3.55**	**0.63**	**656**	**3.34**	**0.64**	**566**	**3.25**	**0.69**	**1811**	**3.38**	**0.66**

Anmerkung. n = Stichprobengröße, M = Mittelwert, SD = Standardabweichung

Betrachtet man Abbildung 46, zeigt sich, dass die ARCS-WBTs augenscheinlich zu allen drei Themenbereichen als am meisten erfolgszuversichtsfördernd wahrgenommen werden.

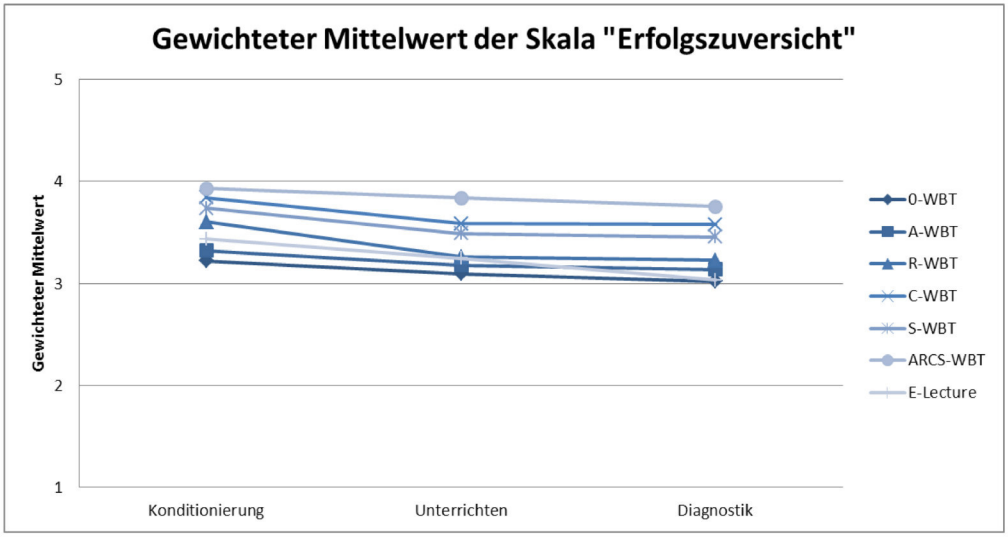

Abbildung 46. Gewichtete Mittelwerte der Skala "Erfolgszuversicht" für die E-Learning-Gruppen zu allen drei Themen

Direkt darunter folgen die Erfolgszuversichts(C)- und Zufriedenheits(S)-WBTs. Am wenigsten förderlich werden die 0- und Aufmerksamkeits(A)-WBTs sowie die E-Lectures

von den Studierenden wahrgenommen. Auch in Bezug auf die Bewertung der Erfolgszuversichtsförderung ist ein leichter Abwärtstrend von der ersten zur dritten Erhebung zu beobachten.

Diese Beobachtungen bestätigen sich, betrachtet man die Ergebnisse der Varianzanalysen (siehe Tabelle 85). Zu allen drei Messzeitpunkten sowie insgesamt ergeben sich signifikante Mittelwertsunterschiede zwischen den Gruppen mit starken Effektstärken.

Tabelle 85
Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Erfolgszuversicht" zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Erfolgszuversicht	**ANOVA**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	12.99	446,6	<.01*	.15	ARCS, C, S, R > 0; ARCS, C, S > A; ARCS, C > EL; ARCS > R
Zufallsauswahl 2	13.76	448,6	<.01*	.16	ARCS, C, S, R > 0; ARCS, C, S > A, EL; ARCS > R
Zufallsauswahl 3	12.10	452,6	<.01*	.14	ARCS, C, S, R > 0; ARCS, C, S > A; ARCS, C > EL
Unterrichten					
Zufallsauswahl 1	14.20	485,6	<.01*	.15	ARCS > EL, 0, A, R, S; C > EL, 0, A; S > EL, 0
Zufallsauswahl 2	13.14	488,6	<.01*	.14	ARCS > 0, A, R, EL, S; C > 0, A, R, EL; S > 0
Zufallsauswahl 3	13.17	490,6	<.01*	.14	ARCS > 0, A, R, EL, S; C > 0, A, R; S > 0
Diagnostik					
Zufallsauswahl 1	12.86	423,6	<.01*	.15	ARCS > EL, 0, A, R; C > EL, 0, A; S > EL, 0
Zufallsauswahl 2	13.76	425,6	<.01*	.16	ARCS > EL, 0, A, R; C > EL, 0, A; S > EL, 0
Zufallsauswahl 3	10.71	423,6	<.01*	.13	ARCS > 0, A, EL, R;C > 0, A, EL; S > 0
Gesamt					
Zufallsauswahl 1	37.59	1368,6	<.01*	.14	ARCS, C, S > 0, A, R, EL; R > 0, ARCS > S
Zufallsauswahl 2	37.81	1375,6	<.01*	.14	ARCS, C, S > 0, A, R, EL; R > 0, ARCS > S
Zufallsauswahl 3	33.79	1379,6	<.01*	.13	ARCS, C, S > 0, A, R, EL; R, EL > 0, ARCS > S

Anmerkung. F = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Dabei zeigen die Posthoc-Vergleiche, dass die ARCS-, Erfolgszuversichts(C)- und Zufriedenheits(S)-WBTs generell von den Studierenden als erfolgszuversichtsfördernder wahrgenommen werden, während die 0- und die Aufmerksamkeits(A)-WBTs sowie die E-Lectures als generell weniger förderlich eingeschätzt werden. Die Relevanz(R)-WBTs werden unterschiedlich beurteilt: zum Thema „Konditionierung" erscheint es im Gegensatz zum 0-WBT erfolgszuversichtsfördernder, zu den Themen „Unterrichten" und „Diagnostik" reiht es sich allerdings in die Reihe der weniger förderlichen E-Learning-Anwendungen ein. Für die Gesamtauswertung erreichen die R-WBTs dann auch eine schlechtere Einstufung als die ARCS-, C- und S-WBTs. Sie können sich allerdings mehrheitlich noch gegenüber den 0-WBTs behaupten. Darüber hinaus ergibt sich in der Gesamtauswertung ein signifikanter Vorteil der ARCS-WBTs gegenüber den S-WBTs.

In der Mehrheit der Analysen erreichen die ARCS-WBTs folglich eine bessere Beurteilung als die 0-, A-, R-WBTs und die E-Lectures, die C-WBTs schneiden besser ab als die 0-, A-WBTs und die E-Lecture und die S-WBTs werden förderlicher wahrgenommen als die 0-WBTs. Obwohl man auch hier – vergleichbar zu den Skalen Aufmerksamkeit und Relevanz

– von tendenziell hypothesenkonformen Ergebnissen sprechen kann, muss die sehr streng formulierte Hypothese 8c damit ebenfalls verworfen werden.

Tabelle 86 gibt erneut die Ergebnisse der abhängigen Varianzanalyse zum Vergleich der Skala „Erfolgszuversicht“ zwischen den drei Erhebungen wieder.

Tabelle 86
Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Erfolgszuversicht“ über alle Gruppen zwischen den Themen

Erfolgszuversicht	**Deskriptive Statistik**		**Abhängige Varianzanalyse**				**Posthoc**
n = 304	***M***	***SD***	***F***	***df***	***p***	**Partielles *Eta²***	
Konditionierung	**3.57**	0.63					
Unterrichten	3.36	0.67	22.22	1.98	<.01*	.07	Kon > Unt, Dia
Diagnostik	3.28	0.65					

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, Partielles *Eta²* = Effektstärke

Auch hier ergibt sich ein mittelstarker signifikanter Effekt zwischen den drei Erhebungen. Dabei werden die E-Learning-Anwendungen zum Thema „Konditionierung“ von den Studierenden erfolgszuversichtsfördernder beurteilt als die Anwendungen zu den Themen „Unterrichten“ und „Diagnostik“.

7.2.4 Skala „Zufriedenheit“

Im Gegensatz zu den drei ersten Skalen variieren die Beurteilungen zur Skala „Zufriedenheit“ sehr deutlich zwischen 2.5 und 4.0 (siehe Abbildung 47 bzw. Tabelle 87). Über alle Erhebungen und Anwendungen hinweg resultiert ein Gesamtmittelwert von *M* = 3.1.

Tabelle 87
Deskriptive Statistik zur gewichteten Skala "Zufriedenheit" für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Zufrieden-	**Konditionierung**			**Unterrichten**			**Diagnostik**			**Gesamt**		
heit	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
0-WBT	64	2.92	0.69	74	2.54	0.74	50	2.51	0.85	188	2.66	0.78
A-WBT	64	3.00	0.72	81	2.83	0.70	78	2.90	0.78	223	2.90	0.73
R-WBT	68	3.17	0.77	66	2.84	0.70	58	2.85	0.78	192	2.96	0.76
C-WBT	56	3.57	0.68	64	3.44	0.54	51	3.43	0.63	171	3.48	0.61
S-WBT	80	3.92	0.60	61	3.72	0.62	53	3.46	0.69	194	3.73	0.66
ARCS-WBT	57	**3.90**	0.57	68	**3.84**	0.65	73	**3.78**	0.66	198	**3.84**	0.63
E-Lecture	200	3.01	0.77	242	2.78	0.74	203	2.59	0.75	645	2.79	0.77
Gesamt	**589**	**3.28**	**0.80**	**656**	**3.03**	**0.81**	**566**	**2.96**	**0.86**	**1811**	**3.09**	**0.84**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Hier ist erneut eine Überlegenheit des ARCS-WBTs zu beobachten. Das Zufriedenheits(S)-WBT liegt zwar im Rahmen der ersten Erhebung gleichauf. Die Beurteilungen der Zufriedenheits(S)-WBTs verschlechtern sich allerdings über die beiden folgenden Erhebungen hinweg deutlich stärker als die der ARCS-WBTs.

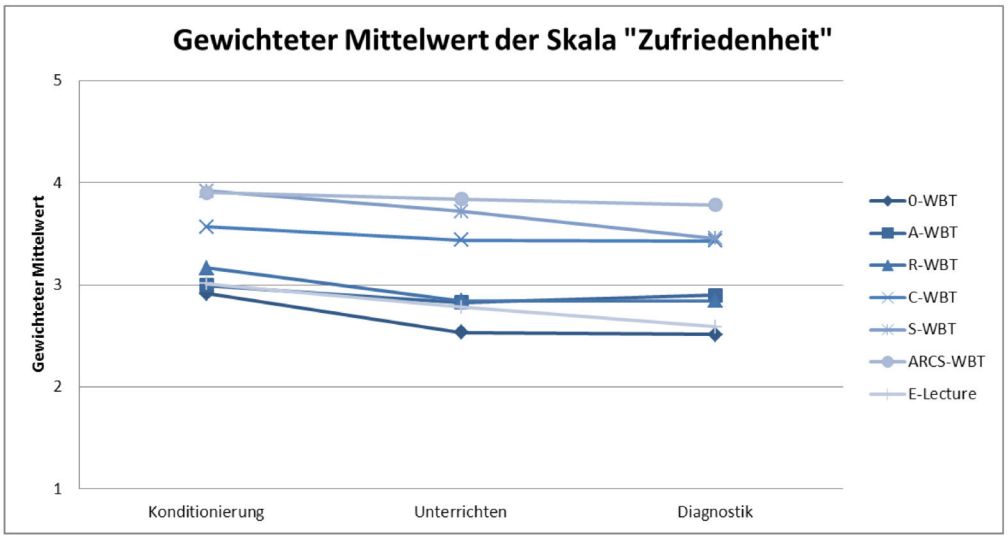

Abbildung 47. Gewichtete Mittelwerte der Skala "Zufriedenheit" für die E-Learning-Gruppen zu allen drei Themen

Auch die Erfolgszuversichts(C)-WBTs erhalten gute und stabile Beurteilungen hinsichtlich der Zufriedenheitsförderung. Am schlechtesten beurteilt werden die 0-WBTs. Überdies erzielen die E-Lectures, Aufmerksamkeits(A)- und Relevanz(R)-WBTs augenscheinlich schlechtere Beurteilungen auf der Skala „Zufriedenheit" durch die Studierenden.

Hinsichtlich der Zufriedenheitsförderung durch die E-Learning-Anwendungen ist ebenfalls erneut eine leicht abfallende Tendenz von der ersten zur letzten Erhebung zu beobachten. Diese augenscheinlichen Beobachtungen können durch einfaktorielle Varianzanalysen bestätigt werden (siehe Tabelle 88).

Tabelle 88
Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Zufriedenheit" zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Zufriedenheit	**ANOVA/Brown-Forsythe-Test**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	26.49	446,6	<.01*	.26	ARCS, S, C > 0, A, EL, R
Zufallsauswahl 2	27.73	448,6	<.01*	.27	ARCS, S, C > 0, A, EL, R
Zufallsauswahl 3	26.02	452,6	<.01*	.26	ARCS, S, C > 0, A, EL, R
Unterrichten					
Zufallsauswahl 1	42.92	485,6	<.01*	.35	ARCS > 0, A, R, EL, C; S, C > 0, A, R, EL
Zufallsauswahl 2	39.31	475.82,6	<.01*	.32	ARCS > 0, A, R, EL, C; S, C > 0, A, R, EL
Zufallsauswahl 3	40.63	490,6	<.01*	.33	ARCS > 0, A, R, EL, C; S, C > 0, A, R, EL
Diagnostik					
Zufallsauswahl 1	28.30	423,6	<.01*	.29	ARCS, S, C > 0, EL, R, A; A > EL
Zufallsauswahl 2	29.94	425,6	<.01*	.30	ARCS, S, C > 0, EL, R, A; A > EL
Zufallsauswahl 3	23.63	423,6	<.01*	.25	ARCS, S, C > 0, EL, R, A
Gesamt					
Zufallsauswahl 1	93.40	1327.95,6	<.01*	.29	ARCS, C, S > 0, A, R, EL; ARCS, S > C; R > EL, 0; A > EL
Zufallsauswahl 2	92.33	1338.80,6	<.01*	.29	ARCS, C, S > 0, A, R, EL; ARCS, S > C; A, R > 0
Zufallsauswahl 3	85.52	1342.65,6	<.01*	.27	ARCS, C, S > 0, A, R, EL; ARCS, S > C; EL, A, R > 0

Anmerkung. F = Teststatistik der Varianzanalyse bzw. des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Zu allen drei Themenbereichen sowie insgesamt sind es jeweils die ARCS-, Erfolgszuversichts(C)- und Zufriedenheits(S)-WBTs, die höhere Werte auf der Skala

„Zufriedenheit“ erreichen als die 0-, Aufmerksamkeits(A)- und Relevanz(R)-WBTs sowie die E-Lectures. Für die Gesamtauswertung ergeben sich darüber hinaus eine signifikant bessere Beurteilung der ARCS- und S-WBTs gegenüber den C-WBTs sowie ein signifikanter Vorteil der A- und R-WBTs gegenüber den 0-WBTs. Dabei sind die Effekte der Mittelwertsunterschiede jeweils als sehr stark einzustufen.

Hypothese 8d kann nach den Gesamtergebnissen folglich beibehalten werden: Insgesamt fördern die ARCS- und S-WBTs die Zufriedenheit der Studierenden signifikant besser als alle anderen Einzelkomponenten-WBTs sowie das 0-WBT.

Tabelle 89
Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Skala „Zufriedenheit“ über alle Gruppen zwischen den Themen

Zufriedenheit	**Deskriptive Statistik**		**Abhängige Varianzanalyse**				**Posthoc**
n = 304	***M***	***SD***	***F***	***df***	***p***	**Partielles *Eta²***	
Konditionierung	**3.30**	0.84					
Unterrichten	3.05	0.85	13.38	1.98	<.01*	.04	Kon > Unt, Dia
Diagnostik	2.98	0.84					

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, Partielles *Eta²* = Effektstärke

Die Vergleiche der Beurteilungen zwischen den drei Themenbereichen über alle E-Learning-Anwendungen hinweg ergeben erneut einen signifikanten – wenn diesmal auch nur geringen – Effekt. Auch hier werden die Anwendungen zum Thema „Konditionierung“ als zufriedenheitsfördernder eingeschätzt als die Anwendungen zu den Themen „Unterrichten“ und „Diagnostik“.

7.2.5 Gesamtskala „ARCS“

Um eine Aussage über die von den E-Learning-Anwendungen ausgehende Lernunterstützung insgesamt zu machen, werden im Folgenden die Ergebnisse der Erhebungen des ARCS-Fragebogens insgesamt besprochen. Dieser hatte einen Wertebereich zwischen 28 und 140 Punkten. Aufgrund der unterschiedlichen Itemzahlen der einzelnen Unterskalen wurde allerdings ein gewichtetes Gesamtergebnis berechnet, welches sich nun zwischen 4 und 20 Punkten bewegt.

Betrachtet man die Ergebnisse zur Gesamtskala (siehe Tabelle 90 sowie Abbildung 48), zeigt sich, dass die ARCS-WBTs augenscheinlich über alle drei Themen hinweg sowie insgesamt (mit einem Skalenwert von M = 15.3) von den Studierenden als am meisten lernunterstützend wahrgenommen wird. Die geringsten Werte erreichen die 0-WBTs (M = 12.8) und die E-Lectures (M = 13.1). Die Einzelkomponenten-WBTs liegen mit Skalenmittelwerten von 13.4 bis 14.7 zwischen diesen Extremen. Es ist erneut eine abfallende Tendenz der Bewertungen von der ersten zur letzten Erhebung zu beobachten.

Tabelle 90
Deskriptive Statistik zur gewichteten Gesamt-ARCS-Skala für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Gesamt ARCS	Konditionierung			Unterrichten			Diagnostik			Gesamt		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
0-WBT	64	13.54	1.99	74	12.53	2.34	50	12.15	3.03	188	12.77	2.49
A-WBT	64	14.23	2.03	81	13.12	2.48	78	12.98	2.29	223	13.39	2.35
R-WBT	68	14.78	2.05	66	13.14	2.32	58	12.78	2.83	192	13.61	2.55
C-WBT	56	15.01	2.42	64	13.96	1.93	51	13.92	2.35	171	14.29	2.27
S-WBT	80	15.51	1.84	61	14.34	2.16	53	13.75	2.51	194	14.66	2.26
ARCS-WBT	57	**16.06**	2.10	68	**15.09**	2.37	73	**14.88**	2.38	198	**15.29**	2.34
E-Lecture	200	14.24	2.09	242	12.83	2.24	203	12.13	2.36	645	13.05	2.39
Gesamt	**589**	**14.65**	**2.18**	**656**	**13.35**	**2.40**	**566**	**12.98**	**2.65**	**1811**	**13.66**	**2.51**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

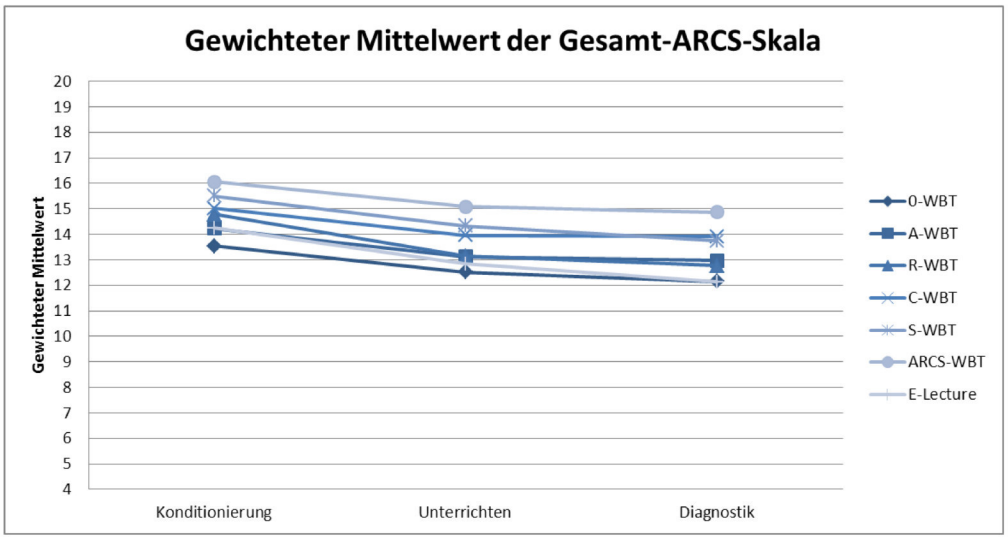

Abbildung 48. Gewichtete Mittelwerte zur gewichteten Gesamt-ARCS-Skala für die E-Learning-Gruppen zu allen drei Themen

Die einfaktoriellen Varianzanalysen ergaben zu allen drei Erhebungen sowie insgesamt signifikante Unterschiede mit mittleren bis starken Effekten (siehe Tabelle 91).

Dabei schneiden die ARCS-WBTs durchgängig signifikant besser ab als die 0-, Aufmerksamkeits(A)- und Relevanz(R)-WBTs sowie die E-Lectures. Auch die Zufriedenheits(S)- und Erfolgszuversichts(C)-WBTs erhalten eine gute Bewertung auf der Gesamt-ARCS-Skala. Die Zufriedenheits(S)-WBTs erreichen zu allen drei Themenbereichen eine signifikant höhere Einschätzung als die 0-WBTs und E-Lectures. Zu den Themenbereichen „Konditionierung“ und „Unterrichten“ ergibt sich ebenfalls eine höhere Bewertung als für die Aufmerksamkeits(A)-WBTs. Die Erfolgszuversicht(C)-WBTs schneiden durchgängig besser ab als die 0-WBTs. Für den Themenbereich „Diagnostik“ ergibt sich zudem ein signifikanter Vorteil gegenüber der E-Lecture.

Tabelle 91
Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Gesamt-ARCS-Skala zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Gesamt ARCS	**ANOVA**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	10.99	446,6	<.01*	.13	ARCS > 0, A, R, EL; S > 0, A, EL; R, C > 0
Zufallsauswahl 2	11.09	448,6	<.01*	.13	ARCS > 0, A, R, EL; S > 0, A, EL; R, C > 0
Zufallsauswahl 3	10.45	452,6	<.01*	.12	ARCS > 0, A, R, EL; S > 0, A, EL; R, C > 0
Unterrichten					
Zufallsauswahl 1	12.65	485,6	<.01*	.14	ARCS > EL, 0, A, R; S > EL, 0, A; C > EL, 0
Zufallsauswahl 2	10.97	488,6	<.01*	.12	ARCS > 0, A, EL, R; S > 0, A, EL; C > 0
Zufallsauswahl 3	10.95	490,6	<.01*	.12	ARCS > 0, A, EL, R; S > 0, A, EL; C > 0
Diagnostik					
Zufallsauswahl 1	10.69	423,6	<.01*	.13	ARCS > EL, 0, R, A; C, S > EL, 0
Zufallsauswahl 2	12.15	425,6	<.01*	.15	ARCS > EL, 0, R, A; C, S > EL, 0
Zufallsauswahl 3	8.73	423,6	<.01*	.11	ARCS > 0, EL, R, A; C, S > 0
Gesamt					
Zufallsauswahl 1	30.80	1368,6	<.01*	.12	ARCS > 0, A, R, C, EL; S > 0, A, R, EL; C > 0, A, EL; R > 0, EL
Zufallsauswahl 2	30.41	1375,6	<.01*	.12	ARCS > 0, A, R, C, EL; S > 0, A, R, EL; C > 0, A, EL; R > 0
Zufallsauswahl 3	26.91	1379,6	<.01*	.11	ARCS > 0, A, R, C, EL; S > 0, A, R, EL; C > 0, A, EL; R > 0

Anmerkung. F = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Über alle Erhebungen hinweg zeigen sich zusätzliche Signifikanzen. Hier erreichen nun die ARCS-WBTs signifikant bessere Einstufungen als die 0-, A-, R-, und C-WBTs sowie die E-Lectures, die S-WBTs werden besser eingestuft als die 0-, A-, R-WBTs und die E-Lectures, die C-WBTs immerhin noch besser als die 0-, A-WBTs und die E-Lectures und die R-WBTs erreichen ein besseres Ergebnis als die 0-WBTs.

Die Hypothesen 8e und 8f werden durch die Ergebnisse fast vollständig bestätigt: Die ARCS-WBTs erreichen ein signifikant besseres Ergebnis als vier der fünf Einzelkomponenten-WBTs, die 0-WBTs ein signifikant schlechteres Ergebnis als drei der vier Einzelkomponenten-WBTs. Lediglich gegenüber den S-WBTs können sich die ARCS-WBT nicht abheben und die 0-WBTs erreichen zumindest ein vergleichbar gutes Ergebnis wie die A-WBTs. Aus diesen Gründen müssen die erneut sehr streng formulierten Hypothesen 8e und 8f dennoch verworfen werden.

Die abhängige Varianzanalyse, die durchgeführt wurde, um das Abschneiden auf der Gesamt-ARCS-Skala über die drei Themenbereiche hinweg zu vergleichen, ergibt einen signifikanten Unterschied mit starkem Effekt (siehe Tabelle 92).

Tabelle 92
Ergebnisse der Mittelwertsanalysen zum Vergleich der gewichteten Gesamt-ARCS-Skala über alle Gruppen zwischen den Themen

Gesamt ARCS	**Deskriptive Statistik**		**Abhängige Varianzanalyse**				**Posthoc**
n = 304	***M***	***SD***	***F***	***df***	***p***	**Partielles *Eta²***	
Konditionierung	**14.80**	2.24					
Unterrichten	13.46	2.49	55.30	1.97	<.01*	.15	Kon > Unt, Dia
Diagnostik	13.11	2.46					

Anmerkung. n = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, Partielles *Eta²* = Effektstärke

Die E-Learning-Anwendungen zum Thema „Konditionierung“ werden signifikant lernunterstützender wahrgenommen als die Anwendungen zu den Themen „Unterrichten“ und „Diagnostik“.

7.2.6 Zusammenfassung und Diskussion

Zwischen den E-Learning-Varianten ergaben sich zumeist keine derart deutlichen Unterschiede, dass alle Hypothesen hätten beibehalten werden können. Dies lag allerdings nicht nur daran, dass die Ergebnisse durchgängig sehr heterogen waren, sondern ist hauptsächlich darin begründet, dass die Hypothesen äußerst streng formuliert worden waren. Dadurch ist es schwierig, einzelne Hypothesen beizubehalten. Natürlich wäre es möglich gewesen, anstatt der hier formulierten Makrohypothesen viele Einzelhypothesen zum Vergleich einzelner E-Learning-Varianten aufzustellen. Dadurch hätten sicherlich einige Hypothesen beibehalten werden können. Dies hätte für die Interpretation der Ergebnisse jedoch zu keinem Mehrwert geführt. Deshalb wurde die Aufstellung von Makrohypothesen präferiert. Die gefundenen Ergebnisse zur ARCS-Skala werden im Folgenden ausführlich diskutiert.

Hinsichtlich der Skala **„Aufmerksamkeit“** konnten sich in der Mehrzahl der Analysen zu den einzelnen Themen lediglich die ARCS-WBTs gegenüber den 0-WBTs und E-Lectures durchsetzen. Über die drei Themen hinweg erreichten die ARCS-WBTs zudem eine signifikant bessere Beurteilung als die R- und C-WBTs. Hier konnten sich die A- und S-WBTs ebenso gegenüber den E-Lectures abheben.
Die ARCS-WBTs erreichten demnach eine bessere aufmerksamkeitsfördernde Wirkung als die inhaltsgleichen 0-, R- und C-WBTs sowie die E-Lectures. Dies spricht gegen eine Überfrachtung der ARCS-WBTs durch zu viele eingesetzte Strategien, da die Studierenden sich durch die ARCS-WBTs in der Aufmerksamkeitserhaltung gut unterstützt fühlten. Auch wenn dieses Ergebnis erst einmal erfreulich ist, ist auffällig, dass sich nicht auch eine deutliche Überlegenheit der Aufmerksamkeits(A)-WBTs ergab. Dieses konnte sich – vergleichbar zu den S-WBTs – nur für die Gesamtstichprobe gegenüber den E-Lectures abheben. Dass sich beide Einzelkomponenten-WBTs hinsichtlich der Skala „Aufmerksamkeit“ insbesondere gegenüber den E-Lectures durchsetzen können, könnte darin begründet sein, dass die Aufmerksamkeitserhaltung im Rahmen der Bearbeitung einer E-Lecture – welche die Inhalte eher flüchtig in gesprochener Sprache darbietet – ggf. schwieriger ist als im Rahmen der Bearbeitung eines WBTs mit schriftlichem Text. In Kombination mit der Tatsache, dass die allermeisten Lernenden E-Lectures zumeist linear durchsehen, ohne sich der Möglichkeiten des Spulens, Zurückspringens oder der konkreten

Suche eines bestimmten Inhaltes zu bedienen (siehe u.a. Glowalla et al., 2002 bzw. Görlich, 2006), spricht dafür, dass sie die Funktionen, die ihnen die Aufmerksamkeitserhaltung und das Verständnis der Inhalte erleichtern würden, nur selten nutzen. Schon in Studie I zeigte sich, dass die Studierenden im Rahmen der E-Lecture-Bearbeitung deutlich häufiger Pausen einlegten als im Rahmen der Bearbeitung der WBTs. Dies könnte ein Zeichen dafür sein, dass die Bearbeitung der E-Lectures für die Studierenden anstrengender ist und sich insbesondere im Vergleich zu dieser Lehrform die Aufmerksamkeitsstrategien (sowie die Zufriedenheitsstrategien) positiv bemerkbar machen können.

Die Aufmerksamkeitsstrategien allein scheinen allerdings nicht auszureichen, um den Studierenden das Gefühl zu vermitteln, dass die Anwendung sie in der Aufmerksamkeitserhaltung besser unterstützt als die 0-WBTs. Es zeigt sich jedoch auch kein signifikanter Einfluss einer anderen Komponente auf die Aufmerksamkeitsförderung in der Art, dass sich ein anderes Einzelkomponenten-WBT gegenüber den 0-WBTs durchsetzte. Zum Teil war es bei der Erstellung der WBT-Varianten schwierig, eine Komponente umzusetzen, ohne dabei auch gleich eine andere zu berücksichtigen. Besonders betroffen hiervon war sicherlich die Aufmerksamkeitskomponente, da jede Umsetzung einer Strategie innerhalb eines WBTs unweigerlich eine Abwechslung zum eintönigen 0-WBT darstellt. Es ist somit nicht überraschend, dass sich die A-WBTs hinsichtlich der Aufmerksamkeitsförderung nicht gegenüber den anderen Einzelkomponenten-WBTs durchsetzen konnten. Schließlich fördern diese durch andere Maßnahmen zumindest ähnlich die Variabilität und damit die Aufmerksamkeitserhaltung. Es ist allerdings überraschend, dass keines der Einzelkomponenten-WBTs gegenüber den 0-WBTs eine bessere Aufmerksamkeitsförderung erreichen konnte.

Allerdings können auch nicht alle Einzelkomponenten-WBTs die Aufmerksamkeit so gut fördern, dass sie genauso gut abschneiden wie die ARCS-WBTs. Zumindest für die Gesamtauswertung schneiden die R- und C-WBTs signifikant schlechter ab als die ARCS-WBTs. Demnach scheinen die Aufmerksamkeits- und Zufriedenheitskomponenten des Keller'schen Modells etwas mehr zur Aufmerksamkeitsförderung beitragen zu können als die Relevanz- und Erfolgszuversichtskomponenten. Nur die Variante, in der alle vier Komponenten umgesetzt wurden – das ARCS-WBT – kann die Aufmerksamkeit insgesamt besser fördern als das annähernd „ARCS-leere" 0-WBT. Es scheint so zu sein, dass je mehr der Strategien zum Einsatz kommen, desto besser fühlen sich die Lernenden durch die Anwendung hinsichtlich ihrer Aufmerksamkeitserhaltung unterstützt.

Interessant wäre hier, weitere Varianten mit Kombinationen der Strategien zu evaluieren, um zu prüfen, ob bereits der Einsatz von zwei Komponenten zum gleichen Effekt führt oder ob drei bzw. alle vier Komponenten eingesetzt werden müssen, um einen solchen Effekt zu beobachten. Bei vollständiger Umsetzung würde dies allerdings die möglichen WBT-

Variationen auf 16 erhöhen (0, A, R, C, S, AR, AC, AS, RC, RS, CS, ARC, ARS, ACS, RCS, ARCS), was eine empirische Prüfung noch deutlich aufwendiger machen würde.

Bezogen auf die Skala **„Relevanz“** ergab sich in der Mehrheit der themenspezifischen Analysen kein signifikanter Unterschied zwischen den E-Learning-Varianten. Lediglich für die Gesamtauswertung konnten sich die ARCS- und R-WBTs gegenüber den E-Lectures und 0-WBTs durchsetzen.
Gründe dafür, dass sich zu den Themen „Unterrichten“ und „Diagnostik“ keine Signifikanzen ergeben haben, könnten in zwei Aspekten liegen: Zum einen sind die Inhalte der WBTs für die Studierenden per se relevant, da sie zum Curriculum ihres Studiengangs gehören und die Studierenden eine Klausur über die Inhalte schreiben müssen. Dies gilt aber auch für das Thema „Konditionierung“, für das signifikante Unterschiede beobachtbar waren. Bei den beiden anderen Themen handelt es sich allerdings um Inhalte, die von sich aus bereits einen hohen Anwendungsbezug haben. Vergleichbar zur Studie von Means et al. (1997), die zeigen konnten, dass der Einsatz von Relevanzstrategien besonders dann sinnvoll und notwendig ist, wenn den Lernenden die Relevanz des Lerninhaltes nicht bewusst ist, ist es möglich, dass die R-Strategien deshalb hier nicht zu einer weiteren Verbesserung des Relevanzgefühls beitragen konnten. Unterstützt wird diese Annahme dadurch, dass die Mittelwerte der Relevanzskala generell etwas höher ausfallen, vergleicht man sie mit den mittleren Einschätzungen zu den anderen drei Skalen. Das Relevanzgefühl scheint folglich innerhalb der Stichprobe schon per se stärker vorhanden gewesen zu sein.
Dass sich zum Thema „Konditionierung“ signifikante Vorteile des R- und ARCS-WBTs gegenüber dem 0-WBT und des R-WBTs gegenüber dem C-WBT ergaben, kann folglich dadurch begründet werden, dass den Studierenden die Relevanz dieses Themas zuvor nicht so deutlich bewusst war wie für die Themen „Unterrichten“ oder „Pädagogische Diagnostik“. Die Anwendung der beiden letzten Themen auf den Schulalltag ist offensichtlich, während für das Thema „Konditionierung“ Anwendungsmöglichkeiten genannt werden müssen. Dies könnte dazu geführt haben, dass die relevanzsteigernden Strategien sich im Rahmen dieses Themas bemerkbar machen konnten. Das Relevanzgefühl scheint folglich im Gegensatz zu den anderen drei Komponenten von Keller (2010) deutlich themenabhängiger zu sein. Dies macht auch Sinn, denn es spielt nicht nur eine Rolle, wie relevant die Themen für das Bestehen einer Prüfung im Allgemeinen sind. Auch die Relevanz für den späteren Berufsalltag ist entscheidend. Dadurch macht es einen Unterschied, ob den Lernenden die Wichtigkeit der Themen von vornherein klar ist – sie dementsprechend keine Hinweise hierzu benötigen, wie dies bei den Themen „Unterrichten“ und „Diagnostik“ der Fall war – oder ob sie für die Einstufung der Relevanz

durch die Lernanwendung Unterstützung benötigen, wie anscheinend für das Thema „Konditionierung“ nötig.
Darüber hinaus ist allerdings darauf hinzuweisen, dass aus verschiedenen Gründen (siehe Abschnitt „6.3.1 Lehrmaterial: E-Lectures und Web-based Trainings“) der Punkt „Motivanpassung“ der Relevanzkomponente nicht umgesetzt werden konnte. Alle anderen Komponenten konnten innerhalb ihrer WBTs vollständig eingesetzt werden. Dadurch ist eine Interpretation des Einflusses der Relevanzstrategien nur in begrenztem Maße möglich. Gegebenenfalls wäre eine Erhöhung des Relevanzgefühls deutlicher nachweisbar gewesen, wenn Strategien zur Motivanpassung hätten umgesetzt werden können.

Für die themenspezifische Auswertung erreichten die ARCS-, Erfolgszuversichts(C)- oder Zufriedenheits(S)-WBTs auf der Skala **„Erfolgszuversicht“** bessere Beurteilungen als die 0-WBTs. Die ARCS- und C-WBTs konnten sich zudem von den A-WBTs und E-Lectures abheben. Die ARCS-WBTs erreichten darüber hinaus ein besseres Ergebnis als die R-WBTs. Über die Themen hinweg ergaben sich weitere Signifikanzen, so dass sich die ARCS-, C- und S-WBTs insgesamt gegenüber den 0-, A-, R-WBTs und den E-Lectures durchsetzten. Die ARCS-WBTs wurden darüber hinaus besser beurteilt als die S-WBTs und die R-WBTs erhielten eine bessere Beurteilung als die 0-WBTs.
Im Vergleich zu den 0-WBTs scheinen folglich nicht nur die Erfolgszuversichtsstrategien eine Erhöhung der Erfolgszuversicht zu bewirken. Auch die Zufriedenheitsstrategien (und zumindest für die Gesamtauswertung die Relevanzstrategien) können die Wahrnehmung von Erfolgszuversicht verbessern. Dies ist auch sinnvoll, wenn man bedenkt, dass das Zufriedenheits-WBT innerhalb der Lernspiele ebenfalls die Möglichkeit bietet, den eigenen Lernerfolg einzustufen und innerhalb des Relevanz-WBTs die Lernziele genannt und am Ende noch einmal zum selbständigen Abgleichen mit dem Gelernten angeboten werden. Die Möglichkeit, ihr Wissen zu testen, scheint den Studierenden wichtig zu sein, was auch in Abschnitt „7.4.3 Einschätzung wichtiger Eigenschaften von E-Learning-Anwendungen“ deutlich wird. Auch hier stellen sich die Eigenschaften, die mit den Übungstests zusammenhängen, als diejenigen heraus, die von den Studierenden als besonders wichtig eingestuft werden.
Welche der S- (oder R-)Strategien sich jedoch tatsächlich auf die Erfolgszuversicht der Lernenden ausgewirkt hat/haben, müsste innerhalb einer weiteren Studie betrachtet werden, in der die einzelnen Strategien in unterschiedlichen Anwendungen miteinander verglichen werden. Dabei scheinen die Strategien der C-Komponente jedoch hinsichtlich der Erfolgszuversicht einen etwas deutlicheren Einfluss zu haben, da sich die C-WBTs bereits in der themenspezifischen Auswertung nicht nur gegenüber den 0-WBTs, sondern auch gegenüber den A-WBTs und E-Lectures abheben. Der Einfluss des S-WBTs auf die

Erfolgszuversicht kann darüber hinaus nicht vollständig durch die Gemeinsamkeiten mit der C-Komponente interpretiert werden (s.o. Lernspiele), da sonst ein gemeinsamer Einsatz beider Komponenten nicht mehr zu einer Erhöhung der Erfolgszuversicht geführt hätte. Gemeinsam eingesetzt – wie dies im ARCS-WBT der Fall ist – scheint sich die Unterstützung durch beide Komponenten auszuzahlen, denn die ARCS-WBTs erreichen themenspezifisch ausgewertet zusätzlich ein besseres Ergebnis als die Relevanz-WBTs, über alle Themen hinweg sogar ein besseres Ergebnis als die Zufriedenheits(S)-WBTs.
In Abschnitt „6.1.4 Personenvariablen" zeigte sich darüber hinaus, dass nach den erhobenen Personenvariablen die Studierendengruppe im Mittel insbesondere hinsichtlich ihrer Erfolgszuversicht unterstützt werden sollte. Dies kann ein Grund dafür sein, dass die Studierenden hier alle Möglichkeiten zur Unterstützung ihrer Erfolgszuversicht dankbar angenommen haben.
Insgesamt scheint es demnach für die Erfolgszuversicht der Lernenden nicht bedeutsam zu sein, ob S- oder C-Strategien im WBT eingesetzt werden. Zumindest schneiden diese Einzelkomponenten-WBTs gleich gut ab wie die ARCS-WBTs und beide besser als die 0-WBTs. Folglich scheinen beide Komponenten die Erfolgszuversicht der Lernenden bedeutsam zu verbessern. Betrachtet man jedoch die Vorteile dieser WBTs gegenüber den anderen WBTs (A-, R- und 0-WBTs), scheint sich die C-Komponente etwas stärker auszuwirken und die Kombination beider Komponenten zudem zu einer besseren Erfolgszuversicht zu führen als der getrennte Einsatz nur einer der Komponenten.

Auch hinsichtlich der Skala **„Zufriedenheit"** fallen die ARCS-, Erfolgszuversichts(C)- und Zufriedenheits(S)-WBTs positiv auf. Alle drei WBT-Varianten werden deutlich zufriedenheitsfördernder eingeschätzt als die 0-, Aufmerksamkeits(A)- und Relevanz(R)-WBTs sowie die E-Lectures. In der Gesamtauswertung über die drei Themen hinweg zeigte sich zudem ein signifikanter Vorteil der ARCS- und S-WBTs gegenüber den C-WBTs sowie der A- und R-WBTs gegenüber den 0-WBTs. Hinsichtlich der Zufriedenheit scheinen folglich erneut die Erfolgszuversichts- und Zufriedenheitskomponenten des ARCS-Modells eine bedeutsame Rolle zu spielen. Hier sind es allerdings erwartungsgemäß die Zufriedenheitsstrategien, die zu einer besseren Einschätzung führen als die Erfolgszuversichtsstrategien. Auch die Aufmerksamkeits- und Relevanzstrategien können bereits im Vergleich zu den 0-WBTs zu einer höheren Zufriedenheit der Studierenden beitragen. Dies zeigt eindrucksvoll, dass die einzelnen Komponenten alle zur Zufriedenheit der Studierenden beisteuern, die Komponente „Zufriedenheit" selbst jedoch nicht überflüssig ist. Im Gegensatz zur Erfolgszuversicht erreichen hier die S-WBTs eine vergleichbar gute Beurteilung wie die ARCS-WBTs. Nach Hinzufügen der

Zufriedenheitsstrategien kommt es folglich durch die Ergänzung weiterer ARCS-Komponenten nicht zu einer weiteren Steigerung der Zufriedenheit der Studierenden.
Die Strategien der C- und S-Komponente von Keller (2010) scheinen folglich eng verbunden zu sein. Es ist nachvollziehbar, dass die Studierenden sehr zufrieden mit der Lernanwendung waren, wenn diese sie dabei unterstützte, während der Bearbeitung erfolgszuversichtlich zu sein. Die Ergebnisse zeigen allerdings auch, dass beide Komponenten nebeneinander eine Daseinsberechtigung haben.

Hinsichtlich des Gesamtergebnisses der **ARCS-Skala** schnitten die ARCS-WBTs im themenspezifischen Vergleich durchgängig signifikant besser ab als die 0-, Aufmerksamkeits(A)- und Relevanz(R)-WBTs sowie die E-Lectures, die Zufriedenheits(S)-WBTs besser als die 0- und Aufmerksamkeits(A)-WBTs und die E-Lectures und die Erfolgszuversichts(C)-WBTs schließlich besser als die 0-WBTs. Im Gesamtvergleich ergaben sich darüber hinaus weitere Signifikanzen: Die ARCS-WBTs schnitten zusätzlich besser ab als die C-WBTs, die S-WBTs besser als die R-WBTs, die C-WBTs als die A-WBTs und E-Lectures und die R-WBTs besser als die 0-WBTs.
Sowohl nach der Betrachtung der Einzel-Skalen als auch der Gesamtskala sind es folglich insbesondere die Erfolgszuversichts(C)- und die Zufriedenheits(S)-Strategien, die die Wahrnehmung der Studierenden hinsichtlich der Lernunterstützung positiv beeinflussen. Dabei scheinen – betrachtet man die Gesamtskala – die Zufriedenheitsstrategien einen stärkeren Einfluss als die Erfolgszuversichtsstrategien zu haben und beide Komponenten unterschiedlich zu wirken. Es ist folglich sinnvoll, innerhalb einer Lernanwendung die C- und S-Strategien gemeinsam umzusetzen, um die Lernenden möglichst gut zu unterstützen. Ob dies allerdings ausreicht, oder ob die A- und R-Strategien nicht doch noch einen Einfluss haben, der hier nur nicht (bzw. bezogen auf die R-Komponente – nur wenig) sichtbar wurde, könnte nur überprüft werden, wenn man einen Vergleich zwischen einer CS- und einer ARCS-Variante unternimmt – ggf. ergänzt durch eine 0-, eine C- und eine S-Variante.
Welche Strategien innerhalb der C- und S-Komponenten jedoch zu den guten Beurteilungen geführt haben, müsste innerhalb einer Laboruntersuchung untersucht werden. Hier müssten Lernende dabei beobachtet werden, wie sie mit den WBT-Varianten arbeiten. Welche der eingesetzten Strategien nehmen sie wirklich wahr und welche werden von ihnen tatsächlich genutzt? Wenn die Studierenden zum Beispiel die Lernspiele oder Übungsaufgaben innerhalb der WBTs gar nicht genutzt haben, können diese nicht für die gute Bewertung der C- und S-WBTs ausschlaggebend gewesen sein. Eine vollständige Aufklärung ist auf der Basis der vorliegenden Untersuchung folglich nicht möglich. Erste

Hinweise ergeben sich aber durch die Auswertung der Studierendenangaben zur Einschätzung wichtiger Eigenschaften von E-Learning-Anwendungen (Abschnitt 7.4.3).

Insgesamt kann geschlussfolgert werden, dass es sinnvoll ist, zur Förderung der Aufmerksamkeit alle ARCS-Strategien umzusetzen. Das Relevanzgefühl konnte – wie oben diskutiert – durch die eingesetzten Relevanzstrategien nur dann merkbar gefördert werden, wenn es inhaltlich Bedarf dazu gab. Für die Unterstützung der Erfolgszuversicht, Zufriedenheit bzw. der Wahrnehmung der Lernunterstützung insgesamt (ARCS-Skala) spielen insbesondere die C- und S-Strategien eine bedeutende Rolle. Dabei scheinen sie unterschiedliche Auswirkungen zu haben bzw. sich in ihrer Förderlichkeit zu ergänzen.
Ob das bessere Abschneiden hinsichtlich der ARCS-Skala auch eine höhere Akzeptanz für die C-, S- und ARCS-WBTs mit sich bringt, zeigt sich im nächsten Abschnitt.

7.3 Akzeptanz

Hypothese 9a: Für das ARCS-WBT ergibt sich eine signifikant höhere Lernerakzeptanz (Akzeptanzsskala) als für alle anderen E-Learning-Anwendungen.
➔ **ARCS > A, R, C, S, 0, EL**

Hypothese 9b: Für die Einzelkomponenten-WBTs sowie die E-Lecture ergibt sich eine signifikant höhere Lernerakzeptanz (Akzeptanzsskala) als für das 0-WBT.
➔ **A, R, C, S, EL > 0**

Hypothese 10a: Für das ARCS-WBT ergibt sich eine signifikant höhere Lernerakzeptanz (Noteneinstufung) als für alle anderen E-Learning- Anwendungen.
➔ **ARCS > A, R, C, S, 0, EL**

Hypothese 10b: Für die Einzelkomponenten-WBTs sowie die E-Lecture ergibt sich eine signifikant höhere Lernerakzeptanz (Noteneinstufung) als für das 0-WBT.
➔ **A, R, C, S, EL > 0**

7.3.1 Skala „Akzeptanz der Lernumgebung“

Um die Hypothesen 9a und 9b zu untersuchen, werden die Ergebnisse zur Skala „Akzeptanz der Lernumgebung“ betrachtet. Der Wertebereich dieser Skala liegt zwischen 9 und 54 Punkten. Ein hoher Wert spricht für eine hohe Akzeptanz der jeweiligen E-Learning-Anwendung.

Tabelle 93
Deskriptive Statistik zur Skala "Akzeptanz der Lernumgebung" für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Akzeptanz	Konditionierung			Unterrichten			Diagnostik			Gesamt			Diff. M „Kond.“/ „Dia.“
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	
0-WBT	64	39.34	8.33	74	37.61	9.11	50	37.76	9.26	188	38.24	8.88	1.58
A-WBT	64	41.92	7.32	81	39.04	7.34	78	38.05	7.59	223	39.52	7.56	3.87
R-WBT	68	**42.94**	7.23	66	39.26	8.64	58	38.14	8.78	192	40.22	8.43	4.80
C-WBT	56	42.32	7.20	64	38.89	7.26	51	38.57	7.06	171	39.92	7.34	3.75
S-WBT	80	42.35	7.45	61	39.11	7.25	53	35.36	9.10	194	39.42	8.34	**6.99**
ARCS-WBT	57	42.91	8.47	68	**41.10**	7.67	73	**40.05**	8.33	198	**41.24**	8.19	2.86
E-Lecture	200	41.67	7.67	242	38.57	7.28	203	37.22	8.18	645	39.11	7.89	4.45
Gesamt	**589**	**41.87**	**7.69**	**656**	**38.93**	**7.71**	**566**	**37.79**	**8.32**	**1811**	**39.53**	**8.07**	**4.08**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Tabelle 93 (veranschaulicht in Abbildung 49) zeigt die Ergebnisse der Akzeptanzerhebung über die drei Themenbereiche. Es ist erfreulich, dass alle E-Learning-Anwendungen Bewertungen im mittleren bis oberen Wertebereich der Akzeptanzskala erreichen (zwischen 35 und 43 Punkten, über alle Themen hinweg durchschnittlich *M* = 39.5 Punkte). Dabei ist es erneut das ARCS-WBT mit insgesamt *M* = 41.2 Punkten, das augenscheinlich die besten Akzeptanzbeurteilungen erhält. Die schlechtesten Bewertungen erhalten für die Themen „Konditionierung“ und „Unterrichten“ sowie insgesamt jeweils die 0-WBTs. Zum Thema „Diagnostik“ erreicht das Zufriedenheits(S)-WBT die geringste Akzeptanz. Generell liegen die Bewertungen für alle E-Learning-Variationen sehr eng beieinander.

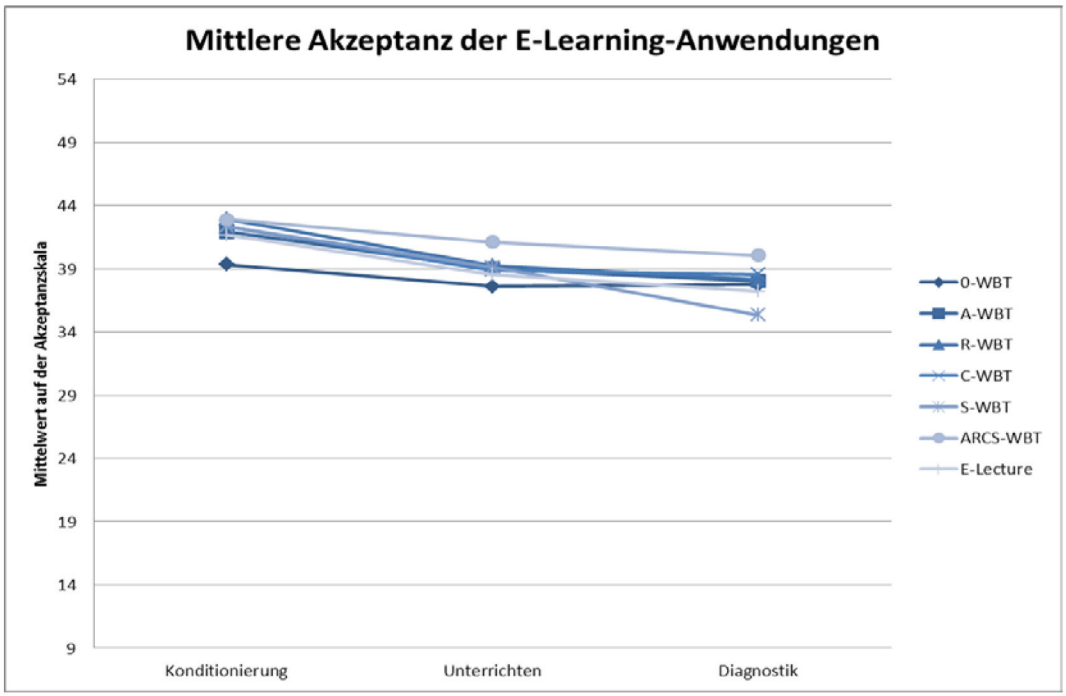

Abbildung 49. Mittelwerte der Skala "Akzeptanz der Lernumgebung" für die E-Learning-Gruppen zu allen drei Themen

Auffällig ist der ähnlich zur ARCS-Skala zu beobachtende Abwärtstrend der Beurteilungen, der – mehr oder weniger – alle E-Learning-Anwendungen betrifft. Im Schnitt verlieren die E-Learning-Anwendungen von der ersten zur letzten Datenerhebung 4 Akzeptanzpunkte.

Tabelle 94
Ergebnisse der Mittelwertsanalysen zum Vergleich der Skala „Akzeptanz der Lernumgebung" zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Akzeptanz Konditionierung	**ANOVA/Brown-Forsythe-Test**				**Posthoc**
	F	***df***	***p***	***Eta²***	
Zufallsauswahl 1	1.66	446,6	.13	-	
Zufallsauswahl 2	1.64	448,6	.14	-	-
Zufallsauswahl 3	1.64	452,6	.13	-	
Unterrichten					
Zufallsauswahl 1	1.34	485,6	.24	-	
Zufallsauswahl 2	1.28	445.91,6	.27	-	-
Zufallsauswahl 3	1.36	490,6	.23	-	
Diagnostik					
Zufallsauswahl 1	1.64	423,6	.13	-	-
Zufallsauswahl 2	2.38	425,6	.03*	.03	ARCS > S
Zufallsauswahl 3	1.72	423,6	.12	-	-
Gesamt					
Zufallsauswahl 1	2.48	1368,6	.02*	.01	ARCS > 0
Zufallsauswahl 2	2.77	1375,6	.01*	.01	ARCS > 0
Zufallsauswahl 3	2.61	1379,6	.02*	.01	ARCS > 0

Anmerkung. F = Teststatistik der Varianzanalyse bzw. des Brown-Forsythe-Tests, df = Freiheitsgrade, p = Signifikanz, Eta^2 = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

In der themenspezifischen Auswertung ergeben die Varianzanalysen lediglich zum Thema „Diagnostik" – und das auch nur in einer der drei Zufallsauswahl-Analysen – einen signifikanten Effekt, der als gering zu beurteilen ist. Hier erreicht das ARCS-WBT eine signifikant bessere Akzeptanzbeurteilung als das inhaltlich vergleichbare Zufriedenheits(S)-WBT. In der Gesamtauswertung ergeben sich durchgängig Signifikanzen mit allerdings schwachem Effekt. Hier zeigt sich ein signifikanter Akzeptanzvorteil der ARCS-WBTs gegenüber den 0-WBTs.
In der Mehrheit der Analysen erreichen die E-Learning-Varianten folglich vergleichbar gute Akzeptanzwerte. Die Hypothesen 9a und 9b sind demnach zu verwerfen.

In der abhängigen Varianzanalyse zwischen den Akzeptanzbeurteilungen über alle E-Learning-Anwendungen zwischen den drei Themen ergibt sich ein signifikanter Unterschied zwischen den Akzeptanzwerten mit einem mittleren bis starken Effekt (siehe Tabelle 95). Dabei schneiden die E-Learning-Anwendungen zum Thema „Konditionierung" signifikant besser ab als die Anwendungen zu den Themen „Unterrichten" und „Diagnostik". Die Anwendungen zum Thema „Unterrichten" erhalten zudem bessere Akzeptanzwerte als die Anwendungen zum Thema „Diagnostik".

Tabelle 95
Ergebnisse der Mittelwertsanalysen zum Vergleich der Skala „Akzeptanz der Lernumgebung“ über alle Gruppen zwischen den Themen

Akzeptanz	**Deskriptive Statistik**		**Abhängige Varianzanalyse**				**Posthoc**
n = 304	***M***	***SD***	***F***	***df***	***p***	**Partielles *Eta²***	
Konditionierung	**42.62**	7.39					
Unterrichten	39.85	7.78	41.84	1.97	<.01*	.12	Kon > Unt > Dia
Diagnostik	38.52	7.60					

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, Partielles *Eta²* = Effektstärke

7.3.2 Schulnote

Neben der feinstufigeren Akzeptanzskala wurde zudem das grobe Akzeptanzmaß „Schulnote“ (1-6) erfasst. Im Gegensatz zur Akzeptanzskala sprechen hier höhere Werte für eine schlechtere Bewertung.

Tabelle 96
Deskriptive Statistik zur Variable „Schulnote“ (1-6) für die E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Note	**Konditionierung**			**Unterrichten**			**Diagnostik**			**Gesamt**		
	n	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
0-WBT	64	2.45	0.69	74	2.96	0.99	50	2.94	0.94	188	2.78	0.91
A-WBT	64	2.20	0.91	81	2.68	0.92	78	2.64	0.88	223	2.53	0.92
R-WBT	68	1.99	0.70	66	2.65	0.83	58	2.83	1.03	192	2.47	0.93
C-WBT	56	2.13	0.76	64	2.72	0.97	51	2.71	0.86	171	2.52	0.91
S-WBT	80	2.09	0.73	61	2.59	0.90	53	2.70	0.93	194	2.41	0.89
ARCS-WBT	57	**1.88**	0.73	68	**2.31**	0.90	73	**2.47**	0.88	198	**2.24**	0.88
E-Lecture	200	2.22	0.80	242	2.76	0.88	203	2.88	0.88	645	2.63	0.90
Gesamt	**589**	**2.15**	**0.78**	**656**	**2.69**	**0.91**	**566**	**2.76**	**0.91**	**1811**	**2.54**	**0.91**

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Tabelle 96 und Abbildung 50 zeigen, dass die E-Learning-Anwendungen mit Noten zwischen 1.9 und 3.0 benotet wurden. Über alle Themen hinweg resultiert eine durchschnittliche Note von *M* = 2.5. Die Anwendungen erhalten von den Studierenden folglich zumeist eine gute Beurteilung.

Dabei werden die Beurteilungen für alle E-Learning-Anwendungen von der ersten zur dritten Erhebung zum Teil deutlich schlechter. Auch hier erreichen die ARCS-WBTs mit einer durchschnittlichen Note von *M* = 2.2 augenscheinlich durchgängig die besten Beurteilungen. Die 0-WBTs (*M* = 2.8) werden zu allen drei Erhebungen am schlechtesten eingestuft. Alle anderen E-Learning-Variationen liegen dazwischen und mit durchschnittlichen Noten von 2.4 bis 2.6 eng beieinander.

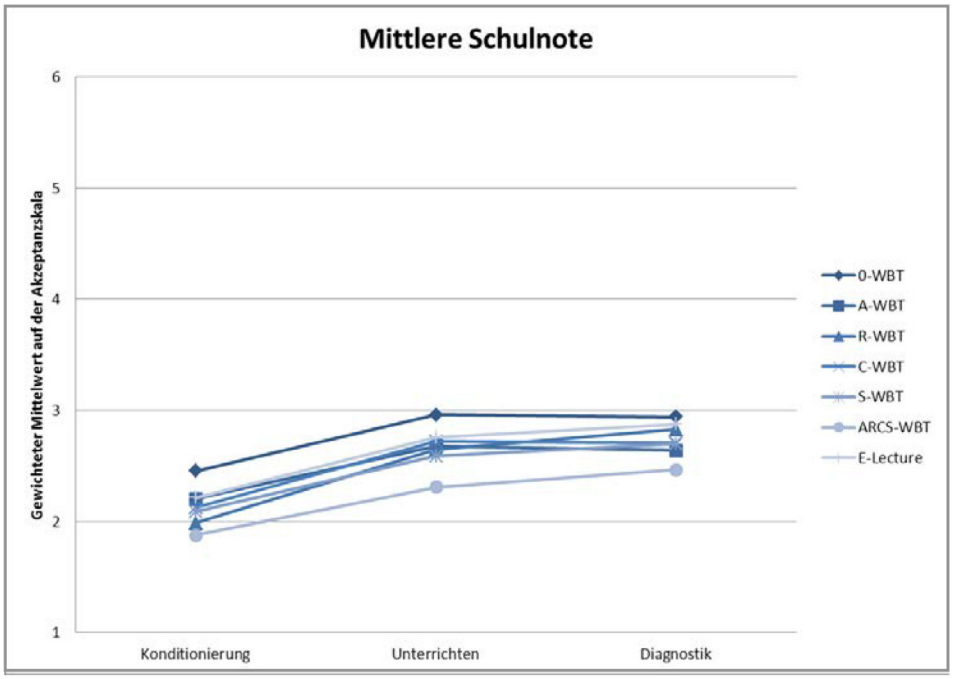

Abbildung 50. Mittlere „Schulnote" der E-Learning-Versionen zu allen drei Themen

Tabelle 97 gibt die Ergebnisse der einfaktoriellen Varianzanalysen wieder. Im Gegensatz zur Skala „Akzeptanz" ergeben sich in Bezug auf die Schulnote zumindest in zwei der drei Erhebungen signifikante Unterschiede. Sie weisen jedoch nur eine geringe Effektstärke auf. Zum Thema „Konditionierung" werden das ARCS- und das Relevanz(R)-WBT signifikant besser bewertet als das 0-WBT, zum Thema „Unterrichten" wird in der Mehrheit der Analysen lediglich das ARCS-WBT besser beurteilt als das 0-WBT. Zum Thema „Diagnostik" ergibt sich zumeist kein Unterschied zwischen den E-Learning-Anwendungen.

Tabelle 97
Ergebnisse der Mittelwertsanalysen zum Vergleich der Variable „Schulnote" zwischen den E-Learning-Gruppen getrennt für die drei Themen sowie insgesamt

Note	**ANOVA**				**Posthoc**
Konditionierung	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	3.74	446,6	<.01*	.05	ARCS, R < 0
Zufallsauswahl 2	3.57	448,6	<.01*	.05	ARCS, R < 0
Zufallsauswahl 3	3.56	452,6	<.01*	.05	ARCS, R < 0
Unterrichten					
Zufallsauswahl 1	3.53	485,6	<.01*	.04	ARCS < 0, EL
Zufallsauswahl 2	3.17	488,6	<.01*	.04	ARCS < 0
Zufallsauswahl 3	3.20	490,6	<.01*	.04	ARCS < 0
Diagnostik					
Zufallsauswahl 1	2.95	423,6	<.01*	.04	ARCS < EL
Zufallsauswahl 2	1.98	425,6	.07	-	-
Zufallsauswahl 3	1.63	423,6	.14	-	-
Gesamt					
Zufallsauswahl 1	7.85	1368,6	<.01*	.03	ARCS < 0, A, EL; S < 0, EL; R < 0
Zufallsauswahl 2	6.62	1375,6	<.01*	.03	ARCS < 0, A, EL; R, S < 0
Zufallsauswahl 3	6.31	1379,6	<.01*	.03	ARCS < 0, A, EL; R, S < 0

Anmerkung. F = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-Lecture-Gruppe vorgenommen.

Anders sieht es für die Gesamtauswertung aus. Hier ergeben sich mehrheitlich Vorteile der ARCS-WBTs gegenüber den 0-, A-WBTs und E-Lectures sowie der R- und S-WBTs gegenüber den 0-WBTs. Obwohl sich hier erneut tendenziell hypothesenkonforme Ergebnisse zeigen, sind die wiederum sehr streng formulierten Hypothesen 10a und 10b zu verwerfen.

Zwischen den Themen ergibt sich abermals ein signifikanter Unterschied (siehe Tabelle 98). Die E-Learning-Anwendungen zum Thema „Konditionierung" erreichen eine signifikant bessere Note als diejenigen zu den Themen „Unterrichten" und „Diagnostik". Die Effektstärke dieses Unterschiedes ist als stark zu bezeichnen.

Tabelle 98
Ergebnisse der Mittelwertsanalysen zum Vergleich der Variable „Schulnote" über alle Gruppen zwischen den Themen

Note	**Deskriptive Statistik**		**Abhängige Varianzanalyse**				**Posthoc**
n = 304	***M***	***SD***	***F***	***df***	***p***	**Partielles *Eta²***	
Konditionierung	2.11	0.76					
Unterrichten	2.63	0.94	50.18	1.97	<.01*	.14	Kon < Unt, Dia
Diagnostik	**2.64**	0.84					

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *F* = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, Partielles *Eta²* = Effektstärke

7.3.3 Zusammenfassung und Diskussion

Die Akzeptanz wurde durch die Skala „*Akzeptanz der Lernumgebung*" von Krause (2007) sowie mittels einer „groben" Einschätzung auf der Schulnotenskala erhoben. Für das feinere Akzeptanzmaß zeigte sich in der themenspezifischen Auswertung, dass alle E-Learning-Anwendungen Akzeptanzbewertungen im mittleren bis oberen Wertebereich der Skala erhalten, die Beurteilungen zu den sieben verschiedenen E-Learning-Varianten sich jedoch nur zum Teil signifikant voneinander unterscheiden. Dabei waren die Akzeptanzbeurteilungen nicht so hoch, dass bereits von einem Deckeneffekt gesprochen werden kann. Für die Auswertung über die drei Themen hinweg ergab sich zumindest ein signifikanter Unterschied zwischen den ARCS- und den 0-WBTs zugunsten der angereicherten E-Learning-Variante.
Hinsichtlich des groben Akzeptanzmaßes „Schulnote" kam es vermehrt zu Signifikanzen. Die gefundenen Unterschiede weisen geringe bis mittelstarke Effekte auf. Für die themenspezifische Auswertung konnten lediglich die ARCS-WBTs in der Mehrheit der Analysen eine signifikant bessere Note erreichen als die inhaltsgleichen 0-WBTs. Über die drei Themen hinweg erreichten die ARCS-WBTs signifikant bessere Ergebnisse als die 0- und A-WBTs sowie die E-Lectures. Die R- und S-WBTs wurden signifikant besser beurteilt als die 0-WBTs.

Die Ergebnisse hinsichtlich der ARCS-Skala schlagen sich folglich nur zum Teil auf die Akzeptanzbeurteilung nieder. Vergleichbar zu den Studien von Astleitner und Lintner (2004) oder Chyung et al. (1999) zeigten sich durch den Einsatz aller ARCS-Komponenten allerdings positive Auswirkungen auf die Akzeptanzbeurteilung. In Studie I konnten hinsichtlich der Akzeptanz keine signifikanten Unterschiede zwischen den E-Learning-Anwendungen festgestellt werden. Es ist deshalb erfreulich, dass die ARCS-WBTs gegenüber den E-Lectures nun signifikant besser beurteilt wurden. Im Vergleich zwischen den WBT-Varianten scheinen hinsichtlich der Akzeptanz die Relevanz- und erneut die Zufriedenheitsstrategien des Keller'schen Modells eine größere Rolle zu spielen. Die Strategien der Aufmerksamkeits- und Erfolgszuversichtskomponente konnten sich – isoliert umgesetzt – nicht positiv gegenüber den 0-WBTs durchsetzen.

Erstaunlich bleibt allerdings, dass die 0-WBTs von den Studierenden immer noch so positiv beurteilt wurden. Dies kann allerdings dadurch begründet sein, dass die Studierenden generell mit dem Wechsel zu einer E-Learning-Anwendung (im Gegensatz zur traditionellen Vorlesung) zufrieden waren, da dies eine Abwechslung zum „üblichen" Hochschulalltag darstellte und sie die allgemeinen Vorteile eines E-Learning-Szenarios schätzten (bequem und zeitflexibel von Zuhause aus in eigenem Tempo arbeiten etc.).

Hinsichtlich der ARCS-Skala und der Akzeptanzbewertungen zeigte sich ein klarer Abfall der Bewertungen von der ersten zur zweiten – und dann zum Teil noch einmal zur dritten Erhebung. Die Bewertungen bezüglich der Gesamt-ARCS-Skala fielen zu den Anwendungen zum Thema „Konditionierung" signifikant besser aus als zu den Anwendungen zu den Themen „Unterrichten" und „Diagnostik". Gleiches gilt für die Schulnote. Für das feinere Akzeptanzmaß ergab sich darüber hinaus ein Unterschied zwischen den beiden letzten Erhebungen. Dabei fiel die Akzeptanz gegenüber den Anwendungen zum Thema „Unterrichten" signifikant höher aus als gegenüber den Anwendungen zum Thema „Diagnostik". Bei allen gefundenen Unterschieden handelt es sich um mittelstarke bis starke Effekte.

Sicherlich können die Unterschiede inhaltlich begründet sein. Das Thema „Konditionierung" ist inhaltlich nicht so komplex wie die beiden anderen Themen, was sich auch auf die Beurteilung der E-Learning-Anwendungen ausgewirkt haben kann. Dies hätte sich allerdings auch in der Lernwirksamkeit niedergeschlagen. Eine abfallende Tendenz lässt sich im Rahmen der Lernwirksamkeit aber nicht feststellen.

Eine weitere Erklärung für den Abfall der Bewertungen könnte darin liegen, dass die Studierenden durch die Nutzung der verschiedenen Varianten erfahrener wurden. Aufgrund der Nutzung einer E-Learning-Variante hatten sie nun Erwartungen an die nächste. Wenn diese nicht erfüllt wurden, resultierte hieraus eine schlechtere Beurteilung der zweiten bzw. dritten E-Learning-Variante. Wenn ein Studierender z.B. in der ersten

Erhebung ein Erfolgszuversichts(C)-WBT genutzt hatte und in der zweiten Erhebung ein anderes Einzelkomponenten- oder 0-WBT zur Verfügung gestellt bekam, war er/sie sicherlich überrascht, dass nun kein Inhaltsverzeichnis oder Übungsaufgaben zur Verfügung standen. Hiernach hätten die Bewertungen für die ARCS-WBTs allerdings nicht in ihrer Bewertung abfallen dürfen, da sie stets alle Aspekte enthielten und die Studierenden dementsprechend nicht enttäuschen sollten.
Daher ist es wahrscheinlicher, dass die E-Learning-Anwendungen zum ersten Thema aufgrund der Abwechslung innerhalb des Universitätskontextes besser beurteilt wurden. Nachdem die Studierenden einmal eine E-Learning-Anwendung bearbeitet hatten, war dies beim zweiten und dritten Mal scheinbar nicht mehr so interessant. Die Studierenden wurden in ihrer Beurteilung kritischer. Die späte Darbietung der E-Learning-Varianten zum Thema „Diagnostik" kurz vor der Prüfungsphase im Wintersemester wird sich neben der Nutzungsdauer vermutlich auch auf die Akzeptanzbeurteilungen niedergeschlagen haben.

Im Folgenden werden die Ergebnisse aus der freiwilligen Nutzung zur Klausurvorbereitung vorgestellt. Dabei wird zunächst auf die Nutzungszeit und -häufigkeiten eingegangen, bevor die Lernleistung der verschiedenen Nutzergruppen dargestellt wird. Zu guter Letzt wird ein Blick auf die Einschätzung wichtiger Eigenschaften von E-Learning-Anwendungen durch die Studierenden geworfen.

7.4 Freiwillige Nutzung im Folgesemester

Um zu überprüfen, mit welchem Material sich die Studierenden hauptsächlich auf die Modulabschlussprüfung vorbereitet haben, wurden sie gebeten, hierzu eine Selbstauskunft zu geben. Da auch Offline-Materialien zur Verfügung standen, erschien dies sinnvoller als nur auf die Aufzeichnung von Nutzungszeiten zu vertrauen.

Tabelle 99
Hauptsächliche Nutzung der Lernmaterialien zur Vorbereitung auf die MAP

Genutztes Lernmaterial	*n*	%
E-WBTs	235	**45.9%**
Vorlesungsskript und eigene Mitschrift	134	26.2%
K-WBTs	86	16.8%
E-Lectures	44	8.6%
Andere Materialien	9	1.8%
Literatur	4	0.8%
Gesamt	**512**	**100.0**%

Anmerkung. *n* = Stichprobengröße, % = Gültige Prozente

784 Studierende konnten die Anwendungen zur Vorbereitung auf die Modulabschlussprüfung nutzen. 639 (81.5%) von ihnen nahmen (nach Aufzeichnungen des Learning-Management-Systems) diese Gelegenheit in Anspruch, 512 machten uns im

Rahmen der Probeklausur eine Angabe dazu, mit welchem Material sie sich hauptsächlich auf die MAP vorbereitet hatten. Tabelle 99 gibt die absoluten und prozentualen Häufigkeiten zu den laut den Studierenden genutzten Materialien an. Es zeigt sich deutlich, dass die meisten der Befragten (45.9%) angaben, vornehmlich die elaborierten WBTs zur Vorbereitung auf die Modulabschlussprüfung zu nutzen.

Die nächstgrößere Gruppe an Studierenden (26.2%) bereitete sich hauptsächlich mit dem Vorlesungsskript und den eigenen Mitschriften auf die Klausur vor. 16.8% der Studierenden gaben an, zumeist die kurzen WBTs zu nutzen und 8.6% lernten vorwiegend mit den E-Lectures. Nur 1.8% der hier befragten Studierenden benannten andere Materialien (z.B. Unterlagen von Freunden, Zusatzlektüren oder bisher noch gar keine Vorbereitung) und lediglich 0.8% bereiteten sich hauptsächlich anhand von Sekundärliteratur auf die Klausur vor. Ein Chi-Quadrat-Test zur Untersuchung, ob sich diese Verteilung der Studierenden auf die Lehrmaterialien signifikant unterscheidet, wird mit $Chi\text{-}Quadrat(5) = 456.09$ und $p < .01$ erwartungsgemäß signifikant.

Von den 365 Studierenden, die angaben, sich hauptsächlich mit den E-Lectures oder den WBTs auf die MAP vorbereitet zu haben, gaben 91.5% an, sich für die jeweilige Anwendung entschieden zu haben, da sie diese besser fanden als die anderen E-Learning-Anwendungen. Nur 8.5% beklagten technische Gründe für die geringe Nutzung der anderen E-Learning-Anwendungen.

Diese durch Selbstauskunft gewonnenen Nutzergruppen wurden nun hinsichtlich ihrer Nutzungsdauer für die verschiedenen E-Learning-Anwendungen verglichen. Im Rahmen der Selbstauskunft sollten die Studierenden lediglich angeben, mit welchem Material sie sich **hauptsächlich** auf die Klausur vorbereitet hatten. Dementsprechend sollte ein solcher Vergleich zeigen, ob die jeweilige Selbstauskunft-Gruppe auch die signifikant längsten Nutzungszeiten hinsichtlich der jeweiligen E-Learning-Anwendung aufweist.

Tabelle 100
Deskriptive Statistik zur Nutzungszeit der E-Learning-Anwendungen für die verschiedenen Selbstauskunft-Gruppen

Gruppe	K-WBT			E-WBT			E-Lecture		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Vorlesungsskript	134	4003.31	9973.39	134	10824.15	19274.49	134	3748.00	15781.69
Literatur	4	0.00	0.00	4	0.00	0.00	4	0.00	0.00
E-Lecture	44	2572.86	6722.66	44	12175.57	28092.46	44	**36918.64**	47235.32
K-WBT	86	**29489.19**	28047.51	86	13953.00	16350.75	86	6102.80	18844.74
E-WBT	235	5760.59	11161.74	235	**46565.39**	32951.31	235	7158.55	28346.08
Andere	9	1942.89	4723.85	9	6022.89	12809.95	9	0.00	0.00
Gesamt	512	8900.28	17465.09	512	27701.54	31820.55	512	8464.36	27544.21

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Tabelle 100 zeigt die Nutzungszeiten der sechs Selbstauskunft-Gruppen zu den drei E-Learning-Anwendungen. Schon auf den ersten Blick zeigt sich sehr deutlich, dass die

Selbstauskunft-Gruppen jeweils die höchste durchschnittliche Nutzungszeit für „ihre“ E-Learning-Anwendung aufweisen. Dabei ist die kleine Gruppe derjenigen, die sich mit der Literatur vorbereitet haben, am konsequentesten. Diese vier Studierenden haben tatsächlich in der Summe keine der E-Learning-Anwendungen genutzt. Alle anderen Gruppen haben alle E-Learning-Varianten mehr oder weniger zur Klausurvorbereitung genutzt – allerdings mit wesentlich geringerem Zeitumfang im Vergleich zur „hauptsächlich“ genutzten Anwendung. Deutlich wird hier bereits, dass die E-WBTs über alle Gruppen hinweg im Schnitt mehr als dreimal so lang genutzt wurden als die K-WBTs und E-Lectures.

Da sich eine Auswertung über die sehr kleinen Gruppen „Literatur“ und „Andere“ verbietet, wurden sie für die folgenden Varianzanalysen generell ausgeklammert. Auch für die anderen Gruppen ist eine Korrektur notwendig. So ist die Selbstauskunft-Gruppe „Vorlesungsskript“ dreimal und die Gruppe „Elaboriertes(E)-WBT“ sogar fünfmal so groß wie die Gruppe „E-Lecture“. Dementsprechend wurden ähnlich zu bereits vorausgegangenen Analysen auch hier mehrfach Zufallsstichproben gezogen. Da es sich um zwei Gruppen handelt, aus denen Zufallsstichproben gezogen werden mussten, wurden die Analysen dementsprechend jeweils sechsfach durchgeführt. Dieses Vorgehen wurde auch bei den weiteren Analysen angewandt. Tabelle 101 gibt die Ergebnisse zum Vergleich der Nutzungszeiten wieder.

Tabelle 101
Ergebnisse der Mittelwertsanalysen zum Vergleich der Nutzungszeiten der drei E-Learning-Anwendungen zwischen der Vorlesungs-, der E-Lecture- und den beiden WBT-Gruppen

Nutzungszeit	**Brown-Forsythe-Test**				**Posthoc**
K-WBTs	***F***	***df***	***p***	***Eta²***	
Zufallsauswahl 1	49.44	154.36,3	<.01*	.30	K > Alle anderen Gruppen
Zufallsauswahl 2	48.55	154.31,3	<.01*	.30	
Zufallsauswahl 3	44.83	171.94,3	<.01*	.28	
Zufallsauswahl 4	55.84	134.50,3	<.01*	.33	
Zufallsauswahl 5	54.92	134.45,3	<.01*	.32	
Zufallsauswahl 6	50.40	151.57,3	<.01*	.30	
E-WBTs					
Zufallsauswahl 1	30.10	194.90,3	<.01*	.26	E > Alle anderen Gruppen
Zufallsauswahl 2	34.27	194.23,3	<.01*	.28	
Zufallsauswahl 3	30.38	194.12,3	<.01*	.26	
Zufallsauswahl 4	35.88	172.49,3	<.01*	.29	
Zufallsauswahl 5	40.56	171.93,3	<.01*	.32	
Zufallsauswahl 6	35.80	172.28,3	<.01*	.29	
E-Lectures					
Zufallsauswahl 1	14.53	79.10,3	<.01*	.18	EL > Alle anderen Gruppen
Zufallsauswahl 2	12.49	94.18,3	<.01*	.15	
Zufallsauswahl 3	10.58	112.70,3	<.01*	.12	
Zufallsauswahl 4	15.13	77.64,3	<.01*	.19	
Zufallsauswahl 5	13.06	92.62,3	<.01*	.16	
Zufallsauswahl 6	11.10	111.08,3	<.01*	.13	

Anmerkung. F = Teststatistik des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-WBT-Gruppe sowie zwei Zufallsziehungen aus der Vorlesungsgruppe vorgenommen und miteinander kombiniert ausgewertet.

Die Analysen ergaben durchgängig starke Effekte. Dabei wurden die kurzen(K)-WBTs in der Summe signifikant am meisten von der Selbstauskunft-Gruppe „kurzes(K)-WBT“ genutzt, die elaborierten(E)-WBTs am längsten von der Gruppe „elaboriertes(E)-WBT“ und die E-Lectures am längsten von der Gruppe „E-Lecture“. Das bestätigt deutlich die Selbstauskünfte der Studierenden, so dass die Studierenden für die folgenden Auswertungen in die Nutzergruppen „Vorlesungsskript“, „E-WBT“, „K-WBT“ und „E-Lecture“ eingeteilt werden können.

Im Folgenden werden zunächst die Nutzungszeiten und -häufigkeiten im Rahmen der freiwilligen Nutzung analysiert. Danach werden die Lernleistungen der verschiedenen Nutzergruppen miteinander verglichen.

7.4.1 Nutzungszeit und -häufigkeit

Hypothese 11a: Die E-Learning-Anwendung, die im Wintersemester durch die Studierenden die höchsten Akzeptanzwerte erreicht, wird im Sommersemester signifikant häufiger genutzt als die anderen E-Learning-Anwendungen.

Hypothese 11b: Die E-Learning-Anwendung, die im Wintersemester durch die Studierenden die höchsten Akzeptanzwerte erreicht, wird im Sommersemester signifikant länger genutzt als die anderen E-Learning-Anwendungen.

Tabelle 102 gibt einen Überblick über die Nutzungszeiten und -häufigkeiten, die für die E-Learning-Anwendungen über die Gesamtzeit der freiwilligen Nutzung hinweg aufgezeichnet wurden.

Tabelle 102
Deskriptive Statistik sowie Summe zur Gesamtnutzungszeit und Nutzungshäufigkeit der E-Learning-Nutzergruppen

Nutzungszeit	***M* in Sek. (in Std.)**	***SD***	**Summe in Sek. (in Std.)**
E-Lecture	12577.57 (3.49)	50546.00	8037067 (2232.52)
K-WBT	10562.78 (2.93)	17814.46	6749616 (1874.89)
E-WBT	**31799.24 (8.83)**	31624.00	**20319714 (5644.37)**
Nutzungshäufigkeit	***M***	***SD***	**Summe**
E-Lecture	1.56	2.99	997
K-WBT	7.05	8.48	4505
E-WBT	**25.61**	26.95	**16365**

Anmerkung. n = 639; n = Stichprobengröße, M = Mittelwert, SD = Standardabweichung

Von den ursprünglich 1232 Studierenden nahmen 902 Studierende an der Modulabschlussprüfung teil. Weitere ca. 150 Studierende waren für die Klausur angemeldet, erschienen aber aus verschiedenen Gründen nicht (vornehmlich wegen einer Erkrankung). Wie oben erwähnt, nutzten 639 Studierende die E-Learning-Anwendungen

zur Vorbereitung auf die Modulabschlussprüfung. Da die Zahl der Nutzer durch die Teilnahme an den Wintersemester-Evaluationen auf 784 eingeschränkt war (von denen sicherlich auch nicht alle an der MAP teilnahmen), handelt es sich hierbei um über 80% der Studierenden. Dies ist sehr erfreulich, bedenkt man den Aufwand für die Erstellung der WBTs. Dabei zeigt sich in Tabelle 102 auch für die Gesamtgruppe der E-Learning-Nutzer erneut schon per Augenschein, dass die elaborierten WBTs im Mittel und in der Summe deutlich länger und häufiger genutzt wurden.

Die Nutzungszeiten sind allerdings in der Summe für alle E-Learning-Anwendungen sehr eindrucksvoll: Die kurzen (K-)WBTs und E-Lectures wurden von den Studierenden in der Summe ca. 2000 Stunden genutzt. Daraus resultiert eine durchschnittliche Bearbeitungszeit der E-Lectures von etwa dreieinhalb Stunden, für die K-WBTs von ca. drei Stunden (wenn jeder der 639 Studierenden die Anwendungen vergleichbar lange genutzt hätte). Die elaborierten WBTs wurden sogar über 5600 Stunden genutzt – dies entspricht ca. $M = 8.8$ Stunden pro Studierendem.

Vergleicht man nun die Nutzungshäufigkeit der 639 Studierenden, die alle drei Anwendungsarten mehr oder weniger umfangreich für ihre Klausurvorbereitung genutzt hatten mittels abhängiger Varianzanalyse mit Greenhouse-Geisser-Korrektur, ergibt sich (mit $F(1.18) = 407.13$) ein signifikanter Unterschied ($p < .01$) mit starkem Effekt ($Eta^2 = .39$). Hinsichtlich der Nutzungshäufigkeit unterscheiden sich alle drei Anwendungen voneinander: Die elaborierten (E-)WBTs werden signifikant häufiger genutzt als die kurzen (K-)WBTs und die E-Lectures, die kurzen (K-)WBTs wiederum signifikant häufiger als die E-Lectures.

Hypothese 11a („Die E-Learning-Anwendung, die im Wintersemester durch die Studierenden die höchsten Akzeptanzwerte erreicht, wird im Sommersemester signifikant häufiger genutzt als die anderen E-Learning-Anwendungen.“) ist nach den Ergebnissen aus dem Wintersemester[53] folgendermaßen anzupassen: „Die E-WBTs (ARCS-WBTs) werden im Sommersemester signifikant häufiger genutzt als die K-WBTs (0-WBTs).“[54] Diese Hypothese ist nach den vorliegenden Ergebnissen beizubehalten.

Vergleichbares ergibt sich in einer abhängigen Varianzanalyse mit Greenhouse-Geisser-Korrektur für die Nutzungszeit. Auch hier zeigt sich (mit $F(1.59) = 65.91$) ein signifikanter Unterschied ($p < .01$) mit mittelstarkem Effekt ($Eta^2 = .09$). Posthoc-Tests ergeben eine signifikant höhere Nutzungszeit für die elaborierten (E-)WBTs im Gegensatz zu den anderen beiden E-Learning-Anwendungsarten. Hypothese 11b („Die E-Learning-

[53] Hier erreichten die ARCS-WBTs (zumindest hinsichtlich des groben Akzeptanzmaßes „Schulnote“) eine bessere Akzeptanzeinschätzung als die 0-WBTs.

[54] Für den Vergleich mit der E-Lecture lässt sich keine überprüfbare Hypothese formulieren, da die Akzeptanzeinschätzung der E-Lecture nicht signifikant von der der anderen E-Learning-Anwendungen abwich.

Anwendung, die im Wintersemester durch die Studierenden die höchsten Akzeptanzwerte erreicht, wird im Sommersemester signifikant länger genutzt als die anderen E-Learning-Anwendungen."), muss ebenfalls nach den Ergebnissen aus dem Wintersemester folgendermaßen umformuliert werden: „Die E-WBTs (ARCS-WBTs) werden im Sommersemester signifikant länger genutzt als die K-WBTs (0-WBTs)." Auch diese Hypothese ist folglich beizubehalten.

7.4.1.1 Nutzungszeiten je Thema und Kalenderwoche

Auch wenn zu diesem Aspekt keine explizite Hypothese aufgestellt wurde, ist es doch interessant zu betrachten, zu welchen der Themen und in welchen Kalenderwochen die Studierenden vermehrt die E-Learning-Anwendungen nutzten.

Betrachtet man Abbildung 51, zeigt sich auch hier auf den ersten Blick die vermehrte Nutzung der elaborierten (E-)WBTs über alle Themen. Dabei ist anzumerken, dass es zu Thema 12 („Rückblick") zwar eine E-Lecture gab, aber keine WBTs.

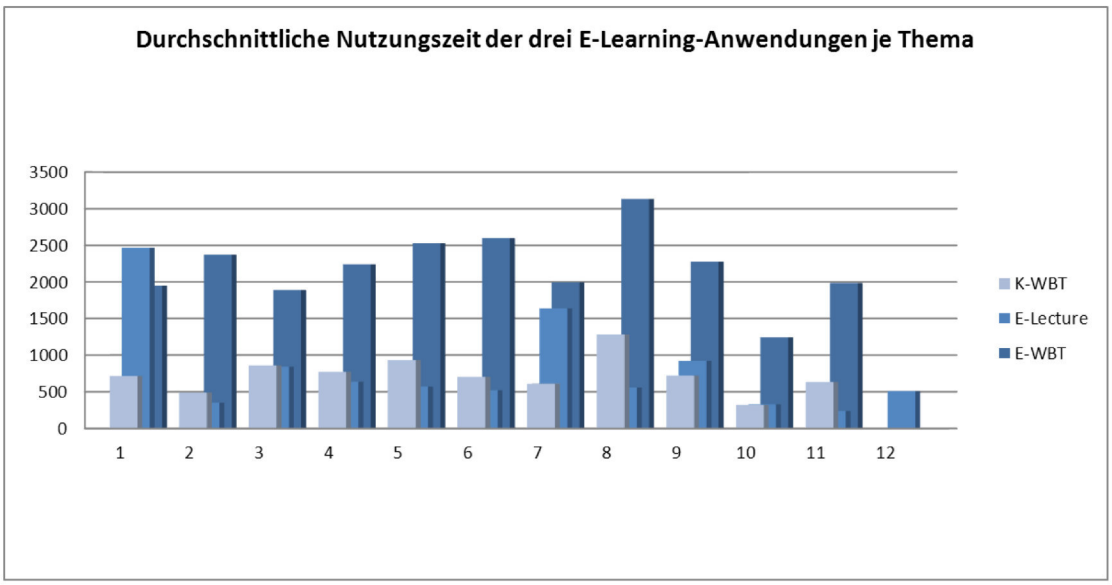

Abbildung 51. Durchschnittliche Nutzungszeit der E-Learning-Anwendungen für die 12 Themen (in Sek.)

Bei genauerer Betrachtung der Ergebnisse für die E-WBTs fällt auf, dass insbesondere das elaborierte WBT zu Thema 10 („Lernen mit Medien") relativ kurz bearbeitet wurde. Dieses Thema ist inhaltlich nicht allzu umfangreich, so dass die geringe Nutzungszeit nicht überrascht. Für alle anderen Themen liegt die durchschnittliche Nutzungszeit der elaborierten (E-)WBTs zwischen etwa 2000 und 3000 Sekunden (etwa 30 bis 50 Minuten). Am längsten bearbeitet wird das Thema 8 („Entwicklungspsychologie").

Anders sieht es bei den E-Lectures aus: Hier ist deutlich zu sehen, dass die Studierenden sich hauptsächlich die E-Lecture zu den Themen „Gedächtnis I" (Thema 1) und „Denken, Problemlösen und Intelligenz" (Thema 7) ansahen. Alle andere Themen wurden im Schnitt

lediglich um die 500 Sekunden (ca. 8 Minuten) bis maximal 1000 Sekunden (17 Minuten) lang bearbeitet.
Auch die Bearbeitungszeiten der kurzen (K-)WBTs liegen im Schnitt zwischen 500 und 1000 Sekunden (8 und 17 Minuten). Lediglich das kurze (K-)WBT zu Thema 8 („Entwicklungspsychologie“) wurde vergleichbar zu den Ergebnissen zum E-WBT etwas länger bearbeitet.

Abbildung 52 stellt nun die durchschnittliche Nutzungszeit der drei E-Learning-Anwendungen für die einzelnen Kalenderwochen der Klausurvorbereitung dar. Die durchschnittliche Nutzungszeit pro Woche wurde hierfür anhand der Anzahl der Tage relativiert, da die Studierenden die Anwendungen in Kalenderwoche 30 aufgrund der anstehenden Klausur nur noch drei Tage lang nutzen konnten.
Es zeigt sich deutlich, dass die Klausurvorbereitung mit den E-Learning-Anwendungen zunächst eher langsam beginnt. Erst ein bis zwei Wochen vor der Modulabschlussprüfung (Kalenderwoche 28 und 29) steigen die durchschnittlichen Nutzungszeiten für alle drei Anwendungsvarianten deutlich an.

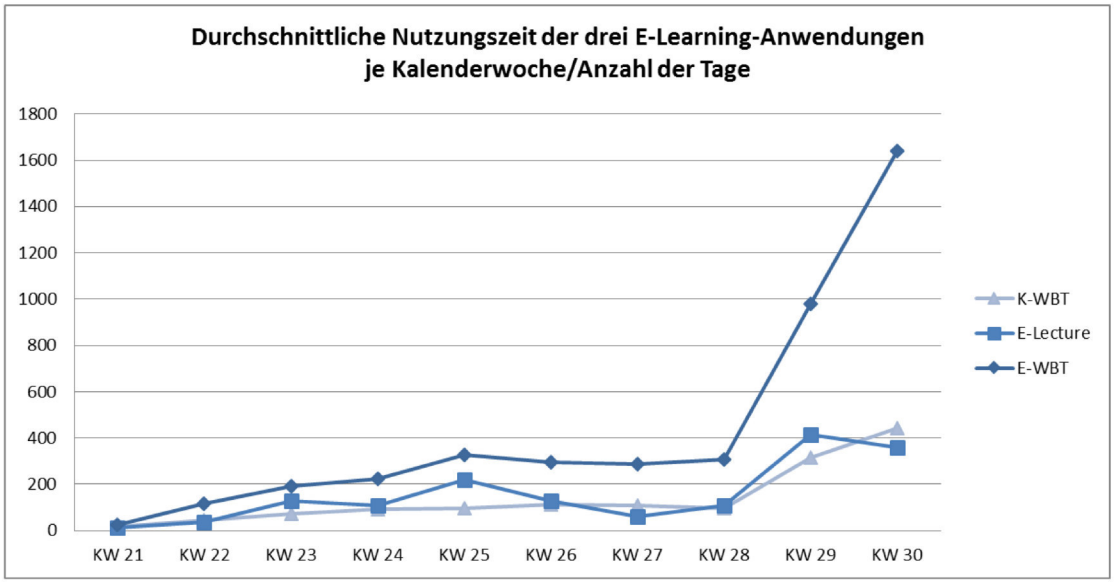

Abbildung 52. Durchschnittliche Nutzungszeit (in Sek.) der E-Learning-Anwendungen getrennt für die Kalenderwochen 21 bis 30 relativiert durch die Anzahl der möglichen Nutzungstage

Es zeigt sich insbesondere, dass sich die Nutzung der WBTs in Woche 30 noch einmal intensiviert, während sie für die E-Lectures etwas sinkt. Dies bestätigt erneut die Beobachtungen aus Studie I, dass die E-Lectures und WBTs zu unterschiedlichen Zwecken genutzt werden: Die E-Lectures vermehrt für das Nachholen verpasster Sitzungen und die WBTs für das konkrete Nachschlagen oder Wiederholen mit Übungsaufgaben. Dies wird sicherlich in den letzten Tagen vor der MAP wichtiger für die Studierenden geworden sein,

während das Nachholen verpasster Sitzungen hier weitestgehend abgeschlossen gewesen sein sollte.

7.4.2 Lernleistung

Zum Vergleich der Lernleistung der Nutzergruppen können zwei verschiedene Maße herangezogen werden: Zum einen das Ergebnis der Studierenden innerhalb der Probeklausur, die kurz vor der Modulabschlussprüfung zur Verfügung gestellt wurde und an der die Studierenden auf freiwilliger Basis teilnehmen konnten, zum anderen ihr Abschneiden in der Modulabschlussprüfung selbst. Im Folgenden wird zunächst die Probeklausur (getrennt für die verschiedenen Themen sowie insgesamt) betrachtet, bevor auf die beiden Klausurteile der Modulabschlussprüfung eingegangen wird.

7.4.2.1 Ergebnisse der Probeklausur vor der MAP

Hypothese 12a: Studierende, die angeben, sich hauptsächlich mit den E-Learning-Anwendungen auf den Teil A der abschließenden Klausur vorbereitet zu haben, schneiden signifikant besser in der Probeklausur ab als Studierende, die angeben, sich hauptsächlich mit Offline-Material (z.B. dem Vorlesungsskript) vorbereitet zu haben.
➔ **E-Learning > Offline-Material**

Hypothese 12b: Die Nutzergruppe, die sich am längsten mit den E-Learning-Anwendungen auf den Teil A der abschließenden Klausur vorbereitet hat, schneidet signifikant besser in der Probeklausur ab.

Tabelle 103 zeigt die Ergebnisse der sechs Nutzergruppen hinsichtlich der zwölf Themen in der Probeklausur. Die Ergebnisse der Gruppen „Literatur“ und „Andere“ sind hier zwar deskriptiv abgetragen, werden in den inferenzstatistischen Analysen allerdings erneut aufgrund der geringen Stichprobengrößen nicht berücksichtigt.
In den Themenbereichen „Konditionierung“, „Unterrichten“ und „Pädagogische Diagnostik“ konnten die Studierenden jeweils maximal acht Punkte erreichen, in allen anderen Themenbereichen waren es lediglich vier Punkte. Insgesamt resultiert so eine Gesamtpunktzahl von 60 Punkten für die Probeklausur.
Die mittleren Punktwerte zwischen den Themenbereichen sind aufgrund der unterschiedlichen Gewichtung nur durch eine Umrechnung in Prozente vergleichbar. Betrachtet man die einzelnen Mittelwerte jedoch einmal im Überblick, ist bereits ersichtlich, dass die Studierenden im Schnitt zumeist mehr als 75% der Punktzahl erreichten (mehr als drei bzw. sechs Punkte).

Tabelle 103
Deskriptive Statistik zu den Testleistungen der Nutzergruppen bezüglich der zwölf Themen der Probeklausur

Probeklausur	Gedächtnis I			Gedächtnis II			Gedächtnis III			Gedächtnis IV		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Vorlesungsskript	107	3.12	1.06	109	2.55	0.99	108	3.09	0.92	109	2.80	1.03
Literatur	3	3.33	1.16	3	**3.67**	0.58	3	3.33	0.58	3	2.00	0.87
E-Lecture	29	**3.52**	0.99	29	2.95	0.89	29	3.34	1.01	29	2.83	1.16
K-WBT	65	3.26	0.85	65	2.93	0.87	65	3.38	0.78	65	3.08	0.93
E-WBT	191	3.47	0.69	190	3.20	0.75	192	**3.58**	0.67	190	**3.34**	0.82
Andere	7	2.14	1.22	7	1.93	0.89	6	3.17	0.75	7	1.71	0.86
Gesamt	402	3.32	0.88	403	2.94	0.90	403	3.39	0.81	403	3.08	0.97
	Entwicklungspsych.			**Konditionierung**			**Modelllernen**			**Denken & Probleml.**		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Vorlesungsskript	106	2.96	1.06	107	6.31	1.40	109	3.40	0.71	111	2.23	1.15
Literatur	3	3.10	0.85	3	**7.33**	0.58	3	3.03	0.91	3	**3.33**	0.76
E-Lecture	28	2.99	1.31	29	6.69	1.45	29	3.51	0.57	28	2.47	1.14
K-WBT	65	3.17	0.95	64	6.74	1.21	65	3.60	0.62	64	2.38	1.09
E-WBT	191	**3.47**	0.70	188	7.15	0.90	191	**3.80**	0.36	188	2.84	1.00
Andere	6	2.13	1.23	6	5.57	2.05	6	2.98	0.93	7	1.66	0.54
Gesamt	399	3.23	0.94	397	6.80	1.22	403	3.62	0.58	401	2.56	1.10
	Intelligenz			**Unterrichten**			**Lernen mit Medien**			**Pädagogische Diagn.**		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Vorlesungsskript	108	2.92	0.89	104	5.67	1.60	106	2.82	0.83	107	5.36	1.59
Literatur	3	3.20	0.85	3	6.30	1.76	3	3.00	0.36	3	6.00	0.66
E-Lecture	27	3.18	1.08	28	5.71	1.60	28	3.01	0.83	28	5.58	1.77
K-WBT	64	2.97	0.92	65	5.94	1.31	64	3.17	0.62	65	5.57	1.74
E-WBT	190	**3.23**	0.74	188	**6.57**	1.17	192	**3.32**	0.66	188	**6.45**	1.36
Andere	6	2.25	1.22	7	5.96	1.52	6	2.60	0.62	6	4.48	1.10
Gesamt	398	3.09	0.86	395	6.15	1.41	399	3.13	0.74	397	5.92	1.59

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung

Das schlechteste Ergebnis erlangten die Studierenden im Themenbereich „Denken und Problemlösen“ (*M* = 2.56 Punkte – entspricht 64%), das beste Ergebnis im Themenbereich „Modelllernen“ (*M* = 3.62 Punkte – entspricht 90.5%).

Vergleicht man die Ergebnisse zu allen zwölf Themenbereichen zwischen den vier Nutzergruppen „Vorlesungsskript“, „elaboriertes (E-)WBT“, „kurzes (K-)WBT“ und „E-Lecture“ mittels einfaktoriellen Varianzanalysen bzw. Brown-Forsythe-Tests (vollständige Ergebnisse siehe Anhang D, Tabelle D3), ergeben sich für zehn der zwölf Themen signifikante Unterschiede zwischen der Gruppe „elaboriertes (E-)WBT“ und „Vorlesungsskript“. Der Effekt dieser Unterschiede ist jeweils als mindestens mittelstark zu beurteilen und fällt zugunsten der Gruppe „E-WBT“ aus. Zum Themenbereich „Diagnostik“ schneidet die Gruppe „kurzes (K-)WBT“ zudem signifikant schlechter ab als die Gruppe „elaboriertes (E-)WBT“. Lediglich für die Themenbereiche „Gedächtnis I“ und „Intelligenz“ ergibt sich kein mehrheitlicher signifikanter Unterschied zwischen den vier Gruppen.

Um die Hypothesen 12a und 12b zu überprüfen, soll das Gesamtergebnis in der Probeklausur betrachtet werden.

Tabelle 104
Deskriptive Statistik zur Lernleistung der Nutzergruppen in der Probeklausur

Gesamtscore	*n*	*M*	*SD*	*%*
Vorlesungsskript	104	43.43	9.16	72.4%
Literatur	3	47.63	6.96	79.4%
E-Lecture	27	46.32	9.68	77.2%
K-WBT	63	46.22	7.59	77.0%
E-WBT	188	**50.47**	5.75	**84.1%**
Andere	6	36.83	9.16	61.4%
Gesamt	391	47.39	8.10	79.0%

Anmerkung. n = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung, *%* = Prozentualer Anteil richtiger Lösungen

Tabelle 104 gibt die mittleren Ergebnisse für die sechs Nutzergruppen wieder. Durchschnittlich erreichen die Studierenden 79% der Gesamtpunktzahl in der Probeklausur. Generell ist das Gesamtergebnis folglich als erfreulich zu beurteilen. Dabei schneidet die Gruppe „Andere" mit ca. 61% am schlechtesten, die Gruppe „E-WBT" mit 84% am besten ab (siehe auch Abbildung 53). Vergleichbar gut können die Gruppen „E-Lecture" und „K-WBT" die Probeklausur abschließen. Die Gruppe „Vorlesungsskript" erreicht ein schlechteres Ergebnis, die Gruppe „Literatur" ein etwas besseres Ergebnis als diese beiden Gruppen.

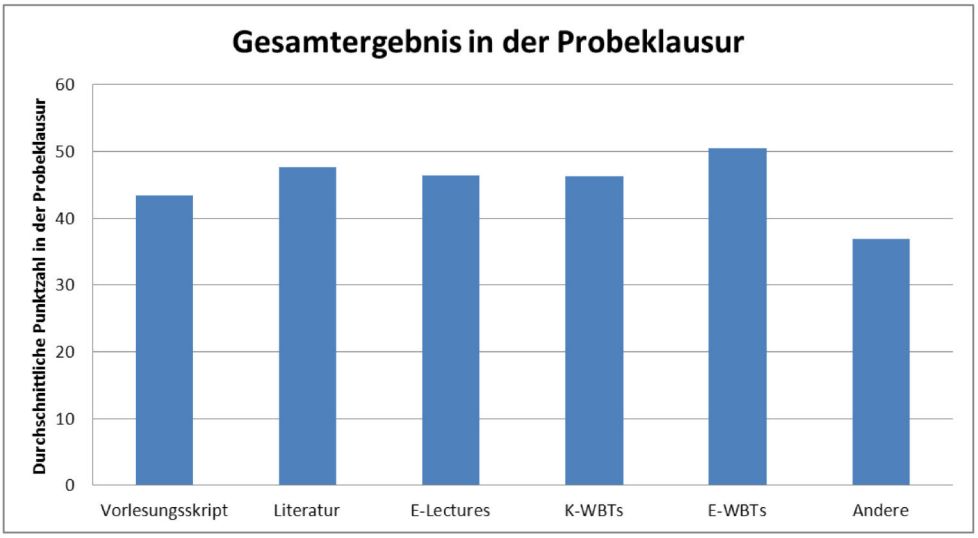

Abbildung 53. Durchschnittliches Gesamtergebnis der sechs Nutzergruppen in der Probeklausur

Brown-Forsythe-Tests (siehe Tabelle 105) ergeben in der Folge erwartungsgemäß signifikante Unterschiede mit mittelstarken bis starken Effekten zugunsten der Gruppe „elaboriertes WBT". Diese Gruppe schneidet signifikant besser in der Probeklausur ab als die Studierenden, die hauptsächlich mit dem Vorlesungsskript oder dem kurzen (K-)WBT gelernt haben.

Tabelle 105
Ergebnisse der Mittelwertsanalysen zum Vergleich der Gesamttestleistung in der Probeklausur der Vorlesungs-, der E-Lecture- und der beiden WBT-Gruppen

Probeklausur	Brown-Forsythe-Test				Posthoc
	F	*df*	*p*	*Eta²*	
Zufallsauswahl 1	7.52	117.53,3	<.01*	.11	E > VLSk, K
Zufallsauswahl 2	7.38	116.66,3	<.01*	.11	
Zufallsauswahl 3	6.94	119.47,3	<.01*	.10	
Zufallsauswahl 4	7.06	116.57,3	<.01*	.11	
Zufallsauswahl 5	6.92	115.69,3	<.01*	.10	
Zufallsauswahl 6	6.49	118.51,3	<.01*	.10	

Anmerkung. F = Teststatistik des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-WBT-Gruppe sowie zwei Zufallsziehungen aus der Vorlesungsgruppe vorgenommen und miteinander kombiniert ausgewertet.

Die erneut sehr streng formulierte Hypothese 12a muss allerdings zurückgewiesen werden: Die Nutzung nicht aller E-Learning-Materialien führt zu besseren Ergebnissen in der Probeklausur als die Vorbereitung mit dem Vorlesungsskript.

Hypothese 12b („Die Nutzergruppe, die sich am längsten mit den E-Learning-Anwendungen auf den Teil A der abschließenden Klausur vorbereitet hat, schneidet signifikant besser in der Probeklausur ab.“) muss nach den Ergebnissen zur Nutzungszeit folgendermaßen angepasst werden: „Die Nutzergruppe „E-WBT“ schneidet signifikant besser in der Probeklausur ab als die Nutzergruppen „K-WBT“ und „E-Lecture“.“ Diese Hypothese muss ebenfalls zurückgewiesen werden, da die E-Lecture-Nutzer nicht schlechter abschneiden als die Nutzer der elaborierten (E-)WBTs.

7.4.2.2 Ergebnisse der Modulabschlussprüfung

Hypothese 13a: Studierende, die angeben, sich hauptsächlich mit den E-Learning-Anwendungen auf den Teil A der abschließenden Klausur vorbereitet zu haben, schneiden hier signifikant besser ab als Studierende, die angeben, sich hauptsächlich mit Offline-Material (z.B. dem Vorlesungsskript) vorbereitet zu haben.
➔ **E-Learning > Offline-Material**

Hypothese 13b: Die Nutzergruppe, die sich am längsten mit den E-Learning-Anwendungen auf den Teil A der abschließenden Klausur vorbereitet hat, schneidet hier signifikant besser ab.

Die Ergebnisse in der Modulabschlussprüfung werden getrennt für die beiden Klausurteile A und B betrachtet. Die Studierenden konnten sich mit den Lernmaterialien lediglich auf den Teil A der Klausur vorbereiten. Klausurteil B überprüfte die Inhalte einer anderen Vorlesung und kann so als Kontrollergebnis zu Rate gezogen werden. Da in beiden Klausurteilen jeweils 34 Punkte erreicht werden konnten, sind die in beiden Klausurteilen erreichten Punkte direkt miteinander vergleichbar. Insgesamt resultiert für die

Modulabschlussprüfung eine Gesamtpunktzahl von 68 Punkten und die Benotung erfolgte auf der Notenskala 0 bis 15.

In Tabelle 106 sind die Ergebnisse der beiden Klausurteile sowie die Gesamtpunktzahl und die Gesamtnote in der Modulabschlussprüfung für alle sechs Nutzergruppen abgetragen.

Tabelle 106
Deskriptive Statistik zu den Leistungen der Nutzergruppen in den beiden Teilen der MAP sowie der MAP insgesamt

MAP	Punkte Teil A			Punkte Teil B			Gesamtpunkte			Gesamtnote		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Vorlesungsskript	123	22.59	4.72	123	22.67	5.57	123	45.25	9.51	123	8.40	3.06
Literatur	3	20.92	0.52	3	20.67	4.16	3	41.58	4.45	3	7.33	1.53
E-Lecture	41	22.83	4.18	41	23.67	5.09	41	46.50	8.42	41	8.73	2.72
K-WBT	78	22.72	4.62	78	22.76	5.22	78	45.48	9.16	78	8.59	2.90
E-WBT	227	**24.58**	4.13	227	**24.29**	5.01	227	**48.87**	8.15	227	**9.57**	2.64
Andere	8	20.47	4.82	8	20.06	6.47	8	40.53	9.94	8	6.75	3.11
Gesamt	480	23.53	4.48	480	23.48	5.27	480	47.00	8.89	480	8.98	2.86

Anmerkung. *n* = Stichprobengröße, *M* = Mittelwert, *SD* = Standardabweichung; MAP = Modulabschlussprüfung, Teil A = Inhalte der Wintersemestervorlesung, Teil B = Inhalte der Sommersemestervorlesung

Insgesamt erreichten die Studierenden in beiden Klausurteilen jeweils *M* = 23.5 Punkte. Dies entspricht 69.1% der Gesamtpunktzahl und verdeutlicht die vergleichbare Schwierigkeit der beiden Klausurteile. Im Mittel erreichten die Studierenden damit *M* = 9.0 Notenpunkte, was einer befriedigenden Leistung (3+) entspricht.

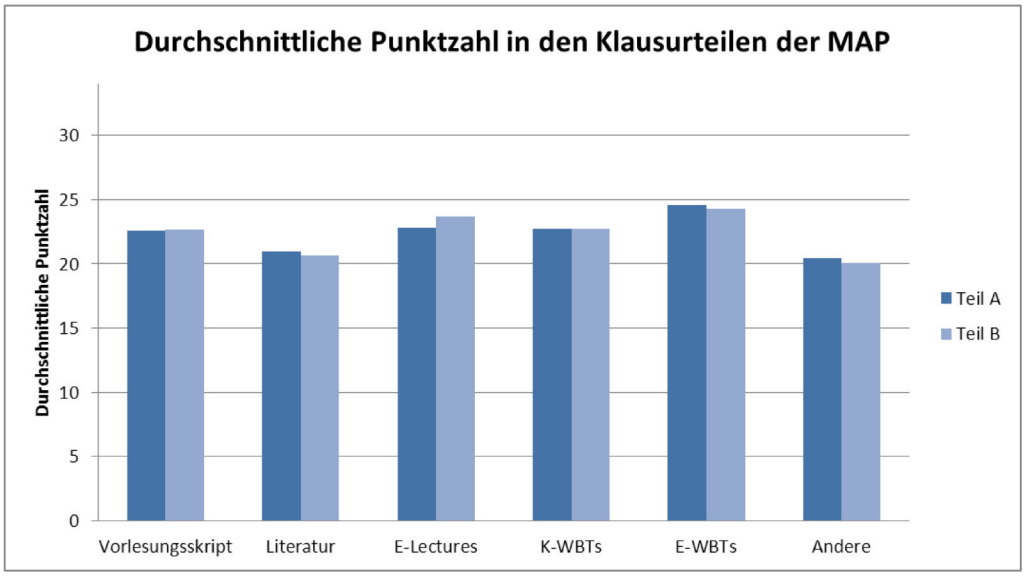

Abbildung 54. Durchschnittliche Punktzahl der Nutzergruppen in den beiden Klausurteilen der Modulabschlussprüfung

Abbildung 54 veranschaulicht die Ergebnisse der sechs Nutzergruppen in den beiden Klausurteilen. Hier ist zu sehen, dass alle Gruppen hinsichtlich der beiden Klausurteile vergleichbar gut abschneiden. Auch die Unterschiede zwischen den Gruppen sind nicht allzu groß. In beiden Teilen erreichen alle Nutzergruppen eine Punktzahl zwischen 20 und

25 Punkten, wobei die Gruppe „E-WBT“ erneut für beide Teile den höchsten, die Gruppe „Andere“ den niedrigsten mittleren Punktwert erreicht.

Tabelle 107
Ergebnisse der Mittelwertsanalysen zum Vergleich der Leistungen der Vorlesungs-, E-Lecture- und der beiden WBT-Gruppen in den beiden Teilen der MAP

Punkte Teil A	**ANOVA / Brown-Forsythe-Test**				**Posthoc**
	F	***df***	***p***	***Eta²***	
Zufallsauswahl 1	3.00	255,3	.03*	.03	-
Zufallsauswahl 2	3.51	255,3	.02*	.04	E > K
Zufallsauswahl 3	3.15	254,3	.03*	.04	-
Zufallsauswahl 4	3.04	216.65,3	.03*	.04	-
Zufallsauswahl 5	3.51	214.19,3	.02*	.04	E > VLSk, K
Zufallsauswahl 6	3.18	249,3	.03*	.04	E > VLSk
Punkte Teil B					
Zufallsauswahl 1	0.66	255,3	.58	-	-
Zufallsauswahl 2	0.97	255,3	.41	-	-
Zufallsauswahl 3	2.08	254,3	.10	-	-
Zufallsauswahl 4	2.05	250,3	.11	-	-
Zufallsauswahl 5	2.54	250,3	.06	-	-
Zufallsauswahl 6	4.14	249,3	<.01*	.05	E > VLSk

Anmerkung. F = Teststatistik der Varianzanalyse bzw. des Brown-Forsythe-Tests, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-WBT-Gruppe sowie zwei Zufallsziehungen aus der Vorlesungsgruppe vorgenommen und miteinander kombiniert ausgewertet.

Tabelle 107 gibt die Ergebnisse der einfaktoriellen Varianzanalysen bzw. Brown-Forsythe-Tests zum Vergleich der Leistungen in Teil A und Teil B zwischen den vier vergleichbaren Nutzergruppen (Vorlesungsskript, E-Lecture, elaboriertes (E-)WBT und kurzes (K-)WBT) wieder.

Die Ergebnisse der Gruppen hinsichtlich Teil A unterscheiden sich in drei der sechs Analysen signifikant, dies allerdings nur mit kleinem bis mittelstarkem Effekt. In zwei der sechs Analysen schneidet die E-WBT-Gruppe signifikant besser ab als die K-WBT-Gruppe, in zweien erreicht sie ein besseres Ergebnis als die Vorlesungsskriptgruppe.

Trotz der tendenziell hypothesenkonformen Ergebnisse muss Hypothese 13a folglich zurückgewiesen werden, da sich die E-Learning-Nutzer nur in zwei – und damit nicht in der Mehrzahl der Analysen – gegenüber der Gruppe „Vorlesungsskript“ durchsetzen können. Dies gilt auch für die Hypothese 13b („Die Nutzergruppe, die sich am längsten mit den E-Learning-Anwendungen auf den Teil A der abschließenden Klausur vorbereitet hat, schneidet hier signifikant besser ab.“). Diese ist erneut umzuformulieren: „Die Nutzergruppe „E-WBT“ schneidet signifikant besser in Teil A der Modulabschlussprüfung ab als die Nutzergruppen „K-WBT“ und „E-Lecture“.“ Auch diese Hypothese muss trotz des tendenziellen Vorteils der E-WBT-Gruppe gegenüber der Gruppe „K-WBT“ zurückgewiesen werden.

Zu Teil B der Klausur sind die Ergebnisse eindeutiger. Hier ergibt sich lediglich für eine Zufallsauswahl eine Signifikanz zwischen den Gruppen. In Teil B der Klausur schneiden alle Nutzergruppen folglich vergleichbar gut ab. Die tendenziellen Vorteile der E-WBT-Gruppe

hinsichtlich Teil A der Klausur kommen dementsprechend nicht nur dadurch zustande, dass diese Gruppe von Studierenden sich generell durch bessere Leistungen auszeichnet.

Ergänzend soll auch ein Blick auf das Gesamtergebnis der Modulabschlussprüfung geworfen werden.
Die Mittelwerte und Standardabweichungen wurden bereits in Tabelle 106 abgetragen und zeigten, dass die Gruppen im Schnitt zwischen 40.5 und 49 Punkten (von 68 möglichen) in der MAP erreichten. Die Durchschnittsnote lag zwischen etwa 7 (noch befriedigend) und 10 Notenpunkten (noch gut).
Betrachtet man die Ergebnisse der einfaktoriellen Varianzanalysen in Tabelle 108, ergibt sich für die Auswertung der Gesamtergebnisse ein ähnliches Bild wie bereits für die Auswertung des Klausurteils A: tendenziell zeigt sich ein Vorteil für die Nutzer des „elaborierten (E-)WBTs", aber auch hier ist keine Mehrzahl an Signifikanzen auszumachen und dementsprechend kein eindeutiger Vorteil einer der Nutzergruppen zu identifizieren.

Tabelle 108
Ergebnisse der Mittelwertsanalysen zum Vergleich der Gesamtpunkte sowie der Note der Vorlesungs-, E-Lecture- und der beiden WBT-Gruppen in der MAP

Gesamtpunkte	**ANOVA**				**Posthoc**
	F	***df***	***p***	***Eta²***	
Zufallsauswahl 1	1.52	255,3	.21	-	-
Zufallsauswahl 2	2.09	255,3	.10	-	-
Zufallsauswahl 3	2.88	254,3	.04*	.03	E > K
Zufallsauswahl 4	2.64	250,3	<.05*	.03	E > VLSk
Zufallsauswahl 5	3.34	250,3	.02*	.04	E > VLSk
Zufallsauswahl 6	4.19	249,3	<.01*	.05	E > VLSk, K
Gesamtnote					
Zufallsauswahl 1	1.33	255,3	.27	-	-
Zufallsauswahl 2	1.81	255,3	.15	-	-
Zufallsauswahl 3	2.57	254,3	.06	-	-
Zufallsauswahl 4	2.65	250,3	<.05*	.03	E > VLSk
Zufallsauswahl 5	3.26	250,3	.02*	.04	E > VLSk
Zufallsauswahl 6	4.09	249,3	<.01*	.05	E > VLSk

Anmerkung. F = Teststatistik der Varianzanalyse, *df* = Freiheitsgrade, *p* = Signifikanz, *Eta²* = Effektstärke; Es wurden jeweils drei Zufallsziehungen aus der E-WBT-Gruppe sowie zwei Zufallsziehungen aus der Vorlesungsgruppe vorgenommen und miteinander kombiniert ausgewertet.

7.4.3 Einschätzung zentraler Eigenschaften von E-Learning-Anwendungen

Je nachdem, mit welchem Lernmaterial sie sich hauptsächlich auf die MAP vorbereitet hatten, sollten die Studierenden in der letzten Befragung auf einer Skala von 1 (sehr unwichtig) bis 10 (sehr wichtig) einige Eigenschaften von E-Learning-Anwendungen (nach dem ARCS-Modell) hinsichtlich ihrer Wichtigkeit beurteilen.
Dabei wurden den K- und E-WBT-Nutzern 20 Items dargeboten, die sie aus den elaborierten (E-)WBTs kennen oder bei den kurzen (K-)WBTs vermisst haben könnten (siehe Tabelle 109). Die E-Lecture-Nutzer beurteilten ebenfalls 16 dieser 20 Items (die restlichen

erschienen hier nicht sinnvoll). Alle anderen Nutzergruppen wurden hierzu nicht befragt, da für sie keine intensive E-Learning-Nutzung angenommen wurde.

Tabelle 109 gibt die durchschnittlichen Einschätzungen der Studierenden in der Reihenfolge von „sehr wichtig“ bis „sehr unwichtig“ wieder. Die Einschätzungen der Studierenden zu den einzelnen Eigenschaften variieren sehr stark zwischen M = 4.7 und M = 9.5. Am wenigsten wichtig ist den Studierenden, dass „Professor Glowalla“ selbst als Lernbegleiter in die E-Learning-Anwendung integriert ist, am wichtigsten ist ihnen das Abschlussfeedback nach dem Übungstest.

Tabelle 109
Durchschnittliche Einschätzung der Wichtigkeit von 20 Eigenschaften von E-Learning-Anwendungen sowie die Angabe, zu welcher ARCS-Komponente diese Eigenschaft nach Keller (2010) gehört

ARCS	Wichtigkeit...	*n*	*M*	*SD*
C	... des Abschlussfeedbacks am Ende des Übungstests	299	**9.52**	1.18
C	... von Übungsaufgaben am Ende des WBTs/die man selbst bearbeiten kann (mit Rückmeldung über den eigenen Lernstand)	333	9.49	1.34
C	... von informierendem Feedback nach jeder Übungsaufgabe	301	9.28	1.53
R	... von Anwendungsbeispielen zu theoretischen Inhalten	347	8.92	1.61
R	... von Hinweisen auf Relevanz des Gelernten für den späteren Berufsalltag	345	8.58	1.90
C	... der Angabe der Gesamtseitenzahl auf jeder Seite des WBTs/der Gesamtdauer der E-Lecture	341	8.22	2.25
S	... von Lernspielen	333	8.09	2.50
C	... eines "anklickbaren" Inhaltsverzeichnisses am Anfang des WBTs/links unterhalb der Videoaufzeichnung	349	7.92	2.57
C	... der Möglichkeit, durch dieses flexible Inhaltsverzeichnis einen eigenen Einstieg/Lernweg zu wählen	346	7.66	2.70
C	... eines Inhaltsverzeichnisses auf jeder Seite des WBTs/der E-Lecture	303	7.46	2.68
S	... einer "Belohnung" am Ende des WBTs/jeder E-Lecture	339	6.99	2.87
A	... von ergänzenden Videos	337	6.74	2.60
A	... von abwechslungsreichen Bildern (auf den Präsentationsfolien)	340	6.61	2.63
A	... von Denkfragen	336	6.21	2.46
R	... der Angabe der Lernziele des jeweiligen WBTs/der jeweiligen E-Lecture	342	6.20	2.82
R	... von persönlicher Ansprache innerhalb des WBTs/durch Prof. Glowalla	341	6.07	2.78
A	... einer Einstiegsgeschichte zu Beginn des WBTs/der E-Lecture	341	5.85	2.71
S	... von weiterführenden Informationen	330	5.79	2.65
C	... einer Einstiegsfolie/eines Einstiegs mit Tipps, wie man sinnvoll mit dem WBT/der E-Lecture lernen kann	344	4.74	2.84
R	... von "Professor Glowalla" als Lernbegleiter	308	4.68	2.83

Anmerkung. n = Stichprobengröße, M = Mittelwert, SD = Standardabweichung; A = Aufmerksamkeit, R = Relevanz, C = Erfolgszuversicht, S = Zufriedenheit

Die Abstufung der Wichtigkeitseinschätzungen wird durch Abbildung 55 noch einmal visualisiert. Interessant ist, dass auf den ersten drei Plätzen jeweils Aspekte liegen, die mit den Übungstests verbunden sind: das Abschlussfeedback, die Übungsaufgaben selbst und das informierende Feedback direkt nach den Übungsaufgaben – alles Aspekte, die zu der Komponente „Erfolgszuversicht“ nach dem ARCS-Modell von Keller (2010) gehören. Danach folgen Aspekte, die das Relevanzgefühl stärken sollen: Anwendungsbeispiele und Hinweise auf die Relevanz des Gelernten.

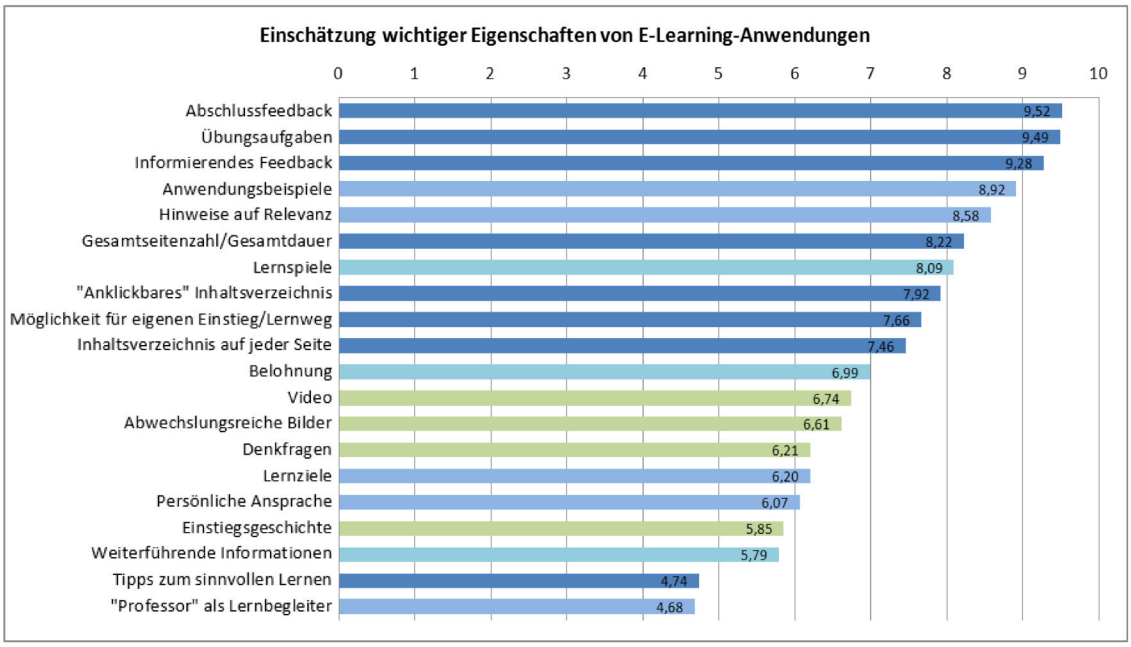

Abbildung 55. Einschätzung der Wichtigkeit der Eigenschaften von E-Learning-Anwendungen
Anmerkung. Bedeutung der Farben: hellgrün = Komponente „Aufmerksamkeit“, hellblau = Komponente „Relevanz“, dunkelblau = Komponente „Erfolgszuversicht“ und türkis = Komponente „Zufriedenheit“

Am deutlichsten abgeschlagen sind die Aspekte „Tipps zum sinnvollen Lernen“ (Komponente „Erfolgszuversicht“) und „Professor“ als Lernbegleiter (Komponente „Relevanz“). Hierbei handelt es sich um Punkte, die bei der vorliegenden Stichprobe vermutlich weniger wichtig sind als ggf. bei jüngeren oder unerfahreneren Lernenden. Weniger wichtig werden von den Studierenden jedoch auch weiterführende Informationen (Komponente „Zufriedenheit“) eingeschätzt, was mit der Motivation der Lernenden zusammenhängen könnte, lediglich die Inhalte für die Klausur zu lernen und nicht unbedingt aus Eigeninteresse heraus in die Tiefe der Materie einzutauchen. Auch solche „Zugaben“ wie die persönliche Ansprache innerhalb der E-Learning-Anwendung (Kompoente „Relevanz“), die Angabe von Lernzielen (Komponente „Erfolgszuversicht“), die Stimulation durch Denkfragen, abwechslungsreiche Bilder, Videos (jeweils Komponente „Aufmerksamkeit“) oder eine Belohnung am Ende des elaborierten (E-)WBTs (Komponente „Zufriedenheit“) scheinen zwar nicht abgelehnt, aber auch nicht als besonders wichtig eingestuft zu werden.

Die Farbverteilung in Abbildung 55 unterstützt dabei, die Aspekte gebündelt für die vier ARCS-Komponenten zu bewerten. Hier zeigt sich, dass die Aspekte in dunkelblau (Komponente „Erfolgszuversicht“) zumeist relativ weit oben angesiedelt sind. Insbesondere die Übungstests, aber auch die Inhaltsverzeichnisse und die dadurch gegebene flexible Bearbeitung der elaborierten (E-)WBTs werden von den Studierenden als sehr wichtig eingeschätzt. Im Gegensatz dazu befinden sich die Aspekte zur Komponente „Aufmerksamkeit“ (hellgrün) allesamt im unteren Bereich. Diese sind den Studierenden

weniger wichtig. Die Aspekte zu den Komponenten „Relevanz“ (hellblau) und „Zufriedenheit“ (türkis) sind über das gesamte Ranking gleichmäßig verteilt. Zu diesen Komponenten gibt es folglich Teilaspekte, die den Studierenden besonders wichtig sind und Teilaspekte, die weniger wichtig erscheinen.

7.4.4 Zusammenfassung und Diskussion

7.4.4.1 Nutzungszeit und -häufigkeit

Von den 784 Studierenden, die die E-Learning-Anwendungen für ihre Vorbereitung auf die MAP zur Verfügung gestellt bekamen, haben 639 diese genutzt. Dies entspricht einer Nutzungsquote von ca. 80% und mehr, da nicht alle der 784 Studierenden auch an der MAP teilgenommen haben. Betrachtet man in diesem Zusammenhang die Nutzungsquote in Studie I (43%)[55], ist die hier beobachtete Nutzung sehr erfreulich. Darüber hinaus ist zu bedenken, dass die Studierenden sicherlich von „älteren Semestern“ den Tipp erhielten, sich mit dem Vorlesungsskript und den eigenen Mitschriften auf die Modulabschlussprüfung vorzubereiten. Dass die Studierenden dennoch in so großer Zahl von dieser Strategie abwichen – oder ihr Material zumindest durch die E-Learning-Anwendungen ergänzten – ist bemerkenswert. Es handelt sich hier allerdings um die ausgewählte Studierendengruppe derer, die bereits im Wintersemester rege an den E-Learning-Evaluationen teilgenommen hatten. Diese Studierenden werden grundsätzlich eine tendenziell positive Einstellung gegenüber der Nutzung von E-Learning-Anwendungen haben.

Hinzu kommen erfreuliche Gesamt-Nutzungszeiten: Die kurzen (K-)WBTs und E-Lectures wurden von den Studierenden insgesamt jeweils ca. 2000 Stunden genutzt, die elaborierten (E-)WBTs sogar 5600 Stunden. Die intensive Nutzung der E-Learning-Anwendungen durch die Studierenden zeigt, dass sich die große Mühe, die in die Erstellung der E-Learning-Anwendungen investiert wurde – insbesondere in die Erstellung der elaborierten (E-)WBTs – gelohnt hat. Dass der Mehraufwand für die Erstellung der E-WBTs sinnvoll ist, zeigt sich auch im Vergleich der Nutzungszeiten und -häufigkeiten der drei E-Learning-Nutzungsgruppen. Die elaborierten WBTs werden signifikant länger und häufiger genutzt als die kurzen WBTs und E-Lectures. Hinsichtlich der Nutzungshäufigkeit ergibt sich darüber hinaus ein signifikanter Unterschied zwischen den kurzen (K-)WBTs und den E-Lectures zugunsten der kurzen WBTs. Betrachtet man die Nutzungszeiten und -häufigkeiten im Vergleich, zeigt sich auch hier ein vergleichbares Bild zu Studie I: Die WBTs wurden häufiger kurz genutzt, die E-Lectures seltener länger. Dies bestätigt erneut

[55] Wobei es sich hier auch um 43% aller zum Modul angemeldeten Studierenden handelt. Von der Gesamtzahl der Studierenden werden ebenfalls nicht alle auch zur MAP angetreten sein, weshalb die wahre Beteiligungsquote höher liegt.

die Annahme aus Studie I, dass die WBTs während der Klausurphase vermehrt zum Nachschlagen und Üben genutzt wurden, während die E-Lectures vornehmlich zum Nachholen ganzer Vorlesungssitzungen bearbeitet wurden.

Ähnliches lässt sich auch aus den Nutzungszeiten je Thema und Kalenderwoche für die drei verschiedenen E-Learning-Anwendungen schließen. Hier zeigte sich, dass die elf elaborierten WBTs zu den einzelnen Themen in der Mehrzahl zwischen ca. 30 und 50 Minuten genutzt wurden, während die E-Lectures und kurzen WBTs zumeist nur zwischen 8 und 16 Minuten bearbeitet wurden. Sicherlich sind die K-WBTs vom Umfang her deutlich kürzer als die E-WBTs – dies kann den extremen Unterschied hinsichtlich der Nutzungszeiten allerdings nicht vollständig erklären.

Dabei zeigt sich sehr deutlich, dass die Anwendungen zwar bereits ab dem Zeitpunkt der Bereitstellung genutzt werden – jedoch nur in sehr geringem Maße. Erst etwa eine Woche vor der Modulabschlussprüfung steigt die durchschnittliche Nutzungszeit deutlich an. Diese erhöht sich für die WBTs bis zum Zeitpunkt der MAP, während sie für die E-Lectures ein paar Tage vor der MAP etwas absinkt. Hieraus lässt sich schließen, dass das Nachholen verpasster Sitzungen ein paar Tage vor der Modulabschlussprüfung keine große Rolle mehr spielt. So kurz vor der Prüfung werden die Studierenden hauptsächlich noch einmal nicht-verstandene Inhalte nachgeschlagen und insbesondere die Inhalte wiederholend geübt haben, um sie zu festigen.

7.4.4.2 Lernleistung der Nutzergruppen

Im Rahmen der freiwilligen Nutzung der E-Learning-Varianten im Sommersemester war es von Interesse, welche Nutzergruppe das beste Ergebnis in der Probeklausur und der Modulabschlussprüfung erreichen konnte. Aufgrund der sehr geringen Stichprobengröße konnten dabei die Gruppen „Literatur“ und „Andere Materialien“ nicht in die Analysen einbezogen werden, so dass ein Vergleich der Lernleistung lediglich für die Gruppen „elaboriertes (E-)WBT“, „kurzes (K-)WBT“, „E-Lecture“ und „Vorlesungsskript“ erfolgte.

Hinsichtlich der Probeklausur ergaben sich in der Mehrheit der Analysen signifikante Unterschiede mit mittelstarken bis starken Effekten zugunsten der Gruppe „elaboriertes (E-)WBT“. Diese Gruppe erreichte ein signifikant besseres Probeklausurergebnis als die Gruppen „Vorlesungsskript“ und „kurzes (K-)WBT“. Im Rahmen der Modulabschluss-prüfung erreichen die Nutzergruppen jedoch hinsichtlich beider Klausurteile ungefähr vergleichbar gute Ergebnisse. Zwar ergeben sich für die Gesamtpunktzahl in der MAP in vier der sechs Analysen signifikante Unterschiede zugunsten der E-WBT-Nutzer, die Posthoc-Analysen ergeben jedoch kein mehrheitlich eindeutiges Bild: Die Gruppe „elaboriertes (E-)WBT“ schneidet in zwei von sechs Analysen signifikant besser ab als die Gruppe „kurzes (K-)WBT“, in drei Analysen besser als die Vorlesungsskript-Gruppe.

In dem Falle, dass sich Signifikanzen in Teil A der Klausur ergeben hätten, wäre es interessant gewesen zu analysieren, ob die gleichen Gruppen auch in Teil B besser abgeschnitten haben. Dies hätte dafür gesprochen, dass die Nutzer bestimmter Lernangebote durchgängig motivierter sind oder effektiver lernen. Ein Ausbleiben von Unterschieden in Teil B hingegen hätte dies ausgeschlossen. Da nun allerdings keine durchgängigen Unterschiede bezüglich der Leistungen in Teil A ausfindig gemacht werden konnten, könnte eine bessere Leistung einer der Gruppen in Teil B für eine schlechtere Leistung durch die Vorbereitung mit der jeweiligen Lernmethode sprechen. Die statistischen Analysen zeigten allerdings, dass die vier betrachteten Gruppen sich ebenfalls nicht hinsichtlich ihrer Leistung in Teil B der Klausur unterscheiden. Die Studierenden können sich folglich mit jedem Lernangebot vergleichbar gut auf die Modulabschlussprüfung vorbereiten.

Wie bereits in Studie I diskutiert, ist es generell schwierig, Lernleistungsunterschiede zwischen Gruppen innerhalb einer bewerteten Wissensüberprüfung zu finden, die für die Lernenden eine wichtige Bedeutung hat. In solchen Fällen gleichen Lernende unterschiedlich gute Instruktion durch ein unterschiedlich ausgeprägtes Maß an Anstrengung aus. Dies ist im Falle der Modulabschlussprüfung gegeben. Dementsprechend ist es nicht verwunderlich, dass sich zwar im Rahmen der Probeklausurergebnisse signifikante Ergebnisse zugunsten der Gruppe „elaboriertes (E-) WBT“ ergeben, diese aber im Rahmen der Modulabschlussprüfung nicht mehr vorhanden sind. Tendenziell scheinen die Studierenden sich mit den elaborierten WBTs besser vorbereiten zu können.

7.4.4.3 Wichtige Eigenschaften von E-Learning-Anwendungen

Zusätzlich sollte betrachtet werden, welche Eigenschaften von E-Learning-Anwendungen die Studierenden besonders wichtig finden. Dabei wurden den Studierenden 20 Eigenschaften dargeboten, die sich auf WBT-ähnliche E-Learning-Anwendungen bezogen und nach dem ARCS-Modell von John Keller förderlich für den Lernprozess wirken sollen.

Die Einschätzungen der Studierenden variierten auf einer Skala von 1 bis 10 deutlich von 4.7 bis 9.5. Spannend ist, dass insbesondere die drei Items, die Aspekte von Übungstests beinhalten, auf den ersten drei Plätzen des Rankings landeten (Abschlussfeedback, Übungsaufgaben, Feedback nach jeder Aufgabe). Betrachtet man die abgefragten Aspekte gebündelt für die einzelnen ARCS-Komponenten, zeigt sich, das insbesondere die Aspekte zur Komponente „Erfolgszuversicht“ häufiger als wichtig eingestuft wurden (insbesondere die Übungstests, die Inhaltsverzeichnisse und die flexible Bearbeitungsmöglichkeit). Die Aspekte zur Komponente „Aufmerksamkeit“ hingegen werden allesamt als weniger wichtig beurteilt (Bilder, Abwechslung, Videos etc.). Aspekte zu den Komponenten „Relevanz“ und

„Zufriedenheit“ verteilen sich über das gesamte Ranking – in diesen sind folglich für die Studierenden wichtige und weniger wichtige Aspekte enthalten.

Klauser, Kim und Born (2002) befragten in ihrer Studie ebenfalls Studierende hinsichtlich der Wichtigkeit von 26 Kriterien für ein gutes E-Learning-Angebot. Hier stellten sich insbesondere Aspekte wie Übersichtlichkeit, Übungsaufgaben, Lernzielformulierung, Lernerfolgskontrolle und eine regelmäßig stattfindende Online-Sprechstunde als besonders wichtig heraus. Die Studierenden wurden zusätzlich gebeten, die drei wichtigsten Kriterien zu nennen. Auch hier nannten sie zumeist die Übersichtlichkeit der Inhalte, eine klare Lernzielformulierung, eine Lernerfolgskontrolle sowie das Vorhandensein von Übungsaufgaben. Technische Aspekte, wie Multimedialität, wurden auch hier als weniger wichtig eingestuft. Zwei der vier am häufigsten genannten Aspekte beziehen sich folglich vergleichbar zur Studie II auf die Übungstests. Die Lernzielformulierung (Komponente „Relevanz“) wird in Studie II als weniger wichtig von den Studierenden eingestuft. Das kann mit dem längerfristigen Lernziel „Bestehen der Modulabschlussprüfung“ zusammenhängen. Aspekte zur Übersichtlichkeit (Gesamtseitenzahl, -dauer, Inhaltsverzeichnis), welche ebenfalls der Komponente „Erfolgszuversicht“ zuzuordnen sind, tauchen im Ranking der Studie II ebenfalls im mittleren bis oberen Bereich auf.

Vergleicht man die Einschätzungen der Studierenden in der Studie von Klauser et al. (2002) sowie in Studie II mit den Ergebnissen zur ARCS-Skala, zeigt sich erneut die Wichtigkeit der Erfolgszuversichts- sowie die geringere Bedeutung der Aufmerksamkeitskomponente. Vergleichbar zur Studie I dieser Arbeit scheinen dabei die Übungsaufgaben eine herausragende Rolle zu spielen.

Im folgenden Kapitel soll nun ein abschließendes Fazit aus den Ergebnissen zur zweiten Studie gezogen werden, bevor im letzten Abschnitt der vorliegenden Dissertation die Ergebnisse beider Studien miteinander verglichen und hieraus abschließende Schlussfolgerungen abgeleitet werden.

8 Studie II: Fazit

Durch die zweite Studie sollte überprüft werden, inwieweit die Akzeptanz und Lernwirksamkeit von E-Learning-Anwendungen durch die Variation verschiedener Gestaltungsmerkmale verbessert werden kann und welche Gestaltungsmerkmale sich hinsichtlich der Steigerung von Lernwirksamkeit und Akzeptanz am deutlichsten auszahlen. Als Leitfaden für die Erstellung der unterschiedlichen Varianten wurde das ARCS-Modell von John Keller (2010) herangezogen. An dieser Stelle sollen die Ergebnisse auch aus einer Kosten-Nutzen-Perspektive heraus interpretiert werden, da die Produktion von E-Learning-Anwendungen mit lern- und akzeptanzförderlichen Aspekten wie Übungsaufgaben, Lernspielen oder Animationen etc. kostenintensiver ist als die von einfacher gestalteten Anwendungen. Abschließend soll resümiert werden, ob sich der Einsatz des Keller`schen Modells gelohnt hat und welche der von ihm postulierten Gestaltungsmerkmale sich konkret ausgezahlt haben. Es stellt sich folglich nicht nur die Frage, ob alle Strategien des Modells eingesetzt werden *dürfen* (im Hinblick auf zum Beispiel die Gefahr der Überfrachtung/Überforderung des Arbeitsgedächtnisses etc.), sondern auch, ob sie denn alle eingesetzt werden *müssen*.

Zunächst einmal ist festzuhalten, dass die Ergebnisse aus Studie II belegen, dass alle E-Learning-Varianten eine gute Lerneffektivität und -effizienz aufweisen. Hinsichtlich der Lernwirksamkeit und -effizienz zeigen sich Vorteile zwischen den evaluierten Lehrmethoden zumeist nur im Vergleich der E-Learning-Varianten mit der Vorlesung – diese fallen immer zugunsten der E-Learning-Anwendungen aus. Alle eingesetzten Varianten führen folglich zu gleich guten und in vielen Vergleichen sogar besseren Lernergebnissen als die herkömmliche Hochschullehre und können folglich auch in Zukunft mit gutem Gewissen eingesetzt werden.

8.1 Lernleistung und Lerneffizienz

Betrachtet man die E-Learning-Varianten im Vergleich hinsichtlich der Lernwirksamkeitsergebnisse im Wintersemester, zeigt sich nur in geringem Maße ein Vorteil der mit ARCS-Strategien angereicherten WBT-Varianten. Durch den Einsatz der ARCS-Strategien konnte hier weder die Lernleistung der Studierenden noch die Effizienz ihrer Lernanstrengungen bedeutsam verbessert werden. Augenscheinlich erreichten die Studierenden nach Bearbeitung der ARCS-WBTs zwar sowohl kurz- als auch längerfristig den höchsten und die E-Lectures sowie 0-WBTs den niedrigsten Lernzuwachs. Der Abstand

zwischen den Anwendungen war allerdings nicht so gravierend, dass ein signifikant besseres Abschneiden durch die Nutzung von mit ARCS-Strategien angereicherten Anwendungen nachgewiesen werden konnte. Dabei können die Lernwirksamkeits-ergebnisse für alle Varianten als gut beurteilt werden. Alle Anwendungen – auch die 0-WBTs – können folglich mit gutem Gewissen in der Lehre eingesetzt werden.

Nach der freiwilligen Nutzung im Sommersemester schnitten in der Probeklausur dann diejenigen, die hauptsächlich mit den elaborierten WBTs gelernt hatten, signifikant besser ab als die Studierenden, die vornehmlich das Vorlesungsskript oder die kurzen WBTs genutzt hatten. Dieses Ergebnis konnte sich zwar für die Modulabschlussprüfung nicht bestätigen, die Ergebnisse von Abschlussprüfungen werden allerdings zumeist durch Kompensationen wie z.B. Anstrengung und/oder Fleiß beeinflusst. Dies wurde oben bereits diskutiert.

Im Hinblick auf die Lernwirksamkeit *dürfen* folglich alle ARCS-Strategien im Rahmen von E-Learning-Anwendungen eingesetzt werden. Die angereicherten Anwendungen schneiden entweder zumindest gleich gut oder sogar besser ab als die nicht-angereicherten 0-(K-)WBTs. Inwieweit eine Umsetzung des ARCS-Modells allerdings zwingend notwendig ist, ist abhängig davon, wie man im Rahmen einer Kosten-Nutzen-Analyse „Nutzen" definiert. Zielt man lediglich auf die Verbesserung des Lernerfolgs bzw. der -effizienz ab, könnte nach den Ergebnissen aus dem Wintersemester die Entscheidung getroffen werden, ein völlig „ARCS-freies" 0-WBT, mit vergleichsweise geringen Erstellungskosten reiche aus. Schließlich konnten die Studierenden mit diesem genauso viel und genauso effizient lernen wie mit den wesentlich aufwendiger gestalteten Einzelkomponenten-WBTs bzw. dem ARCS-WBT. Durch die Ergebnisse des Sommersemesters zeigt sich allerdings, dass dies womöglich nur für die kurzfristige Nutzung der E-Learning-Anwendungen gilt. Längerfristig scheint es im Hinblick auf die Lernwirksamkeit durchaus sinnvoll zu sein, den Studierenden mit ARCS-Strategien angereicherte Anwendungen zur Verfügung zu stellen. Die Bearbeitung der E-WBTs scheint den Studierenden mehr Spaß zu machen. Hieraus resultiert eine längere Nutzung und daraus wiederum unweigerlich eine zumindest in der Probeklausur bessere Lernleistung. Dies zeigt, wie wichtig es im Rahmen von (längerfristigen) E-Learning-Szenarien ist, den Lernenden attraktive Lernanwendungen zur Verfügung zu stellen.

8.2 Akzeptanz

Bezieht man zusätzlich die Variable „Akzeptanz“ mit in die Kosten-Nutzen-Analyse ein – will man folglich die Lernenden nicht nur „schlauer“ sondern auch in Bezug auf die zu nutzende Lernanwendung zufriedener machen – scheint die Umsetzung der ARCS-Strategien auch für den kurzfristigen Einsatz relevant. Hinsichtlich der Akzeptanz wurde im Wintersemester die Einstellungsakzeptanz erhoben, im Sommersemester konnte durch die freiwillige Nutzung der Anwendungsvarianten ARCS-(E-)WBT, 0-(K-)WBT und E-Lecture auf die Verhaltensakzeptanz geschlossen werden. Zwar ergaben sich hinsichtlich der Einstellungsakzeptanz für die E-Learning-Varianten hauptsächlich über alle drei Themenbereiche gemittelt bedeutsame Unterschiede mit schwachem Effekt. Sie zeigen jedoch eine ganz klare Tendenz zugunsten der vollständigen Umsetzung der ARCS-Strategien. Bezogen auf die Akzeptanzskala konnten sich die ARCS-WBTs gegenüber den 0-WBTs durchsetzen und bezüglich des groben Akzeptanzmaßes „Schulnote“ erreichten sie eine bessere Beurteilung als die 0- und die A-WBTs sowie die E-Lectures. Auch die R- und S-WBTs erreichten hier eine bessere Benotung als die 0-WBTs.

Wie oben bereits diskutiert, waren darüber hinaus sowohl die Häufigkeit als auch die Dauer der freiwilligen E-Learning-Nutzungen im Sommersemester äußerst erfreulich. Dabei bevorzugten die Studierenden klar die mit ARCS-Strategien angereicherten WBTs. Sie wurden dreimal so lange genutzt wie die faktischen, „ARCS-leeren“ kurzen WBTs sowie die E-Lectures. Die Ergebnisse zur Verhaltensakzeptanz untermauern damit deutlich die erhobene Einstellungsakzeptanz. Der Einsatz der ARCS-Strategien führte folglich im Rahmen der relativ kurzen Nutzung im Wintersemester zu einer moderaten Verbesserung der Einstellungsakzeptanz, im Rahmen der längerfristigeren freiwilligen Nutzung im Sommersemester dann allerdings zu einer deutlich höheren Verhaltensakzeptanz.

Interessant ist die Frage, welche Eigenschaften der elaborierten WBTs zu einer vermehrten Nutzung dieser E-Learning-Variante geführt haben. Aus der abschließenden Befragung lässt sich vermuten, dass hier insbesondere die Übungstests eine vorrangige Rolle gespielt haben. Aus diesem Grunde wäre es interessant zu analysieren, wie häufig neben 0- und ARCS-WBTs alleinstehende Übungstests genutzt würden. Wenn diese vergleichbar häufig oder sogar häufiger genutzt würden und diese Nutzung zu gleichem Lernerfolg führen würde wie die wesentlich aufwendiger zu erstellenden elaborierten WBTs, würde dies die Kosten-Nutzen-Analyse für die ARCS-Strategien deutlich relativieren. Eine weitere Möglichkeit wäre die Analyse der Lesezeiten der elaborierten WBTs durch ein exaktes User-Tracking für jede Seite der WBTs oder eine Beobachtung der Lernenden bei der Bearbeitung der Anwendungen im Labor. Hier könnte genau untersucht werden, wie viel Zeit die Studierenden mit welchem Teil der WBTs verbringen. Dies würde einen klareren Blick auf

die wichtigen Eigenschaften von E-Learning-Anwendungen bzw. Aspekte des ARCS-Modells ermöglichen.

8.3 ARCS-Skala: Wahrnehmung der Lernunterstützung

Für die Entscheidung für oder gegen eine Umsetzung der Strategien nach dem Modell von John Keller (2010) kann allerdings noch ein dritter Aspekt bedeutsam sein. Durch die ARCS-Skala wurde die subjektive Einschätzung der Studierenden hinsichtlich der Lernunterstützung durch die verschiedenen WBT-Varianten erhoben. Hier ergaben sich eindeutig signifikante Unterschiede zwischen den E-Learning-Anwendungen. Dabei konnte der Einsatz der ARCS-Strategien die Wahrnehmung der Studierenden hinsichtlich der Lernunterstützung durch die E-Learning-Anwendungen grundsätzlich positiv beeinflussen. In keinem Fall wirkte sich der Einsatz der Strategien im Vergleich zu den 0-WBTs negativ aus. Im Gegensatz dazu schnitten die 0-WBTs hinsichtlich jeder Unterskala des ARCS-Fragebogens sowie hinsichtlich der Lernförderung insgesamt deutlich schlechter ab als die ARCS-WBTs – und zumeist auch die skalenspezifischen Einzelkomponenten-WBTs (gemeint ist hier z.B. das R-WBT bezogen auf die Unterskala „Relevanz“). Die ARCS-WBTs erreichten hinsichtlich aller Unterskalen der ARCS-Skala sowie der Gesamtskala die besten Beurteilungen durch die Studierenden. Es kann also mitnichten von einer Überfrachtung der ARCS-WBTs durch zu viele Strategien gesprochen werden. Die einzelnen Strategien scheinen im Gegensatz dazu im Zusammenspiel die Lernförderlichkeitswahrnehmung zu verstärken. Dies zeigt sich deutlich dadurch, dass die ARCS-WBTs auf den Unterskalen (z.B. Skala „Aufmerksamkeit“) grundsätzlich zumindest augenscheinlich besser abschneiden als das jeweils skalenspezifische Einzelkomponenten-WBT (hier z.B. A-WBT). Sie erreichen zwar nie ein signifikant besseres Ergebnis als das skalenspezifische Einzelkomponenten-WBT, können sich allerdings zumeist gegenüber mehr der anderen E-Learning-Varianten durchsetzen als dieses. Die subjektive Einschätzung der Lernförderlichkeit spricht – vergleichbar zu den Akzeptanzergebnissen sowie den Lernwirksamkeitsergebnissen aus dem Sommersemester – für einen ganzheitlichen Einsatz des ARCS-Modells bei der Entwicklung von E-Learning-Anwendungen.

8.4 Betrachtung der Ergebnisse in Bezug auf die vier Komponenten

Will man jedoch wissen, welche der ARCS-Strategien zwingend umgesetzt werden sollten und welche gegebenenfalls vernachlässigt werden können, müssen die Ergebnisse zu den Einzelkomponenten-WBTs noch einmal genauer betrachtet werden. Hinsichtlich der

Lernwirksamkeit und Lerneffizienz ergaben sich hier mehrheitlich keine Vorteile eines Einzelkomponenten-WBTs und auch hinsichtlich der Akzeptanzkriterien konnten sich lediglich für das Akzeptanzmaß „Note“ und über alle Themen hinweg die Relevanz- und Zufriedenheits-WBTs gegenüber dem 0-WBT durchsetzen. Nur diese beiden Komponenten haben folglich einen Einfluss auf die Akzeptanz der hier befragten Studierenden. Die immer noch relativ gute Bewertung der 0-WBTs ist erstaunlich. Hier waren zwar alle zu lernenden Fakten enthalten, mehr allerdings nicht. Es wäre durchaus nachvollziehbar gewesen, wenn eine solche Darstellung der Inhalte die Studierenden furchtbar frustriert hätte – dies war aber nicht der Fall. Hier kann die Abwechslung zum „üblichen Hochschulalltag“ eine Rolle spielen, wie dies bereits oben diskutiert wurde.

Da die einzelnen Komponenten sich hinsichtlich der Beurteilungen auf der ARCS-Skala sehr unterschiedlich auswirkten, soll an dieser Stelle noch einmal ein genauerer Blick auf die Auswirkungen der einzelnen Komponenten des ARCS-Modells auf diese Einschätzungen geworfen werden.

Es zeigte sich, dass sich insbesondere die Strategien zur **Erfolgszuversicht und Zufriedenheit** deutlich auf die Einschätzung der Studierenden hinsichtlich der Skalen „Erfolgszuversicht“, „Zufriedenheit“ und der Gesamt-ARCS-Skala auswirken. Die Wahrnehmung der Relevanz- und Aufmerksamkeitsunterstützung können beide Komponenten im Vergleich zu den 0-WBTs nicht bedeutsam verstärken.

Gefragt nach wichtigen Eigenschaften von E-Learning-Anwendungen schätzen die Studierenden gerade die Aspekte als bedeutsam ein, die zur Komponente „Erfolgszuversicht“ des ARCS-Modells gehören. Im Speziellen sind dies die Übungstests und die darin enthaltenen Rückmeldungen. Dies bestätigt eindrücklich die Ergebnisse aus dem Wintersemester, die zeigen, dass die Wahrnehmung der Lernunterstützung insbesondere durch die Erfolgszuversichts- und Zufriedenheitsstrategien erhöht werden konnte. Auch in den Zufriedenheits(S)-WBTs sind durch die Lernspiele Möglichkeiten enthalten, die es den Studierenden erlauben, ihr Wissen zu überprüfen. Es überrascht allerdings, dass sich die C-WBTs dann nicht auch vergleichbar zu den S-WBTs hinsichtlich der Akzeptanz gegenüber den 0-WBTs abheben konnten.

Die vermehrte Nutzung der elaborierten WBTs (gegenüber den kurzen WBTs) im Sommersemester könnte demnach hauptsächlich darin begründet sein, dass in ihnen Übungstests und Lernspiele enthalten waren. Welche Aspekte es allerdings wirklich waren, die die Studierenden dazu veranlasst haben, die elaborierten WBTs für ihre Klausurvorbereitung auszuwählen, müsste in einer weiteren Studie überprüft werden, in der die Nutzung der E-Learning-Anwendungen durch ein User-Tracking aufgezeichnet

wird. Dadurch könnte genau überprüft werden, welche Seiten die Studierenden besonders intensiv nutzen und welche der Keller'schen Strategien womöglich unbemerkt bleiben.

Der Einfluss der Strategien der Komponente „**Aufmerksamkeit**" auf die Wahrnehmung der Aufmerksamkeitsförderung der Studierenden fiel deutlich schwächer aus. Die A-WBTs konnten sich bezogen auf die Skala „Aufmerksamkeit" lediglich gegenüber den E-Lectures, nicht jedoch gegenüber den 0-WBTs durchsetzen. Dies gelingt dieser WBT-Variante lediglich für die Skala „Zufriedenheit".
Die Studierenden scheinen sich in ihrer Aufmerksamkeitserhaltung erst dann (im Gegensatz zu den 0-WBTs) bedeutsam unterstützt zu fühlen, wenn alle ARCS-Strategien zum Einsatz kommen. Hier bedarf es weiteren Studien mit verschiedenen WBT-Varianten, um zu prüfen, wie viele und welche Komponenten des ARCS-Modells konkret eingesetzt werden müssen, um eine merkliche Verbesserung hinsichtlich der Aufmerksamkeits-erhaltung zu erzielen. Tendenziell scheinen sowohl die Aufmerksamkeits- als auch die Zufriedenheitsstrategien hier wirksam zu sein – ggf. ist eine Kombination dieser beiden Komponenten zur Förderung der Aufmerksamkeitserhaltung ausreichend und keine vollständige „ARCS-Version" notwendig.
Neben dem nicht ganz eindeutigen Einfluss der Aufmerksamkeitsstrategien auf die Skala „Aufmerksamkeit" kann die Aufmerksamkeitskomponente allerdings die Zufriedenheit der Studierenden deutlich verbessern. Zwar ergibt sich für die A-WBTs (im Gegensatz zu allen anderen Einzelkomponenten-WBTs) keine Verbesserung im Vergleich zu den 0-WBTs im Hinblick auf die Gesamt-ARCS-Skala, aber zumindest ist auf der Unterskala „Zufriedenheit" ein positiver Einfluss der Aufmerksamkeitsstrategien festzustellen.

Die Strategien der Komponente „**Relevanz**" können sowohl zur Förderung der Relevanzwahrnehmung beitragen als auch zur Steigerung der Erfolgszuversicht (beides allerdings nur für das Thema „Konditionierung" bzw. über alle Themen hinweg) und der Zufriedenheit der Studierenden (über alle Themen hinweg). Die R-WBTs schneiden damit auch hinsichtlich der Gesamt-ARCS-Skala signifikant besser ab als die 0-WBTs (ebenfalls nur für das Thema „Konditionierung" bzw. über alle Themen hinweg). Die Beurteilungen der R-WBTs zeigen damit eindeutig die inhaltliche Abhängigkeit der Relevanzkomponente. Nur für das Thema „Konditionierung" konnten die Relevanzstrategien die Wahrnehmung der Relevanzförderung, der Erfolgszuversicht sowie der Gesamt-Lernförderlichkeit bedeutsam beeinflussen. Und auch hinsichtlich der Akzeptanz können sich die R-WBTs lediglich für das Thema „Konditionierung" und über alle Themen hinweg durchsetzen. Es handelt sich hierbei um das Thema, dessen Relevanz für den späteren Berufsalltag den Studierenden ohne Anwendungsbeispiele zumeist nur schwer klar wird. Dies ist bei den

anderen beiden Themen („Pädagogische Diagnostik“ und „Unterrichten“) so nicht der Fall. Hier erschließt sich die Anwendbarkeit für die Studierenden bereits ohne konkrete Beispiele. Der Einsatz der Relevanzstrategien sollte folglich je nach den zu vermittelnden Inhalten abgewogen werden. Ihr Einsatz erscheint dann notwendig, wenn Inhalte vermittelt werden sollen, die keinen direkten Anwendungsbezug haben, da mit den Relevanzstrategien die Lernenden unter anderem darin unterstützt werden, diesen Anwendungsbezug herzustellen.

8.5 Abschließendes Fazit zur Studie II

Nach Sichtung all dieser Ergebnisse scheint eine ganzheitliche Umsetzung der Strategien nach dem Modell von John Keller (2010) sinnvoll, aber nicht in jedem Fall notwendig. Gegen einen vollständigen Einsatz des gesamten ARCS-Modells spricht lediglich die Kosten-Nutzen-Analyse, da sich die Strategien grundsätzlich positiv auswirken und keine negativen Folgen für die Lernenden zu haben scheinen. Möchte man Lernenden zum Beispiel nur einmalig eine E-Learning-Anwendung anbieten und ist dabei lediglich darauf bedacht, dass die Lernenden die Inhalte gut verinnerlichen, erscheinen die Kosten für die Umsetzung des Modells zu hoch. Schließlich zeigen die Ergebnisse von Studie II, dass die Studierenden auch mit der wesentlich einfacher gestalteten 0-Version erfolgreich lernen konnten. Soll E-Learning jedoch häufiger und längerfristiger eingesetzt werden, sollten neben der Lernwirksamkeit auch die Akzeptanz und das Gefühl der Lernunterstützung mit in die Kosten-Nutzen-Analyse einfließen. Dies ist allein schon deshalb notwendig, da eine geringe Akzeptanz im Rahmen eines E-Learning-Szenarios zu hohen Abbruchquoten führen kann – eine Variable, die im Rahmen der vorliegenden Studie leider nicht explizit erhoben wurde, die aber für den langfristigen Einsatz von E-Learning eine bedeutende Rolle spielt.
Wenn man Akzeptanz und Wahrnehmung von Lernunterstützung in die Entscheidung für oder gegen den Einsatz der Gestaltungsmerkmale einbezieht, muss man sich für eine vollständige Umsetzung des Modells entscheiden. In beiderlei Hinsicht erreichte das WBT, in dem alle Strategien nach dem ARCS-Modell umgesetzt worden waren, die besten Beurteilungen. Möchte man jedoch aufgrund von ökonomischen Überlegungen zumindest auf einen Teil der Komponenten verzichten, sollte man nach den Ergebnissen zur ARCS-Skala auf keinen Fall auf die Umsetzung der Erfolgszuversichts- und Zufriedenheitsstrategien verzichten. Die Strategien zur Förderung der Relevanz-wahrnehmung könnten je nach Thematik abgewogen werden, die Umsetzung der Komponente „Aufmerksamkeit“ könnte aufgrund ihrer nicht vorhandenen Auswirkungen auf die Lernförderlichkeitswahrnehmung insgesamt, die Lernwirksamkeit sowie die

Akzeptanz am ehesten vernachlässigt werden. Auch die Abfrage zu den wichtigen Eigenschaften von E-Learning-Anwendungen zeigte, dass die Strategien aus der Komponente „Aufmerksamkeit“ für die Studierenden insgesamt nicht besonders bedeutsam zu sein scheinen.

Diese Empfehlungen gelten zumindest für den Einsatz von E-Learning im Bereich der Hochschullehre, denn es stellt sich in diesem Zusammenhang die Frage, inwieweit die gegebenen Umstände die Einstufungen der Lernenden beeinflusst haben. Es wäre interessant zu prüfen, ob die ARCS-Komponenten in einer Situation, in der sich Lernende freiwillig und ohne extern vorgegebenes Lernziel (Bestehen der Klausur) mit einem Inhalt auseinandersetzen, anders eingestuft werden. Hier könnte die Komponente „Erfolgszuversicht“ eine geringere Rolle spielen, da die Lernenden keinen „Druck“ verspüren, die Inhalte zu verinnerlichen. Aspekte aus der Komponente „Aufmerksamkeit“ könnten im Gegensatz dazu mehr in den Vordergrund treten, da den Lernenden ggf. die Gestaltung und Multimedialität der E-Learning-Anwendung (Bilder, Videos, Animationen) beim intrinsisch gesteuerten Lernen in ihrer Freizeit wichtiger wird.

Des Weiteren wäre spannend zu untersuchen, welche der Maßnahmen innerhalb der ARCS-Komponenten konkret besonders bedeutsam sind (Übungstests, Inhaltsverzeichnis, Lernspiele etc.). Die abschließende Befragung der Studierenden lässt jedoch bereits erkennen, dass für die Studierenden zumindest subjektiv die Übungstests und die in diesen integrierten Zwischen- und Abschlussfeedbacks eine bedeutende Rolle spielen.

8.6 Limitationen von Studie II und Anregung für weitere Studien

Zu der vorliegenden Studie sind zunächst einige Punkte kritisch anzumerken, die zum Teil bereits im Rahmen der ersten Studie diskutiert wurden.
Zunächst ist hier erneut die Schwierigkeit der Konstanthaltung der Inhalte über alle Versuchsbedingungen zu nennen. In Studie II wurde die inhaltliche Gleichheit der Anwendungen dadurch angestrebt, dass für die WBT-Variationen zunächst aus dem Transkript der E-Lecture eine inhaltliche Grundstruktur (entspricht dem 0-WBT) erstellt wurde. Dieses wurde darauffolgend für die Erstellung der anderen Variationen durch die jeweiligen Strategien angereichert. Dadurch ist die inhaltliche Konsistenz zwischen den WBTs gut gesichert. Im Vergleich zu den E-Lectures sind kleinere Abweichungen durchaus denkbar, da zur Erstellung des „ARCS-freien“ 0-WBTs einzelne Aspekte der E-Lecture, wie z.B. Anwendungsbeispiele oder Fragen an das Plenum aus dem Transkript entfernt wurden,

die zum Teil in der Form dann nicht wieder in den Einzelkomponenten-WBTs ergänzt wurden. Solche Abweichungen wurden im Rahmen der Erstellung der Wissensdiagnostik allerdings berücksichtigt. Auch durch die dreifache Replikation der Studie im Wintersemester sollten solche kleineren Abweichungen nicht allzu stark ins Gewicht fallen. Dabei konnte die Inhaltsgleichheit über die verschiedenen E-Learning-Variationen in Studie II deutlich leichter sichergestellt werden als in Studie I. Es ist nachvollziehbar einfacher, inhaltsgleiche E-Learning-Varianten zu erstellen als eine Präsenzsitzung mit E-Learning-Anwendungen vergleichbar zu halten. Eine „live"-Sitzung bietet naturgemäß mehr Möglichkeiten für spontane inhaltliche Abweichungen, durch zum Beispiel Fragen aus dem Plenum, Exkurse des Dozenten oder den begrenzten Zeitrahmen. Diese „Gefahren" bestanden in Studie II nicht, da die Vorlesung dreimal durch E-Learning-Anwendungen ersetzt wurde. Dadurch war allerdings keine inhaltsgleiche Gegenüberstellung zwischen Präsenz und E-Learning möglich. Die Ergebnisse des Vergleichs zwischen den drei „E-Learning-Themen" mit den neun „Vorlesungsthemen" wurden deshalb mit Bedacht interpretiert und diskutiert.

Zudem muss sich auch Studie II, die vergleichbar zur Studie I eine Felduntersuchung darstellt, erneut die Kritik gefallen lassen, dass es sicherlich eine Reihe von nicht-kontrollierbaren Störfaktoren gibt, die die Aussagekraft der Vergleiche einschränken können. Die Vor- und Nachteile der Felduntersuchung wurden im Rahmen von Studie I ausführlich diskutiert, weshalb hier nicht nochmals darauf eingegangen wird (siehe Abschnitt „4.3 Einschränkungen von Studie I"). Das gewählte Evaluationsszenario bot ganz klare Vorteile, die die Nachteile definitiv aufwiegen: es ermöglichte inhalts- und personinvariante Vergleiche durch mehrere Replikationen und stellte eine hohe ökologische Validität der Ergebnisse sicher. Dies wäre im Rahmen einer Laboruntersuchung so nicht möglich gewesen.
Wie oben bereits angeklungen, wäre es allerdings interessant, einzelne Aspekte wie die Nutzung der verschiedenen Maßnahmen der ARCS-Komponenten oder die Kompensation von instruktional schlechtem Material durch Lernende durch ein User-Tracking oder im Rahmen einer Laboruntersuchung noch einmal tiefergehend zu betrachten.

Auch in Studie II wurde den Studierenden ein Anreiz für die Beteiligung an der Evaluation im Wintersemester geboten. Dieser stellte jedoch im Gegensatz zu Studie I, in der sich die Studierenden Bonuspunkte zur MAP verdienen konnten, diesmal die freie Nutzung der E-Learning-Anwendungen zur Klausurvorbereitung dar. Dadurch, dass die Vorlesung im Gegensatz zur Studie I in den Wochen der E-Learning-Evaluationen entfiel, wäre die Teilnahme der Studierenden hinsichtlich der E-Learning-Nutzung sicherlich auch ohne

zusätzlichen Anreiz relativ gut ausgefallen. Um die Studierenden jedoch auch dafür zu gewinnen, die Evaluationsfragebögen vollständig auszufüllen, erschien ein Anreiz sinnvoll. Sicherlich kann ein solcher Anreiz erneut dazu führen, dass manche Studierende nicht der Sache wegen teilnahmen und dementsprechend die E-Learning-Anwendungen und Evaluationen nur halbherzig bearbeiteten. Eine völlig freiwillige Teilnahme hätte allerdings den Nachteil eines deutlich stärkeren Drop-outs haben können. Auch dies hätte die Aussagekraft der Studie eingeschränkt.

Im Hinblick auf die Schwierigkeiten mit einzelnen Aspekten der Evaluationsbefragung in Studie I wurde diese für Studie II überarbeitet und inhaltlich angepasst. Dabei wurden diesmal explizit Skalen ausgewählt, die bereits durch andere Autoren auf ihre Eignung hin überprüft worden waren. Lediglich für die „ARCS-Skala“, die angelehnt an den *IMMS* von Keller (2006) entwickelt wurde, mussten neben der Übersetzung ins Deutsche auch einige Anpassungen vorgenommen werden. Der ursprüngliche *IMMS* wies zum einen methodische Mängel auf, wie z.B. das Abfragen zweier Aspekte innerhalb eines Items. Zum anderen passten die Formulierungen zum Teil nicht zum vorliegenden Setting. Die Anpassung des *IMMS* auf die jeweilige Lernsituation wird bereits durch Huang et al. (2006) nahegelegt. Die Autoren hatten den Fragebogen im Original auf seine Güte hin überprüft.
Die aus der umfangreichen Überarbeitung entstandenen Skalen wurden im Rahmen der Studie mittels Faktorenanalyse überprüft und validiert. Die resultierenden Skalen weisen eine akzeptable bis gute Reliabilität auf. Sicherlich wäre eine Multitrait-Multimethod-Validierung des ARCS-Fragebogens (ggf. sogar am besten im Rahmen einer Vorstudie) methodisch angemessener gewesen als die Validierung ausschließlich über Fragebogendaten im Verlauf der Studie. Dies war allerdings in dem gegebenen Setting aus organisatorischen Gründen nicht sinnvoll durchführbar.

Für beide Studien der vorliegenden Dissertation gilt darüber hinaus, dass es sich bei den befragten Nutzern um eine sehr homogene Stichprobe handelt. Dies schränkt die Aussagekraft der Studien auf den Hochschulkontext ein. Ihre Ergebnisse können nur mit Bedacht auch auf andere Gruppen von Lernenden übertragen werden. Durch die Beobachtungen anhand der Studierendengruppe können allerdings Hypothesen über die Wirkung der ARCS-Komponenten für andere Lernendengruppen abgeleitet werden. Die hier untersuchten Studierenden haben ihr „Können als Lerner“ bereits so gut ausgebaut, dass sie den höchsten schulischen Bildungsabschluss erreichen konnten. Es handelt sich bei der Stichprobe folglich um relativ erfahrene Lernende. Sie sind es durchaus gewohnt, auch mit instruktional weniger geeignetem Lehrmaterial umzugehen und die Schwächen dieses durch verschiedene Strategien zu kompensieren. Dies ist bei Lernenden mit geringerem

Bildungsabschluss oder jüngeren bzw. deutlich älteren Lernenden, deren „Lerngeschichte“ weniger erfolgreich, wesentlich kürzer oder schon länger her ist, sicher so nicht der Fall. Aus diesem Grunde ist anzunehmen, dass die Wirkung der ARCS-Komponenten auf die Wahrnehmung der Lernunterstützung, die Akzeptanz sowie die Lernwirksamkeit für diese Lernenden deutlicher ausfallen sollte, da sie stärker auf Lernunterstützung angewiesen sein sollten. Es wäre spannend, dies in weiteren Studien zu überprüfen.

Eine andere forschungsrelevante Stichprobe stellt die der Lernenden dar, die sich in der Freizeit aus Eigeninteresse mit Lerninhalten auseinander setzen möchten. Auch diese könnten – wie oben bereits diskutiert – anders über die einzelnen ARCS-Komponenten urteilen als die hier untersuchte Stichprobe. Auch dies wäre im Rahmen weiterer Studien zu untersuchen.

9 Abschließende Diskussion und Fazit

Ziel der ersten Studie war ein Vergleich zwischen zwei prominenten E-Learning-Anwendungen und der traditionellen Präsenzlehre. Als E-Learning-Methoden wurden die E-Lecture und das Web-based Training ausgewählt, da sie beide häufig in der Praxis eingesetzt werden und in der Darstellung der Inhalte sehr unterschiedlich sind. Während die E-Lecture Inhalte vergleichbar zur Vorlesung in gesprochener Form darbietet, werden diese im Rahmen des Web-based Trainings vornehmlich in schriftlicher Form präsentiert – ergänzt durch Bilder, Videos, Audios, Übungsaufgaben etc. Studie II zielte darauf ab, durch verschiedene Gestaltungselemente nach dem ARCS-Modell von John Keller (2010) die Akzeptanz und Lernwirksamkeit von Web-based Trainings zu verbessern. Dabei sollte untersucht werden, welche Komponenten des ARCS-Modells insbesondere zur Förderung dieser Aspekte sinnvoll bzw. zwingend erforderlich sind.

Beide Studien haben einen ähnlichen Versuchsplan. Es handelt sich jeweils um umfangreiche Feldexperimente, an denen insgesamt über 2500 Studierende teilnahmen. Dies ermöglichte es in beiden Studien, mehrere Replikationen und dadurch inhalts- und materialinvariante Vergleiche durchzuführen. Die Studierenden konnten die Lehrmethoden in ihrer gewohnten Lernumgebung nutzen, was eine hohe ökologische Validität der Ergebnisse nahelegt. Darüber hinaus war es möglich, dass jeder Studierende an allen (Studie I) bzw. an drei von sieben Bedingungen (Studie II) teilnahm, so dass personinvariante Vergleiche realisiert werden konnten.

Die Ergebnisse aus beiden Studien sprechen für eine gute Lernwirksamkeit und Akzeptanz aller eingesetzten Lehrmethoden. Sogar das puristische 0-(bzw. K-)WBT, das in Studie II zum Einsatz kam, erzielte recht gute Lern- und Akzeptanzergebnisse. Hier spielt sicherlich auch die Abwechslung zur traditionellen Präsenzlehre an der Hochschule eine bedeutende Rolle.

Alle Lehrmethoden können folglich mit gutem Gewissen in der Hochschullehre eingesetzt werden. Dabei ergaben sich hinsichtlich der Lernwirksamkeit häufig keine signifikanten Unterschiede zwischen den Lehrmethoden. Personenbezogene Aspekte wie Potential, Interesse, Motivation, Lernstrategien etc. scheinen hier eine bedeutsamere Rolle zu spielen als die Lehrmethode. Wenn sich allerdings bedeutsame Unterschiede ergaben, dann zumeist in kurzfristigen Tests oder in Wissensüberprüfungen, in denen es für die Studierenden „um nichts ging“ – also in den Situationen, in denen die Studierenden nicht

kompensieren konnten oder mussten, um eine möglicherweise unterschiedlich gute Instruktion auszugleichen. Die vorliegende Lernendengruppe hat sicherlich bereits einige Erfahrung mit unterschiedlich gutem Lehrmaterial und weiß, wann und wie sie das durch Lernstrategien und Fleiß ausgleichen kann und muss. Ggf. wären bei einer heterogeneren Lernendengruppe oder einer Stichprobe von weniger erfahrenen Lernenden deutlichere Lernwirksamkeitsunterschiede beobachtbar gewesen.
Darüber hinaus ergaben sich – wenn es zu signifikanten Unterschieden kam – diese zumeist im Vergleich zwischen den E-Learning-Anwendungen im Allgemeinen und der Präsenzlehre. Dabei schnitten die E-Learning-Anwendungen generell besser ab. Diese Vorteile lassen sich in beiden Studien hinsichtlich der Lerneffektivität und -effizienz nachweisen sowie in Studie I auch für die meisten Akzeptanzvariablen[56] belegen.
Ähnliches gilt für den Vergleich zwischen Web-based Training und E-Lecture. Wenn signifikante Unterschiede hinsichtlich Akzeptanz oder Lernwirksamkeit zwischen diesen beiden E-Learning-Varianten auftraten, schnitten die Web-based Trainings – in Studie II die mit ARCS-Strategien angereicherten WBT-Varianten – signifikant besser ab als die E-Lectures. Dies gilt insbesondere für die Phasen der Klausurvorbereitung innerhalb beider Studien. Insbesondere in diesen Phasen werden die WBTs – sogar die K-(bzw. 0-)WBTs – deutlich häufiger und länger von den Studierenden genutzt als die E-Lectures. Darüber hinaus zeigt die Art der Bearbeitung dieser beiden sehr unterschiedlichen E-Learning-Anwendungen, dass sie in der von den Studierenden in Studie I gewünschten Form genutzt werden: Die E-Lecture-Nutzer nutzen nur einzelne E-Lectures, diese aber sehr lange. Die WBT-Nutzer bearbeiten hingegen viele verschiedene WBTs recht kurz. Hieraus lässt sich schließen, dass die E-Lecture-Nutzer ganze Vorlesungssitzungen ansehen/nachholen, die WBT-Nutzer im Gegensatz dazu einzelne Inhalte nachschlagen und vermutlich die Übungsaufgaben zum Wiederholen nutzen. Für eine solche Interpretation spricht auch, dass die WBT-Nutzung in Studie II bis zur Klausur immer weiter ansteigt, während die E-Lecture-Nutzung ein paar Tage vor der MAP deutlich abnimmt. Das Nachholen verpasster Sitzungen sollte ein paar Tage vor der Klausur beendet sein und die Wiederholung/Übung der Inhalte sollte in diesen letzten Tagen vor der MAP in den Vordergrund rücken.

Die Akzeptanz- und Nutzungsbefragungen aus beiden Studien zeigen eindrucksvoll, dass die Studierenden die Stärken und Schwächen **aller** eingesetzten Lehr- und Lernmethoden gut einschätzen können und sie deshalb je nach gerade vorherrschendem Ziel (erste Wissensaufnahme im Wintersemester oder Klausurvorbereitung im Sommersemester) unterschiedlich einsetzen möchten bzw. sie auch konkret unterschiedlich nutzen. Das hat

[56] Ein Akzeptanzvergleich zwischen E-Learning und Vorlesung wurde im Rahmen der zweiten Studie nicht durchgeführt.

sicherlich Einfluss auf die unterschiedlich gute Beurteilung der Lehrmethoden in den verschiedenen Phasen des Lernprozesses. Eine generelle Aussage dazu, welche Lehrmethode die geeignetste ist, verbietet sich folglich. Je nach Lernziel gilt es stattdessen, die jeweils geeignetste Methode auszuwählen. Studie I hat deutlich gezeigt, dass den Lernenden diese Passung bewusst und wichtig ist. So wünschen sich die Studierenden zur Wissensaneignung und zum Austausch mit anderen Lernenden die Vorlesung und zum Nachholen verpasster Vorlesungssitzungen die E-Lectures. Zur Prüfungsvorbereitung möchten sie hingegen WBTs, Übungsaufgaben und das Vorlesungsskript nutzen.

Die Ergebnisse beider Studien sprechen folglich weder für den alleinigen Einsatz der Präsenzlehre noch für eine vollständige Ersetzung dieser durch E-Learning-Anwendungen. Schon die Befragung zu den Nutzungswünschen in Studie I zeigt, dass die Vorlesung von den Studierenden am wenigsten „nicht gewünscht“ wird. Trotz aller Zusatzangebote möchten sie folglich am wenigsten auf die gewohnte Präsenzveranstaltung verzichten. Am sinnvollsten erscheint daher ein Blended Learning-Konzept, in dem Vorteile beider Lehrvarianten genutzt werden. Die Wünsche der Studierenden sprechen dementsprechend für die Fortsetzung der traditionellen Vorlesung sowie die Unterstützung dieser durch E-Learning-Maßnahmen. Dadurch wird die Prüfungsvorbereitung für die Studierenden erleichtert, Flexibilität ermöglicht und Chancengleichheit für diejenigen hergestellt, die aus familiären oder beruflichen Gründen nicht oder nicht vollständig an der Vorlesung teilnehmen können.

Eine solche Ergänzung der traditionellen Hochschullehre durch E-Learning verursacht allerdings zusätzliche Kosten. Kann oder will man diese nicht durch Einsparungen im Rahmen der Präsenzlehre auffangen, müssen sich diese zusätzlich investierten Ressourcen auszahlen – hinsichtlich der Akzeptanz der Studierenden sowie ihrer Lernerfolge. Es stellt sich demgemäß die Frage, wie man möglichst ökonomisch eine sehr lernförderliche und durch die Lernenden gut akzeptierte E-Learning-Anwendung erstellt. Die Ergebnisse aus Studie II zeigen hier deutlich, dass die vollständige Umsetzung des ARCS-Modells sinnvoll ist und nicht zu einer Überfrachtung der Anwendungen oder einer Überforderung der Lernenden führt. Die Ergebnisse legen die vollständige Umsetzung der Strategien insbesondere bei langfristigem Einsatz von E-Learning nahe. Der vollständige Einsatz der vier Komponenten erscheint jedoch nicht immer notwendig. Es zeigte sich, dass insbesondere die Komponenten „Erfolgszuversicht“ und „Zufriedenheit“ umgesetzt werden sollten, um die Akzeptanz und Wahrnehmung der Lernförderlichkeit einer E-Learning-Anwendung zu optimieren. Der Einsatz der Komponente „Relevanz“ erwies sich im Gegensatz dazu als stark inhaltsabhängig. Die hierunter zusammengefassten Strategien

können scheinbar bei der Vermittlung von praxisnahen Inhalten vernachlässigt werden. Sie sollten allerdings immer dann zum Einsatz kommen, wenn die praktische Anwendung sich nicht direkt aus den Inhalten heraus ergibt, wie dies in Studie II beim Thema „Konditionierung“ der Fall war.
Die Komponente „Aufmerksamkeit“ erwies sich in Studie II als schwächste (und von den Studierenden als am wenigsten wichtig eingestufte) Komponente und könnte ggf. bei der Erstellung von E-Learning-Anwendungen vernachlässigt werden. Um die Notwendigkeit ihrer Umsetzung endgültig zu klären, wäre es interessant, ein ARCS-WBT mit einem RCS-WBT zu vergleichen – also einem WBT, in dem alle Komponenten außer der Aufmerksamkeitskomponente umgesetzt wurden. Die bisher beobachteten Ergebnisse sprechen dafür, dass die drei Komponenten „Erfolgszuversicht“, „Zufriedenheit“ und „Relevanz“ die Akzeptanz und Lernwirksamkeit vermutlich ähnlich gut fördern können wie dies durch eine E-Learning-Version der Fall ist, in der alle vier Komponenten umgesetzt wurden. Dies ist deshalb naheliegend, da durch die Umsetzung von Lernspielen, Übungsaufgaben, einem pädagogischen Agenten etc. ebenfalls Wahrnehmungsvielfalt, Variation und Abwechslung erzeugt wird – Aspekte, die nach Keller (2010) zur Förderung der Aufmerksamkeit genannt werden. Diese Hypothese könnte in einer weiteren Studie überprüft werden.

Aus den Ergebnissen der zweiten Studie können zwar konkrete Aussagen zur Notwendigkeit der vier ARCS-Komponenten insgesamt abgeleitet werden. Fraglich bleibt dennoch, welche Maßnahmen (z.B. Lernspiele, Übungsaufgaben, Inhaltsverzeichnis etc.) innerhalb der Komponenten „Erfolgszuversicht“ und „Zufriedenheit“ es konkret sind, die einen bedeutenden Einfluss auf die Lernförderlichkeitswahrnehmung sowie die Einstellungs- und Verhaltensakzeptanz haben. Auch dies könnte durch weitere Studien – mittels konkreter Beobachtung im Labor oder durch ein User-Tracking – näher beleuchtet werden.
Klar gezeigt hat sich allerdings bereits in beiden Studien die bedeutende Rolle von Übungsaufgaben. Die Möglichkeit, das Gelernte wiederholt zu überprüfen, scheint für die Studierenden besonders wichtig zu sein. Bereits in Studie I zeigte sich eine hohe Akzeptanz der Übungsaufgaben und ein deutlicher Wunsch der Studierenden, diese im Rahmen des Hochschulkontextes nutzen zu wollen. Sie nutzen die WBTs mit Übungsaufgaben häufiger als die ohne Übungsaufgaben und zeigen eine höhere Akzeptanz für diese WBTs. In Studie II werden die ARCS-, S- und C-WBTs, in denen Übungsaufgaben und/oder Lernspiele enthalten sind, besser hinsichtlich ihrer Lernförderlichkeit und Akzeptanz beurteilt und die elaborierten (bzw. ARCS-)WBTs werden etwa dreimal häufiger zur Klausurvorbereitung genutzt als die kurzen (bzw. 0-)WBTs. Bei der Einschätzung wichtiger E-Learning-

Eigenschaften beurteilen die Studierenden insbesondere die Übungsaufgaben und Feedbacks als sehr wichtig.

Es können allerdings nicht nur die Übungsaufgaben sein, die für die bessere Beurteilung und häufigere Nutzung der WBTs mit Übungsaufgaben (Studie I) bzw. der ARCS-(bzw. E-)WBTs (Studie II) verantwortlich sind. Dagegen spricht zum Beispiel die Beobachtung, dass der Unterschied hinsichtlich der Nutzungshäufigkeit zwischen WBTs und E-Lectures in Studie I deutlich größer ist als der Unterschied zwischen WBTs mit und ohne Übungsaufgaben. Die ARCS-WBTs müssen folglich weitere Aspekte beinhalten, die sie für die intensive freiwillige Nutzung im Rahmen der Prüfungsvorbereitung attraktiv machen. Hier spielt sicherlich auch eine Rolle, dass sie es ermöglichen, unklare Inhalte schnell und gezielt nachzuschlagen. Zudem sind die Studierenden es sicherlich eher gewöhnt, mit geschriebenem Text zu lernen als mit gesprochener Sprache. All dies bieten die kurzen (0-)WBTs allerdings auch. Dementsprechend wäre es spannend zu prüfen, welche Eigenschaften es neben den Übungsaufgaben sind, die die meisten Studierenden dazu bewogen haben, die ARCS-WBTs für die Klausurvorbereitung auszuwählen.

Dass gerade die Übungsaufgaben für die hier befragten Lernenden eine große Rolle spielen, kann eng mit der ausgewählten Lernendenstichprobe und dem Lernsetting zusammenhängen. Wie in Studie II berichtet, scheinen die befragten Studierenden tendenziell eher weniger erfolgszuversichtlich zu sein. Darüber hinaus wurde die Untersuchung in einem Kontext mit deutlichem Lerndruck durchgeführt. Die Studierenden waren vermutlich nicht ausschließlich intrinsisch motiviert, die Inhalte zu lernen, sondern taten dies, um die abschließende Klausur zu bestehen. Dies könnte auch erklären, warum für die Stichprobe die Erfolgszuversicht und Zufriedenheit so eng verknüpft zu sein scheinen. Es handelt sich hier um eine eher homogene Lernendengruppe, die bereits auf eine lange Lerngeschichte zurückblickt. Diese Lerngeschichte muss darüber hinaus relativ erfolgreich gewesen sein, da alle Studienteilnehmer über ein Abitur verfügen. Aus den Ergebnissen zu dieser Stichprobe ist es möglich, Hypothesen für andere Lernendengruppen aufzustellen, die in weiteren Untersuchungen überprüft werden können. So zeigte sich in Studie I beispielsweise, dass sich die Studierenden für andere Studienmodule ebenfalls Übungsfragen wünschen, für interessante Themen außerhalb der Universität jedoch lieber Web-based Trainings und E-Lectures nutzen möchten – am seltensten Übungsaufgaben. Dies zeigt erneut eindrücklich, dass die Beurteilung der Lehrmethoden stark von dem Lernziel und der Lernmotivation der Lernenden abhängt. Ggf. finden freiwillig Lernende die Multimodalität (Komponente „Aufmerksamkeit“) enorm wichtig für ihren Lernprozess. Evtl. sind solchen Lernenden im Gegensatz dazu Übungsaufgaben (Komponente „Erfolgszuversicht“) gar nicht so wichtig, da sie keinen externen Lerndruck verspüren. Für

weniger geübte Lernende spielen die ARCS-Strategien darüber hinaus ggf. eine viel stärkere Rolle als für die hier untersuchte Stichprobe, da sie die Lernunterstützung mehr benötigen sollten. Es ist folglich durchaus denkbar, dass andere Lernendengruppen in anderen Lernkontexten andere Aspekte des ARCS-Modells bevorzugen.

10 Literaturverzeichnis

Adobe Systems GmbH (2012). *Adobe Connect*. Abgerufen am 11.06.2012 von http://www.adobe.com/de/products/connect/

Astleitner, H. (2000). Designing emotionally sound instruction: The FEASP-approach. *Instructional Science*, *28*(3), 169–198.

Astleitner, H. (2005). *Motivationsförderung im E-Learning: Stand der Forschung zum ARCS-Modell*. *Beitrag zur NET-ELC Jahrestagung. ETH Zürich.* Abgerufen am 27.03.2011 von http://www.uni-salzburg.at/pls/portal/ docs/1/499428.PDF

Astleitner, H. & Hufnagl, M. (2003). The effects of situation-outcome-expectancies and of arcs-strategies on self-regulated learning with web-lectures. *Journal of Educational Multimedia and Hypermedia*, *12*(4), 361–376.

Astleitner, H. & Lintner, P. (2004). The effects of arcs-strategies on self-regulated learning with instructional texts. *E-Journal of Instructional Science and Technology*, 7(1). Abgerufen am 08.07.2016 von http://files.eric.ed.gov/fulltext/EJ850349.pdf

Astleitner, H., Pasuchin, I. & Wiesner, C. (2006). Multimedia und Motivation – Modelle der Motivationspsychologie als Grundlage für die didaktische Mediengestaltung. *MedienPädagogik*, 1-19. Abgerufen am 08.07.2016 von http://www.medienpaed.com/globalassets/medienpaed/2006/astleitner0603.pdf

Beer, D., Hamburg, I., Lindecke, C. & Terstriep, J. (2003): *E-Learning: Kollaboration und veränderte Rollen im Lernprozess*. Gelsenkirchen: Institut Arbeit und Technik. Projektbericht des Instituts Arbeit und Technik, 2003-04. Abgerufen am 10.02.2011 von http://iat-info.iatge.de/aktuell/veroeff/2003/pb2003-04.html

Bell, T., Cockburn, A., McKenzie, B. & Vargo, J. (2001). *Digital lectures: If You make them, will students use them? Constraints of effective delivery of flexible learning systems*. Abgerufen am 25.11.2015 von http://imej.wfu.edu/articles/2001/2/06/

Bellon, T. & Oates, R. (2002). *Best practices in cyberspace: Motivating the online learner*. Abgerufen am 31.07.2014 von http://home.fau.edu/musgrove/web/ communication/Best%20 Practices%20in%20Cyberspace%20-%20Motivating% 20the%20online%20learner .pdf

Bendel, O. (2004). Merkmale, Ziele und Funktionen pädagogischer Agenten. In B. Bekavac, J. Herget & M. Rittberger (Hrsg.), *Informationen zwischen Kultur und Marktwirtschaft. Proceedings des 9. Internationalen Symposiums für Informationswissenschaft (ISI 2004)* (S. 213–226). Konstanz: UVK Verlagsgesellschaft mbH.

Bescherer, C. & Spannagel, C. (2011). *Fragebogen zur computerbezogenen Selbstwirksamkeit – reduziert*. Abgerufen am 04.10.2011 von www.ph-heidelberg.de/wp/spannagel/cuse/CUSE-D-r.doc

Bligh, D. A. (2000). *What's the use of lectures?*. San Francisco, California: Jossey-Bass Publishers.

Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Aufl.). Heidelberg: Springer Verlag.

Breiter, A., Aufenanger, S., Averbeck, I., Welling, S. & Wedjelek, M. (2013). Medienintegration in Grundschulen. *Schriftenreihe Medienforschung der Landesanstalt für Medien Nordrhein-Westfalen, Band 73*. Berlin: VISTAS Verlag GmbH.

Brömme, R. (2010). *Vergleich der Nutzerakzeptanz von Vorlesung, E-Lecture und Web-based Training* (Unveröffentlichte Bachelorarbeit). Justus-Liebig-Universität Gießen, Gießen.

Bühner, M., Schmidt-Atzert, L., Richter, S. & Grieshaber, E. (2002). Selbstbeurteilung der Aufmerksamkeit: Ein Vergleich zwischen Hirngeschädigten und Gesunden. *Zeitschrift für Neuropsychologie, 13*(4), 263–296.

Bürg, O. & Mandl, H. (2005). Akzeptanz von E-Learning in Unternehmen. *Zeitschrift für Personalpsychologie*, *4*(2), 75–85.

Bundesministerium für Bildung und Forschung. (2000). *Förderprogramm Neue Medien in der Bildung*. Abgerufen am 01.05.2012 von www.bmbf.de/pub/neue_medien_bildung.pdf

Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz. (2016). *Fischerei in Deutschland – Portal des Bundes und der Länder*. Abgerufen am 22.01.2016 von http://www.portal-fischerei.de/bundeslaender/

Bundesministerium für Wirtschaft und Arbeit & Bundesministerium für Bildung und Forschung. (2006). *Informationsgesellschaft Deutschland 2006*. Abgerufen am 09.05.2010 von www.bmbf.de/pub/aktionsprogramm_informationsgesellschaft_2006.pdf

Bundesministerium für Wirtschaft und Technologie. (2008). *E-Learning in KMU – Markt, Trends, Empfehlungen*. Abgerufen am 31.08.2011 von http://www.bmwi.de/BMWi/Redaktion/PDF/Publikationen/ Dokumentationen/dok-575-e-learning-in-kmu

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1118–1133.

Buzzetto-More, N. A. (2008). Student perceptions of various e-learning components. *Interdisciplinary Journal of E-Learning and Learning Objects*, *4*, 113–135.

Buzzetto-More, N. A. & Sweat-Guy, R. (2006). Incorporating the hybrid learning model into minority education at a historically black university. *Journal of Information Technology Education, 5*, 153–164.

Carey, J. M. (2001). Effective student outcomes: A comparison of online and face-to-face delivery modes. *DEOSNEWS*, *11*(9). Abgerufen am 08.07.2016 von http://learningdesign.psu.edu/assets/uploads/deos/deosnews11_9.pdf

Chu, G. C. & Schramm, W. (1967). *Learning from television: What the research says*. Washington, D.C.: National Association of Educational Broadcasters.

Chyung, Y., Winiecki, D. & Fenner, J. A. (1999*). Evaluation of effective interventions to solve the dropout problem in adult distance education*. Abgerufen am 10.05.2011 von http://coen.boisestate .edu/ychyung/edmedia.htm

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, *53*(4), 445–459.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Hillsdale, Michigan: Lawrence Erlbaum Associates.

Connolly, T. & Stansfield, M. (2007). Developing constructivist learning environments to enhance e-learning. In N. A. Buzzetto-More (Hrsg.), *Principles of effective online teaching* (S. 19–38). Santa Rosa, California: Informing Science Press.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. San Francisco, California: Jossey-Bass Publishers.

Dalsgaard, C. (2005). Pedagogical quality in e-learning. Designing e-learning from a learning theoretical approach. *E-learning and education*. Abgerufen am 04.05.2011 von http://eleed.campussource.de/archiv/78/index_html

Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, *39*(2), 223–238.

Deimann, M. (2007). *Entwicklung und Erprobung eines volitionalen Designmodells*. Abgerufen am 08.07.2016 von https://www.db-thueringen.de/servlets/MCRFileNodeServlet/dbt_derivate_00010831/deimann.pdf

Demetriadis, S. & Pombortsis, A. (2007). E-Lectures for flexible learning: a study on their learning efficiency. *Educational Technology & Society*, *10*(2), 147–157.

DeSouza, E. & Fleming, M. (2003). A comparison of in-class and online quizzes on student exam performance. *Journal of Computing in Higher Education*, *14*(2), 121–134.

Diehl, J.M. & Arbinger, R. (1992). *Einführung in die Inferenzstatistik*. Eschborn: Verlag Dietmar Klotz GmbH.

Diener, E., Emmons, R. A., Larsen, R. J. & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *46*(1), 71–75.

Dittler, U. & Jechle, T. (2009). eLearning in der Aus- und Weiterbildung. In P. Klimsa & L. J. Issing (Hrsg.), *Online-Lernen: Handbuch für Wissenschaft und Praxis* (S. 419–426). München: Oldenbourg Verlag.

Dittler, U. & Jechle, T. (2011). eLearning in der Aus- und Weiterbildung. In P. Klimsa & L. J. Issing (Hrsg.), *Online-Lernen - Planung, Realisation, Anwendung und Evaluation von Lehr- und Lernprozessen online* (2. Aufl., S. 419–426). München: Oldenbourg Verlag.

Döring, N. (2002). Online-Lernen. In P. Klimsa & L. J. Issing (Hrsg.), *Information und Lernen mit Multimedia und* Internet*. Lehrbuch für Studium und Praxis* (3. überarb. Aufl., S. 247–266). Weinheim: Beltz/PVU.

Domagk, S. & Niegemann, H.M. (2008). Motivationsdesign im Hochschulunterricht. In J. Zumbach & H. Mandl (Hrsg.), *Pädagogische Psychologie in Theorie und Praxis. Ein fallbasiertes Lehrbuch* (S. 205–211). Göttingen: Hogrefe Verlag.

Driscoll, M. (1998). *Web-based training. Using technology to design adult learning experiences*. SanFrancisco, California: Jossey-Bass Pfeiffer.

Dziuban, C. & Moskal, P. (2001). Evaluating distributed learning at metropolitan universities. *Educause Quarterly*, *24*(4), 60–61.

Ellis, L. & Mathis, D. (1985). College student learning from televised versus conventional classroom lectures: a controlled experiment. *Higher Education*, *14*(2), 165–173.

Farahat, T. & Astleitner, H. (2004). The effectiveness of web-based instruction in educating teachers – A quasi-experimental field study. *Salzburger Beiträge zur Erziehungswissenschaft*, *8*(1), 53–60.

Frenger, R. P. (2010). *E-Lecture: Was ist das?* Abgerufen am 10.03.2011 von http://fss.plone.uni-giessen.de/fss /fbz/svc/hrz/svc/medien/el/file/flyer-electures/file/Flyer-E-Lectures.pdf

Gagné, R. M. (2011). Die Bedingung des menschlichen Lernens. In D. H. Rost (Hrsg.), *Standardwerke aus Psychologie und Pädagogik Reprints* (5. Aufl., Band 6). Münster: Waxmann Verlag.

Ginsberg, M. B. & Wlodkowski, R. J. (2009). Professional learning to promote motivation and academic performance among diverse adults. In D. Bamford-Reese, B. Doyle, B. Klein-Collins & J. Wertheim (Hrsg.), *CAEL FORUM and NEWS: Learning Never Ends* (S. 23–32). Chicago, Illinois: Council for Adult & Experiential Education.

Glowalla, U. (2003). E-Learning im Einsatz: Befunde und Implikationen. In M. Franzen (Hrsg.), *Mensch und E-Learning. Beiträge zu E-Didaktik und darüber hinaus* (S. 39–51). Aarau: Sauerländer Verlage AG.

Glowalla, U. (2004). Utility und Usability von E-Learning am Beispiel von Lecture-on-demand Anwendungen. In C. Steffens, M. Thüring & L. Urbas (Hrsg.), *Entwerfen und Gestalten. 5. Berliner Werkstatt Mensch-Maschine-Systeme* (Band 18, S. 603–621). Düsseldorf: VDI Verlag GmbH.

Glowalla, U. & Glowalla, G. (2005). *Instruktionspsychologie für Lehramtsstudierende. Eine interaktive NetLecture DVD* (2. Aufl.). Gießen: Lerndesign Publishing.

Glowalla, U., Glowalla G. & Görlich, S. (2004). Verbessern von Vorlesungen durch ELearning Komponenten. *i-com: Zeitschrift für interaktive und kooperative Medien, 3*(2), 57–62.

Glowalla, U., Glowalla, G. & Kohnert, A. (2001). *EVALIS - Evaluation interaktiven Studierens*. Gütersloh: Verlag Bertelsmann Stiftung.

Glowalla, U., Glowalla, G. & Kohnert, A. (2002). Studierverhalten in Online-Bildungsangeboten. In P. Klimsa & L. J. Issing (Hrsg.), *Information und Lernen mit Multimedia* (3. vollst. überarb. Aufl., S. 415–434). Weinheim: Beltz Psychologie Verlags Union.

Glowalla, U. & Häfele, G. (1997). Einsatz elektronischer Medien: Befunde, Probleme und Perspektiven. In P. Klimsa & L. J. Issing (Hrsg.), *Information und Lernen mit Multimedia* (2. überarb. Aufl., S. 415–434). Weinheim: Beltz Psychologie Verlags Union.

Glowalla, U., Herder, M., Süße, C. & Koch, N. (2011). Methoden und Ergebnisse der Evaluation elektronischer Lernangebote. In P. Klimsa & L. J. Issing (Hrsg.), *Online Lernen* (2. Aufl., S. 309–328). München: Oldenbourg Wissenschaftsverlag.

Glowalla, U., Kohnert, A. & Schneider, S. (2003). Entwicklung wissensdiagnostischer Module für E-Learning Anwendungen. In A. Bode, J. Desel, S. Rathmayer & M. Wessner (Hrsg.), *DeLFI 2003: Die 1. e-Learning Fachtagung Informatik, 16.-18.9.2003, Garching b. München., Lecture Notes in Informatics (LNI).* Bonn: Gesellschaft für Informatik.

Glowalla, U. & Schoop, E. (1996). *Deutscher Multimedia Kongress `96 – Perspektiven multimedialer Kommunikation*. Heidelberg: Springer-Verlag.

Görlich, S. (2006). *Fundierung und Integration von E-Learning-Komponenten in die Präsenzlehre.* Inaugural Dissertation. Justus-Liebig-Universität Gießen. Abgerufen am 29.08.2011 von http://geb.uni-giessen.de/geb/volltexte/2007/4535/pdf/GoerlichSascha-2007-02-07.pdf

Hamaker, C. (1986). The effects of adjunct questions on prose learning. *Review of Educational Research*, *56*(2), 212–242.

Hamburg, I. & Terstriep, J. (2003). Verteilt und doch gemeinsam – Portale als Instrument kollaborativer Lernprozesse. In D. Beer, I. Hamburg, C. Lindecke & J. Terstriep (Hrsg.), *E-Learning: Kollaboration und veränderte Rollen im Lernprozess*. Gelsenkirchen: Institut Arbeit und Technik. Abgerufen am 10.02.2011 von http://iat-info.iatge.de/aktuell/veroeff/2003/pb2003-04.html

Hasebrook, J. P. (2002). Co-operative and interactive distance learning: Application of team-oriented and selective learning strategies in a european bank. *Journal of Universal Computer Science*, *8*(9), 837–847.

Heckhausen, H. & Rheinberg, F. (1980). Lernmotivation im Unterricht, erneut betrachtet. *Unterrichtswissenschaft, 8*(1), 7–47.

Helic, D., Maurer, H. & Scherbakov, N. (2000). *Web based training: What do we expect from the system*. Abgerufen am 28.02.2010 von http://coronet.iicm.tugraz.at/denis/pubs/icce2000.pdf

Hesse, F. W. & Mandl, H. (2000). Neue Technik verlangt neue pädagogische Konzepte. In Bertelsmann Stiftung & Heinz Nixdorf Stiftung (Hrsg.), *Studium online. Hochschulentwicklung durch neue Medien* (S. 31–50). Gütersloh: Verlag Bertelsmann Stiftung.

Hessische Staatskanzlei (2013). *Verordnung über den Umfang der Lehrverpflichtung des wissenschaftlichen und künstlerischen Personals an den Hochschulen des Landes (Lehrverpflichtungsverordnung)*. Melsungen: A. Bernecker Verlag GmbH.

Hochschuldidaktisches Kompetenzzentrum der JLU (2016). *"Die Einheit von Forschung und Lehre leben": Das ist Ziel des hochschuldidaktischen Kompetenzzentrums im ZfbK der JLU.* Abgerufen am 26.02.2016 von https://www.uni-giessen.de/fbz/zentren/zfbk/didaktik/hdk

Holzinger, A. (2000). *Basiswissen Multimedia. Band 2: Lernen*. Würzburg: Vogel Fachbuch.

Huang, W., Huang, W., Diefes-Dux, H. & Imbrie, P. K. (2006). A preliminary validation of attention, relevance, confidence and satisfaction model-based instructional material motivational survey in a computer-based tutorial setting. *British Journal of Educational Technology, 37*(2), 243–259.

Ihringer, S. (2000). E-Learning: Modetrend oder neue Potenziale netzgestützer Medien für das Lernen im öffentlichen Dienst?. *SpLITer - IT-Nachrichten der Berliner Verwaltung*, *10*(2).

Institut für Medien- und Kompetenzforschung (2010). *Telefonische Befragung zum Einsatz von eLearning in deutschen Großunternehmen*. Abgerufen am 13.07.2016 von http://www.mmb-institut.de/projekte/digitales-lernen/Einsatz-von-E-Learning-in-deutschen-Grossunternehmen.pdf

International Telecommunication Union (2015). *Global ICT developments*. Abgerufen am 08.01.2016 von http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx

Jacobs, B. (2001). *Aufgaben stellen und Feedback geben*. Abgerufen am 02.12.2015 von http://psydok.sulb.uni-saarland.de/volltexte/2004/438/pdf/feedback.pdf

Johnson, S. D., Aragon, S. R., Shaik, N. & Palma-Rivas, N. (2000). Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments. *Journal of Interactive Learning Research, 11*(1), 29–49.

Justus-Liebig-Universität Gießen (2012). *Stud.IP*. Abgerufen am 19.06.2012 von http://www.uni-giessen.de/cms/fbz/svc/hrz/org/mitarb/abt/2/el/werkzeuge/studip

Kaltenbaek, J. (2009). Hochschule online – Online Lehren und Lernen in der Hochschule. In L.J. Issing & P. Klimsa. (Hrsg.), *Online-Lernen. Handbuch für Wissenschaft und Praxis* (S. 367–388). München: Oldenbourg Wissenschaftsverlag.

Keller, J. M. (2006). *Development of two measures of learner motivation*. Florida State University. Abgerufen am 09.07.2016 von https://www.google.de/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwj70ImtjufNAhXEtBoKHeWQAAoQFggjMAA&url=https%3A%2F%2Folpcorps.wikispaces.com%2Ffile%2Fview%2FARCSMEA%2BPartial%2BDraft%2B060222.doc&usg=AFQjCNEfxRdXY0iZJLwluNtkD6w884HAWg&sig2=A5lneDmiFa6kFqcSsI1ZEA&cad=rja

Keller, J. M. (2010). *Motivational design for learning and performance. The arcs model approach*. New York: Springer Verlag.

Keller, J. M. & Back, A. (2004). *Blended-Learning-Projekte im Unternehmen. Arbeitsberichte des Learning Center der Universität St. Gallen* (S. 1–31). Abgerufen am 22.01.2010 von http://web.iwi.unisg.ch/org/iwi/iwi_pub.nsf/wwwPublRecentGer/A4DB5FE8C6CF 50BAC125759A004869FA/$file/AB_Blended_Learning.pdf

Keller, J. M. & Kopp, T. W. (1987). An application of the arcs model of motivational design. In C. M. Reigeluth (Hrsg.), *Instructional theories in action. Lessons illustrating selected theories and models* (S. 289–320). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Kerres, M. (2009). *Multimediale und telemediale Lernumgebungen* (2. vollst. überarb. Auflage). München: Oldenbourg Wissenschaftsverlag.

Kerres, M. & Stratmann, J. (2007). E-University: Zur systematischen Integration von IT-Anwendungen in Kernprozessen der Hochschule. In R. Keil, M. Kerres & R. Schulmeister (Hrsg.), *eUniversity – Update Bologna* (S. 33–44). Münster: Waxmann Verlag.

Klauer, K. J. (1985). Framework for a theory of teaching. *Teaching & Teacher Education*, *1*(1), 5–17.

Klauser, F., Kim, H. & Born, V. (2002). *Erfahrung, Einstellung und Erwartung der Lernenden – entscheidende Determinanten netzbasierten Lernens*. Abgerufen am 27.03.2014 von http://www.bwpat.de/ausgabe2/ klauser-kim-born_bwpat2.html

Kleimann, B., Özkilic, M. & Göcks, M. (2008). *Studieren im Web 2.0 – Studienbezogene Web- und E-Learning-Dienste*. Abgerufen am 07.07.2011 von https://hisbus.his.de/ hisbus/docs/hisbus 21.pdf

Kleimann, B. & Schmid, U. (2006). *eReadiness an deutschen Hochschulen – Ergebnisse einer Umfrage zum Stand von IT-Management und E-Learning*. Abgerufen am 27.03.2014 von http://www.mmkh.de/fileadmin dokumente/Publikationen/eReadiness_der_ deutschen_ Hochschulen_Kleimann_Schmid.pdf

Klimsa, P. & Issing, J. L. (2011). *Online-Lernen* (2. Aufl.). München: Oldenburg Wissenschaftsverlag GmbH.

Knop, K., Hefner, D., Schmitt, S. & Vorderer, P. (2015). *Mediatisierung mobil. Handy- und Internetnutzung von Kindern und Jugendlichen*. *Schriftenreihe Medienforschung der Landesanstalt für Medien Nordrhein-Westfalen (LfM)* (Band 77). Leipzig: Vistas.

Köhler, T., Kahnwald, N. & Reitmeier, M. (2008). Lehren und Lernen mit Multimedia und Internet. In B. Batinic & M. Appel (Hrsg.), *Medienpsychologie* (S. 478–501). Heidelberg: Springer Verlag.

Krapp, A. & Weidenmann, B. (2001). *Pädagogische Psychologie* (4. vollst. überarb. Aufl.). Weinheim: Beltz Psychologie Verlags Union.

Krause, U.-M. (2007). *Feedback und kooperatives Lernen.* Münster: Waxmann Verlag.

Kühn, A. (2009). Lehramtsstudierende in Baden-Württemberg. *Statistisches Monatsheft Baden-Württemberg, 12*, 3–6. Abgerufen am 05.07.2016 von https://www.statistik-bw.de/Service/Veroeff/Monatshefte/PDF/Beitrag09_12_01.pdf

Kulik, C.-L. C. & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior*, *7*(1-2), 77–94.

Larsen, D. P., Butler, A. C. & Roediger III, H. L. (2009). Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Medical Education, 43*(12), 1174–1181.

LaRose, R., Gregg, J. & Eastin, M. (1998). Audiographic telecourses for the web: An experiment. *Journal of Computer-Mediated Communication*, *4*(2). Abgerufen am 13.07.2016 von http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.1998.tb00093.x/full

Lindecke, C. & Beer, D. (2003). Vom dozieren zum beraten – veränderte Rollen im Lernprozess der betrieblichen Weiterbildung. In D. Beer, I. Hamburg, C. Lindecke & J. Terstriep (Hrsg.), *E-Learning: Kollaboration und veränderte Rollen im Lernprozess* (S. 14–27). Gelsenkirchen: Institut Arbeit und Technik. Abgerufen am 05.07.2016 von http://iat-info.iatge.de/aktuell/veroeff/2003/ pb2003-04.html

Lyon Jr., H. C., Healy, J. C., Bell, J. R., O'Donnell, J. F., Shultz, E. K., Moore-West, M. et al. (1992). PlanAlyzer, an interactive computer-assisted program to teach clinical problem solving in diagnosing anemia and coronary artery disease. *Academic Medicine*, *67*(12), 821–828.

Maki, R. H., Maki, W. S., Patterson, M. & Whittaker, P. D. (2000). Evaluation of a web-based introductory psychology course: I. Learning and satisfaction in on-line versus lecture courses. *Behavior Research Methods, Instruments, & Computers*, *32*(2), 230–239.

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science, 5*(4), 333–369.

Malone, T. W. & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Hrsg.), *Aptitude, learning, and instruction. Volume 3: Conative and affective process analyses* (S. 223–253). London: Lawrence Erlbaum Associates.

McKeachie, W. J., Pintrich, P. R., Lin, Y.-G. & Smith, D. A. F. (1986). *Teaching and learning in the college classroom: a review of the research literature*. Ann Arbor, Michigan: National center for research to improve postsecondary teaching and learning.

McLeish, J. (1968). *The lecture method*. Cambridge, U.K.: Cambridge Institute of Education.

Means, T. B., Jonassen, D. H. & Dwyer, F. M. (1997). Enhancing relevance: Embedded arcs strategies vs. purpose. *Educational Technology Research & Development*, *45*(1), 5–17.

Müller-Böling, D. & Müller, M. (1986). *Akzeptanzfaktoren der Bürokommunikation*. München: Oldenbourg Wissenschaftsverlag.

Niegemann, H. M., Domagk, S., Hessel, S., Hein, A., Hupfer, M. & Zobel, A. (2008). *Kompendium multimediales Lernen*. Heidelberg: Springer Verlag.

Norton, W. (2000). *Designing web-based training*. New York, New York: John Wiley & Sons.

Nungester, R. J. & Duchastel, P. C. (1982). Testing versus review: Effects on retention. *Journal of Educational Psychology*, *74*(1), 18–22.

Olson, T. M. & Wisher, R. A. (2002). The effectiveness of web-based instruction: An initial inquiry. *International Review of Research in Open and Distance Learning, 3*(2). Abgerufen am 13.07.2016 von http://www.irrodl.org/index.php/irrodl/article/view/103/182

Overmann, M. (2002). *Multimediales Lernen – Perspektiven, Forderungen, Probleme*. Abgerufen am 05.04.2011 von http://www.lehrer-online.de/dyn/bin/284792-285103-1-multimediales_lernen.pdf

Prenzel, M. (1996). Bedingungen für selbstbestimmt motiviertes und interessiertes Lernen im Studium. In J. Lompscher & H. Mandl (Hrsg.), *Lehr- und Lernprobleme im Studium. Bedingungen und Veränderungsmöglichkeiten* (S. 11–22). Bern: Hans Huber.

Prenzel, M., Drechsel, B. & Kramer, K. (1998). Lernmotivation im kaufmännischen Unterricht: Die Sicht von Auszubildenden und Lehrkräften. In K. Beck & R. Dubs (Hrsg.), *Kompetenzentwicklung in der Berufserziehung* (S. 169–187). Stuttgart: Steiner.

Questback. (2016). *Feedback Management.* Abgerufen am 07.07.2016 von https://www.questback.com/de/

Rasch, B., Friese, M. & Hofmann, W. J. (2009). *Quantitative Methoden. 2. Einführung in die Statistik für Psychologen*. Berlin, Heidelberg: Springer Verlag.

Rasch, B., Friese, M., Hofmann, W. J. & Naumann, E. (2008). *Quantitative Methoden. 1. Einführung in die Statistik für Psychologen*. Berlin, Heidelberg: Springer Verlag.

Rindermann, H. & Amelang, M. (1994). Entwicklung und Erprobung eines Fragebogens zur studentischen Veranstaltungsevaluation. *Zeitschrift zu Theorie und Praxis erziehungswissenschaftlicher Forschung, 8*(2), 131–151.

Rivera, J. C., McAlister, M. K. & Rice, M. L. (2002). A comparison of student outcomes & satisfaction between traditional & web based course offerings. *Online Journal of Distance Learning Administration*, *5*(3). Abgerufen am 13.07.2016 von http://www.westga.edu/~distance/ojdla/fall53/rivera53.html

Roediger III, H. L. & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*(3), 181–210.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, *76*(4), 647–658.

Sanders, D. W. & Morrison-Shetlar, A. I. (2001). Student attitudes toward web-enhanced instruction in an introductory biology course. *Journal of Research on Computing in Education, 33*(3), 251–262.

Scheele, N., Wessels, A. & Effelsberg, W. (2004). *Die interaktive Vorlesung in der Praxis*. Abgerufen am 10.04.2010 von http://subs.emis.de/LNI/Proceedings/ Proceedings52/GI.-.Proceedings.52-25.pdf

Scheer, A. (2009). *Vortrag im Rahmen der E-Learning Pressekonferenz bei der CeBIT 2009*. Abgerufen am 06.04.2011 von http://www.bitkom.org/files/documents/BITKOM_ Vortrag_Prof__ Scheer_E-Learning_PK_04_03_ 2009.pdf

Schilbach, L., Wohlschläger, A. M., Krämer, N. C., Newen, A., Shah, N. J., Fink, G. R. & Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, *44*(5), 718–730.

Schlag, B. (2013). *Lern- und Leistungsmotivation* (4. überarb. und aktual. Aufl.). Wiesbaden: Springer Verlag.

Schlageter, G. & Feldmann, B. (2002). E-Learning im Hochschulbereich: der Weg zu lernzentrierten Bildungssystemen. In P. Klimsa & L. J. Issing (Hrsg.), *Information und Lernen mit Multimedia und Internet* (3. Aufl., S. 347–357). Weinheim: Beltz Psychologie Verlags Union.

Schramm, W. (1962). What we know about learning from instructional television. In Institute for Communication Research (Hrsg.), *Educational television. The next ten years* (S. 52–76). Stanford, California: Stanford University.

Schüler, H. (2002). *Einsatz von eLearning-content*. Abgerufen am 03.05.2011 von http://www. contentmanager.de/magazin/artikel_246_einsatz_von_elearning-content.html

Schuler, H., Prochaska, M. & Frintrup, A. (2001). *LMI - Leistungsmotivationsinventar*. Göttingen: Hogrefe Verlag.

Schulmeister, R. (2007). *Grundlagen hypermedialer Lernsysteme: Theorie – Didaktik – Design*. München: Oldenbourg Wissenschaftsverlag.

Schumacher, J. (2003). SWLS – Satisfaction with life scale. In J. Schumacher, A. Klaiberg & E. Brähler (Hrsg.), *Diagnostische Verfahren zu Lebensqualität und Wohlbefinden* (S. 305–309). Göttingen: Hogrefe Verlag.

Schutte, J. G. (1997). Virtual teaching in higher education: The new intellectual superhighway or just another traffic jam? Abgerufen am 14.12.2015 von http://ddi.cs.uni-potsdam.de/HyFISCH/Teleteaching/VirtualTeachingSchutte.htm

Shomaker, T. S., Ricks, D. J. & Hale, D. C. (2002). A prospective, randomized controlled study of computer-assisted learning in parasitology. *Academic Medicine*, *77*(5), 446–450.

Siegert, M., Schlieck, A., Jäger, T., Gotthardt, M. & Glowalla U. (2003). Der Einfluss von Übungsaufgaben auf das Lernen mit einer computerbasierten Lerneinheit – eine experimentelle Evaluationsstudie. In F. Puppe, J. Albert, J. Bernauer, M. Fischer, R. Klar & J. Leven (Hrsg.), *Rechnergestützte Lehr- und Lernsysteme in der Medizin*. Aachen: Shaker-Verlag.

Soliman, I. (1999). *Lecturing to large groups. Teaching and learning centre university of New England*. Abgerufen am 18.04.2012 von http://amec.glp.net/c/document_ library/get_file? p_l_id=841375& folderId=1077184&name=DLFE-20797.pdf

Song, S. H. & Keller, J. M. (2001). Effectiveness of motivationally adaptive computer-assisted instruction on the dynamic aspects of motivation. *Educational Technology, Research and Development*, *49*(2), 5–22.

Spickard III, A. W., Alraje, N., Cordray, D. & Gigante, J. (2002). Learning about screening using an online or live lecture. *Journal of General Internal Medicine*, *17*(7), 540–545.

Spinath, B., Stiensmeier-Pelster, J., Schöne, C. & Dickhäuser, O. (2002). *SELLMO – Skalen zur Erfassung der Lern- und Leistungsmotivation*. Göttingen: Hogrefe Verlag.

Spitzer, D. R. (1995). *SuperMotivation – A blueprint for energizing your organization from top to bottom*. New York, New York: American Management Association.

Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Hochschulzulassung und Kapazitätsermittlung*. Abgerufen am 04.06.2012 von http://www.kmk.org/wissenschaft-hochschule/studentische-angelegenheiten/hochschulzulassung-und-kapazitaetsermittlung.html

Statista (2014). *Online-Aktivitäten am PC bzw. Laptop in der Deutsch- und Westschweiz im Jahr 2014.* Abgerufen am 17.02.2016 von http://de.statista.com/statistik/daten/studie/459672/umfrage/online-aktivitaeten-am-pc-laptop-in-der-deutsch-und-westschweiz/

Statista (2015). *Anteil der Jugendlichen in Deutschland, die einen eigenen Computer oder Laptop besitzen in den Jahren 2007 bis 2015.* Abgerufen am 17.02.2016 von http://de.statista.com/statistik/daten/studie/29432/umfrage/anteil-der-jugendlichen-der-einen-computer-oder-laptop-besitzt/

Streitenberger, E. (nicht datiert). *Zahl der Lehramtsstudierenden nach Geschlecht*. Abgerufen am 14.08.2015 von http://www.uni-heidelberg.de/imperia/md/content/studium/kontakt/zlb/la-statistik/tabellen/lehramtsstatistik_1998ff_gender.pdf

Teo, T. (2010). A structural equation modelling of factors influencing student teachers' satisfaction with e-learning. *British Journal of Educational Technology*, *41*(6), 150–152.

Tergan, S.-O. (2004). Was macht Lernen erfolgreich? Die Sicht der Wissenschaft. In S.-O. Tergan & P. Schenkel (Hrsg.), *Was macht E-Learning erfolgreich?* (S. 15–28). Berlin: Springer Verlag.

Trivantis. (2016). *lectora16 - The Mobile Learning Game Changer.* Abgerufen am 07.07.2016 von www.trivantis.com

Visser, L. (1998). *The development of motivational communication in distance education support* (Dissertation). Den Haag: CIP – Gegevens Koninklijke Bibliotheek

Vogt, M. (2002). *Lernen mit Maus und Modem – Effekte verschiedener Lehrmedien auf Lernerfolg und Akzeptanz* (Unveröffentlichte Diplomarbeit). Justus-Liebig-Universität Gießen, Gießen.

Vogt, M., Görlich, S. & Glowalla, U. (2006). Learning from face-to-face or computer-supported courses: Differences with respect to user acceptance and learning outcomes. In T. Reeves & S. Yamashita (Hrsg.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2006* (S. 2459 –2464). Chesapeake, Virginia: Association for the Advancement of Computing in Education (AACE).

Vollmeyer, R. & Rheinberg, F. (1998). Motivationale Einflüsse auf Erwerb und Anwendung von Wissen in einem computersimulierten System. *Zeitschrift für Pädagogische Psychologie*, *12*(1), 11–23.

Wache, M. (2003). *E-Learning – Bildung im digitalen Zeitalter*. Abgerufen am 13.07.2011 von http://www.bpb.de/files/FWQFK9.pdf

Weidenmann, B. (2000). Medien und Lernmotivation: Machen Medien hungrig oder satt? In U. Schiefele & K.-P. Wild (Hrsg.), *Interesse und Lernmotivation. Untersuchungen zu Entwicklung, Förderung und Wirkung* (S. 117–132). Münster: Waxmann Verlag.

Weinmann, G. (1998). Lebenslanges Lernen in der Informationsgesellschaft. In H. Kubicek, H.-J. Braczyk, D. Klumpp, G. Müller, W. Neu, E. Raubold & A. Roßnagel (Hrsg.), *Lernort Multimedia. Jahrbuch Telekommunikation und Gesellschaft 1995* (S. 332–336). Heidelberg: R. v. Deckers Verlag.

Wernet, S. P., Olliges, R. H. & Delicath, T. A. (2000). Postcourse evaluations of WebCT (web course tools) classes by social work students. *Research on Social Work Practice, 10*(4), 487–504.

Wikipedia (2016). *Skype*. Abgerufen am 15.07.2016 von http://de.wikipedia.org/wiki/Skype

Wlodkowski, R. J. (2003). Fostering motivation in professional development programs. *New directions for adult and continuing education*, *98*. 39–48.

Wlodkowski, R. J. (2004). Creating motivating learning environments. In M. W. Galbraith (Hrsg.), *Adult learning methods: A guide for effective instruction* (3. Ausg., S. 141–164). Malabar, Florida: Krieger Publishing Company.

Wofford, M. M., Spickard III, A. W. & Wofford, J. L. (2001). The computer-based lecture. *Journal of General Internal Medicine*, *16*(7), 464–467.

11 Anhang

ANHANG A: STUDIE I – FRAGEBOGENENTWICKLUNG UND FRAGEBÖGEN III
ANHANG B: STUDIE I – ERGÄNZENDE ERGEBNISSE XXXVIII
ANHANG C: STUDIE II – FRAGEBÖGEN LIII
ANHANG D: STUDIE II – ERGÄNZENDE ERGEBNISSE LXXV