

UNIVERSITÄT DES SAARLANDES
PHILOSOPHISCHE FAKULTÄT III
EMPIRISCHE HUMANWISSENSCHAFTEN

**Theoretisieren für die Praxis: Förderung anwendbaren
bildungswissenschaftlichen Wissens in der
Lehramtsausbildung anhand instruktionaler Fehler**

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Philosophie
der Philosophischen Fakultät III
der Universität des Saarlandes

vorgelegt von
David Martin Klein
aus Ottweiler

Saarbrücken, 2015

Der Dekan:

Univ.-Prof. Dr. phil. Roland Brünken

Berichterstatter/innen:

Univ.-Prof. Dr. phil. Robin Stark

Univ.-Prof. Dr. phil. Franziska Perels

Tag der Disputation: 27.07.2015

„Es ist ein großer Vorteil im Leben, die Fehler, aus denen man lernen kann, möglichst früh zu begehen.“

Winston Churchill

„I wanted to save Gotham. I failed.“ - „Why do we fall, Sir? So that we can learn to pick ourselves up.“

Bruce Wayne & Alfred Pennyworth , *Batman Begins*

Danksagung

Mein Dank gilt natürlich an erster Stelle meinem Doktorvater, Prof. Robin Stark, ohne den diese Arbeit nicht möglich gewesen wäre. Nach dem kürzesten Vorstellungsgespräch aller Zeiten und einer kurzen Eingewöhnungszeit an einem eher ungewöhnlichen Lehrstuhl war unsere Zusammenarbeit immer sehr positiv und vor allem von viel Vertrauen geprägt. Ja, ich hätte bei manchen Dingen öfter nachfragen können, aber mir kam es sehr entgegen, auch einfach mal machen zu können und zu wissen, dass man mir zutraut, es gut zu machen. Danke dafür und für viele Chancen - und natürlich auch für ein jederzeit offenes Ohr, wenn ich wirklich mal nicht weiterwusste.

Dank gilt auch meinen Kollegen am Lehrstuhl, vor allem Kai Wagner und Eric Klopp, für fachliche Beratung, Austausch von Ideen, Projektplanung, Feedback, gegenseitige Tritte in den Hintern und die Möglichkeit, innerhalb von Sekundenbruchteilen jegliches hochgeistige Thema zugunsten der wirklich wichtigen Dinge im Leben („Gehn ma ball mo Middach esse?“) fallen zu lassen.

Meiner Kollegin und sehr guten Freundin Kathrin Kaub möchte ich vor allem für ihre Freundschaft und Unterstützung in diversen beruflichen und auch persönlichen Krisen danken. Und natürlich auch für geschickte Appelle an meinen ausgeprägten Trotz, der am Ende das Schreiben der Dissertation *wesentlich* beschleunigt hat.

Unerlässlich war dabei von Anfang an auch die Hilfe unserer studentischen Hilfskräfte - allen voran Thorsten Mai und Tobias Stahl, deren Unterstützung bei der Durchführung der Experimente, in der Dateneingabe und später sogar bei der Auswertung von unschätzbarem Wert war. Besonders erwähnt seien auch noch Claudia Buck und Sinah Auctor, die beim Korrekturlesen des Literaturverzeichnisses schier Übermenschliches geleistet haben.

Vielen anderen Kolleginnen und Kollegen aus der Fachrichtung danke ich ebenfalls für den fachlichen Austausch, aber viel mehr noch für gute Gesellschaft und gemeinsame Erlebnisse auf diversen Konferenzen, Fachrichtungsfeiern („Heute Sekt ab elf!“) und einfach beim Treffen auf dem Flur. Davon können sich andere Fachrichtungen echt was abschneiden.

Bei Prof. Franziska Perels möchte ich mich noch gesondert für die Übernahme des Zweitgutachtens und speziell dafür bedanken, wie unkompliziert so etwas von Statten gehen kann.

Meiner Partnerin Maike Lauterbach danke ich ganz besonders - dafür, dass sie mich in den letzten paar Wochen und Monaten ausgehalten hat, in denen mein Nervenkostüm zusehends dünner wurde, aber auch dafür, dass sie immer da und einfach ein wunderbarer Mensch ist. Den Rest sage ich ihr persönlich ;).

Zu guter Letzt, von Anfang an und immer wieder möchte ich mich bei meiner Familie bedanken. Bei meinem großen kleinen Bruder, der mich am Ende bei der Promotion sogar überholt und mir grade dadurch den nötigen Motivationsschub gegeben hat. Ich bin stolz auf dich, Tobi! Und vor allem natürlich - bei meinen Eltern, ohne die ich nichts wäre, die mir eine gute Erziehung, Bildung und Wissensdurst (und den Dickkopf) mitgegeben haben, die mir immer mit Rat und Tat zur Seite standen und es auch dann immer gut meinten, wenn man es als „pubertärer Panz“ nicht mitbekommen hat. Danke für alles.

Inhaltsverzeichnis

1. Überblick über die der Arbeit zugrundeliegenden Aufsätze	1
2. Problemstellung und Ziel des Forschungsprojekts	2
2.1 Anwendung wissenschaftlichen Wissens im Lehramt	2
2.2. Probleme und Fehler bei der Anwendung wissenschaftlichen Wissens	5
2.3 Ursachen für Fehler bei der Anwendung wissenschaftlichen Wissens.....	6
3. Wissensarten	10
3.1 Wissenschaftliches vs .Alltagswissen	10
3.2 Wissensarten und Qualitäten von Wissen	12
3.2.1 Das Modell von DeJong und Ferguson-Hessler: Arten und Qualitäten von Wissen.....	12
3.2.2 Vergleich mit anderen Konzeptionen	13
3.2.3 Merkmale anwendbaren Wissens	14
3.2.4 Träges Wissen.....	16
3.3 Theorieartikulation: Theorieanwendung zur Erklärung von Sachverhalten	18
4. Lernen aus Fehlern	21
4.1 Begriffsbestimmung: Was ist ein Fehler?	21
4.2. Fehlerkultur: Produktiver Umgang mit Fehlern.....	23
4.3 Negatives Wissen als Produkt des Lernens aus Fehlern	25
4.4. Negatives Wissen als zusätzliche Wissensdimension aus der Perspektive der Conceptual Change-Forschung	27
4.5. Lernen aus eigenen versus Lernen aus fremden Fehlern	29
4.6. Bedingungsfaktoren beim Lernen aus Fehlern.....	31
5. Umsetzung des Lernens aus Fehlern anhand integrierter Lernumgebungen	35
5.1 Begriffsbestimmung: Problem- und Instruktionsorientierung	35
5.2 Problemorientierung.....	37
5.2.1 Problemorientierte Ansätze	37
5.2.2 Kritik am problemorientierten Lernen.....	39
5.3 Instruktionsorientierung	39
5.3.1 Instruktionsorientierte Ansätze	40
5.3.2 Kritik am instruktionsorientierten Lernen	41

5.4 Synthese: Designprinzipien integrierter Lernumgebungen.....	42
6. Studie 1: Lernen aus eigenen vs. Lernen aus fremden Fehlern	45
6.1 Productive Failure	45
6.2 Advokatorisches Lernen aus Fehlern	47
6.3 Fostering theory application competences in student teachers by learning from errors: a comparison of productive failure and advocatory error learning	51
7. Studie 2: Instruktionale Unterstützung beim Lernen aus fremden Fehlern in integrierten Lernumgebungen.....	77
7.1 Problemdarstellung	78
7.2 Theoretischer Hintergrund	79
7.4 Fragestellung und Hypothesen.....	86
7.5 Methode	87
7.6 Ergebnisse	96
7.7 Diskussion.....	101
8. Studie 3: Implementation des Lernens aus Fehlern in Seminaren anhand von Blended learning und Fading	112
8.1. Stabilität der Lerneffekte	113
8.2 Strukturierung der Seminarkonzeption anhand von Blended Learning	113
8.3 Fading instruktionaler Unterstützung.....	115
8.4 Förderung anwendbaren bildungswissenschaftlichen Wissens bei Lehramtsstudierenden anhand fehlerbasierten kollaborativen Lernens: Eine Studie zur Replikation und Stabilität bisheriger Befunde und zur Erweiterung der Lernumgebung.....	117
9. Gesamtdiskussion	145
9.1 Übergreifende Diskussion aller Studien mit Bezug zu den theoretischen Grundlagen	146
9.3 Limitationen	151
9.4 Pädagogische Konsequenzen	152
10. Literatur Synopse	154

Tabellen- und Abbildungsverzeichnis

Tabellen

Synopse

Tabelle 1: <i>Auszug aus dem Kompetenzkatalog der Standards für Lehrerbildung</i>	3
Tabelle 2: <i>Unterschiede zwischen wissenschaftlichem Wissen und Alltagswissen</i>	11

Studie 1: Lernen aus eigenen vs. Lernen aus fremden Fehlern

Table 1: <i>Means and standard deviations of the learning outcomes in all groups</i>	66
Table 2: <i>Means and standard deviations of the subjective dimensions in all groups</i>	67
Table 3: <i>Correlations between subjective dimensions and learning outcomes</i>	67

Studie 2: Instruktionale Unterstützung beim Lernen aus fremden Fehlern in integrierten Lernumgebungen

Tabelle 1: <i>Versuchsdesign – experimentelle Variation des Ausmaßes instruktionaler Unterstützung</i>	87
Tabelle 2: <i>Struktur der Trainings szenarien</i>	90
Tabelle 3: <i>Korrelationsmatrix der Wissenstests zu t2</i>	95
Tabelle 4: <i>konzeptuelles Fehlerwissen -- Mittelwerte und Standardabweichungen</i>	97
Tabelle 5: <i>strategisches Fehlervermeidungswissen -- Mittelwerte und Standardabweichungen</i>	97
Tabelle 6: <i>prozedurales Fehleridentifikationswissen -- Mittelwerte und Standardabweichungen</i>	98
Tabelle 7: <i>Erklärungswissen -- Mittelwerte und Standardabweichungen</i>	99
Tabelle 8: <i>Subjektiver Lernerfolg - Mittelwerte und Standardabweichungen</i>	99
Tabelle 9: <i>Akzeptanz - Mittelwerte und Standardabweichungen</i>	100
Tabelle 10: <i>Korrelationsmatrix zu t2</i>	100

Studie 3: Implementation des Lernens aus Fehlern in Seminaren anhand von Blended learning und Fading

Tabelle 1 <i>Versuchsablauf in der Experimental- und Kontrollgruppe</i>	128
Tabelle 2 <i>Seminartermine Phase 2 - Blended Learning</i>	130
Tabelle 3 <i>Korrelationen der Wissenstests zu allen Messzeitpunkten</i>	133
Tabelle 4 <i>Vorwissen; Mittelwerte und Standardabweichungen</i>	135
Tabelle 5 <i>Intraklassenkorrelation</i>	135
Tabelle 6 <i>Mittelwerte und Standardabweichungen gesamt zu zwei Messzeitpunkten</i>	136
Tabelle 7 <i>Mittelwerte und Standardabweichungen der Versuchsgruppen zu t3</i>	137
Tabelle 8 <i>Mittelwerte und Standardabweichungen der Versuchsgruppen zu t2 und t3</i>	138

Abbildungen

Synopse

Abbildung 1: *Negatives Wissen als zusätzliche Wissensdimension* 28

Studie 1: Lernen aus eigenen vs. Lernen aus fremden Fehlern

Figure 1: *Model of knowledge types involved in theory articulation* 54

Figure 2: *Experimental conditions* 60

Figure 3: *Training and testing procedure* 61

Studie 3: Implementation des Lernens aus Fehlern in Seminaren anhand von Blended learning und Fading

Abbildung 1: *Ausschnitt aus einer fehlerhaften Erklärung* 122

Abbildung 2: *Beispiel zur Korrektur eines Fehlers* 122

Abbildung 3: *Test zum Prozeduralen Fehleridentifikationswissen* 132

Abbildung 4: *Test zum Erklärungswissen* 133

1. Überblick über die der Arbeit zugrundeliegenden Aufsätze

Die vorliegende Arbeit beruht auf drei Studien, die in den folgenden Aufsätzen beschrieben werden:

Studie 1

Klein, M., Wagner, K., & Stark, R. (eingereicht). Fostering Theory Application Competences in Student Teachers by Learning from Errors: A Comparison of Productive Failure and Advocatory Error Learning. *Journal of teacher education*.

Studie 2

Wagner, K., Klein, M., Klopp, E. & Stark, R. (2014). Instruktionale Unterstützung beim Lernen aus advokatorischen Fehlern in der Lehramtsausbildung: Effekte auf die Anwendung wissenschaftlichen Wissens. *Psychologie in Erziehung und Unterricht*, 61, 287 - 301.

Studie 3

Klein, M., Wagner, K., Klopp, E. & Stark, R. (angenommen). Förderung anwendbaren bildungswissenschaftlichen Wissens bei Lehramtsstudierenden anhand fehlerbasierten kollaborativen Lernens: Eine Studie zur Replikation bisheriger Befunde sowie zur Nachhaltigkeit und Erweiterung der Trainingsmaßnahmen. *Unterrichtswissenschaft*

2. Problemstellung und Ziel des Forschungsprojekts

2.1 Anwendung wissenschaftlichen Wissens im Lehramt

Nach der Veröffentlichung der ersten Pisa-Studien um die Jahrtausendwende (z.B. Baumert et al., 2001; Prenzel et al., 2004) ist in Deutschland die Debatte um die Qualität der Lehrerbildung neu entbrannt. Es gab und gibt zahlreiche Versuche, die Lehrerbildung neu zu ordnen und zu verbessern. Insbesondere sollte klarer definiert werden, welche Fähigkeiten Lehrpersonen mitbringen müssen, um im Beruf erfolgreich zu sein. Studien zeigen, dass erfolgreiche Lehrerinnen und Lehrer nicht nur über umfassendes Fachwissen, sondern auch über ein evidenzbasiertes und an Standards orientiertes Repertoire an pädagogischen Fähigkeiten und Fertigkeiten verfügen, die sich im schulischen Arbeitsfeld als effektiv erwiesen haben (Ingvarson & Rowe, 2008; Meier 2006; Neuweg 2007; Sherin 2007; Swinkels, Koopman & Beijaard, 2013). Verschiedene Ansätze etwa aus der Expertiseforschung (z.B. *Bewusstmachung und Reflexion*; Bromme, 1992; s.a. Bromme, 2008; Smith, 2005) wurden herangezogen, um diese Forderungen zu konkretisieren. Dabei werden v.a. die Anwendbarkeit des pädagogischen Wissens und der Bezug auf Erkenntnisse der Bildungsforschung (Evidenzbasierung) betont.

Deutlich wird das in der Neufassung des Kerncurriculums Erziehungswissenschaft der DGfE von 2004, welche die Grundlage zahlreicher Curricula aus der universitären Lehrerbildung darstellt (DGfE, 2004). Gemäß deutscher Universitätscurricula umfassen die Aufgaben zukünftiger Lehrerinnen und Lehrer die Anregung, Unterstützung und Beurteilung von Lernprozessen sowie Erziehung und Beratung. Um die Lehramtsstudierenden auf diese Aufgaben vorzubereiten, erwerben diese im Laufe ihres Studiums umfangreiches psychologisches und pädagogisches Wissen. Dies soll ihnen ermöglichen, im späteren Berufsalltag aus einem reichen Fundus an Theorien und Methoden diejenigen wählen zu können, die am besten zur Erreichung ihrer Lehrziele geeignet sind. Zudem soll dadurch die eine umfassendere Perspektive auf Lehr-Lernprozess vermittelt werden, die die Reflexion eigener Handlungen erlaubt (vgl. z.B. Universität des Saarlandes, Kerncurriculum Bildungswissenschaften 2007). Die Studierenden sollen die Fähigkeit erwerben,

- „Vorgehensweisen für pädagogisches Handeln in der Schule vor einem wissenschaftlichen Hintergrund (theoretische Fundierung, empirische Bestätigung) zu entwerfen, erproben und zu analysieren;

- Voraussetzungen, Bedingungen und Risikofaktoren für Erziehungs- und Bildungsprozesse zu erfassen; und
- Kognitionen, Emotionen und Handeln von Kindern und Jugendlichen wahrzunehmen und zu verstehen“

(UDS, 2007; S. 571f).

Die inhaltlichen Grundlagen für diese Fähigkeiten werden in den Standards der Lehrerbildung (KMK, 2004; 2014) in Form von Kompetenzbereichen definiert. Dort werden als inhaltliche Schwerpunkte der Bildungswissenschaften in der Ausbildung von Lehrerinnen und Lehrern u.a. Wissen und Kompetenzen in der Gestaltung von Unterricht und Lernumgebungen, aber auch Wissen über Lernprozesse von Kindern und Jugendlichen innerhalb und außerhalb der Schule sowie die Diagnose und Förderung individueller Lernprozesse genannt. Diese werden ergänzt durch Wissen über motivationale Grundlagen der Lern-, Leistungs- und Kompetenzentwicklung sowie Wissen über Ziele und Methoden der Bildungsforschung und Kompetenzen in der Interpretation und Anwendung ihrer Ergebnisse.

Ausgehend von diesen Schwerpunkten werden konkrete Kompetenzen formuliert, die in verschiedene Bereiche (z.B. „Unterrichten“, „Erziehen“, „Innovieren“, ebd., S 7f.) unterteilt werden. In den folgenden Auszügen aus dem Kompetenzkatalog wird der Bezug auf wissenschaftliches Wissen besonders deutlich:

Tabelle 1: *Auszug aus dem Kompetenzkatalog der Standards für Lehrerbildung*

Kompetenz-

bereich	Kompetenz: Die Absolventinnen und Absolventen ...
Unterrichten	<p>...kennen die einschlägigen Erziehungs- und Bildungstheorien, verstehen bildungs- und erziehungstheoretische Ziele sowie die daraus abzuleitenden Standards und reflektieren diese kritisch</p> <p>...kennen Lerntheorien und Formen des Lernens</p> <p>...kennen Theorien der Lern- und Leistungsmotivation und Möglichkeiten, wie sie im Unterricht angewendet werden</p>
Erziehen	<p>...kennen pädagogische, soziologische und psychologische Theorien der Entwicklung und der Sozialisation von Kindern und Jugendlichen</p> <p>...verfügen über Kenntnisse zu Kommunikation und Interaktion unter besonderer Berücksichtigung der Lehrer-SchülerInteraktion</p>

Beurteilen ...wissen, wie unterschiedliche Lernvoraussetzungen Lehren und Lernen beeinflussen und wie sie im Unterricht in heterogenen Lerngruppen positiv nutzbar gemacht werden können

...kennen unterschiedliche Formen und Wirkungen der Leistungsbeurteilung und -rückmeldung, ihre Funktionen und ihre Vor- und Nachteile

Innovieren ...rezipieren und bewerten Ergebnisse der Bildungsforschung
...reflektieren die eigenen beruflichen Haltungen, Erfahrungen und Kompetenzen sowie deren Entwicklung und können hieraus Konsequenzen ziehen

...nutzen Erkenntnisse der Bildungsforschung für die eigene Tätigkeit

(KMK, 2004, 2014; S. 7f)

Als entscheidend wird also neben Wissen um Didaktik und Lehr-Lern-Theorien auch Wissen über Motivations- und Selbstregulierungstheorien sowie darüber angesehen, wie die psychosoziale Entwicklung Jugendlicher von Statten geht und die übrigen Bereiche beeinflusst. Zur Vermittlung der bildungswissenschaftlichen Inhalte werden dabei u. a. der Situationsansatz, eine Fall- und Praxisorientierung, die Vermittlung von Problemlösestrategien, biographisch-reflexive Ansätze sowie Kontext-, Phänomen- und Forschungsorientierung vorgeschlagen.

Der den Standards zugrundeliegende Kompetenzbegriff orientiert sich an Weinert (2001, S.27) und versteht Kompetenzen als „[...] die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, die Problemlösung in variablen Situationen erfolgreich nutzen zu können“. Diesem Kompetenzbegriff und seiner Verankerung in kognitiven Dimensionen wird von manchen Autoren vorgeworfen, Handeln zum „bloßen Anhängsel eines im Prinzip bestimmbar und vermittelbaren Wissens“ (Neuweg, 2011, S. 34, vgl. auch Ickler, 1994) zu machen und das eigentlich relevante Phänomen „pädagogischer Könnerschaft“ als „irgendwie zum Laufen gebrachtes Wissenschaftswissen“ (ebd., S. 42) misszudeuten. Er bietet jedoch gerade im Bereich anwendbaren Wissens eine gute Handhabe, da er die *Erlernbarkeit* der kognitiven Fähigkeiten und Fertigkeiten betont. Zwar schließt er auch einstellungsbezogene Dimensionen ein, mit Bezug auf die Diskussion um Haltungen und Einstellungen und deren tatsächliche Handlungsleitung (z.B. Limón, 2006) konzentrieren sich die Studien in der

vorliegenden Arbeit aber in erster Linie auf die kognitiven Dimensionen. Der Fokus liegt damit auf der Förderung *anwendbaren Wissens*. Problematisch ist dabei, dass trotz diverser Kompetenzkataloge, Umstrukturierungen und Forderungen im Kontext der Lehrerbildung immer noch keine systematische Förderung der Kompetenzen zur Nutzung und Bewertung bildungswissenschaftlicher Theorien (Meier, 2006; Neuweg, 2007; Stark, Herzmann & Krause, 2010) oder Evidenz (Bauer et al., 2010; Terhart, Lohmann & Seidel, 2010; Weber & Achtenhagen, 2009) erfolgt. Die Konsequenz sind zahlreiche Probleme Lehramtsstudierender bei der tatsächlichen Anwendung bildungswissenschaftlichen Wissens (Stark, 2005; Stark et al., 2010).

2.2. Probleme und Fehler bei der Anwendung wissenschaftlichen Wissens

Voraussetzung für die Anwendung bildungswissenschaftlichen Wissens im schulischen Alltag ist eine solide Basis an gut strukturiertem wissenschaftlichem Wissen über pädagogische Theorien, Ansätze, Konzepte und empirische Befunde (Bortz & Döring, 2006). Problematisch dabei ist, dass Lehramtsstudierende zwar bildungswissenschaftliches Wissen erwerben und in Prüfungen reproduzieren, es aber nicht zu Anwendung auf konkrete pädagogischen Problemstellungen nutzen können (Gräsel & Mandl, 1999; Seidel & Prenzel, 2008; Star & Strickland, 2008; Stark et al., 2010). Ihr Wissen bleibt also „träge“ (Gruber & Renkl, 2000; Renkl, 1994). Lehramtsstudierende haben dementsprechend gravierende Probleme mit der Anwendung wissenschaftlichen Wissen bei der Beurteilung relevanter pädagogischer Situationen (z.B. *rückschauendes Begreifen*; vgl. Beck & Krapp, 2006).

Diese Problematik zeigt sich auf verschiedenen Ebenen, etwa bei der für professionelles Lehrerhandeln zentralen *theoriegeleiteten Wahrnehmung* (Neuweg, 2007). Gerade Lehramtsstudierende und Berufseinsteiger beschreiben Schulsituationen häufig oberflächlich und neigen entweder zur Übergeneralisierung oder im Gegenteil dazu, sich zu stark auf einzelne Schüler/innen zu konzentrieren und den Gesamtüberblick zu vernachlässigen (Berliner, 1992; Bromme, 2008; Seidel & Prenzel, 2008; Star & Strickland, 2008). Das eng mit der theoriegeleiteten Wahrnehmung verwandte Konzept der *professionellen Unterrichtswahrnehmung* beschreibt die Fähigkeit Lehrender, Unterricht professionell zu beobachten und zu interpretieren (Sherin, 2007, Sherin & van Es, 2009). Sie umfasst die Komponenten „*noticing*“, also die Identifikation pädagogisch relevanter Situationen, sowie „*knowledge-based reasoning*“, also die Beschreibung, Erklärung und Prognose dieser Situationen auf Basis bildungswissenschaftlichen Wissens (Berliner & Carter, 1989; van Es & Sherin 2002). Untersuchungen, in denen Lehramtsstudierende mit Unterrichtsvideos

konfrontiert wurden, zeigen auch diesbezüglich deutliche Defizite (Star & Strickland, 2008; Sherin & van Es 2009; van ES & Sherin, 2002).

In der Konsequenz erfolgt die Beurteilung komplexer schulischer Situationen und Phänomene meist auf der Basis von Alltagswissen und subjektiven Theorien (vgl. Kagan, 1992; Krause & Stark, 2009; Pajares 1992; Stark, 2005; Stark et al., 2010). Subjektive Theorien basieren auf eigenen Erfahrungen oder episodischen Berichten Dritter, haben sich jedoch im Alltag vielfach bewährt und sind dementsprechend tief in der Wissensbasis der Lernenden verankert (Groeben et al., 1988; Smith, diSessa & Roschelle, 1993). Sie stimmen allerdings häufig nicht mit wissenschaftlichem Wissen überein. Die damit vorhandenen und den Lernenden häufig nicht bewussten Fehlkonzepte erschweren zusätzlich den Erwerb und die Anwendung wissenschaftlichen Wissens (Chi & Roscoe, 2002; Krause, 2007; Stark, 2003). Daraus resultieren konkrete Fehler bei der Anwendung wissenschaftlichen Wissens etwa zur Erklärung eines Phänomens. Beispiele sind der Bezug auf verschiedene Formen von Pseudo- oder Nicht-Evidenz, Zirkelerklärungen oder die Anpassung der Interpretation empirischer Befunde an eine vorgefasste Erklärung (anstatt umgekehrt; Kuhn, 1989; Kuhn & Reiser, 2005). Derartige Fehler (Bezug auf Alltagswissen, fehlerhafte Auswahl oder Interpretation wissenschaftlicher Theorien und Evidenzen) fanden sich auch in einer Studie von Stark (2005) im Bereich der Lehramtsausbildung.

2.3 Ursachen für Fehler bei der Anwendung wissenschaftlichen Wissens

Für Lehramtsstudierende ist die Verbindung von Theorie und Praxis im Studium nur schwer zu erkennen (Le Cornu & Ewing, 2008). Dies führt dazu, dass sie sich wenig mit der wissenschaftlichen Basis ihres späteren Berufsfeldes beschäftigen oder diese ablehnen (Ritchie & Wilson, 2001). Entsprechend werden Wissen und Einstellungen, die während des Studiums erworben werden, spätestens in der zweiten Phase der Lehramtsausbildung und beim Berufseinstieg wenig beachtet oder verworfen (Terhart, Czerwenka, Ehrich, Jordan & Schmidt, 1994). Stattdessen werden dort subjektive Erfahrungen oder das Beispiel von Kollegen oder eigenen Lehrpersonen handlungsleitend, ohne dass eine kritische Reflexion erfolgt. Dementsprechend schätzen auch erfahrene Lehrpersonen die Relevanz der pädagogischen Anteile ihres Studiums für ihr professionelles Handeln im Rückblick eher gering ein (Terhart, 2006). Diese Befunde verweisen auf eine gewisse Zurückhaltung vieler Lehramtsstudierender gegenüber Inhalten, deren Praxisrelevanz sie nicht unmittelbar erkennen können. Aus dieser Reserviertheit kann sich eine Distanz gegenüber

bildungswissenschaftlichen Erkenntnissen entwickeln, die sich als Aspekt berufsbezogener Habitusformen (Bohnsack, 2000) verfestigen kann.

Die Ursachen dafür liegen u.a. im motivationalen Bereich und resultieren aus übersteigerten Erwartungen und Fehlkonzepten bezüglich der Eigenschaften wissenschaftlicher Theorien. Diese basieren beispielweise auf einer nicht trennscharfen Verwendung des Theoriebegriffs in Schulen (vgl. Driver, Newton & Osborne 2000; Gordon & O'Brien, 2007). Driver et. al (2000) argumentieren, dass Studierende aufgrund der Verwendung des (oft nur im naturwissenschaftlichen Bereich gebrauchten) Theoriebegriffs in ihrer eigenen schulischen Ausbildung übersteigerte Erwartungen an die Gültigkeit, Reichweite oder Aussagekraft wissenschaftlicher Theorien haben. Wissenschaftliche Erkenntnisse werden als Ergebnis einer Reihe erfolgreicher Versuche und Entdeckungen dargestellt. Diese scheinbar lineare Vorgehensweise lässt wesentliche Bestandteile der Erkenntnisgewinnung außer Acht, so werden Rückschritte, aber auch die kritische Auseinandersetzung mit der zugrundeliegenden Methodik, den Forschungsfragen, oder der Interpretation der Daten (zusammenfassend zur Wissenschaftstheorie z.B. Rook, Irle & Frey, 2001) oft zu wenig beachtet.

Zudem wird so der Eindruck vermittelt, die Wissenschaft sei durch ihre Methodik unfehlbar und böte Vorgehensweisen und Methoden, die bei korrekter Anwendung in jedem Fall zum gewünschten Ziel führen. Konsequenterweise hält sich bei Lehramtsstudierenden das Fehlkonzept, dass pädagogische Theorien feste Fakten oder unbestreitbare Wahrheiten darstellen, die direkt auf die Situation in einer Klasse übertragbar sind (Gordon & O'Brien, 2007). Die zugrundeliegenden epistemologischen Überzeugungen (z.B. „Gute“ Theorien sind auf pädagogische Problemstellungen wie eine bewährte Arznei bei einer Krankheit anwendbar, ohne dass die Theorien modifiziert oder ihre Passung auf die Situation überprüft werden müsste“, vgl. Gordon & O'Brien, ebd., S. XI, Übersetzung: Autor) verbleiben auf einer sehr niedrig entwickelten Stufe und können als „naiv“ bezeichnet werden (vgl. Schommer, 1990; Stahl & Bromme, 2007). Spätestens beim Kontakt mit sozialwissenschaftlichen Theorien lernen Studierende eine sehr viel weniger deterministische Auffassung von Theorien kennen (Driver et. al, 2000; Terhart, 2006). Bildungs- und sozialwissenschaftliche Theorien können die übersteigerte Erwartungshaltung, Schule, Schülerverhalten und Unterricht präzise und umfassend abzubilden, nicht erfüllen. Stattdessen werden Studierende oft mit verschiedenen Modellen konfrontiert, die erst in ihrer Gesamtheit ein Phänomen erklären können. Die anfänglich zu hohe Erwartung schlägt infolgedessen häufig in Resignation oder Ablehnung um. Lehramtsstudierende fühlen sich in Bezug auf die Anwendbarkeit pädagogischer Theorien auf zukünftige Problemstellungen oft allein gelassen

(Gordon & O'Brien, 2007). Theorien werden entweder als abstrakte, idealisierte Philosophien, die mit der komplexen, chaotischen Realität wenig gemein haben (ebd.), oder aber aufgrund der scheinbar großen Überschneidung zwischen Alltags- und wissenschaftlichem Wissen im pädagogischen Bereich als komplizierte Paraphrasierung allgemein bekannter Alltagskonzepte (vgl. Schneewind 1977, S.11) wahrgenommen.

Zudem sind viele Studierende mangels Kenntnis von Strukturierungsmöglichkeiten durch den Wissenserwerb überfordert (Levin, 2010). Dadurch sind theoretische Konzepte oft zu wenig elaboriert und kaum verankert, so dass ihre Bedeutung in der zweiten Ausbildungsphase oder in den begleitenden Praktika nur schwer salient werden kann (Terhart, 2006). Neue Theorien werden von Studierenden bestenfalls „auswendig gelernt“, wenn sie bezugslos zu ihrem verfügbaren Wissen sind (Lind, 2001) oder kein Bezug zum späteren Beruf hergestellt werden kann (*fallacy of miscellaneous information*; Cohen, 1952; Driver et al., 2000). Studierende weisen zudem häufig ungünstige Einstellungen und geringes Interesse gegenüber wissenschaftlichen Theorien auf (Bainbridge, 2010; Stark, 2005; Stark et al., 2010), beispielsweise betrachten sie Wissen über Theorien nur als Mittel zum Bestehen der jeweils anstehenden Klausuren (Driver et al., 2000).

Diese zahlreichen Probleme erscheinen auf den ersten Blick kaum zu bewältigen und weisen auf umfassende Mängel im gesamten System der Lehramtsausbildung (nicht nur in Deutschland) hin. Ihre relativ umfassend erforschten Ursachen bieten jedoch gleichzeitig Ansatzpunkte für Lösungen.

Ein geeigneter Lösungsansatz muss Studierenden vermitteln, inwiefern bildungswissenschaftliches Wissen bei allen Einschränkungen wichtige Handlungsimpulse für alle Arten schulischer Situationen und zur Lösung schulischer Probleme vermitteln kann. Um diese Verbindung zwischen Theorien und Praxis aufzuzeigen, eignen sich problemorientierte Ansätze (z.B. *Anchored Instruction*; Cognition and Technology Group at Vanderbilt (CTGV), 1992). Zudem müssen die Studierenden bei der Strukturierung und dem Aufbau von konzeptuellem (Fach)Wissen unterstützt werden, hier eignen sich instruktionsorientierte Ansätze (z.B. *Instructional Design*; Kirschner, Sweller & Clark, 2006). Um eine systematische Förderung der Anwendung bildungswissenschaftlicher Theorien und die Nachhaltigkeit der erworbenen Kenntnissen und Fähigkeiten über die universitäre Lehramtsausbildung hinaus zu sichern, sollte zudem ein Format gewählt werden, das unter praktischen Gesichtspunkten ohne größeren Aufwand in die bestehenden Strukturen der Lehramtsausbildung implementiert werden kann.

In erster Linie muss eine geeignete Intervention aber auf die spezifischen Probleme der Lehramtsstudierenden mit der Anwendung bildungswissenschaftlichen Wissens eingehen. In diesem Zusammenhang bieten die o.g. Fehler (Stark, 2005; s. auch Kapitel 4.1) einen Ansatzpunkt, um den Lernenden Fehlkonzepte aufzuzeigen und ihnen Alternativen anzubieten. Gleichzeitig können so Unterschiede zwischen einer Alltags- und einer auf wissenschaftlichem Wissen basierenden Argumentation gezeigt und verdeutlicht werden, in welcher Hinsicht sich wissenschaftliches Wissen von Alltagswissen unterscheidet und inwiefern ersteres daher zu treffende Entscheidungen besser absichern kann.

Um diesen Anforderungen gerecht zu werden, wurden im Projekt „Theoretisieren für die Praxis“ verschiedene integrierte Lernumgebungen (Reinmann & Mandl, 2006) entwickelt und auf ihre Lernwirksamkeit in Bezug auf die effektive Förderung der Theorienanwendung in komplexen pädagogischen Situationen untersucht. Integrierte Lernumgebungen kombinieren systematisch problem- und instruktionsorientierte didaktische Designprinzipien und haben sich insbesondere für die Förderung anwendbaren Wissens als lernwirksam erwiesen (z.B. Wagner et al., 2013; Stark et. al, 2010).

Die vorliegende Synopse soll den Bogen über das gesamte Forschungsprogramm „Theoretisieren für die Praxis“ spannen. Ziel ist die Einordnung und Diskussion dreier Studien vor einem gemeinsamen theoretischen Hintergrund. Die drei der Arbeit zugrundeliegenden Aufsätze dokumentieren anhand dreier Studien die Entwicklungsschritte hin zur finalen Version der integrierten Lernumgebung, wie sie mittlerweile in Seminaren eingesetzt wird. Studie 1 diente zur Auswahl eines geeigneten Lehr-Lern-Ansatzes (Lernen aus eigenen Fehlern vs. Lernen aus fremden Fehlern). In Studie 2 wurde die Umsetzung des ausgewählten Ansatzes in Form einer integrierten Lernumgebung vorgenommen und untersucht, wie viel instruktionale Unterstützung dabei notwendig ist. Studie 3 untersuchte Bedingungen der Implementation in ein Seminar der Lehramtsausbildung. Dabei wurden verschiedene Seminarkonzeptionen mit Bezug auf die Lernumgebung verglichen.

3. Wissensarten

Ziel des hier entwickelten Forschungsprogramms ist die effektive Förderung der Theorienanwendung in komplexen pädagogischen Situationen. Theorieanwendung basiert auf der *Anwendbarkeit* von Wissen und erfordert das *Zusammenspiel* verschiedener Wissensarten sowie eigenständiges, reflektiertes Denken (vgl. Gruber, Mandl & Renkl, 2000; Krause, 2007). Der Fokus liegt dabei auf den kognitiven Dimensionen des Kompetenzbegriffes nach Weinert (2001, s. Kap. 2.1). Lehrpersonen sollen in die Lage versetzt werden, Theorien, empirische Befunde, Konzepte und Modelle aus den Bildungswissenschaften zur Erklärung pädagogischer Phänomene und Situationen zu nutzen.

Im folgenden Kapitel wird zunächst der Begriff des wissenschaftlichen Wissens auf Grundlage der Taxonomie nach Bortz und Döring (2006) näher definiert und von Alltagswissen abgegrenzt. Darauf folgt die Darstellung des diesem Forschungsprogramm zugrundeliegenden Verständnisses verschiedener Wissensarten. Diese lehnt sich an eine Taxonomie nach De Jong und Ferguson-Hessler (1996) an und erhebt keinen Anspruch auf Vollständigkeit. Verglichen mit anderen Wissenstaxonomien (Anderson & Krathwohl, 2001; Rittle-Johnson, Siegler & Alibali, 2001) hat sich diese Konzeption jedoch für die vorliegenden Studien als zielführend erwiesen. Anschließend wird das Konzept der Theorieartikulation (Ohlsson, 1992) als mögliche Operationalisierung der Anwendung (bildungs)wissenschaftlichen Wissens zur Erklärung komplexer schulischer Situationen dargestellt.

3.1 Wissenschaftliches versus Alltagswissen

Wissenschaftliches Wissen im Bereich der Lehramtsausbildung umfasst Theorien, Ansätze, Konzepte und empirische Evidenzen aus der Pädagogik und pädagogischen Psychologie sowie anderen verwandten Domänen wie der Entwicklungspsychologie. Um das Verständnis und die kritische Bewertung empirischer Studien aus einer praktischen Perspektive zu ermöglichen, ist zudem die Vermittlung grundlegender Kenntnisse wissenschaftlicher Methoden Teil der Lehramtsausbildung (vgl. Driver et al., 2000). Beispielsweise erwerben Lehramtsstudierende Kenntnisse über grundlegende Lehr-Lerntheorien (z.B. Kognitivismus vs. Konstruktivismus) und darauf aufbauende Lehr-Lern-Ansätze (z.B. *Instructional Design*, Kirschner et al., 2006 vs. *Anchored Instruction*, CTGV, 1992) und lernen empirische Studien zu diesen Ansätzen kennen.

Wissenschaftliches Wissen kann von sogenanntem Alltagswissen abgegrenzt werden. Alltagswissen umfasst subjektive Theorien (z.B. Groeben et al., 1988), Heuristiken (Tversky & Kahneman, 1974) und alle Arten von indirekt und unsystematisch erworbenem Wissen und Erfahrungen. Nach Bortz und Döring (2006, S.31) unterscheiden sich wissenschaftliches und Alltagswissen vor allem in Hinsicht darauf, wie Wissen generiert und wie damit umgegangen wird. Die Dimensionen, auf denen sich diese Unterschiede im Wesentlichen zeigen, finden sich in Tabelle 2.

Tabelle 2. *Unterschiede zwischen wissenschaftlichem Wissen und Alltagswissen*

	Wissenschaftliches Wissen	Alltagswissen
Systematik und Dokumentation	systematischer Gewinn von Evidenz auf Grundlage standardisierter Verfahren und sorgfältige Dokumentation	unsystematische, beiläufige Erfahrungen, selten Dokumentation
Präzision der Terminologie	präzise definiert	unpräzise, umgangssprachlich
Art der Auswertung und Interpretation von Informationen	standardisierte Auswertungsverfahren	erfahrungsbasierte subjektive Bewertung
Überprüfung von Gültigkeitskriterien	kritische Prüfung der internen und externen Validität	intuitive, individuumszentrierte Weltsicht
Umgang mit Theorien	wiederholter, institutionalisierter Evaluationsprozess, Kritik und Qualitätsmanagement	unkritischer Umgang mit subjektiven Theorien, kein Versuch der Falsifizierung

Diese Dichotomie ist notwendigerweise stark vereinfacht. Als didaktische Reduktion erscheint sie jedoch sinnvoll, da die Unterscheidung für Novizen (als die Lehramtsstudierende angesichts der in Kap. 2.2 geschilderten Probleme betrachtet werden müssen) aufgrund einer gewissen Überlappung zwischen den beiden Wissensbereichen ohnehin schwierig ist. Diese scheinbare Überlappung stellt ein inhärentes Merkmal der Gesellschaftswissenschaften dar und führt insbesondere in der Wahrnehmung von Novizen zu einer gewissen Unschärfe wissenschaftlicher Konstrukte (vgl. Schneewind, 177, S. 11). Daraus resultiert, dass zur Förderung der Anwendung wissenschaftlichen Wissens auch Wissen über Unterschiede zwischen Alltags- und wissenschaftlichem Wissen vermittelt werden muss.

3.2 Wissensarten und Qualitäten von Wissen

3.2.1 Das Modell von DeJong und Ferguson-Hessler: Arten und Qualitäten von Wissen

DeJong und Ferguson-Hessler (1996) verwenden ein zweidimensionales Modell, das zwischen *Arten* und *Qualitäten* von Wissen unterscheidet. Die *Wissensarten* *konzeptuelles*, *prozedurales* und *strategisches* Wissen beschreiben die Form, in der Wissen vorliegt (einzelne Konzepte, konkretes Handlungswissen, Pläne bzw. Sequenzen von Handlungen). Die Qualitäten beziehen sich auf weitere Eigenschaften wie Verarbeitungstiefe, Struktur, Modalität oder Automatisierungsgrad von Wissen. Für die vorliegende Arbeit ist in erster Linie der Automatisierungsgrad der Wissensarten von Bedeutung, da dieser direkt die Anwendbarkeit von Wissen beeinflusst (z.B. Stark, 2001). Dabei unterscheiden DeJong und Ferguson-Hessler (1996) zwischen *deklarativem* und *kompiliertem* Wissen (vgl. Anderson, 1983) und beschreiben damit die Endpunkte eines Kontinuums der *Automatisierung* von Wissen. Allerdings werden die einzelnen Qualitäten nicht als voneinander unabhängig konzeptualisiert. Der automatische Abruf von Wissen erfordert einen hohen Strukturierungsgrad, der durch eine tiefere Wissensverarbeitung erreicht wird. Strukturiertheit von Wissen bezieht sich dabei auch auf Zusammenhänge zwischen den einzelnen Wissensarten. Besitzt also eine Qualität von Wissen eine höhere bzw. als höherwertig betrachtete Ausprägung, so bedeutet das im Allgemeinen auch eine höhere Ausprägung der übrigen (vgl. DeJong & Ferguson-Hessler, ebd., z.B. S.108).

DeJong und Ferguson-Hessler (1996) verstehen *konzeptuelles Wissen* als Wissen über Fakten, Konzepte und Prinzipien in einer Domäne. Häufig enthält eine Problemstellung gerade dieses Wissen nicht selbst, die Lernenden müssen in diesem Fall auf ihr Vorwissen zurückgreifen. Konzeptuelles Wissen deklarativer Qualität ist unverbunden, unstrukturiert und nicht automatisiert, d.h. es muss aktiv abgerufen werden, um ein Problem zu lösen. Es ist damit i.d.R. explizit, d.h. bewusst abrufbar und verbalisierbar (Krause, 2007). Ist es bereits kompiliert, so besteht ein intuitives, implizites Verständnis bzw. ein größerer Überblick über eine Domäne. Ein Beispiel aus dem bildungswissenschaftlichen Bereich ist das Wissen über Konzepte innerhalb pädagogisch-psychologischer Theorien, wie etwa das Wissen um die Definitionen, Ursachen und Bedingungen von Leistungsangst (Rost & Schermer, 2006).

Als *prozedurales Wissen* beschreiben DeJong und Ferguson-Hessler (1996) Wissen über erfolgreiche, adäquate Handlungen in einer Domäne. Dabei umfasst prozedurales Wissen einzelne Handlungsschritte, die Übergänge zwischen verschiedenen Phasen im Prozess der Problemlösung realisieren („...knowing how to“). Bei geringer Automatisierung, also Wissen deklarativer Qualität, werden die einzelnen Handlungen schrittweise abgerufen und

ausgeführt, kompiliertes Wissen ermöglicht den automatischen Abruf und Routine in der Ausführung. Beispielsweise entwickelt sich beim Fremdsprachenerwerb aus einer anfänglichen wortweisen Satzkonstruktion im Laufe des Lernprozesses ein intuitives Sprachgefühl bzw. -verständnis (vgl. Bärenfänger, 2002).

Die einzelnen Handlungsschritte werden im Rahmen einer Problemlösung schließlich durch *strategisches Wissen* anhand von Handlungsplänen organisiert. Strategisches Wissen ist damit metakognitiver Natur und ermöglicht die Organisation der anderen Wissensarten in einer sinnvollen Sequenz. Diese schließt auch die *Bedingungen der Situation* ein (an anderer Stelle als *konditionales Wissen* aufgeführt, z.B. Krause, 2007; van Gog, Paas & Van Merriënboer, 2004). Bei strategischem Wissen mit geringer Automatisierung wird dabei der jeweils folgende Schritt bewusst ausgewählt („step by step choices and planning“; DeJong & Ferguson-Hessler, 1996, S. 111). Kompiliertes Wissen erlaubt die automatisierte Analyse und Planung bei gleichzeitiger Überwachung der Ausführung („parallel checking“, ebd.). So könnte im bildungswissenschaftlichen Bereich strategisches Wissen schrittweise bzw. im Rahmen eines vollständigen Handlungsplans dazu genutzt werden, den Unterricht gezielt so zu gestalten, dass er leistungängstlichen Schülern weniger Schwierigkeiten bereitet, indem bekannte Auslöser von Leistungsangst mit den tatsächlichen Unterrichtsmerkmalen abgeglichen und diese beseitigt werden (Jakobs & Strittmatter, 1979; Rost & Schermer, 2006).

3.2.2 Vergleich mit anderen Konzeptionen

Andere Konzeptionen unterscheiden sich insbesondere hinsichtlich ihres Verständnisses von prozeduralem Wissen vom Modell von DeJong und Ferguson-Hessler (1996). Anderson und Krathwohl (2001) bezeichnen das Wissen darüber, wann welche Prozesse angewendet werden, als prozedurales Wissen. Rittle-Johnson et al. (2001) verstehen prozedurales Wissen als die Fähigkeit, Handlungsabläufe zur Problemlösung auszuführen. Dabei werden das bei DeJong und Ferguson-Hessler (1996) getrennt aufgeführte strategische und prozedurale Wissen vermischt. Strategisches Wissen ist damit bei Rittle-Johnson et al. (2001) bereits Bestandteil prozeduralen Wissens, zudem werden Wissensarten und Qualitäten konfundiert.

Dies gilt auch für die *ACT-R-Theorie* (Anderson, 1983; Anderson et al., 2004). Dort ist prozedurales Wissen in Form von *Produktionsregeln* repräsentiert. Diese entsprechen Verfahren zur Konstruktion, Verknüpfung und Anwendung von *propositional vorliegendem deklarativem Wissen*. Als *Kompilierung* bezeichnet die ACT-Theorie die *Prozeduralisierung* deklarativen Wissens durch Einübung, also die Überführung in Produktionsregeln. Diese enthalten sowohl eine Bedingungskomponente („wenn“) als auch eine Aktionskomponente

(„dann“). Eine übergeordnete Wissensart, die die korrekte Sequenz der Handlungsschritte organisiert, ist damit bereits Bestandteil der Produktionsregel. Anderson (1983) verdeutlicht dies am Beispiel einer Telefonnummer: Nach ausreichender Übung der Ziffernfolge wird eine spezifische Produktionsregel entwickelt. Als Resultat werden die einzelnen deklarativen Informationen (die einzelnen Ziffern) nicht mehr zum Wählen benötigt, da die Produktionsregel die Abfolge der einzelnen Ziffern bestimmt. Selbst wenn einzelne Ziffern nicht erinnert werden, kann anhand der Produktionsregel die Nummer gewählt werden

Eine Unterscheidung im *Grad der Automatisierung* im Sinne von DeJong und Ferguson-Hessler (1996) wird dabei nicht vorgenommen. Im Gegensatz zu De Jong und Ferguson-Hessler (ebd.) beinhaltet die Definition der *Wissensart* prozedurales Wissen in der ACT-R-Theorie damit gleichzeitig die *Qualität* der Automatisierung. Dabei wird besonders die psychomotorische Komponente in der Ausführung einer Handlung betont, die oft zum großen Teil automatisiert ist (z.B. beim Fahrradfahren). Entsprechend wird hier prozedurales Wissen als implizit und damit nur schwer verbalisierbar bzw. im Sinne von DeJong und Ferguson-Hessler (ebd.) in seine einzelnen Handlungskomponenten zerlegbar verstanden (vgl. Krause, 2007). Es wird angenommen, dass Wissen in der entsprechenden Situation („wenn“) automatisch und unbewusst aktiviert wird (vgl. auch *tacit knowing view*, Neuweg, 2005; Ummel, 2010). In diesem Zusammenhang weist jedoch Bromme (1992) unter dem Begriff des *Denkstils des Experten* auf die Wichtigkeit von Bewusstmachung und Reflexion von Wissen hin. Obwohl Wissen im Handlungsvollzug implizit angewendet wird, wird es dort als prinzipiell rechtfertigungsfähig angesehen (vgl. Fenstermacher, 1994).

Das Modell von DeJong und Ferguson-Hessler (1996) ermöglicht also verglichen mit den o.g. Konzeptionen durch die Unterscheidung von Wissensarten und Qualitäten sowie die Trennung von prozeduralem und strategischem Wissen zumindest theoretisch eine trennschärfere Konzeptualisierung der einzelnen Wissensarten (auch wenn aufgrund der in allen vorgestellten Modellen mehr oder weniger expliziten Verknüpfung der Wissensarten eine isolierte Operationalisierung bei komplexen Problemstellungen schwierig sein dürfte).

3.2.3 Merkmale anwendbaren Wissens

Anwendbarkeit als Eigenschaft von Wissen basiert auf dem Zusammenwirken verschiedener Wissensarten (Krause, 2007), das durch höhere Ausprägungen verschiedener Qualitätsdimensionen erleichtert werden kann.

Die Notwendigkeit des Zusammenwirkens der einzelnen Wissensarten wird offensichtlich, wenn man schema- oder skriptbasierte Modelle (Bartlett, 1932; Kopp & Mandl, 2005; Mandl,

Friedrich & Hron, 1988; Schank & Abelson, 1977) zugrundelegt. Die kognitive Struktur wird dort als assoziatives Netzwerk verstanden, in dem prozedurales und strategisches Wissen die Verbindungen zwischen einzelnen konzeptionellen, propositional repräsentierten Knotenpunkten herstellt. Objekte, Situationen, Ereignisse und Handlungen werden dabei so verarbeitet, dass ihre einzelnen Komponenten kognitiv als zusammenhängendes Konzept abgebildet werden (Kopp & Mandl, 2005). Wichtig ist dabei, dass die Zusammenhänge zwischen den einzelnen Wissenseinheiten spezifiziert werden (Anderson & Pearson, 1984). Die daraus resultierenden Schemata sind abstrakte Wissensrepräsentationen, in denen Wissen über typische Zusammenhänge in einem Bereich organisiert ist (Bartlett 1932; Mandl et al., 1988).

Schemata können auch Handlungssequenzen und -abläufe in Form sogenannter Handlungsschemata oder Skripte beinhalten (Schank & Abelson, 1977). Durch Wahrnehmung bestimmter Reize (z.B. Situationsmerkmale) können Skripte aktiviert werden (*Instantiierung*, Rumelhart & Ortony, 1977, vgl. auch *Bottom-up Processing*, Kopp & Mandl, 2005). Der Aufbau von Skripten entspricht der *Kompilierung* (vgl. Anderson, 1983) von Wissen und damit einer zunehmenden Automatisierung des Wissensabrufs in einer bestimmten Situation (und damit auch tieferer Verarbeitung und besserer Strukturierung, s. Kap. 3.2.1).

Informationsverarbeitungstheorien wie die *Cognitive Load-Theorie* (Sweller, van Merriënboer & Paas, 1998) oder die *Cognitive Theory of Multimedia Learning* (Mayer, 2005) postulieren auf Grundlage dieser Annahmen zur kognitiven Architektur des Menschen das Vorhandensein eines Langzeitgedächtnisses mit prinzipiell unbegrenzter Kapazität, in dem Schemata verschiedener Komplexität und unterschiedlicher Automatisierungsgrade abgelegt sind. Dem gegenüber steht das Arbeitsgedächtnis (als „Arbeitsspeicher“), dessen Kapazität grundsätzlich begrenzt ist und die Anzahl der Konzepte limitiert, die gleichzeitig verarbeitet bzw. abgerufen werden können. Eine einzelne, unverbundene Information belegt dabei die gleiche Menge an Ressourcen wie ein Schema bzw. Skript (Mayer, ebd.; Sweller et al., 1998).

Vor diesem Hintergrund wird ersichtlich, warum der Grad der Automatisierung einer Handlung ein wichtiger Faktor bei der Wissensanwendung ist (vgl. Stark, 2001). Schemata als kompiliertes Wissen können das Arbeitsgedächtnis entlasten und damit beispielsweise das Lernen neuer Abläufe erleichtern (Kirschner et al., 2006; Sweller et al., 1998). Analog dazu gibt Automatisierung kognitive Kapazität zur Anwendung von Wissen bei einer Problemlösung frei. Dagegen beansprucht der aktive Abruf von Konzepten, Prozeduren oder

Sequenzen bei geringer Automatisierung in hohem Maße kognitive Kapazität, so dass weniger Ressourcen zur Problemlösung zur Verfügung stehen.

Belege dafür finden sich z.B. in der Forschung zur *Wissensenkapsulierung* in der Medizin. *Wissensenkapsulierung* bedeutet, dass biomedizinisches Wissen mit klinischer Erfahrung angereichert und in Schemata generalisiert, also kompiliert wird (Boshuizen & Schmidt, 2008). Dabei kommt es zur Bildung von sogenannten *Illness Scripts*, die Wissen klinischer Natur wie Wissen um Hintergrundfaktoren bzw. Patientencharakteristika direkt mit Wissen über Symptome verknüpfen und so das frühe Erkennen von Krankheiten ermöglichen. Enkapsulierte domänenspezifische Schemata haben das gleiche Erklärungspotenzial wie elaborierte biomedizinische Konzepte, sind dabei aber ökonomischer (Boshuizen, Schmidt, Custers & Van de Wiel, 1995). Studien von Schmidt und Boshuizen (1993, s. a. Custers, Boshuizen & Schmidt, 1999) zeigten, dass Experten zur Diagnose auf kompiliertes Wissen zurückgreifen, während Novizen sich eher auf deklaratives, biomedizinisches Wissen, also einzelne Fakten und Konzepte stützen.

Dies ist beispielsweise beim Transfer von Wissen von Bedeutung, bei dem bekannte Konzepte auf neuartige Problemstellungen angewendet werden müssen. Verschiedene Autoren unterscheiden nach Art der Problemstellung zwischen nahem, mittlerem und weitem (z.B. Krause, 2007; Stark, 2001) oder nach Art des Problemlösungsansatzes zwischen einfachem (low-road) und höherem (high-road) Transfer (z.B. Bereiter, 1995; Bransford & Schwartz, 1999). Naher und weiter Transfer unterscheiden sich darin, dass strukturelle Merkmale der Problemstellung der Lernsituation entsprechen bzw. nicht entsprechen (Stark, 2001). Einfacher Transfer bezeichnet den spontanen und automatischen Transfer hochgeübter Fertigkeiten (das entspricht der Wissenskompilierung nach Anderson, 1983), während höherer Transfer die Anwendung von in einer Situation gelerntem abstraktem Wissen auf eine neuartige Situation umfasst (Woolfolk, 2008). Einen Überblick über zahlreiche weitere verwandte Konzeptionen bieten z.B. Mähler und Stern (2006). Die vorliegende Arbeit fasst die verschiedenen Transferarten aus pragmatischen Gründen jedoch unter dem Begriff der Wissensanwendung zusammen (vgl. Wagner, 2015).

3.2.4 Träges Wissen

Liegt Wissen überwiegend in deklarativer Qualität vor (vgl. DeJong & Ferguson-Hessler, 1996), kann es bei der Wissensanwendung aufgrund der oben angesprochenen Überlastung der kognitiven Kapazität zu Problemen bei der Wissensanwendung kommen. Gelingt zwar der Abruf, nicht aber die Anwendung von Wissen zur Problemlösung, spricht man von

trägem Wissen (z.B. Gruber & Renkl, 2000). Dieses ist speziell im Zusammenhang mit den in Kap. 2.1 erwähnten Problemen Lehramtsstudierender bei der Wissensanwendung von Bedeutung. Unter *trägem Wissen* versteht man Wissen, das zwar in Prüfungen reproduziert werden kann, aber bei der Bearbeitung von Problemstellungen nicht zur Anwendung kommt (*inert knowledge*, Whitehead, 1929; Renkl, 2001; 2006). Das entspricht auch dem Konzept des Nulltransfers (z.B. Bendorf, 2013). Damit ist Trägheit gewissermaßen die konträre Eigenschaft zur Anwendbarkeit von Wissen. Gruber und Renkl (2000) sprechen von *trägem Wissen*, wenn drei Bedingungen erfüllt sind:

- „1. Es kann gezeigt werden, dass eine Person über ein bestimmtes Wissen verfügt
2. Es gilt die begründete Annahme, dass das Vorhandensein des Wissens zu kompetenter Handlung befähigen sollte.
3. es kann gezeigt werden, dass die Person, die über das Wissen verfügt, die kompetente Handlung nicht ausführt, also das Wissen nicht anwendet.“ (S. 155)

Renkl (2006) nennt verschiedene mögliche Erklärungen für die Entstehung *trägen Wissens*. Die *Metaprozesserklärung* geht davon aus, dass *träges Wissen* entsteht, wenn Wissen zwar vorhanden ist, aber aufgrund fehlenden metakognitiven Wissens über die Anwendungsbedingungen nicht genutzt wird. Dieses ist vor allem im Bereich des Transfers von Bedeutung. Eine neue Situation muss als geeignet erkannt werden, um einen Transfer leisten zu können, was metakognitives Differenzierungs- und Generalisierungsvermögen sowie vorausschauende Zweck-Mittel-Abwägung erfordert (Woolfolk, 2008).

Die *Strukturdefiziterklärung* benennt verschiedene Ursachen: *Kompartimentalisierung* von Wissen liegt vor, wenn Wissen in unterschiedlichen Kontexten erworben und daher unterschiedlich kompartimentalisiert, d.h. redundant unter getrennten Kategorien abgespeichert wird (z.B. Vorlesung an der Universität vs. Schulpraktikum). Diese kategoriale Trennung erschwert den Transfer. Ebenso kann Wissen in einer *Form* abgespeichert sein, die keinen Transfer erlaubt, etwa wenn mangels Kenntnis von Strukturierungsmöglichkeiten des Wissenserwerbs Schemata unzureichend ausgebildet sind oder keine Prozeduralisierung bzw. Kompilierung von Wissen stattgefunden hat (s. dazu Kap. 2.3.; Levin, 2010). Insbesondere die mangelnde Wissenskompilierung, also die Überführung deklarativen Wissens in Produktionsregeln, wird dabei im universitären Kontext als Ursache *trägen Wissens* benannt (z.B. Stark, 2001).

Eine weitere Erklärung setzt bei der situativen Bindung von Wissen an. Die *Situiertesklärung* geht davon aus, dass sich Wissen in der Interaktion zwischen der Situation und ihren Gegebenheiten sowie der Person und ihren Erfahrungen (Wissen) konstituiert. Dem liegt die radikale konstruktivistische Perspektive zugrunde, dass Wissen weder abstrakt noch kontextfrei gelernt werden kann. Demzufolge hängt die erfolgreiche Wissensanwendung von der Ähnlichkeit zwischen der Lern- und Anwendungssituation ab. Unterscheiden sich diese zu stark, kann Wissen nicht angewendet werden und bleibt damit träge (s.a. Fölling-Albers, Hartinger & Mörtl-Hafizovic, 2004; Gruber & Renkl, 2000). DeJong und Ferguson-Hesslers (1996) Konzept des situationalen Wissens als kontextgebundenes Wissen, das aber in abstrakte Prinzipien gefasst werden kann, stellt eine moderatere Konzeption dar.

Ähnlich wie in Kap. 2.2 bei der Darstellung der Fehler und ihrer Ursachen bieten auch die Erklärungen zum trägen Wissen Ansatzpunkte, um seine Entstehung zu verhindern. Renkl (2006) empfiehlt zur didaktischen Gestaltung von Lernumgebungen bzw. Lernsituationen etwa, die Lernsituation möglichst ähnlich zur Anwendungssituation zu gestalten. Dies soll Wissenskompartmentalisierung verhindern und gleichzeitig Wissen über die Anwendungsbedingungen vermitteln (vgl. strategisches Wissen; Kap. 3.2.1). Dazu sollte der Wissenserwerb anhand komplexer, realitätsnaher und authentischer Problemstellungen erfolgen, wie sie in *situated cognition*-Ansätzen (Reinmann & Mandl, 2006; z.B. *Anchored Instruction*, CTGV, 1992) vorgestellt werden.

3.3 Theorieartikulation: Theorieanwendung zur Erklärung von Sachverhalten

Als Theorieartikulation bezeichnet Ohlsson (1992) die „Anwendung einer Theorie auf einen Sachverhalt oder ein Phänomen und die Ableitung dessen, was die Theorie implizit oder explizit darüber aussagt“ (S. S.182, Übersetzung d. Autors). Theorieartikulation kann im Sinne von Krause (2007) als Wissensanwendung beschrieben werden, die das Zusammenspiel verschiedener Wissensarten sensu DeJong und Ferguson-Hessler (1996) umfasst und *anwendbares Wissen* erfordert. Zur Theorieanwendung muss damit nicht nur Wissen um den *Inhalt* einer Theorie, sondern auch prozedurales Wissen über das dazu nötige Vorgehen vermittelt werden.

Der Prozess selbst ist komplex und in der Person verankert, die die Theorie „benutzt“, also die Aussagen einer Theorie mit den beobachteten Fakten in Beziehung setzt (Ohlsson, 1992). Das Vorgehen ist damit nicht unabhängig von der Person, aber unabhängig von der jeweiligen Theorie, da es üblicherweise nicht Inhalt der Theorie ist: „Die abstrakten Prinzipien einer

Theorie sagen etwas über die Welt, sie sagen nichts darüber, wie man sie nutzen soll“ (Ohlsson, ebd., S.183; Übersetzung: Autor). Konzeptuelles Wissen über die Prinzipien einer Theorie reicht also zur Anwendung nicht aus, gerade hier besteht im Gegenteil die Gefahr, dass es isoliert bzw. kompartimentalisiert zu trägem Wissen wird. Da Lernende sich das zur Anwendung benötigte prozedurale und strategische Wissen nicht aus der Theorie ableiten und damit nur schwer selbst erschließen können, muss der Erwerb dieses Wissens systematisch gefördert werden (Ohlsson, ebd.).

Die Erklärung von Sachverhalten benennt Ohlsson (1992) dabei explizit als die wichtigste Funktion dieser Form der Theorieanwendung. Sie kann neben anderen Funktionen der Theorieanwendung wie Prognose und Ableitung von Konsequenzen in das Modell von Beck und Krapp (2006) eingeordnet werden und entspricht dem dortigen Konzept des *rückschauenden* Begreifens. Ohlssons (1992) Konzept wissenschaftlicher Erklärungen entspricht dabei der deduktiven Ableitung aus nomologischen Aussagen (Hempel & Oppenheim, 1984, Westermann, 2000). Dabei wird anhand des logisch-systematischen Ansatzes (Westermann, ebd.) die Beziehung zwischen dem Erklärenden und dem zu Erklärenden ausgearbeitet. Dies erfordert auch die Beschreibung des Phänomens, da „einen Vorgang kausal [zu] erklären, heißt, einen Satz, der ihn beschreibt, aus Gesetzen und Randbedingungen logisch abzuleiten“ (Popper, 1994, S. 31). Nach dem deduktiv-nomologischen Modell von Hempel und Oppenheim (1984) besteht die daraus resultierende Erklärung also aus dem Explanans (dem Erklärenden, Gesetze, Prinzipien und Gültigkeitsbedingungen der Gesetze) und dem Explanandum (das zu Erklärende, die Beschreibung des Sachverhalts). Dabei folgt bei einer logisch korrekten Erklärung das Explanandum deduktiv aus dem Explanans. Bei der Beurteilung komplexer schulischer Situationen muss diese Forderung allerdings eingeschränkt werden. Hier kann eine Erklärung zwar dem o.g. Muster folgen, wird aber aufgrund des Fehlens absoluter Gesetzmäßigkeiten, das bildungswissenschaftlichem Wissen zugrundeliegt, eher auf einem probabilistischen Niveau bleiben. Immerhin sollte aber der Bezug auf wissenschaftliches Wissen dabei die Wahrscheinlichkeit einer korrekten Erklärung erhöhen (vgl. Kap. 2.3).

Aus den Erklärungen können sich sogenannte Erklärungsmuster (Ohlsson, 1992) entwickeln, die als Schemata (Schank, 1986, Sweller et al., 1998) aufgefasst werden können. Dies ermöglicht wiederum die in Kapitel 3.2.2 angesprochene Automatisierung der Anwendung und den schnellen Zugriff auf solche Informationen. Im vorliegenden Forschungsprojekt entspricht das einer schnellen, wissenschaftlich fundierten Analyse einer schulischen

Situation, die im Hinblick auf die ökologische Validität des Ansatzes einer ausgedehnten Erklärungssuche meist vorzuziehen ist.

Im Sinne des in Kapitel 2.3 angesprochenen Ansatzes, eine Intervention auf den spezifischen Problemen Lehramtsstudierenden bei der Wissensanwendung aufzubauen, kann dabei das Wissen um mögliche Fehler genutzt werden, um das richtige Vorgehen zusätzlich abzusichern und so durch die Vermeidung von Fehlern im schulischen Alltag wertvolle Zeit zu sparen. Das folgende Kapitel beschäftigt sich daher mit dem Lernen aus Fehlern, seinen Eigenschaften und Bedingungen und seinem Produkt, dem sogenannten „negativen Wissen“.

4. Lernen aus Fehlern

4.1 Begriffsbestimmung: Was ist ein Fehler?

Es gibt verschiedenste Konzeptionen darüber, welche Handlungen aus welchem Grund einen Fehler darstellen und wie diese zustande kommen und demzufolge verschiedenste Fehlertaxonomien (z.B. Kobi, 1994; Norman, 1981; Reason, 1990; Rasmussen, 1982; 1987; zusammenfassend z.B. Weingardt, 2004). Ihnen allen gemein ist die Betrachtung von Fehlern als nicht absichtsgeleiteten Handlungen, in Abgrenzung zu Fälschung und Täuschung als Konsequenz absichtlicher Handlungen (Weimer, 1925; 1926). Fehler resultieren oft aus gut gemeinten Lösungsansätzen, die aus verschiedenen Gründen (Fehler in der Zielbildung, dem Denkprozess, der Ausführung) unangemessen, unökonomisch und ineffizient sind. Dabei orientiert sich die Einstufung der Handlung als „richtig“ oder „falsch“ häufig an gesellschaftlichen Normen als Bezugssystem, ohne die eine solche Unterscheidung nicht möglich wäre (Oser, Hascher & Spychiger, 1999; Rollet, 1999). Fehler können also als Handlungen, Entscheidungen und Unterlassungen definiert werden, die in der Verfehlung eines intendierten Handlungsziels resultieren (Bauer & Mulder, 2008; Frese & Zapf, 1994).

Fehler können in verschiedenen Bereichen auftreten, etwa im fachlichen, moralischen oder sozialen Bereich. Da für die vorliegende Arbeit die beiden letzteren Kategorien jedoch nicht von Bedeutung sind, soll dieser Verweis auf sie genügen. In der Schule liegt das Hauptaugenmerk zumeist auf den fachlichen Fehlern (Althof, 1999). Hier gibt es in verschiedenen Fächern wie der Mathematikdidaktik (Schmassmann, 1992) sowie dem Fremd- und Zweitsprachenunterricht (Demme, 1993; Nickel, 1972) eine lange Tradition der Fehleranalyse. So werden beim Spracherwerb etwa Kompetenz- und Performanzfehler unterschieden (Nickel, 1972), in der Mathematik unterscheidet Schmassmann (1992) zwischen Fehlern beim Umgang mit Symbolen (Schnittstellenfehler), Verständnisfehlern (beim Vernetzen von Wissen), Automatisierungsfehlern (bei der Kombination von Symbolen) und Umsetzungsfehlern beim Transfer. Ebenso werden in diesen Bereichen der Umgang mit Fehlern sowie Umsetzungen des Lernens aus Fehlern diskutiert (z.B. Schweizerische Mathematik-Konferenz; Wieland, 1991). Mehl (1993) unterscheidet zusätzlich auf lernpsychologischer Ebene zwischen „guten“ und „schlechten“ Fehlern, wobei gute oder produktive Fehler den Lernprozess voranbringen können und für die Lernenden von Nutzen sind (Kobi, 1994). In Bezug auf den Lernprozess der Klasse kann hier auch noch zwischen Fehlern unterschieden werden, die möglicherweise für den Einzelnen produktiv sind, aber für

Andere, die den Lernstoff bereits verinnerlicht haben, wenig Neues bieten (oder umgekehrt; vgl. Oser et al., 1999).

Trotz vorhandener Studien wird allerdings in der Lehramtsausbildung selbst der richtige Umgang mit Fehlern wenig beachtet (Demme, 1993). Mindnich, Wuttke und Seifried (2008) weisen in diesem Zusammenhang auf das Fehlen systematischer Analysen möglicher Fehlerarten, deren Ursachen und Wirkungen in vielen Domänen hin. Zwar gibt es durchaus Ausnahmen (z.B. Gschwendtner, Geißel & Nickolaus, 2007; Große & Renkl 2007; Minnameier, 2008; Oser et al., 1999; Oser & Spychiger, 2005, Stark, 2005, Stark, Kopp & Fischer, 2011), Mindnich et al (2008) kommen jedoch zu dem Schluss: „Alles in allem weiß man zum Lernen aus Fehlern bislang erstaunlich wenig.“ (S. 154). Einen Schritt zur Behebung des Problems in Form der Klassifizierung der Formen und des Umgangs mit Fehlern unternahmen Mindnich et al. (ebd.) im Bereich des kaufmännischen Unterrichts. Dabei wurde auf Grundlage der Taxonomie von Anderson und Krathwohl (2001) auf Schülerseite nach Fehlern bei der kognitiven Verarbeitungsleistung bzgl. Reproduktion, Verständnis, Anwendung und Informationserzeugung unterschieden, des weiteren wurde die Qualität der Rückmeldung durch die Lehrpersonen in der sogenannten Fehlersituation beurteilt („nachfassen“ bzw. kein „nachfassen“ bei hoher bzw. niedriger Elaboration der Rückmeldung). Es zeigte sich, dass die meisten Fehler auf relativ niedrigem Niveau (Reproduktion bzw. Verständnis) gemacht wurden und auch die Rückmeldung durch die Lehrpersonen meist eher ungünstig ausfiel (kein Nachfassen bzgl. des Fehlers, geringe Elaboration).

Die bereits in Kapitel 2.2 erwähnte Studie von Stark (2005) stellt eine Klassifizierung von typischen Fehlern bei der Anwendung bildungswissenschaftlichen Wissens im Bereich der Lehrerbildung vor. Stark (ebd.) identifizierte bei der qualitativen Auswertung von Antworten Studierender auf verschiedene Aufgaben aus dem akademischen Kontext drei Arten typischer Fehler: Die Analyse komplexer schulischer Situationen und Phänomene auf der Basis von Alltagswissen und subjektiven Theorien (Fehlertyp 1), die suboptimale Auswahl wissenschaftlicher Theorien, Modelle und Konzepte oder unangemessene Anwendung passender Konstrukte (Fehlertyp 2) und die fehlerhafte Bezugnahme auf empirische Evidenz (Fehlertyp 3) bei der Anwendung wissenschaftlichen Wissens (inklusive verschiedener Subtypen und Mischformen). Die Studie 2 zugrundeliegende *Fehlermatrix* beruht auf dieser Klassifikation (s. Kap. 7).

Andere Studien untersuchen Eigenschaften von Fehlern wie etwa den Schweregrad oder die Neuartigkeit. Das sehr stark mathematisch gefasste Modell von Martínez-Legaz und

Soubeyran (2003) postuliert, dass Lernen aus Fehlern nicht nur die Vermeidung einer einzigen fehlerhaften Handlung in einem hypothetischen Raum ermöglicht, sondern auch „benachbarte“ (fehlerhafte) Handlungen so vermieden werden können. Konsequenterweise werden schwerwiegende Fehler als besonders geeignet für initiales Lernen betrachtet, da sie in diesem Ansatz ein weites Feld zu vermeidender Handlungen umschließen, das nach und nach in kleinere Einheiten (konkretere Fehler) verfeinert werden kann. Aus dem Bereich der Neuropsychologie stammt die Untersuchung von Wills, Lavric, Croft und Hodgson (2007), in der gezeigt wurde, dass Informationen schneller verarbeitet wurden, wenn sie überraschend waren, d.h. die ursprüngliche Erwartung der Lernenden sich als inkorrekt erwiesen hatte. Zudem lenkten Lernende ihre Aufmerksamkeit länger auf derartige Informationen.

In diesen Studien wird eine wichtige Funktion des Fehlers deutlich: Werden Fehler als Lerngelegenheiten genutzt, können sie Grenzen und Probleme im Denk- oder Problemlöseprozess aufzeigen. So können Probleme, Schwächen und Mängel identifiziert werden, die im Ernstfall nicht auftreten sollen. Daraus ergibt sich eine der wichtigsten Funktionen des Fehlermachens: Es kann dazu beitragen, das korrekte Vorgehen besser vom falschen abzugrenzen und damit abzusichern. Oser et al. (1999; s.a. Oser, 2007) bezeichnen diese Art von fehlerbezogenem Wissen daher auch als „Schutz-“ oder „Unterscheidungswissen“. Implizit ist diese Idee in vielen lernpsychologischen Ansätzen verankert, wie etwa beim problemlösenden Lernen (CTGV, 1992), dem Lernen durch Einsicht oder auch dem entdeckenden Lernen (Bruner, 1961; Weibell, 2011). Auch das Expertiseparadigma (z.B. Bromme, 2008; Custers et al., 1999) bezieht sich indirekt auf diese Thematik, wenn dargestellt wird, wie sich die komplexen Wahrnehmungs- und Informationsverarbeitungssysteme von Experten anhand einer langjährigen Fehler- und Verbesserungskultur entwickeln (Oser et al., 1999). Schließlich bauen auch Modelle aus der Entwicklungspsychologie auf der Beobachtung von Kindern auf, die bei der Lösung von Aufgaben Fehler machen und durch Veränderung ihrer Strategie Korrekturen und Sicherung des Wissens vornahmen (vgl. Montada, 1968). Dazu muss es aber erlaubt sein, Fehler zu machen, was eine bestimmte Auffassung über den Umgang mit Fehlern voraussetzt - eine produktive Fehlerkultur.

4.2. Fehlerkultur: Produktiver Umgang mit Fehlern

Es gibt heutzutage zahlreiche verschiedene Ansätze zum Umgang mit und zum Lernen aus Fehlern, etwa in der Medizin (*Incident Reporting*, z.B. BMJ Case Reports, 2015; Frankel, Knoblock & Berwick, 2015; Stefanko, Lancashire, Coombes & Fassett, 2009) aber auch in

verschiedenen Bereichen der Pädagogik (s.o.; auch z.B. *Advokatorisches Lernen*, Oser, 2007; *Productive Failure*, Kapur & Bielaczyc, 2012). Voraussetzung für die heutige Betrachtung und Nutzung von Fehlern in verschiedensten Formen als Lerngelegenheiten war eine sich über Jahrzehnte verändernde Fehlerkultur in Wissenschaft und Gesellschaft. Die Auffassung von Fehlern hat sich insbesondere im 20. Jahrhundert stark verändert. Herrschte zunächst noch eine auch im Vokabular stark defizitorientierte Perspektive vor (z.B. Kießling, 1925; Skinner, 1974; Weimer, 1925; 1926), so werden Fehler heutzutage als Ressourcen im Lernprozess betrachtet (z.B. Booth, Lange, Koedinger & Newton, 2013; Durkin & Rittle-Johnson, 2012; Kapur, 2014; Oser, 2007; Oser & Spychiger, 2005; Stark et al., 2011; Weingardt, 2004).

Weimer (1925; 1926) führte Fehler etwa auf das Versagen psychischer Funktionen (Aufmerksamkeit, Gedächtnis und Denken) zurück. Fehler stellen in seiner Sichtweise Handlungen dar, die „gegen den Willen ihres Urhebers vom richtigen abweich[en]“ und als „Äußerungen und Kennzeichen tiefer liegender Schäden“ bei „entsprechenden erzieherisch-unterrichtlichen Maßnahmen durchaus vermeidbar“ sind (Weimer, 1930, zitiert nach Weingardt, 2004, S. 45). Kießling (1925) vertrat eine ähnlich defizitgeprägte Position, erweiterte diese jedoch um den situativen Kontext. Seine Konzeption bezog den Einfluss situativer und Umweltfaktoren auf die Disposition einer Person, Fehler zu begehen, ein und erweiterte so Weimers Auffassung um wichtige Überlegungen zur Komplexität der Entstehung von Fehlern.

Nachdem Fehler in den folgenden Jahrzehnten in der Pädagogik vornehmlich als unangenehme, aber unvermeidliche Begleiterscheinung des Unterrichts betrachtet wurden (etwa beim *programmierten Lernen*; z.B. Skinner, 1974), entwickelte sich in der Pädagogik der 90er Jahre eine ressourcenorientierte Perspektive in der Fehlerforschung. Kobi (1994) betrachtet Fehler nicht als inhärent falsch (vgl. auch Ohlsson, 1996), sondern als Abweichungen von einer individuell, durch Konvention oder Argumentation gesetzten Norm (s.a. Rollet, 1999). Diese Perspektive betrachtet Fehler als unangemessene, unökonomische oder ineffiziente *Versuche* einer „situations- und systemgerechten Lösung“ (zitiert nach Weingardt, 2004, S. 63). Daraus ergibt sich aber lernpsychologisch eine Möglichkeit, Fehler als Lerngelegenheiten zu nutzen oder in ihnen sogar neue Lösungsansätze zu entdecken.

Im Bereich der Medizin wird dem Lernen aus Fehlern anhand der Kasuistik schwieriger oder uneindeutiger Fälle (z.B. Loke, Price, Derry & Aronson, 2006; Stefanko et al., 2009; spezialisiertes Journal: *BMJ Case reports*, 2015) und den daraus resultierenden inadäquaten Behandlungsmethoden an höchster Stelle eine große Bedeutung zugemessen. So

fordert die WHO (2008) im Rahmen der „World Alliance for Patient Safety“ seit 2004 die Erfassung der Häufigkeit und die Klassifizierung medizinischer Fehler in einem leicht zugänglichen Erfassungssystem, um Strategien zur Vermeidung von Fehlern in der Behandlung von Patienten zu entwickeln. Es existieren zahlreiche Datenbanken, Lernprogramme und spezialisierte Journals zu Behandlungsfehlern (z.B. *Institute for Healthcare Improvement*; Frankel et al., 2015; BMJ Case reports, 2015; Loke, Price, Derry & Aronson, 2006; Stefanko et al., 2009). Derartigen Ansätzen zum Lernen aus Fehlern direkt am Arbeitsplatz als Spezialfall des *Workplace Learning* (Billet, 2004) wird ein besonderes Lernpotential zugesprochen. Am Arbeitsplatz ermöglicht fehlerbasiertes Lernen die Entwicklung beruflicher Kompetenzen durch die Reflektion dysfunktionaler Handlungsrountinen (Bauer & Mulder, 2008; Harteis, Bauer & Gruber, 2008).

Das Lernpotential von Fehlern wird in aktuelleren Konzeptionen also klar herausgestellt. Die frühere Defizitperspektive wird durch eine ressourcenorientierte Perspektive ersetzt. Die resultierende Veränderung der Fehlerkultur führte dazu, dass Fehler heutzutage als explizit erlaubte und wichtige Schritte im Lern- und Optimierungsprozess betrachtet werden. Eine ähnliche Perspektive findet sich auch in der Forschung zum Conceptual Change, die fehlerhaftes Vorwissen („Fehlkonzepte“) explizit in ihre Lernmodelle miteinbezieht (z.B. Posner, Strike, Hewson, & Gertzog, 1982; Pintrich, Marx & Boyle, 1993). Der Einbezug und die Aktivierung fehlerhaften oder vorwissenschaftlichen Vorwissens wird im Lernprozess als unerlässlich betrachtet (z.B. Vosniadou, Ioannides, Dimitrakopoulou & Papademetriou, 2001), da die Fehlkonzepte durch Lernen nicht gelöscht werden, sondern im Gedächtnis verbleiben und dort lediglich als fehlerhaft gekennzeichnet werden (wie das geschieht, wird in unterschiedlichen Ansätzen auch sehr unterschiedlich operationalisiert, zusammenfassend s. Stark, 2003).

4.3 Negatives Wissen als Produkt des Lernens aus Fehlern

Ist eine derartige Fehlerkultur in einem Lernumfeld gegeben, kann das Lernpotential von Fehlern nutzbar gemacht werden, wenn eine detaillierte Fehleranalyse durchgeführt wird. Ziel des Lernens aus Fehlern ist der Aufbau *negativen Wissens* (Oser, 2007; Oser & Spychiger, 2005). Negatives Wissen umfasst deklaratives, prozedurales und strategisches Wissen (Oser, 2007), das an den Definitionen *positiver* Wissensarten von DeJong und Ferguson-Hessler (1996, s. Kapitel 3.2.1) kontrastiert werden kann. Negatives deklaratives Wissen ist *beschreibendes* Wissen über Konzepte, Prinzipien und Modelle und enthält Informationen darüber, wie oder was etwas *nicht* ist, beispielsweise das Wissen, dass die Erde nicht flach ist.

Negatives prozedurales Wissen enthält Wissen darüber, wie etwas nicht funktioniert oder welche Handlungsschritte nicht möglich sind. Negatives strategisches Wissen ist Wissen darüber, welche Strategien *nicht* zur Lösung eines Problems führen (Oser, 2007; Oser & Spychiger, 2005; vgl. auch Guldemann & Zutavern, 1999). Das Wissen um effektive Strategien zur Fehlervermeidung gilt auch als wesentliches Kennzeichen von Experten (Gruber, 1999; Minsky, 1994).

Die *Funktion negativen Wissens* liegt primär in der Abgrenzung des Falschen vom Richtigen und damit im Schutz positiven Wissens (s. Kap. 4.1): Je mehr negatives Wissen erworben wird, desto genauer können falsche von richtigen Konzepten abgegrenzt werden und desto wahrscheinlicher wird richtiges Handeln (Oser, 2007; Oser et al., 1999). Das so erworbene *Schutzwissen* (Oser & Spychiger, 2005) beinhaltet durch den Kontrast am Richtigen Informationen darüber, welche Möglichkeiten zur Korrektur vorliegen und welche Konsequenzen sich aus dem Fehler für den Handelnden ergeben. Besonders lernwirksam sind damit Fehler, bei denen die zur Lösung notwendigen Schritte zumindest implizit schon in der falschen Handlung erkennbar sind (ebd.). Diese Art von Fehlern in Lernmaterialien wird als *instrukionaler Fehler* bezeichnet (z.B. Stark et al., 2011).

Damit liegt der wesentliche *Lernmechanismus* des Lernens aus Fehlern in der Elaboration der Kontrastierung des Falschen am Richtigen (z. B. Curry, 2004; Oser & Spychiger, 2005). Um aus Fehlern zu lernen, müssen Lernende den Fehler identifizieren und bewusst als Lernanlass nutzen. Die Bereitschaft dazu wird als entscheidender individueller Einflussfaktor unter dem Begriff der *Fehlerorientierung* etwa im Bereich des Lernens am Arbeitsplatz diskutiert. (Rybowiak, Garst, Frese & Batinic, 1999). Das Konzept der Fehlerorientierung bezeichnet ein multi-dimensionales Konstrukt, das die individuelle Bewertung von Fehlersituationen, den Umgang mit derartigen Situationen und die Einschätzung des Nutzens für die eigene Praxis umfasst. Eine Fehlerorientierung, die Fehler nicht nur als negative Leistungsindikatoren bzw. zu vermeidende Unfälle (Harteis et al., 2008; Rybowiak et al., 1999; Weimer, 1925) wahrnimmt, sondern einen aktiven und offenen Umgang mit Fehlern ermöglicht, begünstigt ein ertragreiches, fehlerbasiertes Lernen.

Lernende müssen beim aktiven Nachvollzug der Kontrastierung des Fehlers an der Lösung und dessen Korrektur unterstützt werden (z.B. Durkin & Rittle-Johnson, 2012; Stark et al., 2011). Es wird diskutiert, ob als Bedingung für den Erwerb negativen Wissens neben Identifikation und Verständnis auch die Möglichkeit zur eigenständigen Korrektur des Fehlers gegeben sein muss (Große & Renkl, 2007; Kapur & Bielaczyc, 2012; Oser et al., 1999) oder

eine elaborierte Rückmeldung zum korrekten Vorgehen hinreichend ist (z.B. Stark et. al, 2010).

Die Grundlage für den Lernmechanismus sehen verschiedene Autoren in der Auslösung von Selbsterklärungen. VanLehn (1999) postuliert in seiner CASCADE-Theorie, dass Fehler Reflexionen einleiten, die zu einem tieferen Verständnis führen. In Studien zum *impasse-driven learning*-Ansatz von VanLehn, Siler, Murray, Yamauchi und Baggett (2003) führt die Präsentation von instruktionalen Fehlern Lernende in eine Art kognitive Sackgasse, die zu aktiver Elaboration bzw. Selbsterklärung (im Sinne metakognitiver „Selbstoprompts“) anregt. Damit wird der Aufbau einer Wissensstruktur unterstützt, die mit Expertiseentwicklung assoziiert wird (Boshuizen et al., 1995). Auch der auf Piaget zurückgehende kognitive Konflikt (z.B. Piaget, 1968) kann durch instruktionale Fehler ausgelöst werden. In Kolodners (1983) Modell zum Lernen aus Erfahrung wird ein Lernprozess durch die Abweichung vom Sollzustand ausgelöst, die Analyse von Fehlern ist dabei einer der zentralen Lernmechanismen. Eine Studie von Maxwell, Masters, Kerr und Weedon (2010) zu Lernstrategien beim Erwerb psychomotorischer Fähigkeiten beim Golfspielen zeigten, dass Lernende in einer Bedingung mit Fehlern einen „selektiven Lernmodus“ verwendeten, in dem explizit Strategien zur Vermeidung von Fehlern formuliert und getestet wurden. Dies wirkte sich positiv auf den Erwerb verbalisierbaren prozeduralen Wissens aus. Lernende ohne Fehler (unselektiver Lernmodus) formulierten dagegen weniger explizite und verbalisierbare Strategien.

Fehler werden dabei nicht „gelöscht“, sondern als Kontraste zu positivem Wissen im episodischen Gedächtnis enkodiert (Stark, Kopp & Fischer, 2009). Das negative Wissen wird so in Beziehung zu *positivem* Wissen darüber gesetzt, wie der Fehler in Zukunft zu vermeiden ist. Diese Verknüpfung von negativem und positivem Wissen dient zum Aufbau von Strategien zur Vermeidung der Fehler (Oser & Spychiger, 2005).

4.4. Negatives Wissen als zusätzliche Wissensdimension aus der Perspektive der Conceptual Change-Forschung

Negatives Wissen kann als zusätzliche Wissensdimension betrachtet werden, die allen positiven Wissensarten gegenübergestellt werden kann (s. Abb. 1). Damit kann der Ansatz des Lernens aus Fehlern potentiell an jedes erdenkliche Lernprogramm angeschlossen werden, um dieses mit einer zusätzlichen Wissensdimension anzureichern. In Bezug auf die Wissensanwendung kann so das korrekte Vorgehen zusätzlich abgesichert werden, zudem

können Lernende auf Fehlkonzepte aufmerksam gemacht werden, die ihnen möglicherweise nicht bewusst waren.

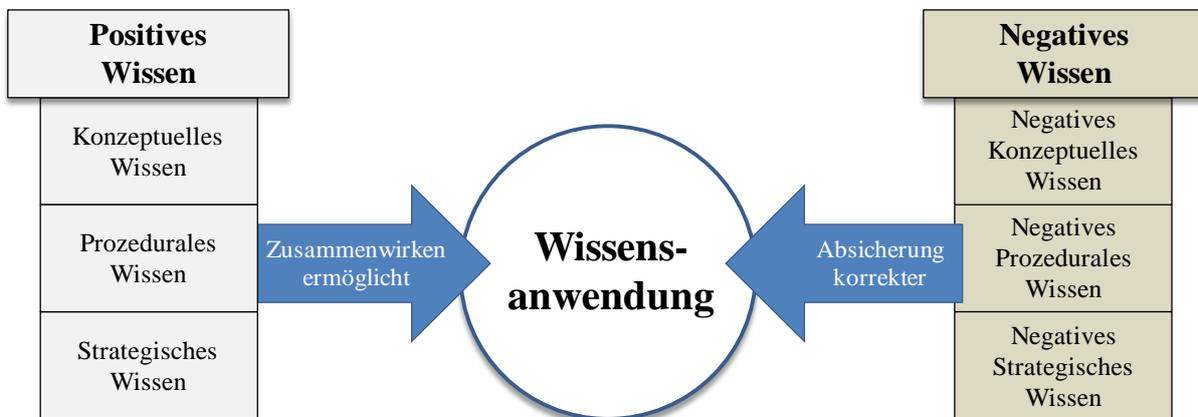


Abbildung 1: Negatives Wissen als zusätzliche Wissensdimension

Die Konzeptionen der Conceptual Change-Forschung eignen sich dabei sehr gut zur Beschreibung des Lernens aus Fehlern. Gerade in diesem Bereich wird das Vorwissen, anders als beispielweise in vielen Modellen aus dem Bereich des *instructional design* (e.g. *Cognitive Load Theory*, Sweller, 2005; Sweller et al., 1998; *Cognitive Theory of Multimedia Learning*, Mayer, 2005), nicht unbedingt als lernförderlich betrachtet. Im Gegenteil können vorhandene Fehlkonzepte den Lernprozess behindern (Chi & Roscoe, 2002). Das ursprüngliche Konzept des *cold Conceptual Change* basiert auf der rationalen Erkenntnis, dass ein Ansatz oder vorhandenes Vorwissen ungeeignet oder defizitär ist, was die Suche nach Erklärungen und Lösungsansätzen auslöst (Posner et al, 1982). Pintrich et al. (1993) kritisierten, dass diese Suche von persönlichen Einstellungen und Interessen, aber auch situativen Zuständen Lernender beeinflusst wird, dies bezeichnet er als *hot conceptual change* (s.a. Sinatra & Mason, 2008). Widerspricht nun vorhandenes Wissen (das sich möglicherweise bewährt hat, vgl. *subjektive Theorien* in Kap. 2.2) neuen Konzepten, kann es zur Ablehnung des neu zu erwerbenden Wissens kommen oder es entstehen parallele Wissenstrukturen, in denen altes (fehlerhaftes) Wissen unvernetzt neben korrekten Konzepten existiert (Chi & Roscoe, 2002; Schnotz, Vosniadou & Carretero, 1994). Lernende sind dabei oft nicht in der Lage, ihren eigenen Wissenstand und dessen Widersprüche zu wissenschaftlichem Wissen einzuschätzen (Merenluoto & Lehtinen, 2004). Dies ist insofern problematisch, als dass ein wesentlicher Teil des Lernprozesses (analog zum Lernen aus Fehlern) gerade in diesem Kontrast besteht (Calik, Okur & Taylor, 2010). Dieser führt dazu, dass Fehlkonzepte als solche gekennzeichnet

werden und dem korrekten Wissen gegenübergestellt werden (Rekategorisierung von Fehlkzepten; Chi & Roscoe, 2002).

4.5. Lernen aus eigenen versus Lernen aus fremden Fehlern

Generell kann beim Thema Lernen aus Fehlern zwischen Ansätzen unterschieden werden, die auf Rückmeldungen zu eigenen Fehlern Lernender basieren (z.B. *Productive Failure*, Kapur, 2008; *impasse-driven learning*, VanLehn, 1999; VanLehn et al., 2003) und Ansätzen, bei denen anhand von Fehlern Anderer gelernt wird („*umgekehrtes Modelllernen*“, Spychiger, 2004, S.3; z.B. *Advokatorisches Lernen*, Oser & Spychiger, 2005; Oser, 2007; *Lernen anhand fehlerhafter Lösungsbeispiele*, Stark et al., 2010). Die Befundlage ist dabei nicht ganz eindeutig. Während Studien aus dem Bereich der Mathematik, Lehrerbildung und Medizin (z.B. Durkin & Rittle-Johnson, 2012; Klein, Wagner & Stark, 2012; Stark et al., 2011) positive Effekte des Lernens aus fremden Fehlern auf den Wissenserwerb nachweisen, zeigen andere Studien Vorteile des Lernens aus eigenen Fehlern, wenn dieses adäquat instruktional unterstützt wird.

Diesen Ansatz verfolgen etwa die Theorie des *Constructive Failure* (Clifford, 1984), das Konzept der *Desirable Difficulties* von Schmidt und Bjork (1992) oder das bereits genannte *Impasse-driven Learning* (VanLehn et al., 2003). In diesen Ansätzen werden Lernende mit offenen Problemstellungen konfrontiert, die bewusst zur Überforderung führen und damit Lernende auf Wissenslücken und Fehlkzepten aufmerksam machen können. Grundlage ist im Prinzip das entdeckende Lernen nach Bruner (1961; s.a. Weibell, 2011), sowie neuere Ansätze von Bransford und Schwartz (1999), die die Rolle der selbständigen Entwicklung von Lösungsansätzen beim Lernen betonen (s.a. Schwartz & Martin, 2004). Auch hier wird beim eigenständigen Problemlöseversuch eine Hypothese aufgestellt, die in einer (unzureichenden) Strategie mündet. Der kognitive Konflikt tritt erst nach dem initialen Problemlöseversuch und dem Erlebnis des Scheiterns auf.

Eine derartige Konzeption liegt etwa dem *productive failure*-Ansatz (Kapur, 2008; Kapur & Bielaczyc, 2012) zugrunde. Dabei werden Lernende mit neuartigen Problemstellungen konfrontiert, die sie nur auf Grundlage ihres Vorwissens oder ihrer Intuition bearbeiten müssen. Dies führt in dem meisten Fällen zum Scheitern (failure), es werden Fehler gemacht und das Ergebnis ist unbefriedigend. Der Erwerb korrekten Wissen wird durch eine im folgenden Schritt bereitgestellte instruktionale Unterstützung (Erklärung, Instruktion oder Feedback) sichergestellt. Studien im Bereich der Statistik zum Thema Standardabweichung

zeigten positive Effekte dieses Ansatzes u. a. auf konzeptuelles Verständnis und Transfer im Vergleich mit direkter Instruktion (Kapur, 2010; 2012) und dem Lernen aus fremden Fehlern (productive failure vs. „*vicarious learning*“; Kapur, 2014). Mit Bezug auf die Wissensarten liegt der Fokus hier jedoch eindeutig auf positivem Wissen, da dieser Ansatz keine explizite Fehleranalyse vorsieht. Oser und Spychiger (2005) beschreiben jedoch eine Studie, in der Kinder auch ohne explizite Benennung von Fehlern falsche Handlungen identifizieren konnten und leiten daraus die Empfehlung ab, Lernende durchaus auch die Erfahrung des eigenen Scheiterns machen zu lassen (S.57).

Einen direkten Vergleich zwischen dem Lernen aus eigenen und fremden Fehlern stellten Ivancic und Hesketh (2000) beim Fahrtraining in einem Fahrsimulator an. Durch einen schwierigen Trainingsparcours wurden Fahrfehler induziert, diese resultierten in Kollisionen oder Strafzetteln. Das Erleben eigener Fehler führte zu besserer Fahrleistung in ähnlichen Situationen und effektiverer Nutzung von Strategien zur Fehlervermeidung in einer neuartigen Situation. Gleichzeitig verringerte sich aber das Vertrauen in die eigenen fahrerischen Fähigkeiten im Vergleich zu einer Bedingung, in der Fahrfehler nicht rückgemeldet wurden. Wurden Fahrfehler nur in einem Video, d.h. durch Dritte, vorgeführt, zeigte sich in ähnlichen Situationen nur eine schwache Überlegenheit gegenüber einer fehlerfreien Vorführung und keine Effekte auf den Transfer; dabei hatte keine der Bedingungen Einfluss auf die Einschätzung der eigenen Fähigkeit.

Das Lernen anhand der Fehler Dritter stellt dabei eine mögliche Operationalisierung des advokatorischen Lernens aus Fehlern dar (Oser & Spychiger, 2005; Oser, 2007). Dabei wird negatives Wissen aufgebaut, indem typische Fehler in einer bestimmten Domäne anhand von Beispielen fehlerhafter Handlungen Anderer vorgestellt und von den Lernenden lediglich nachvollzogen werden. Da Lernende dabei keine eigenen Fehler machen, stellt dieser Ansatz konsistent mit den Ergebnissen von Ivancic und Hesketh (2000) keine Bedrohung des Selbstkonzepts dar. Der zentrale Lernmechanismus ist dabei die Kontrastierung der Fehler an korrekten Konzepten (Oser & Spychiger, 2005). Um aus Fehlern Anderer zu lernen, müssen Lernende sich mit den Protagonisten identifizieren, damit die bei VanLehn et al. (2003) beschriebenen Selbsterklärungsprozesse in Gang gesetzt werden. Hierzu muss der Protagonist möglichst authentisch sein, auch muss der Fehler in einem für die Lernenden relevanten Kontext stattfinden (Oser & Spychiger, ebd.). Dieser Ansatz wurde in zahlreichen Studien im Bereich der Mathematik (z.B. Durkin & Rittle-Johnson, 2012; Große & Renkl, 2004; 2007; Rach, Ufer, & Heinze, 2012) und der Medizin (z.B. Stark, Kopp & Fischer, 2009; Stark et al.,

2011) erfolgreich eingesetzt. Auch in der Lehrerbildung gibt es erste Studien zu diesem Ansatz (Heitzmann, 2014; Heitzmann, Fischer & Fischer, 2013; Klein et al., 2012).

4.6. Bedingungsfaktoren beim Lernen aus Fehlern

Neben den schon in den Abschnitten 4.3 und 4.5 genannten grundlegenden Bedingungen des Erwerbs negativen Wissens haben weitere Faktoren Einfluss darauf, ob Fehler tatsächlich als Lerngelegenheiten genutzt werden. Oser und Spychiger (2005) weisen in diesem Zusammenhang etwa auf die Emotionalität des Lernens aus Fehlern hin. Grundsätzlich versuchen Lernende eher, Fehler zu vermeiden, da diese häufig mit negativen Reaktionen des Umfeldes oder negativen Konsequenzen und damit auch mit einem negativen Affekt verbunden sind (z.B. Rohrbach, 1999; Zink, 2002). Dabei wird insbesondere der Moment der Bewusstwerdung des Fehlermachens hervorgehoben, der zu person- und situationsabhängig unterschiedlichen Reaktionen führen kann (z.B. Unsicherheit, Entrüstung) und auch die Verarbeitungstiefe des erlebten Falschen beeinflusst. Intensive Enttäuschung oder Beschämung werden noch lange erinnert (Oser & Spychiger, 2005) und können in extremen Fällen zur Verweigerung oder ungeeigneten, unvorteilhaften Denkmustern in ähnlichen Situationen führen. Ein Beispiel stellen die aus der Forschung zur Leistungsängstlichkeit bekannten *worry-Kognitionen* dar (Rost & Schermer, 2006).

Negative Emotionen können jedoch beim Erwerb negativen Wissens sehr nützlich sein, da sie eine Verankerung im episodischen Gedächtnis ermöglichen. Oser und Spychiger bezeichnen das Emotionale als „Garant für Schutzwissen“ (2005, S.74). Eine Studie von Klein et al. (2007) stützt diese Argumentation. Die Autoren wiesen mit Hilfe neurologischer bildgebender Verfahren einen Zusammenhang zwischen der Dopaminausschüttung als Reaktion auf negatives Feedback und der Lernleistung in Aufgaben zur Wahrscheinlichkeitsrechnung nach. Lernende mit geringerer Dopaminausschüttung zeigten eine schwächere emotionale Reaktion und waren weniger effizient darin, Handlungen mit negativen Konsequenzen zu vermeiden.

Emotionen werden durch die eigene oder die Reaktion Dritter auf Fehler beeinflusst. Oser und Spychiger (ebd.) unterscheiden hier zwischen mehr oder weniger geeigneten Handlungsweisen in Form „produktiver“ und „destruktiver Beschämer“ (S.74.). *Produktive Beschämer* entstehen in der Person selbst, die sich über einen Fehler ärgert, oder Trauer oder Scham empfindet und es deshalb beim nächsten Mal besser machen möchte (nicht von ungefähr ist hier eine begriffliche Verwandtschaft zum *Productive Failure* festzustellen). Diese Reaktion kann auf das eigene Entdecken eines Fehlers zurückgehen oder auf den Hinweis durch Dritte, wichtig ist dabei jedoch eine grundsätzlich akzeptierende

Basisbeziehung, die die Person, die den Fehler macht, nicht bloßstellt. Eine solche Haltung ist in Ausbildungs- oder Trainingskontexten (wie etwa in einem Fahrtraining bei der Studie von Ivancic & Hesketh, 2000) üblicherweise gegeben, im beruflichen Kontext können jedoch auch weniger geeignete Reaktionen auftreten (vgl. Meurier, Vincent & Parmar, 2008)

Entsprechend stellen Bloßstellung, Zynismus, Vorwürfe oder Zornausbrüche als Beispiele für *destruktive Beschämer* einen nicht adäquaten Umgang mit Fehlern Dritter dar, der den Aufbau von Schutzwissen blockieren kann (Oser & Spychiger, 2005). Eine solche Reaktion auf Fehler ist historisch durchaus verbreitet (s. Abschnitt 4.2), führt aber dazu, dass der Zugang zu negativem Wissen durch emotionale Befindlichkeiten erschwert wird. Stattdessen werden negative emotionale Merkmale der Situation enkodiert, die aber nicht das zu erwerbende Wissen, sondern etwa die beteiligten und als negativ erlebten Personen betreffen. Dies kann zu einer Vermeidungshaltung gegenüber ähnlichen Personen oder Situationen führen, gleichzeitig kann das negative Wissen keine Schutzfunktion erwerben. Oser und Spychiger (ebd.) zitieren dazu Studien zu Lernproblemen, Schulangst und anderen Problemen im Berufsleben, basierend auf ablehnenden, inadäquaten Reaktionen auf Fehler (Mahler-Jeckelmann, 1995; Schohaus, 1933). Eine Studie von Meurier et al. (2008) zum Umgang mit Fehlern bei medizinischem Betreuungspersonal zeigte, dass bei einer Reaktion auf Fehler im Sinne einer „produktiven Beschämung“ durch Verantwortliche und Vorgesetzte die damit verbundene Verantwortungsübernahme und planvolles Problemlösen die Fehlerhäufigkeit verringerte. Waren diese Voraussetzungen nicht gegeben, wurde eher eine Tendenz berichtet, Fehler nicht einzugestehen. Aus Fehlern wird dann nicht mehr gelernt, sondern sie werden verdrängt, abgewehrt oder sie werden als Schwäche gedeutet, wodurch „ein Schatz an Lernmöglichkeiten“ (Oser & Spychiger, 2005, S. 76) verlorengeht. Dementsprechend bezeichnen die Autoren es als einen Standard des (Lehrer)Handelns, mit Fehlern so umzugehen, dass aus ihnen gelernt werden kann.

Bestehen solche Vorerfahrungen bei Lernenden, kann beim Lernen aus Fehlern, insbesondere wenn dies ausdrücklich als Teil des Lernprozesses erwünscht ist und daher etwa in einer Instruktion explizit genannt wird, die dadurch salient werdende Erwartung des Fehlermachens starke negative Emotionen und im Extremfall eine Vermeidungshaltung auslösen (Oser & Spychiger, 2005). Sehen Lernende ihr Selbstkonzept durch eigene Fehler oder die befürchtete Reaktion Anderer darauf bedroht, wird der Lernprozess aus diesen Fehlern zumindest erschwert (Ivancic & Hesketh, 2000; Meurier et al., 2008). Beispielsweise kann es in einem solchen Fall zur Rechtfertigung kommen, wenn das eigene Handeln aufgrund situationaler Merkmale fälschlicherweise doch als geeignet empfunden wird. Auch die gegenteilige

Konsequenz des Erlebens eines eigenen völligen Scheiterns ist im Lernprozess aus motivationalen Gründen nicht erwünscht (nichtsdestotrotz jedoch Bestandteil einiger Ansätze wie dem *Productive Failure* (Kapur, 2008) oder dem *impasse-driven learning* (VanLehn, 1999)). Das advokatorische Lernen aus Fehlern (Oser, 2007) dagegen senkt durch die Distanzierung des eigenen Selbst von der Person, die im Beispiel einen Fehler begeht, die Hemmschwelle beim Nachvollzug des Fehlers und ermöglicht, negatives Wissen wertfrei aufzubauen.

Der Einfluss metakognitiver Fähigkeiten auf das Lernen aus Fehlern wurde in einer Studie von Mathan und Koedinger (2010) untersucht. Metakognition beim Lernen aus Fehlern umfasst die durchgehende Überwachung der eigenen Fähigkeiten Lernender, Fehler zu entdecken und zu korrigieren, aber auch wie Lernende über diese Fehler reflektieren und aus ihnen lernen. Daraus leiten sie die Notwendigkeit eines breiteren Feedbacks ab, als es in zahlreichen Studien vorgesehen ist (üblicherweise *Feedback of correct response* oder *elaboriertes Feedback*, z.B. Jacobs, 2010; Klopp, Stark, Kopp, & Fischer, 2013; Stark et al., 2011). Die Autoren postulieren, dass adäquates Feedback auch inkorrekte, aber nachvollziehbare Handlungen umfassen sollte, um die o.g. metakognitiven Prozesse (die im Prinzip einer Selbsterklärung entsprechen, vgl. Chi, Bassok, Lewis, Reimann & Glaser, 1989) anzuregen. Dies soll Lernende nicht nur auf Situationen vorbereiten, in denen Expertenwissen nicht anwendbar ist oder nicht ausreicht, sondern durch das Nachdenken über Ursachen und Konsequenzen von Fehlern einen besseren Überblick über das Handeln in einer Domäne verschaffen.

Die Selbsterklärung als ein postulierter Wirkmechanismus des Lernens aus Fehlern wurde in einer Laut-Denk-Studie (Stark & Fischer, 2008) überprüft. Die Ergebnisse weisen darauf hin, dass instruktionale Fehler zumindest in Verbindung mit elaboriertem Feedback die Tiefe der Beispielelaboration (Stark, 2001) und damit die Lernwirksamkeit der Informationsverarbeitung unterstützen (vgl. VanLehn, 1999). Welche Prozesse durch Fehler im Einzelnen unterstützt werden und wie sich unterschiedliche Arten von Fehlern auf der Prozessebene auswirken, wurde bislang noch nicht untersucht. In der o.g. Studie finden sich zudem Anhaltspunkte dafür, dass die Lernenden vom elaborierten Feedback gegen Ende der Lernphase nicht mehr profitierten, es lassen sich im Gegenteil sogar ungünstige kognitive und motivationale Effekte erkennen (Stark & Fischer, 2008). Dies könnte darauf zurückzuführen sein, dass die Bereitstellung elaborierten Feedbacks über die gesamte Lernphase eine eher passive, mit fehlerbasiertem Lernen inkompatible Lernhaltung unterstützt.

Die hier dargestellten Studien und Modelle bieten Anhaltspunkte für eine adäquate Umsetzung des Lernens aus Fehlern im Bereich der Lehrerbildung. Wichtig ist dabei die Ähnlichkeit von Lern- und Anwendungssituation sowie beim Lernen aus fremden Fehlern eine Rollenübernahme. Dies erfordert authentische Problemstellungen und für die Lernenden relevante Aufgabenstellungen. Um sicherzustellen, dass der Fehler verstanden wurde und gleichzeitig den Erwerb positiven Wissens zu sichern, muss jedoch, insbesondere beim Lernen aus eigenen Fehlern, instruktional sichergestellt werden, dass die korrekte Lösung oder ein Lösungsansatz präsentiert werden. Eine detaillierte Fehleranalyse kann sicherstellen, dass verstanden wurde, was warum in der gegebenen Situation falsch war.

Das folgende Kapitel beschäftigt sich auf der Grundlage der hier diskutierten Prinzipien und Wirkmechanismen des Lernens aus Fehlern mit möglichen Umsetzungen. Im Vordergrund steht dabei eine ressourcenschonende, ohne großen Aufwand in vorhandene Curricula implementierbare Konzeption.

5. Umsetzung des Lernens aus Fehlern anhand integrierter Lernumgebungen

Das Lernen aus Fehlern setzt direkt bei den konkreten Problemen Lehramtsstudierender bei der Anwendung bildungswissenschaftlichen Wissens an (Kap. 2.2). Diese hängen zu einem großen Teil damit zusammen, dass das vorhandene konzeptuelle Wissen aus verschiedenen Gründen nicht *anwendbar* wird, also *träge* bleibt (Gruber & Renkl, 2000; vgl. Kap. 3.2.2). In Anlehnung an Renkl (2006, s. Kap. 3.2.2) lässt sich aus der Situiertheitserklärung ableiten, dass die Lernsituation möglichst ähnlich zur Anwendungssituation gestaltet werden sollte, um Wissenskompartimentalisierung zu verhindern und gleichzeitig Wissen über die Anwendungsbedingungen zu vermitteln (vgl. strategisches Wissen; Kap. 3.2.1).

Auch in den Standards der Lehrerbildung (KMK, 2004; 2014) wird gefordert, konzeptuelles Wissen in einen authentischen Kontext zu setzen, um die Lücke zwischen Theorie und Praxis zu schließen. Insbesondere die fall- und problembasierten Ansätze setzen direkt an den in Abschnitt 2.3 erwähnten Ursachen für die Probleme Lehramtsstudierender bei der Anwendung wissenschaftlichen Wissens an (z.B. fehlende Anknüpfung an Vorwissen, kein Bezug zum späteren Beruf, vgl. Driver et al., 2000; Lind, 2001).

Zur Umsetzung des Wissenserwerbs anhand komplexer, realitätsnaher und authentischer Problemstellungen bieten sich Ansätze aus der *situated cognition*-Forschung (Reinmann & Mandl, 2006; z.B. *Anchored Instruction*, CTGV, 1992) an. Zur Sicherung des Lernerfolgs muss aber auch eine geeignete instruktionale Unterstützung erfolgen, da offene Problemstellungen Lernende oft überfordern (eine ausführliche Diskussion dazu findet sich etwa bei Kirschner et al., 2006; s. auch Sweller, 2005). Authentizität und instruktionale Unterstützung stellen aber auch beim Lernen aus Fehlern, unabhängig davon, ob es um Lernen aus eigenen oder fremden Fehlern geht, unabdingbare Voraussetzungen für eine adäquate Umsetzung dar (s. Kapitel 4.5) Hier laufen also die Erfordernisse der unterschiedlichen Lösungsansätze zusammen.

5.1 Begriffsbestimmung: Problem- und Instruktionsorientierung

Vorab werden die verwendeten Begriffe geklärt, da diese in der Literatur nicht einheitlich benutzt werden. Die inkonsistente Verwendung der Begriffe instruktions- und beispilsbasiertes Lernen im Gegensatz zu fall- oder problembasierten oder auch situierten Ansätzen oder Lernumgebungen, die oft auch synonym gebraucht oder vermischt werden (s.

Reinmann & Mandl, 2006) erschwert auch die Einschätzung der Befundlage zur Lernwirksamkeit der jeweiligen Konzeption. Diese ist uneinheitlich, zahlreiche Studien belegen die jeweilige Überlegenheit des einen Ansatzes gegenüber dem anderen (z.B. Anderson, Reder & Simon, 1996; CTGV; 1997; Fölling-Albers et al., 2004; Hmelo-Silver, Duncan & Chinn, 2007; Kirschner et al., 2006; Köller, 2008; Krause, Stark & Herzmann, 2011). Mittlerweile finden sich verschiedene Metaanalysen, die versuchen, Ordnung im Chaos der unterschiedlichen Operationalisierungen und Untersuchungen zu schaffen (z.B. Albanese & Mitchell, 1993; Berkson, 1993; Dochy, Segers, van den Bossche & Gijbels, 2003; Schmidt, Rotgans & Yew, 2011; Walker & Leary, 2009). Dazu kommt, dass viele Studien ohnehin Mischformen problem- und instruktionsorientierter Ansätze verwenden. So wird beispielsweise selten eine offene Problemstellung ohne minimale Instruktion gegeben, ebenso bedient sich Instruktion häufig anschaulicher Beispiele und verbleibt nicht auf einem abstrakten Level. „Puristische“ Formen instruktions- oder problemorientierten Lernens sind also eher die Ausnahme als die Regel (Stark et al., 2010).

In der vorliegenden Arbeit werden zur klaren Unterscheidung nur die Begriffe *Instruktionsorientierung* und *Problemorientierung* verwendet. Instruktionsorientierung bezieht sich auf Ansätze, denen eine kognitivistisch-technologische Perspektive auf Lehren und Lernen zugrundeliegt. Problemorientierung bezieht sich auf konstruktivistisch fundierte Ansätze. Wird von integrierten Lernumgebungen gesprochen, so bezieht sich der Begriff explizit auf eine systematische Kombination der beiden o.g. Ansätze. Die zugrundeliegende Definition des Begriffs Lernumgebung von Reinmann und Mandl (2006) ist dabei sehr allgemein gehalten, um alle möglichen Lernarrangements einzuschließen:

„Eine durch Unterricht hergestellte Lernumgebung besteht aus einem Arrangement von Unterrichtsmethoden, Unterrichtstechniken, Lernmaterialien, Medien. Dieses Arrangement ist durch die besondere Qualität der aktuellen Lernsituation in zeitlicher, räumlicher und sozialer Hinsicht charakterisiert und schließt letztlich auch den jeweiligen kulturellen Kontext mit ein.“ (S. 615f).

Diese sehr umfassende Definition kann durch Wollring (2009) ergänzt werden, um sie für die Operationalisierung in der vorliegenden Arbeit zielführender zu machen:

„Eine Lernumgebung ist im gewissen Sinne eine natürliche Erweiterung dessen, was man im Mathematikunterricht traditionell eine ‚gute Aufgabe‘ nennt. [...] Sie besteht aus einem *Netzwerk kleinerer Aufgaben, die durch bestimmte Leitgedanken zusammen gebunden werden.*“ (S. 13, Hervorhebungen im Original)

Eine Lernumgebung umfasst also ein Arrangement didaktischer Prinzipien (Methoden, Techniken, Medien), in dem Lernende unter einem bestimmten Leitgedanken ausgesuchte Aufgaben unter bestimmten zeitlichen, räumlichen und sozialen Bedingungen bearbeiten. Dabei kann die Ausgestaltung der Aufgaben stärker problem- oder stärker instruktionsorientiert sein. In diesem Kapitel werden daher zunächst Ansätze zum problemorientierten und instruktionsorientierten Lernen vorgestellt. Anschließend erfolgen die Synthese der beiden Ausrichtungen und ein Überblick über Studien zu den daraus resultierenden „integrierten Lernumgebungen“ und ihren möglichen Formen.

5.2 Problemorientierung

Problemorientierung basiert auf der aktiven Konstruktion von Wissen. Indem Wissen in praktisch bedeutsamen Kontexten erworben wird, soll träges Wissen (Gruber & Renkl, 2000) vermieden werden. Zudem sollen Lernende zu Aktivität und Eigenverantwortung motiviert werden (vgl. Reinmann & Mandl, 2006). Ziel problemorientierten Lernens ist es, Lernenden anhand authentischer und für sie relevanter Problemstellungen den Erwerb handlungsnahen, anwendbaren Wissens zu ermöglichen (z.B. Hmelo-Silver et al., 2007). Lernen erfolgt situationsgebunden, wobei die häufig kooperative aktive Auseinandersetzung mit authentischen und relevanten Problemstellungen von zentraler Bedeutung ist (z.B. Gräsel, 1997; Gräsel & Mandl, 1999). Da Wissen aus einer moderaten konstruktivistischen Perspektive von Lernenden aktiv konstruiert wird, wird es von Vorwissen, Erfahrungen und Überzeugungen Lernender beeinflusst und ist somit kein Abbild des Wissens des Lehrenden (Krause, 2007). Auf dieser Grundlage gibt es verschiedene Ansätze zur Umsetzung problemorientierten Lernens.

5.2.1 Problemorientierte Ansätze

Die *Situated-Cognition-Bewegung* (z.B. Resnick, 1987) geht davon aus, dass Lernen als aktiver, konstruierender Prozess in einem bestimmten Kontext erfolgt und damit situiert ist. Das bedeutet, dass das Gelernte an die Inhalte der Situation gebunden ist (vgl. Situiertheitserklärung, Kap. 2). Daraus ergibt sich, dass Lernsituationen möglichst realitätsnah gestaltet sein müssen, um die spätere Anwendbarkeit des erworbenen Wissens sicherzustellen. Hauptaufgabe der Lehrperson ist die Konzeption entsprechender Lernsituationen. Lernende arbeiten oft allein oder in Gruppen, die Lehrperson hat meist unterstützende Funktion und stellt bei Bedarf Problemlösestrategien zur Verfügung (vgl. Dochy et al., 2003; Hmelo, 1998; Hmelo & Lin, 2000; Hmelo-Silver et al., 2007; Leinhardt, 1993; Schmidt, Loyens, Van Gog &

Paas, 2007). Hier geht es weniger um das Ergebnis, sondern vielmehr um den Prozess der Wissenskonstruktion. Zusätzlich wird oft die Beteiligung der Lernenden an der Evaluation des Lernprozesses gefordert (Jones, 1992), was dem Bedürfnis Lernender nach Autonomie und Selbstbestimmung (Deci & Ryan, 1985; 1993) entgegenkommt. Ein solches Format ermöglicht die aktive Auseinandersetzung der Lernenden mit den präsentierten Inhalten und fördert so auch motivationale Aspekte (s. a. Reinmann & Mandl, 2006). Problemorientiertes Lernen eignet sich vor allem zum Erwerb *anwendbaren, komplexen* Wissens, da es die Entwicklung einer kontextsensitiven Wissensbasis fördern kann, die die Anforderungen komplexer Situationen erfüllt (Wagner, Stark, Krause, Gutenberg & Klein, 2013).

Die Problemstellungen des problemorientierten Lernens sind häufig in ein narratives Format eingebettet. Den Ansätzen liegt die gemeinsame Annahme zugrunde, dass anhand bedeutungshaltiger, authentischer Probleme gelernt werden sollte, sie nähern sich diesem Anspruch aber aus verschiedenen Perspektiven. Der Situated Cognition sind dabei u.a. die *Cognitive Flexibility-Theorie* (Spiro, Collins & Thota, 2003; Spiro & Jehng, 1990), der *Anchored Instruction-Ansatz* (CTGV, 1992; 1993) und der *Cognitive Apprenticeship-Ansatz* (Collins, Brown & Newmann, 1989) zuzuordnen (Reinmann & Mandl, 2006).

Die Cognitive Flexibility-Theorie und die Anchored Instruction (CTGV, 1992, 1993) sehen vor, eine reale Anwendungssituation in ihrer ganzen Komplexität abzubilden (Spiro & Jehng, 1990). Durch *Landscape Criss-crossing* (ebd.) wird dasselbe Konzept unter multiplen Perspektiven und Kontexten betrachtet, um multiperspektivisches Lernen zu ermöglichen und das erworbene Wissen flexibel anzuwenden. Zentrales didaktisches Prinzip der Anchored Instruction (CTGV, 1992; 1993) ist der *narrative Anker*. In fiktiven Rahmengeschichten helfen Lernende einem Protagonisten beim Lösen verschiedene Probleme, z.B. in den videobasierten Jasper-Woodbury-Stories der CGTV (ebd.). Alle zur Lösung notwendigen Informationen sind in die Geschichte eingebunden, trotzdem muss kontextspezifisches Vorwissen wie mathematische Formeln zur Problemlösung aktiviert werden. Dies soll träges Wissen verhindern (Renkl, 1996).

Der Cognitive Apprenticeship-Ansatz basiert auf der Strukturierung der Komplexität passend zur Fähigkeitsstufe der Lernenden. Zentrales Designprinzip ist die Einführung Lernender in eine Expertenkultur über soziale Interaktionen im Rahmen authentischer Lernaktivitäten (Collins et al., 1989). Diese umfassen eine praxisnahe Anleitung in realitätsnahen Problemstellungen, wobei die Aufgabenschwierigkeit ansteigt. Auch hier werden unterschiedliche Kontexte und Perspektiven präsentiert, letzere kommen jedoch anhand eines im Vergleich zu den anderen Ansätzen stark strukturierten Austauschs mit Experten dazu.

Den Einstieg bildet das kognitive Modelling, in dem die Lehrperson ihr Vorgehen zur Problemlösung modellhaft expliziert. In den folgenden Phasen des coaching, scaffolding und fading wird die Unterstützung durch die Lehrperson schrittweise verringert, was die Lernenden zur Artikulation (articulation) und Reflexion (reflection) ihrer Denkprozesse und Problemlösestrategien im Austausch mit Experten führen soll. In der abschließenden Phase der Exploration lösen Lernende selbständig Probleme und gelangen so zu mehr Autonomie und Selbstbestimmung als beim rezeptiven Lernen (Deci & Ryan, 1985; 1993).

5.2.2 Kritik am problemorientierten Lernen

In Bezug auf anwendbares Wissen werden problemorientierte Ansätze im Allgemeinen als überlegen gegenüber instruktionsorientierten Ansätzen betrachtet (zusammenfassend Dochy et al., 2003; Walker & Leary, 2009). Kritik an den Ansätzen konzentriert sich auf Schwierigkeiten beim Erwerb konzeptuellen Wissens (Albanese & Mitchell, 1993; Strobel & Barnefeld, 2009), insbesondere bei Lernenden mit geringem Vorwissen (Kirschner et al., 2006; Köller, 2008). Problemlöseaufgaben können aufgrund der hohen Anforderungen durch das offene Aufgabenformat insbesondere bei Lernenden mit geringem Vorwissen zum sogenannten *Cognitive Overload* (Mayer, 2005) führen. Dabei übersteigen die Anforderungen der Aufgabe aufgrund der fehlenden Passung von Problemlöseaufgaben mit der kognitiven Struktur des Menschen die kognitive Kapazität des Arbeitsgedächtnisses beim Lernenden, wodurch Verstehensprozesse behindert werden (vgl. Kirschner et al., 2006; Mayer, 2006; Sweller et al. 1998). Insbesondere bei ungünstigen Lernvoraussetzungen und geringem Vorwissen können daher instruktionsorientierte Ansätze zur Vermittlung konzeptuellen Wissens besser geeignet sein.

5.3 Instruktionsorientierung

Instruktionsorientierte Ansätze des Lehrens und Lernens beschäftigen sich mit der Optimierung der Instruktion, wie also Unterricht geplant, gesteuert und organisiert werden muss, damit Lernende Inhalte verstehen können. Grundlage ist der Kognitivismus, Ziel ist eine gegenstandszentrierte Lernumgebung, deren Struktur Lernende beim Wissenserwerb folgen (z.B Reinmann & Mandl, 2006). Im Gegensatz zur Problemorientierung erfolgt der Wissenserwerb also rezeptiv, Wissen wird als vermittelbar betrachtet.

5.3.1 Instruktionsorientierte Ansätze

Beim stark technologisch geprägten Ansatz des *Instructional Design* werden alle Aspekte des Lehrens und Lernens systematisch geplant und durchgestaltet. Entsprechende Instruktionspläne geben vor, unter welchen Bedingungen welche Instruktionsstrategien oder Lehrmethoden eingesetzt werden (z.B. Lowyck, 1991; Reigeluth & Stein, 1983). Typisch ist die Präsentation umfassender Analysen von Anfangs- und Endzuständen sowie der dazwischenliegenden Übergangsphasen. Verschiedene Ansätze sehen auch eine prozessbegleitende Anpassung der Lerninhalte an den Wissensstand Lernender vor (z.B. *Programmierter Unterricht*, Skinner, 1958). Lernende erhalten systematisch Informationen darüber, wie Übergänge vom Anfangs- zum Endzustand zu bewältigen sind (z.B. Snow, 1989, auch *ausgearbeitete Lösungsbeispiele*, Renkl, 1997; Stark, 1999, 2001). Dabei ist keine eigene Strukturierung des Stoffes erforderlich, Lernen ist also ein vorwiegend rezeptiver Prozess. Der Lehrende überwacht als *didactic leader* den Wissensfortschritt (z.B. Krause, et al., 2011). Der Wissenserwerb selbst wird als regelhaft ablaufender, steuerbarer Prozess der Informationsverarbeitung betrachtet, der anhand systematisierter didaktischer Konzepte (Anleiten, Darbieten, Erklären) strukturiert und organisiert wird (Lehner, 1979).

Reinmann und Mandl (2006) nennen als typische Ansätze auf Basis des Instructional Design u.a. das *Expository teaching* (Ausubel, 1974), das *Modell kumulativen Lernens* nach Gagné (1973) und die *Elaborationstheorie* (Reigeluth, 1979). Das Modell kumulativen Lernens ist eng verwandt mit dem *Programmierten Unterricht* (Skinner, 1958) und verbindet Elemente des Kognitivismus und des Behaviorismus, die in einer hierarchischen Sequenz von Lernstufen angeordnet werden (Gagné, 1973). Die Bewältigung einer Lernstufe ist Voraussetzung für die Bewältigung der Folgenden. Die Sequenzierung und Systematisierung der Inhalte in einer didaktisch angemessenen Hierarchie ist Aufgabe des Lehrenden. Eine Sequenz startet dabei üblicherweise mit einem Element operanter Konditionierung (Skinner, 1958), darauf folgt eine Sequenz sogenannten Diskriminationslernens, in der Lernende auf ähnliche, aber trotzdem verschiedene Reize unterschiedlich reagieren müssen. Daran schließen sich mit dem Begriffs- und Regellernen zwei Phasen zum Erwerb von Wissen über Fakten und Regeln an, die Lernende zum Problemlösen befähigen sollen. Die Reihenfolge der einzelnen Phasen kann nicht variiert werden.

Der *Expository teaching-Ansatz* (Ausubel, 1974) betrachtet die Erweiterung der kognitiven Struktur Lernender als den wichtigsten Lernfaktor. Zentrales Designprinzip ist ein dem Lernprozess vorangestellter *Advance Organizer*, der Lernenden als Strukturierungshilfe für die nachfolgenden Informationen dient. In der anschließenden Phase der progressiven

Differenzierung werden Inhalte mit ansteigendem Komplexitätsgrad sequenziert; darauf folgt in der Konsolidierungsphase die Wiederholung und Festigung des Lernstoffs durch den Lehrenden.

Auch die *Elaborationstheorie* (Reigeluth, 1979) zielt auf Aufbau und Festigung kognitiver Strukturen in einer Domäne. Unter Aktivierung des Vorwissens werden neue Inhalte verstanden und in Zusammenhang gebracht, Inhalte werden in einer Sequenz mit steigendem Anforderungsgrad angeordnet. Die Elaborationstheorie sieht die Entwicklung eines Instruktionsgerüsts vor, das die Einordnung von Inhalten unterschiedlicher Elaborationsniveaus zu passenden Unterrichtseinheiten erlaubt. Die Ähnlichkeit des Ansatzes mit Ausubels (1974) Advance Organizer ist vorhanden, dieser dient allerdings in erster Linie zur Unterstützung Lernender, während das Instruktionsgerüst bei der Planung, also auf Seiten des Lehrenden, ansetzt.

Beim Lernen mit ausgearbeiteten Lösungsbeispielen (z.B. Renkl, 1997) wird Lernenden eine Problemstellung zusammen mit der Lösung und allen erforderlichen Zwischenschritten präsentiert. Die Aufgabe der Lernenden besteht im Nachvollzug der Lösungsschritte. Analog zum Lernen aus Fehlern hängt der Lernerfolg von der Qualität der Selbsterklärungen Lernender ab, zusätzlich spielt jedoch auch die Qualität der *Beispielelaboration* (Stark, 2001) eine wichtige Rolle. Diese kann durch instruktionale Maßnahmen wie Prompts (z.B. Stark et al., 2011) unterstützt werden, um ein Scheinverständnis durch die weitgehend passive Rezeption der Beispiele zu verhindern. Insbesondere bei Lernenden mit geringem Vorwissen konnten Vorteile dieser Methode gegenüber Problemlöseaufgaben nachgewiesen werden (*worked example-effect*, Clark & Mayer, 2003).

5.3.2 Kritik am instruktionsorientierten Lernen

Instruktionsorientierte Ansätze haben sich in Studien insbesondere für den Erwerb konzeptuellen Wissens und beim initialen Wissenserwerb als effektiv erwiesen (Albenese & Mitchell, 1993; Dochy et al., 2003; Krause et al., 2011; Stark, Mandl, Gruber & Renkl, 2000). Die Kritik am Instructional Design setzt bei der strikten Strukturierung der Lerninhalte an, die selbstgesteuerte Lernvorgänge ausschließt und zusammenhängende Teile eines Inhaltsgebiets getrennt vermittelt (z.B. Winn, 1993). Die potentiell resultierende Wissenskompartimentalisierung kann träges Wissen fördern (vgl. Kap. 3.2.2). Neben der Frage, inwieweit die Strukturierung und Reduktion von Inhalten eine komplexe Realität abbilden kann (Resnick, 1987), kann die Wirkung einzelner Instruktionsmaßnahmen und -phasen nicht explizit vorhergesagt werden (Duffy & Jonassen, 1991). Hier spielen Merkmale

Lernender eine wichtige Rolle (im Sinne einer *Aptitude-Treatment-Interaktion*; Cronbach & Snow, 1977). Die passive Rezeption des Lernstoffs birgt die Gefahr eines Scheinverständnisses, das zu einer oberflächlichen Rezeption des Lerninhalts führen kann (*Verstehensillusion*, Stark, 2001), zudem kann es durch die stark asymmetrische Rollenverteilung zur Reduktion von Eigeninitiative und Selbstverantwortung und damit der intrinsischen Motivation kommen (Deci & Ryan, 1993). Studien von Kollar und Fischer (2013) und Wecker (2012) weisen darauf hin, dass ein Übermaß an Instruktion Lernende ebenfalls überfordern bzw. bei steigender Kompetenz mit eigenen Problemlösestrategien Lernender interferieren kann (vgl. auch *Expertise Reversal-Effect*; Kalyuga, Rikers & Paas, 2012).

5.4 Synthese: Designprinzipien integrierter Lernumgebungen

Aus den dargestellten Überlegungen lässt sich ableiten, dass verschiedene Ansätze zum Lehren und Lernen immer auch in Abhängigkeit von der *Art* des zu vermittelnden Wissens wirksam sind: Während instruktionsorientiertes Vorgehen eher den Erwerb *konzeptuellen* Wissens unterstützt (vgl. z.B. Dochy et al., 2003; Albanese & Mitchell, 1993), scheint problemorientiertes Lernen vor allem den Erwerb *anwendbaren, komplexen* Wissens zu fördern (vgl. Dochy et al., 2003; Hmelo, 1998; Hmelo & Lin, 2000; Hmelo-Silver et al., 2007; Schmidt et al., 2007). Obwohl die Ansätze in der Literatur häufig diametral entgegengesetzt werden, schließen sich ihre Wirkmechanismen keineswegs gegenseitig aus. Im Gegenteil können Prinzipien eines Ansatzes sehr gut genutzt werden, um die Schwächen des jeweils anderen auszugleichen, da Instruktion und Konstruktion ohnehin meist zeitgleich stattfinden und eng miteinander verknüpft sind (Shuell, 1993). Aus dieser Perspektive kann die jeweilige Problem - oder Instruktionsorientierung durchaus wörtlich genommen und die Ansätze auf einem Kontinuum verortet werden, das je nach didaktischer Ausgestaltung oder den Bedürfnissen Lernender näher an einem der beiden Pole liegt, also eher Instruktion oder Konstruktion betont.

Hier setzen sogenannte integrierte Lernumgebungen (Reinmann & Mandl, 2006) an. Sie zeichnen sich durch eine systematische Integration problem- und instruktionsorientierter Designprinzipien aus. Dies unterstützt Lernende beim Erwerb einer Wissensbasis, die konzeptuelles und prozedurales/strategisches Wissen so verknüpft, dass eine Wissensanwendung ermöglicht wird. Reinmann und Mandl (ebd.) formulieren in Ergänzung der Prinzipien der Cognitive Flexibility, Anchored Instruction und des Cognitive Apprenticeship mit instruktionaler Unterstützung fünf Designprinzipien für integrierte

Lernumgebungen. Demzufolge sollten die *Eigenschaften von Problemstellungen* einen Realitätsbezug umfassen, sie müssen also komplex und authentisch sein; außerdem aktuell, um für die Lernenden relevant zu sein und sie neugierig zu machen. Um erworbenes Wissen transferier- und damit anwendbar zu machen, sollte außerdem in *multiplen Kontexten* und unter *multiplen Perspektiven* gelernt werden. Letzteres lässt sich sehr gut mit der Forderung nach Lernen in *sozialen Kontexten* kombinieren, hier soll die Kooperation der Lernenden untereinander und mit der Lehrperson gefördert werden (vgl. *Enkulturation*, Kap. 5.2.2). Alle diese Prinzipien werden schließlich systematisch mit *instrukionaler Unterstützung* kombiniert, um die Überforderung der Lernenden durch die komplexen, offenen Problemstellungen zu vermeiden.

Wird eine Lernumgebung unter diesen Gesichtspunkten als *integriert* definiert, so lassen sich tatsächlich zahlreiche Befunde zum problem- oder instruktionsorientierten Lernen als Befunde zu integrierten Lernumgebungen interpretieren. Beispiele dafür umfassen Studien von Hmelo-Silver et al. (2007), Hmelo & Lin (2000) oder Schmidt et al. (2007).

Ebenso können zahlreiche Befunde aus der Cognitive Load-Forschung (Sweller et al., 1998; z.B. Studien von Leppink, Paas, van Gog, van der Vleuten & van Merriënboer, 2014; Park, Flowerday & Brünken, 2015) und der Forschung zur *Cognitive Theory of Multimedia Learning* (Überblick z.B. in Mayer, 2005) dahingehend interpretiert werden. Diese Studien basieren häufig auf Lernumgebungen, in denen zur Induktion lernförderlicher Arten kognitiver Belastung (z.B. Seufert, Schütze & Brünken, 2009) oder zur Reduktion unerwünschter Belastung (Klein, Park, Seufert & Brünken, 2009) verschiedene Maßnahmen zur instruktionalen Unterstützung eingesetzt werden. Im Bereich der Lehrerbildung finden sich entsprechende Befunde z.B. bei Fölling-Albers et al. (2004; Anreicherung einer situierten Lernumgebung mit einem Advance Organizer), Stark et al. (2010; Vergleich integrierter Lernumgebungen mit unterschiedlicher Schwerpunktsetzung) und Wagner et al. (2013; u.a. sequentielle Anordnung problem- und instruktionsorientierter Designprinzipien). Diese Studien konnten positive Effekte auf den Erwerb verschiedener Wissensarten nachweisen, üblicherweise erzielte die systematische Integration problem- und instruktionsorientierter Designprinzipien den höchsten Lernerfolg und positive motivationale Effekte.

Die Konzeption von Unterrichtssequenzen oder Seminaren anhand der Designprinzipien integrierter Lernumgebungen (z.B. Krause et al., 2011; Stark, et al., 2010; Stark, Puhl & Krause, 2009; Wagner et al., 2013) unterscheidet sich von in sich geschlossenen Trainingsprogrammen (z.B. *WALe*, Stark & Koch, 2007; Stark, Tyroller, Krause & Mandl, 2008) vor allem durch die zeitlichen und räumlichen Anforderungen der jeweiligen

Unterrichtsform, die etwa die Auswahl der Sozialform beeinflussen. Die selbständige Bearbeitung einer didaktisch ausgeklügelten Lernumgebung durch die Lernenden ermöglicht eine detaillierte instruktionale Steuerung, gleichzeitig ist auch die Situierung anhand problemorientierter Designprinzipien gut umsetzbar. Durch die selbstgesteuerte Bearbeitung kann auch die Motivation gefördert werden. Das Lernen in sozialen Kontexten kommt dabei jedoch unter Umständen zu kurz. Diesem Problem kann etwa durch die Bearbeitung eines Trainingsprogramms in Dyaden (z.B. Wagner et al., 2015, angenommen) oder aber die Konzeption einer ganzen Unterrichtseinheit als integrierte Lernumgebung begegnet werden. In Seminarform bietet sich etwa die Arbeit in Kleingruppen (Krause et al., 2011) oder stärker strukturierter Formen der Gruppenarbeit (z.B. Gruppenpuzzle; Stark, et al., 2010) an. Kombinationen in sich geschlossener Trainingsprogramme mit einer anschließenden aufbauenden Seminarphase sind ebenfalls denkbar. Dabei können zur Strukturierung beispielsweise Ansätze zum Blended Learning (Mandl & Kopp, 2006; Steffens & Reiß, 2009) herangezogen werden.

6. Studie 1: Lernen aus eigenen vs. Lernen aus fremden Fehlern

Klein, M., Wagner, K. & Stark, R. (eingereicht). Fostering Theory Application Competences in Student Teachers by Learning from Errors: A Comparison of Productive Failure and Advocatory Error Learning.

In Studie 1 wurde vor dem Hintergrund der in Kap. 3 dargestellten Überlegungen untersucht, ob das Lernen aus eigenen Fehlern oder das Lernen aus fremden Fehlern im Kontext der Theorieanwendung in der Lehramtsausbildung zielführender ist. Dazu wurde eine integrierte Lernumgebung nach dem *Productive Failure-Ansatz* (Kapur & Bielaczyc, 2012) mit einer integrierten Lernumgebung basierend auf dem *Advokatorischen Lernen aus Fehlern* nach Oser (2007; Oser & Spychiger, 2005) hinsichtlich ihrer Lernwirksamkeit verglichen. Erhoben wurden deklaratives Wissen über Unterschiede zwischen wissenschaftlichem und Alltagswissen, die Qualität der Fehleridentifikation in einer fehlerhaften Analyse und die Qualität der Theorieartikulation, also der Anwendung wissenschaftlichen Wissens zur Erklärung einer schulischen Situation. Im Folgenden werden die zugrundeliegenden Ansätze nach der allgemeinen Betrachtung in Kap. 4.5 ausführlicher dargestellt.

6.1 Productive Failure

Der Ansatz des Productive Failure stammt aus einer Forschungstradition, die sich mit kognitiver Belastung beim Lernen (*Cognitive Strain*, hier nicht im Sinne der Cognitive Load-Theory nach Sweller et al., 1998 verstanden) und Schwierigkeiten und Hindernissen im Lernprozess beschäftigt (s.a. Kap. 4.5 zum *Constructive Failure*, Clifford, 1984; *Desirable Difficulties*, Schmidt & Bjork, 1992; *Impasse-driven Learning*, VanLehn et al., 2003). Lernende bearbeiten für sie neuartige Problemstellungen, die sie auf Grundlage ihres Vorwissens oder ihrer Intuition lösen müssen. Dies führt in den meisten Fällen dazu, dass Fehler gemacht werden und die Aufgabe nicht gelöst werden kann. Mit Bezug auf Studien von VanLehn (1999, 2003) und McNamara (2001) wird angenommen, dass diese Erfahrung die Suche nach geeigneten Lösungsansätzen auslöst, ein „Verlangen danach, zur korrekten Lösung zu gelangen und diese zu verstehen“ („owning“, Kapur & Bielaczyc, 2012). Diese Lösungsansätze bzw. die Lösung selbst müssen im folgenden Schritt präsentiert werden, um den Lernerfolg zu sichern.

Entsprechend werden beim Design von Aufgaben im Productive Failure-Ansatz zwei Phasen unterschieden: Phase 1 besteht aus komplexen Problemstellungen, die kollaborativ bearbeitet werden können, um problembezogenes Vorwissen und Problemlösestrategien zu aktivieren. In Phase 2 findet die Konsolidierung des Wissens statt, indem eine gut strukturierte instruktionale Unterstützungsmaßnahme (z.B. Problemlösung, Instruktion oder elaboriertes Feedback) präsentiert wird. Als primäre Wirkmechanismen nennt Kapur (2008; Kapur & Bielaczyc, 2012) die Aktivierung und Differenzierung des Vorwissens durch die Problemstellung sowie die durch das Scheitern provozierte Selbsterklärung in Form von Kritik, Erklärung und Elaboration. Voraussetzung ist auch hier eine für die Lernenden relevante Problemstellung - kann kein Bezug zur Lebenswelt hergestellt werden, dürfte auch das postulierte intrinsische Interesse an einer korrekten Lösung gering sein.

Dabei wird auch die Rolle instruktionaler Unterstützung diskutiert. Kapur und Rummel (2009) zitieren hier das sogenannte *Assistance Dilemma* (Koedinger & Alevan, 2007; Koedinger, Pavlik, McLaren & Alevan, 2008), das insbesondere der Frage nachgeht, *an welchem Punkt* des Lernprozesses Lernende durch Instruktion unterstützt werden sollten und wann diese zugunsten selbstgesteuerten Problemlösens zurückgehalten werden sollte (McLaren, Lim & Koedinger, 2008). Hier gehen die Positionen durchaus weit auseinander, verschiedene Forscher insbesondere aus der Cognitive Load-Forschung (z.B. Kirschner et al., 2006) befürworten zu jedem Zeitpunkt ein *hohes* Maß an Unterstützung, während andere für *minimale* Unterstützung oder Anpassung an den Lernfortschritt plädieren (z.B. Steffe & Gale, 1995; Wecker, 2012; zusammenfassend: Koedinger & Alevan, 2007). In der Forschung zu Productive Failure wird instruktionale Unterstützung als unerlässlich betrachtet (Kapur & Bielaczyc, 2012).

Productive Failure wurde in erster Linie im Kontext des Mathematik- und Naturkundeunterrichts untersucht (z.B. Kapur, 2008; 2010; 2014; Pathak, Kim, Jacobson, & Zhang, 2011). Zur Umsetzung müssen Lehrende sich einen Überblick darüber verschaffen, was Lernende bereits über ein noch nicht im Unterricht behandeltes Konzept wissen. Auf dieser Grundlage sind Aufgaben so zu gestalten, dass sie die Aktivierung und Differenzierung von Vorwissen unterstützen und Möglichkeiten bieten, Problemlösestrategien zu entwickeln, zu erforschen, zu kritisieren und zu verfeinern. Dies schließt ausdrücklich Fehlschläge und unpassende Lösungsversuche ein. Entscheidend beim Lernen aus Fehlern auf Basis von Productive Failure ist die unmittelbar auf das Scheitern folgende instruktionale Unterstützung, die die Lösung des Problems ermöglicht.

Studien zur Lernwirksamkeit derartiger Aufgaben im Physikunterricht (zur Newtonschen Kinematik) zeigten, dass Lernende anders als eine Kontrollgruppe multiple Repräsentationen eines Problems (also verschiedene Blickwinkel) und mehr Problemlösestrategien generierten. Wurden die Probleme in Gruppen bearbeitet, ergaben sich zudem komplexe Interaktionen in Form von Erklärungen, Kritik und Elaboration. Den Diskussionen mangelte es allerdings an Struktur und Stringenz, was oft eine Einigung verhinderte und zu schlechterer Gruppenperformanz führte. In Einzeltests waren die Lernenden jedoch einer Kontrollgruppe überlegen, was auf die Interaktionen in der Gruppe zurückgeführt wurde (Kapur, 2008).

Weitere Studien verglichen die Effekte einer auf Productive Failure basierenden Unterrichtskonzeption auf konzeptuelles Verständnis und Transfer mit denen direkter Instruktion im Mathematikunterricht (Durchschnittsgeschwindigkeit, Kapur, 2010; Standardabweichungen, Kapur & Bielaczyc, 2012). In der Bedingung mit direkter Instruktion erklärten die Lehrer relevante Konzepte, präsentierten ausgearbeitete Lösungsbeispiele und fungierten als Modell bei der Problemlösung. Zudem boten sie Übungsgelegenheiten und Feedback. In der Bedingung mit Productive Failure unternahmen die Lernenden eigene Lösungsversuche und erhielten anschließend instruktionale Unterstützung. Dabei zeigte sich, dass Lernende in der Konzeption mit Productive Failure Lernende mit direkter Instruktion in Bezug auf konzeptuelles Verständnis und Transfer übertrafen. Die Anzahl der durch die Lernenden selbst generierten Repräsentationen und Lösungsstrategien korrelierte signifikant mit dem Lernzuwachs. Auch Schüler, die große Unterschiede in der allgemeinen und mathematischen Begabung aufwiesen, generierten vergleichbar viele Repräsentationen (Kapur, 2010; Kapur & Bielaczyc, 2012). Ein direkter Vergleich zwischen Productive Failure und dem Lernen aus fremden Fehlern („*vicarious learning from errors*“ Kapur, 2014) erbrachte geringe Vorteile beim konzeptuellen Verständnis und Transfer zugunsten der Bedingung mit eigenen Fehlern. Lernende in der Bedingung mit *vicarious learning* bearbeiteten ausgewählte fehlerhafte Lösungsansätze anderer Schüler.

6.2 Advokatorisches Lernen aus Fehlern

Diesen positiven Effekten stehen neben der in Kap. 4.6 diskutierten Emotionalität des Lernens aus Fehlern auch andere Bedenken gegenüber. So ist etwa die Analyse der Fehler, die an anderer Stelle als entscheidend zum Verständnis und zum Erwerb korrekten Wissens betrachtet wird (Curry, 2004; Oser & Spychiger, 2005), beim Productive Failure nicht explizit vorgesehen. Laut Oser et al. (1999) erfordert das Lernen aus Fehlern das Erkennen eines Fehlers (Was ist falsch?), die Analyse des Fehlers (Warum ist das falsch?) und die Korrektur

des Fehlers anhand einer Kontrastierung am korrekten Vorgehen (Wie ist es richtig?). Erfolgt eine Fehleranalyse dennoch in Form eines Feedbacks zum Lösungsversuch eines Lernenden, ist sie mit erheblichem Aufwand verbunden, wenn sie für jeden Einzelfall erfolgt. Soll stellvertretend für Fehler verschiedener Lernender ein Feedback auf einen Fehler eines Lernenden gegeben werden, stellt sich das Problem, dass dieser Fehler möglicherweise nicht von allen Schülern gemacht wurde oder für diese nicht relevant ist (vgl. Lernprozess der Klasse, Kap. 4.1) All dies erschwert den Aufbau des im Zusammenhang mit dem Lernen aus Fehlern wichtigen *negativen Wissens* (s. Kap. 4.3).

Osers (2007) Ansatz des *advokatorischen Lernens aus Fehlern* basiert dagegen auf der Vorgabe fremder Fehler, die Lernende nachvollziehen müssen. Dies hat gegenüber eigenen Fehlern insbesondere den Vorteil, dass die Lehrperson bestimmen kann, welche Fehler als relevant betrachtet und behandelt werden; dadurch ist eine gezieltere didaktische Steuerung möglich. Um eine oberflächliche Bearbeitung der Fehler zu vermeiden, muss sichergestellt werden, dass Lernende sich mit den Protagonisten identifizieren, so dass eine Rollenübernahme und damit auch ein Verständnis *für* das Machen des jeweiligen Fehlers möglich werden (Spychiger, 2004). Eine detaillierte Fehleranalyse fördert den Aufbau negativen Wissens in einem Prozess, den Spychiger (ebd., S.3) als *umgekehrtes Modelllernen* bezeichnet. Die Kontrastierung an der korrekten Lösung sichert den Erwerb von Wissen über das korrekte Vorgehen.

In entsprechenden Untersuchungen förderte die Präsentation fehlerhafter Konzepte das Wissen um deren Korrektur und damit auch korrekte Konzepte durch die Induzierung eines kognitiven Konflikts, etwa in einem Online-Lernsystem über Dezimalzahlen (Huang, Liu, & Shiu, 2008) oder bei der Präsentation wissenschaftlicher Texte bei Lernenden mit bekannten Fehlkonzepten (Van den Broek & Kendeou, 2008). Booth, Lange, Koedinger und Newton (2013) fanden positive Effekte des Lernens aus inkorrekten Lösungsbeispielen im Bereich der Algebra. Die Ergebnisse einer Studie von Große und Renkl (2004) im Bereich Wahrscheinlichkeitsrechnung weisen darauf hin, dass vor allem Lernende mit hohem domänenspezifischem Vorwissen von der Vorgabe instruktionaler Fehler profitieren: Sie zeigten hinsichtlich weiten Transfers bessere Leistungen als Lernende mit Lösungsbeispielen ohne Fehler. Dieser Befund verdeutlicht die Relevanz einer umfassenden Vorwissensdiagnostik und konnte in einer weiteren Studie repliziert werden (Große & Renkl, 2007). Aus methodischen Gründen wurden die instruktionalen Fehler in der Studie von Große und Renkl (2007, 2004) nicht aufgeklärt. Die Lernenden mussten sich selbst erklären, was warum falsch an den Lösungen war.

Mit der Wirkung unterschiedlicher Rückmeldungen beim Lernen aus Fehlern beschäftigen sich Studien zur Förderung der Diagnosekompetenz in der Medizin. Dabei erhielten Lernende bei der Bearbeitung fallbasierter fehlerhafter Lösungsbeispiele unterschiedliche Arten von Rückmeldung auf die vorgegebenen Fehler (Klopp et al., 2013; Kopp, Stark, & Fischer, 2008; Kopp, Stark, Heitzmann & Fischer, 2009; Kopp, Stark, Kühne-Eversmann & Fischer, 2009; Stark et al., 2011). In zwei Laborstudien wurden instruktionale Fehler in Lösungsbeispielen (mit vs. ohne Fehler) und die Reichhaltigkeit des Feedbacks (elaboriert, d.h. ausführliches Feedback mit der richtigen Lösung sowie weiteren Erklärungen vs. *knowledge of correct result*, d.h. knappes Feedback mit nur wenigen Informationen über die Korrektheit des Ergebnisses) in zwei medizinischen Domänen experimentell variiert. In der ersten Laborstudie in der Domäne Bluthochdruckerkrankungen wurde keine Haupteffekte der instruktionalen Fehler und des Feedbacks nachgewiesen, es zeigte sich jedoch ein signifikanter Interaktionseffekt. Lernende, die mit instruktionalen Fehlern und knappem Feedback lernten, erwarben mit Abstand am wenigsten Diagnosekompetenz. Hingegen war die erfolgreichste Gruppe die, die mit instruktionalen Fehlern und elaboriertem Feedback arbeitete. Dies zeigte sich insbesondere bei komplexen diagnostischen Problemsequenzen, die die Anwendung handlungsnahen Wissens („strategisches und konditionales“ Wissen, entspricht in der hier verwendeten Terminologie prozeduralem und strategischem Wissen) erforderten. Die instruktionalen Effekte dieses Ansatzes erwiesen sich als nachhaltig und konnten in einer ergänzenden Studie, in der nur der Fehlerfaktor experimentell variiert wurde, beim selbstgesteuerten Lernen im Feld bestätigt werden (Kopp et. al., 2009a). In der zweiten Laborstudie (Domäne Schilddrüsenerkrankungen) zeigte sich dagegen weitgehend unabhängig davon, ob korrekte Lösungsbeispiele oder solche mit instruktionalen Fehlern vorgegeben wurden, ein Effekt des elaborierten Feedbacks (Kopp et. al., 2009b, Stark et al., 2011).

Als Erklärung für das unterschiedliche Befundmuster wurden domänenspezifische Aspekte und damit assoziierte Unterschiede in der Komplexität der Lösungsbeispiele angeführt. Eine alternative Erklärung geht von der Unterschiedlichkeit der Umsetzung instruktionaler Fehler in beiden Studien aus. Während in der ersten Laborstudie häufig Fehler bei der konkreten Vorgehensweise (strategisches bzw. prozedurales Wissen) in Kombination mit Fehlern bei der *Begründung* des Vorgehens (konditionales Wissen) präsentiert wurden, überwogen in der zweiten Laborstudie deutlich Fehler im strategischen Wissen. Die Kombination beider instruktionaler Fehlerarten erwies sich als besonders lernwirksam, vor allem in Bezug auf die Förderung konditionalen (strategischen) Wissens.

Studien zum direkten Vergleich des Lernens aus eigenen und fremden Fehlern sind selten (z.B. Ivancic & Hesketh, 2000; Kapur, 2014), im Bereich der Lehramtsausbildung sind dem Autor keine weiteren Studien bekannt. Daher wurden in der folgenden Studie 1 die beiden genannten Ansätze in Form zweier integrierter Lernumgebungen umgesetzt und auf ihre Lernwirksamkeit in Bezug auf den Erwerb positiven und negativen Wissens untersucht.

6.3 Fostering theory application competences in student teachers by learning from errors: a comparison of productive failure and advocatory error learning

Abstract

Two case-based learning environments aiming at fostering competences in the theory-guided explanation of complex pedagogical situations in teacher education were developed. A learning environment based on productive failure (PFLE) was compared to a learning environment based on advocatory errors (EBLE) regarding three aspects of theory application: metatheoretical knowledge, quality of theory articulation and quality of error identification. While both experimental conditions improved and outperformed the control group regarding quality of theory articulation, the PFLE did not foster error identification and the EBLE, while still outperforming the control group, produced significantly worse theory articulations than the PFLE. Subjective learning success and acceptance of both experimental conditions were high. Overall, the EBLE was considered more effective, since the knowledge acquired in this approach („negative knowledge”) enabled learners to avoid errors in theory-based explanations to a greater extent than the acquisition of „positive” knowledge provided in the PFLE fostered error identification.

Keywords: teacher education; theory-practice-divide; theory application; advocatory errors; productive failure

Theorizing In Practice

Scientific pedagogical knowledge as an evidence- and standard-based repertoire of pedagogical skills is a central component of several teaching competence models (Baumert & Kunter, 2006; Bransford, Darling-Hammond, & LePage, 2005; Ingvarson & Rowe, 2008; Shulmann, 1986) and a basic requirement in teacher expertise research (e.g. Bromme, 2008; Stamouli, Gruber, & Schmid, 2010; Smith, 2005). Theory-based reflection of school situations is seen as a key skill for teachers (e.g. Meier, 2006; Neuweg, 2007; Swinkels, Koopman & Beijaard, 2013).

Student teachers therefore acquire scientific knowledge concerning a variety of pedagogical and psychological topics during their university training (e.g. Bainbridge, 2011). However, how the theories taught at universities relate to and how they can be applied to school practice is rarely demonstrated (c.f. Le Cornu & Ewing, 2008). The often lamented tendency of student teachers „to reject or not engage with the [scientific] knowledge base of their chosen profession“ (e.g. Bainbridge, 2011, p. 25; Ritchie & Wilson, 2001) is at least to some extent a consequence of this current instructional practice (Neuweg, 2007). As a result, even basic applications of theory to school practice like the reflective explanation of every day school situations on the basis of pedagogical theories („retrospective understanding“, Beck & Krapp, 2006) are difficult for students (Kuhn, 2010; Stark, Herzmann, & Krause, 2010).

To utilize the potential of the students' ensuing errors in theory application as learning opportunities, two learning environments based on learning from errors were developed. These learning environments explicitly aimed at fostering student teachers' abilities to explain complex pedagogical situations based on scientific knowledge. The current study compared two different approaches to learning from errors, namely productive failure (Kapur & Bielaczyc, 2012) and advocatory errors (Oser & Spychiger, 2005) to investigate the effects of learning from own errors and learning from other persons' errors.

1. Scientific knowledge in Teacher Education

The application of scientific knowledge requires the development of distinct knowledge types (Krause, 2007). Scientific knowledge in teacher education encompasses conceptual knowledge (De Jong & Ferguson-Hessler, 1996) of theories, approaches and empirical evidence in the domains of pedagogy and psychology. It can be differentiated from conceptual „*metatheoretical knowledge*“, i.e. epistemic knowledge about the differences

between scientific and everyday knowledge. Metatheoretical knowledge is an important factor in the application of scientific knowledge, since the epistemic situation of the social sciences causes a seeming vagueness of scientific constructs and an overlap of scientific and everyday knowledge (c.f. Schneewind, 1977). Bortz and Döring (2006, p.31) elaborate on the differences between these knowledge types on several dimensions including 1) the precision of the applied terminology, 2) the way of gathering and documenting empirical evidence 3) the procedure of decision making), 4) evaluation of the validity of propositions and 5) handling of theories.

Apart from conceptual knowledge, scientific knowledge application requires action-related procedural and strategic knowledge („knowledge about appropriate actions and in which order to perform these actions“; c.f. De Jong & Ferguson-Hessler, 1996, p.7).

Ohlsson (1992) provides a suitable operationalization to teach and assess the necessary procedures with his concept of *theory articulation*: Students need to be taught „[...] the activity of applying a theory to a particular situation, to decide how, exactly, the theory should be mapped onto that situation, and to derive what the theory implies or says about that situation.“ (p.182). Theory articulation represents a specific aspect of scientific argumentation, namely explanation, i.e. using the abstract principles of a theory to account for specific phenomena (ibid). Explanations as reflections on how complex school situations have actually come to pass are regarded as the very basis of professional classroom actions („retrospective understanding“, Beck & Krapp, 2006).

In terms of knowledge types, theory articulation combines (1) domain-specific (here: pedagogical) scientific knowledge, (2) metatheoretical knowledge to differentiate between everyday and scientific knowledge and (3) procedural and strategic knowledge on how and when to perform each step of the articulation process (see fig. 1).

1.2 Errors in Theory Articulation

The process of articulating a theory with respect to a particular situation or phenomenon is a complex task, since learners have to establish and integrate distinct knowledge types. As the procedures of theory articulation are not part of the theories themselves, they are neither learned by studying their content, nor can students be expected to figure them out on their own (Ohlsson, 1992). Given that these procedures are rarely demonstrated in university lectures or text books (ibid.), frequent errors occur in the process of theory articulation.

Qualitative analyses of different academic tasks (Stark, 2005) resulted in three types of typical errors, including different subtypes and various hybrid forms: Reference to elements of non-scientific everyday knowledge (error type 1; c.f. *subjective Theories*, Groeben, Wahl, Schlee, & Scheele, 1988), suboptimal choice of scientific theories, approaches, models and concepts (error type 2; c.f. McNeill et al., 2004) and erroneous reference to empirical evidence (error type 3; c.f. Kuhn, 2010).

The frequency of these errors, however, makes them ideal learning opportunities to acquire „negative“ knowledge. The original concept of theory articulation (Ohlsson, 1992) only refers to knowledge about the correct application of scientific knowledge, i.e. which concepts, steps and procedures are necessary for a scientific explanation. Based on Oser (2007), knowledge about correct concepts and procedures is labelled „positive“ knowledge. In contrast, „negative“ knowledge denotes knowledge on how something *is not* and how something *is not done* (Oser, 2007; Oser & Spychiger, 2005).

As depicted in our theoretical model (see fig. 1), negative knowledge can provide an additional knowledge dimension that is not included in „normal“ lessons focusing on positive knowledge. It should therefore help students to identify errors and avoid them when articulating a theory. Metatheoretical knowledge is seen here as an additional resource that enables the distinction between appropriate and inappropriate concepts within a scientific explanation.

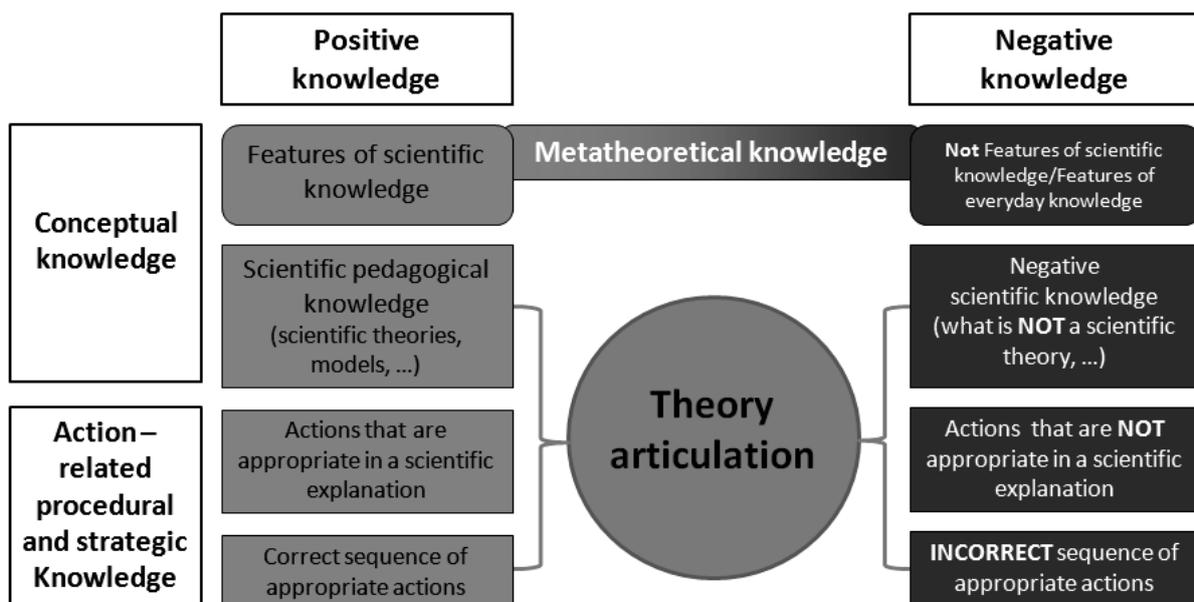


Figure 1: Model of knowledge types involved in theory articulation

An important question that arises in the context of learning from errors is whether learning from one's own errors or learning from someone else's errors is more effective. Approaches of the first kind, building on a learner's own attempts at problem solving (e.g. productive failure), may lead to deeper elaboration, given the right kind of instructional support (e.g. feedback on the correct solution or on the locus of the error(s) committed in the process). However, if a certain type of error is to be demonstrated, the learner would have to be led into committing that error or else it might never occur during learning (Ohlsson, 1996). This is where learning from others' errors (e.g. advocacy error learning) comes in, which of course brings its own set of special requirements. The following paragraph outlines a possible operationalization of each approach.

1.2.1 Productive Failure

Productive failure describes a form of constructivist learning in which learners must solve a novel problem based only on their prior or intuitive knowledge. This often leads to failure (Kapur & Bielaczyc, 2012; see also *impasse-driven learning*; VanLehn et al., 2003). Learning can be fostered by failure if adequate instructional support (consolidation, worked example, feedback or explanation) is provided subsequently. The core mechanisms of productive failure include activation and differentiation of prior knowledge and attention to critical conceptual features of a problem or phenomenon (Kapur, 2008), e.g. the structure of an argument or relevant situational features. This requires learning settings that are authentic and relevant to the learner (ibid.).

The efficacy of this approach has been demonstrated in science teaching (Pathak, Kim, Jacobson, & Zhang, 2011) and mathematics (e.g. Kapur, 2010; 2014). While based on a learner's own errors, productive failure focuses primarily on the acquisition of positive knowledge, since the learner's failure evokes a demand to acquire and understand the correct solution (Kapur, 2008).

Regarding *theory articulation*, the *productive failure-based learning environment* (PFLE) in the present study was therefore expected to foster especially the acquisition of *positive knowledge* about the features of scientifically accurate explanations of complex school situations. As learners noticed their own shortcomings, they could also (unsystematically) acquire negative knowledge.

1.2.2 Advocatory errors

Approaches based on *learning from advocatory errors* address typical errors in a given domain by presenting 'intentional' errors in the learning materials, while the learners' task is to detect the errors (e.g. Durkin & Rittle-Johnson, 2012; Stark, Kopp, & Fischer, 2011) and thus acquire negative knowledge. To profit from advocatory errors, learners must understand *what* is wrong (the error itself), *why* something is wrong in a given situation and what would be the *correct solution* (Curry, 2004). Since learners do not commit errors themselves, but study another person's errors (see also „vicarious learning“, Kapur, 2014) they have to identify with the protagonist and be emphatic, which in turn requires the protagonist and his situation to be authentic and relevant to the learner (Oser & Spychiger, 2005).

A core concept of learning from advocatory errors is contrasting errors to correct solutions (c.f. Durkin & Rittle-Johnson, 2012). This comparison can trigger self-explanations (Siegler & Chen, 2008), leading to a deeper understanding and supporting the development of „error avoidance strategies“, i.e. strategic knowledge on how to avoid an error that can protect learners against future errors (Oser, 2007).

The effectiveness of learning from advocatory errors has been shown in several studies (e.g. Durkin & Rittle-Johnson, 2012; Huang, Liu, & Shiu, 2008; Stark et al., 2011). However, just like productive failure, learning from advocatory errors requires adequate instructional support. Große and Renkl (2004, 2007) showed that only learners with high prior knowledge profited from advocatory errors when no additional instructional support was given. In a study in medicine, erroneous worked-out examples required elaborated feedback-sequences to be effective (Stark et al., 2011). Durkin and Rittle-Johnson (2012) demonstrated the effectiveness of reflection prompts supporting the comparison of correct and incorrect examples.

The *error-based learning environment* (EBLE) in the present study was expected to especially foster *negative knowledge* about what is *not* considered a feature of a scientifically correct explanation, i.e. what is considered an error in theory articulation. At the same time students were expected to acquire error avoidance strategies that should improve their own theory articulations.

1.3 Research Questions and Hypotheses

While there is plenty of research on productive failure and learning from advocatory errors, effects on theory application competences in the domain of teacher education have not been investigated yet. Based on the theoretical considerations, the present study examines which

approach is superior in enhancing the students' *quality of theory articulation* (QTA) when asked to formulate scientific explanations of complex classroom situations. Additionally, effects on the *quality of error identification* (QEI) in a given erroneous explanation were assessed, as well as effects on metatheoretical knowledge (MTK). Subjective learning success and acceptance of the training approaches were investigated to gain further insight into evidence-based optimization of the implemented learning environments. The following research questions were formulated:

1. To what extent does the training foster learning progress (pre-post test comparison)? Assuming no differences in prior knowledge, do the learning progress and thus the post test scores differ between groups? Are there group differences in the individual test dimensions MTK, QTA and QEI?

1a. (Overall learning progress and group differences) A significant main effect of the repeated measures factor (= training) was expected for the experimental groups as well as an interaction effect between the training and the group factor. Both experimental groups should improve substantially, while the control group should not.

1b. (Group differences in individual test dimensions): A significant learning progress for MTK in both experimental groups was expected, as well as an interaction with the group factor, since the control group was not expected to improve. Interactions were also expected in the QTA and QEI tasks. Learners in the PFLE were expected to have a higher learning progress regarding QTA, while learners in the EBLE were expected to show a greater increase regarding QEI; however, both experimental groups should outperform the control group.

2. To what extent do the three groups differ regarding subjective learning success? Do the experimental groups differ in their acceptance of the training approach? Are their self-assessments correlated to the objective learning outcomes?

2a. The experimental groups were expected to show significantly higher means regarding subjective learning success. Students in the control group were expected to be made aware of their problems by the pretests, without being subsequently supported in eliminating them by a specific training.

2b. Regarding acceptance of the training approach, no differences between the experimental groups were expected.

2c. Subjective learning success and acceptance were expected to be significantly correlated to the learning outcomes and to each other.

2. Method

2.1 Sample and Design

$N = 117$ secondary school student teachers from Saarland University (80 female; age: $M = 23.0$, $SD = 3.48$, semesters: $M = 5.87$, $SD = 2.31$) participated in a quasi-experimental field study as part of a course assignment. The students were recruited in four parallel teacher-education seminars on the planning and analysis of school lessons held by one lecturer. Subjects in three seminars were randomly assigned to two experimental conditions: Productive Failure-based learning (PFLE): $n = 49$ (33 female; age: $M = 23.0$, $SD = 3.32$); advocacy error-based learning (EBLE): $n = 43$ (33 female; age: $M = 22.5$, $SD = 2.69$). Since the pilot study had shown that without a dedicated training, virtually no specific theory articulation competences are present in the students (Klein, Wagner & Stark, 2012), larger effects were expected between the control group and both experimental groups than between the PFLE and the EBLE group. To ensure a maximum of participants in the experimental groups* and because of obvious differences between the learning materials of the control and experimental groups, a separate seminar ($n = 26$; 15 female; age: $M = 23.7$, $SD = 3.73$) was chosen for the control group. All data was anonymised.

2.2 Learning Environments

Two *integrated learning environments* (Reinmann-Rothmeier & Mandl, 2006) were designed to incorporate the shared requirements of both error learning approaches. A problem-based learning setting (c.f. Dochy, Segers, Van den Bossche, & Gijbels, 2003; Dabbagh & Dass, 2013) was realized by principles of anchored instruction (narrative format, authentic problems, and relevant exercises; e.g. CTGV, 1993), and combined with the instructional support measures necessary for learning from errors (*direct instruction*; Kirschner, Sweller & Clark 2006; *elaborate feedback*; Krause, 2007; and *reflection prompts*; Durkin & Rittle-Johnson, 2012). The cover stories serving as anchors were presented as short texts only since video-based trainings have been shown to be more effective for experienced teachers with extensive prior teaching knowledge and experiences in video analyses (Seidel & Prenzel, 2008; Seidel, Stürmer, Blomberg, Kobarg, & Schwindt, 2011).

* Power analyses were conducted in G*Power for a minimum effect size of $\eta^2 = .3$, $\alpha < .05$ and a test power of .95 for three groups to be 40 subjects per group for an ANOVA comparing the

experimental groups. Since all other analyses included repeated measurements and therefore allowed for smaller sample sizes, this number was used as a reference.

Additional didactic measures include multiple contexts and perspectives (c.f. CTGV, 1993) which can foster cognitive flexibility in poorly structured domains (e.g. Spiro, Collins, & Thota, 2003). All learners were provided with booklets of 18 - 24 pages that included all information and instructions. In both experimental conditions, four cover texts (about 400 words) about complex classroom problem cases were presented, addressing school bullying, specific group phenomena, achievement motivation, and discipline problems. An authentic problem setting in multiple contexts was realized by having a fictional protagonist (the „recently graduated teacher Roman“) witness multiple complex situations in his class, which he subsequently tries to explain, either erroneously by himself or by having the learner construct an explanation.

Instructional support and a scientific perspective were implemented by a fictional educational psychologist who Roman turns to for help and who supplies a theory collection script (summaries of 4-6 theories, around 600 words overall, based on four introductory teacher education lectures) and an expert explanation.

For example, in the training scenario on group phenomena, when a group of foreign students enters the class, Roman (or the learner) erroneously attributes the hostile reaction of some pupils to racism acquired from their parents via social learning. The psychologist, however, points out that the situation could be explained more conclusively by group phenomena such as intergroup conflicts (e.g. Sherif, 1966) and would thus require Roman to act differently in solving the problem.

In the PFLE, students had to develop their own explanations, based on their prior knowledge (step 1). In step 2, they were instructed to study the theory collection script and write a revised version of their initial explanation based on the theories. In step 3, the students received an expert explanation of the situation, i.e. a feedback measure on the explanation task, and had to compare it to their own revised analysis. The comparison was prompted by five guiding questions. Students were shown an example of a correct theory articulation, but the deficits of their own initial explanations were not explicitly discussed.

In the EBLE, students were asked to study the same theory collection script and expert explanation (Step 1) to compensate for a lack of prior knowledge (c.f. Große & Renkl, 2007). In step 2 the students were confronted with an erroneous explanation and had to identify typical errors related to the application of scientific knowledge (c.f. Stark, 2005), supported by an error glossary provided in the booklet. In step 3, students were prompted to compare

their own error analysis to an expert error analysis, i.e. a feedback measure on the error identification task. Theory articulation was not directly trained; instead, learners were expected to derive error-avoidance strategies from contrasting the erroneous explanations to the expert explanations and identifying the differences.

Metatheoretical knowledge was not explicitly trained in either condition. Learners were expected to abstract knowledge about features of everyday and scientific knowledge from the comparison of the correct and incorrect explanations in the EBLE and their own deficient explanation draft and the model explanation in the PFLE. For a complete overview of both trainings see Fig. 2.

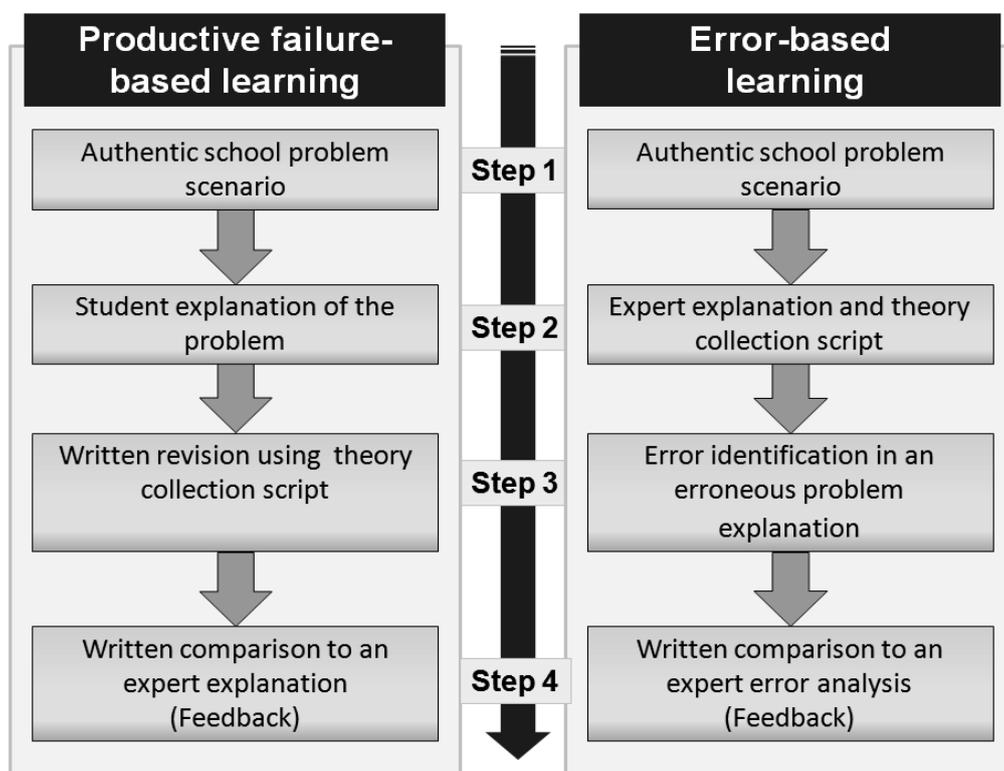


Figure 2: Experimental conditions

2.3 Procedure

After the pretests in the first seminar session (see fig. 3), students attended two whole seminar sessions of 90 minutes each to work on two training scenarios per session during the training phase. In the fourth session, questionnaires on subjective learning success and acceptance were handed out *before* the post-tests. By keeping the learning materials identical in all experimental seminars and giving no further instruction, a high standardization was realized.

The *control group* completed the pre- and post-tests at the same time as the experimental groups and worked on standard written assignments on lesson planning based on psychological and pedagogical theories during the training phase. Training scenarios were accessible to them after the post tests.

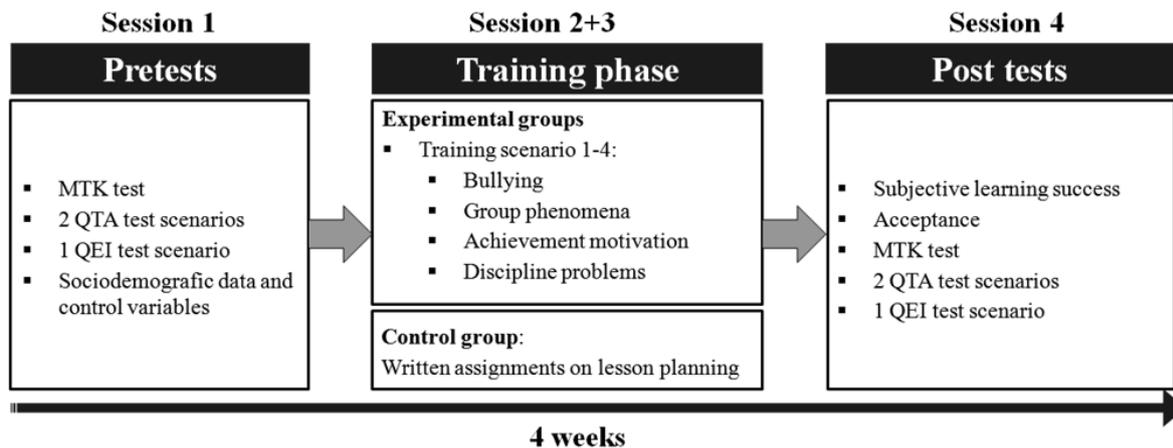


Figure 3: Training and testing procedure

2.4 Instruments

2.4.1 Control variables and biographical data

To ensure internal validity, prior knowledge (Kirschner et al., 2006; Stark et al., 2010) was assessed as a control variable. Age, gender and semesters were recorded by a questionnaire.

2.4.2 Prior knowledge and learning outcomes (pre- and post-tests)

Metatheoretical knowledge was assessed by an identical pre- and post-test with five constructed response questions about differences between scientific and everyday knowledge based on Bortz & Döring's (2006) dimensions (e. g. „Name the differences between an everyday and a scientific explanation of a phenomenon or situation.”).

Procedural and strategic knowledge (*QTA/QEI*) were measured in the pre- and post test by scenario-based tests similar to key feature tests in medicine. Key features represent the critical, or essential, steps in the identification and management of a clinical problem (Kopp, Möltner & Fischer, 2006). Analogously, in the QTA tasks students had to perform the critical steps in the scientific explanation of a complex school problem, i.e. they had to identify the relevant facts, find the appropriate theory or model and explicate the relation between situational facts and theory (the latter being the exact process of theory articulation as

described by Ohlsson, 1992). In the QEI task, students had to identify errors in performing these steps. Due to the complexity of the post test tasks and time-on-task-data in a pilot study (Klein et al., 2012), only three tests were presented.

Quality of theory articulation (QTA) was measured by two problem-solving tasks. In the first one, a QTA *application* task, theories and concepts acquired in the training session had to be applied without additional support. Essentially, students had to construct a theory articulation following the model of the expert explanation in the first training scenario, since the problems in both tasks were structurally analogous. The second QTA task required transfer. Students were presented with a problem setting not previously discussed in the training (school and performance anxiety) and provided with summaries of theories on these topics equal in length to those in the theory collection scripts in the training scenarios (around 600 words overall).

The *quality of error-identification* (QEI) task described a situation that was structurally similar to the second training scenario on group phenomena; however, the erroneous explanation referred to aggression against foreigners. No additional support was given. Learners were instructed to only identify errors related to theory articulation in the erroneous explanation, but not to correct them.

2.4.3 Scoring procedure

Performance in all tasks was rated by two trained educational researchers comparing the students' solutions to prefabricated expert solutions that could be expanded by original student solutions judged to be appropriate by both raters. The raters did not know which condition the students had been in.

In the MTK test, each correct proposition scored one point, with a maximum of 19 points. Cronbach's α was .76 in the pre-test and .82 in the post test. After a short coding training, the mean inter-rater-reliability was .83 for the pre-test and .89 for the post-test (Cohen's kappa).

In the two problem solving (QTA) tasks, the students' answers were rated with regard to the identification of relevant situational facts, explicit reference to scientific theories, and the quality of the connection of situational facts and theoretical propositions to construct a scientific explanation. A maximum of 10.5 points were assigned when the students' explanations corresponded to or expanded the expert solution. Cohen's kappa (mean inter-rater-reliability) was .81 for the pre-test and .91 for the post-test.

In the error-identification task, the rating dimensions included identification of the errors, the relation to situational facts, and a short discussion on why the given explanations were erroneous. However, points were not assigned for attempts at writing a corrected version. Since students in the PFLE could not know the exact error definitions from the error glossary in the EBLE, points were awarded for any proposition that resembled the error description satisfyingly. A maximum of 6.5 points was possible. After a short coding training, the mean interrater-reliability was .79 for the pre-test and .88 for the post-test (Cohen's kappa).

2.4.4 Subjective dimensions

Subjective training success was assessed by a 15-item scale. The learners had to evaluate their progress concerning different aspects of theory application (e. g., „I now have a better concept of how theories are associated with school practice.“; Cronbach's alpha = .90).

Acceptance of the training approach was measured by a five item scale focusing on the learners' evaluation of the case scenarios (e. g., „The learning cases were well chosen“; Cronbach's alpha = .75). Since these questions specifically addressed the training materials, they were only presented to the experimental groups.

2.5 Statistical Analyses

Internal validity tests regarding prior knowledge and the control variables were performed by a MANOVA and a Chi square test. A repeated measures MANOVA was performed to analyze the overall learning progress. Since the experimental conditions were hypothesized to foster the individual knowledge dimensions to different degrees and outperform the control group in each dimension, additional univariate repeated measures ANOVAs with planned contrasts were included. Contrasts were defined to test differences between all three groups according to the hypotheses (see 1.4). Effect sizes are reported as η^2 (0.01 = small, 0.06 = medium, 0.14 = strong effect; Cohen, 1988) for ANOVAs and Cohen's d (0.2 = small, 0.5 = medium, 0.8 = strong effect; *ibid.*) for planned contrasts. Relations between the learning outcomes in the pre- and post and the subjective assessments were examined by bivariate correlations using Pearson's r .

3. Findings

3.1 Internal Validity

The pretests showed no significant differences between the groups ($F(16,212) = 1.08$; $p = .38$; $\eta_p^2 = .08$). Additional univariate analyses showed that effects were $p > .20$ for age, semester, and all prior knowledge measures. There was no difference in gender ($\chi^2(2; n = 117) = 3.90$; $p = .24$).

3.2 Research Question 1: Learning Progress

1a. Overall learning progress and group differences

A MANOVA with repeated measurements for all tasks showed a significant effect of the training ($F(1,111) = 201.9$; $p < .001$, $\eta_p^2 = .645$). Both experimental groups showed descriptive learning progress in all tasks except for the QEI test (see table 1), the control group did not improve. Accordingly, there was an interaction effect of the training and group factor ($F(2,111) = 4.66$; $p = .011$, $\eta_p^2 = .077$).

1b. Learning progress and group differences in individual test dimensions

Regarding MTK, there was a significant main effect of the training ($F(1, 112) = 30.2$; $p < .001$; $\eta_p^2 = .212$), and an interaction effect between training and group factor ($F(1, 112) = 5.37$; $p = .006$; $\eta_p^2 = .087$). As hypothesized, both experimental groups profited substantially from the training (PFLE: $F(1,46) = 28.8$, $p < .001$, $\eta_p^2 = .385$; EBLE: $F(1,41) = 17.3$, $p < .001$, $\eta_p^2 = .296$). The control group did not improve ($F(1,25) = .164$, $p = .689$, $\eta_p^2 = .007$; see table 1).

Planned contrasts showed significant differences between the experimental groups and the control group in the post test (PFLE: $T(113) = 4.47$; $p < .001$; $d = 1.06$; EBLE: $T(113) = 3.65$; $p < .001$; $d = 1.01$; see table 1), but not between the experimental groups ($T(113) = .84$; $p = .403$; $d = .17$). Both groups obtained significantly higher means in the post test than the control group.

QTA improved descriptively in all groups (see table 1). The training had a significant main effect on the QTA application task ($F(1,113) = 61.5$, $p < .001$, $\eta_p^2 = .353$). In addition, there was a medium significant interaction between the training and group factor ($F(2, 113) = 15.5$, $p < .001$, $\eta_p^2 = .216$). Both experimental groups improved significantly and substantially (PFLE: $F(1,48) = 79.9$, $p < .001$, $\eta_p^2 = .625$; EBLE: $F(1,40) = 17.6$, $p < .001$, $\eta_p^2 = .306$), but the PFLE group profited more from the training in this task. Again, there was no significant

learning progress for the control group ($F(1,26) = 1.65, p = .210, \eta_p^2 = .062$). As expected, planned contrasts revealed that the PFLE group significantly outperformed both the EBLE ($T(113) = 2.1; p = .038, d = .42$) and the control group ($T(113) = 5.24; p < .001, d = 1.3$) in the post test (see table 1). The EBLE group was, however, still significantly and substantially better than the control group ($T(113) = 3.27; p = .001, d = .90$).

There was no significant effect of the training in the QTA transfer task ($F(1, 113) = 1.58; p = .212, \eta_p^2 = .014$). However, there was a significant interaction between the training and group factor ($F(2, 113) = 8.23, p < .001, \eta_p^2 = .127$). Descriptively, both experimental groups improved, while the control group performed actually worse than in the pretests. Again, the PFLE group obtained higher scores than the EBLE group (see table 1). In this task, neither the EBLE ($F(1, 40) = 1.90; p = .175; \eta_p^2 = .045$) nor the control group ($F(1, 25) = 2.76; p = .109; \eta_p^2 = .099$) achieved a significant learning progress, while the effect of the training for the PFLE group was substantial ($F(1, 48) = 17.8; p < .001; \eta_p^2 = .271$). Again planned contrasts showed both experimental groups to be substantially more successful than the control group (PFLE: $T(113) = 4.47; p < .001; d = 1.07$; EBLE: $T(113) = 2.52; p = .013; d = .56$) in the post test (see table 1), though the effect in the EBLE was noticeably smaller. Accordingly, there was another significant difference between the experimental groups ($T(113) = 2.14; p = .034, d = .47$); as expected, in this task the PFLE group outperformed the EBLE group.

In the QEI task, another significant effect of the training ($F(1, 114) = 11.2; p = .001; \eta_p^2 = .090$) became apparent. There was a significant interaction between training and group factor ($F(1, 88) = 24.9; p < .001; \eta_p^2 = .304$). In this task, the EBLE group was superior (see table 1) and obtained a significant and substantial learning progress ($F(1, 41) = 53.1; p < .001; \eta_p^2 = .564$). The PFLE and the control group were descriptively even worse than in the pretests, however, the effects were not significant (PFLE: $F(1, 48) = 1.27; p = .265; \eta_p^2 = .026$; control: $F(1, 46) = .335; p = .568; \eta_p^2 = .013$). Planned contrasts revealed that in the post test, the EBLE group was more successful than both the control ($T(114) = 5.37; p < .001, d = 1.26$), and the PFLE group ($T(114) = -6.60; p < .001, d = 1.35$). The latter did not differ significantly from the control group ($T(114) = -.20; p = .84; d = .06$; see table 1).

Table 1: Means and standard deviations of the learning outcomes in all groups

	T1				T2							
	PFLE		EBLE		Control		PFLE		EBLE		Control	
	<i>n</i>	<i>M(SD)</i>										
MTK	47	6.5 (2.3)	41	6.1 (2.6)	26	6.0 (2.3)	47	8.6 (2.5)	41	8.2 (2.1)	26	6.2 (1.9)
QTA application	49	2.1 (1.1)	41	2.5 (1.0)	26	2.1(.9)	49	4.4(1.8)	41	3.6(1.8)	26	2.3(1.2)
QTA transfer	49	4.4 (1.5)	41	4.2 (1.1)	26	4.0 (1.1)	49	5.2(1.6)	41	4.5(1.5)	26	3.5(1.7)
QEI	49	1.9 (0.9)	42	1.9 (0.8)	26	1.9 (0.7)	49	1.7(1.2)	42	3.6(1.6)	26	1.8(1.1)

MTK max. 19 points; QTA (post tests 1&2) max 10.5; QEI (post test 3) max 6.5 points.

3.3 Research question 2: Subjective dimensions

2a. Subjective learning success

Both experimental groups showed higher means than the control group (see table 3), they were in the upper third of the scale. There was a significant effect of the group factor ($F(2, 114) = 11.1$; $p < .001$; $\eta_p^2 = .163$). Significant differences were shown between the EBLE group and the control group ($T(114) = 4.69$; $p < .001$; $d = 1.22$) as well as between the two experimental groups ($T(114) = 2.43$; $p = .017$; $d = .50$), in favour of the EBLE group.

2b. Acceptance of the training approach

Means in both experimental groups were in the upper third of the scale. There were no significant differences between the experimental groups ($F(1, 84) = 2.19$; $p = .14$; $\eta_p^2 = .025$; see table 2).

Table 2: Means and standard deviations of the subjective dimensions in all groups

	PFLE		EBLE		Control	
	<i>n</i>	<i>M(SD)</i>	<i>n</i>	<i>M(SD)</i>	<i>n</i>	<i>M(SD)</i>
Subjective learning success	48	4.2 (.57)	43	4.5 (.54)	26	3.8 (.50)
Acceptance	48	4.4 (.44)	43	4.5 (.48)	-	-

* Max. 6 points

2c. Correlations between subjective dimensions and learning outcomes

As expected, acceptance and subjective learning success were significantly correlated. Also, subjective learning success correlated with all learning outcome measures except the MTK test, while acceptance was not correlated to the outcome measures (see table 3).

Table 3: Correlations between subjective dimensions and learning outcomes

	Subjective learning success	MTK	QTA		
			applicatio n	QTA transfer	QEI
Acceptance	.287**	-.038	.076	.003	.182
Subjective learning success		.162	.342**	.211*	.197*

**= $p < .001$; *= $p < .05$

4. Discussion

4.1 Learning Outcomes

Internal validity was secured, as there were no significant differences in the control variables and in prior knowledge. Overall, both training approaches were effective and fostered competences in the application of scientific knowledge. After the training, both experimental groups outperformed the control group and had significantly improved in a comparison between pre- and post test in all but one task.

However, different types of knowledge as depicted in the theoretical model were fostered to different degrees. Overall the EBLE can be regarded as more effective, since learners in this condition acquired negative as well as positive knowledge on theory articulation, while learners in the PFLE did not establish a negative knowledge base.

With regard to the QTA scores of the EBLE group, negative knowledge lead to a substantial improvement, however, their theory articulations were not quite of the same quality as those in the PFLE. With respect to further optimizations of the learning materials, a direct contrast between correct and incorrect examples (e.g. Durkin & Rittle-Johnson, 2012) might yield better learning outcomes than a detailed error description. Other possible approaches include the explicit training of error avoidance strategies as a form of positive knowledge, or the use of metacognitive prompts (c.f. Tyroller, 2005) to support self-explanations.

In the PFLE, the additional presentation of advocatory errors (e.g. in an elaboration script; Wagner et al., 2013) might prepare learners better for the QEI task. However, at least a pilot study of our research group did not find advantages for a combined condition (Klein et al., 2012).

Overall, the learning outcomes were still far from the theoretical maximum. From a content-based perspective, presenting each theory in the training phase only once may not have been enough to secure unsupported recall in the post-test. Some students reverted to everyday explanations in the QTA application task, possibly because they were unable to recall the theories from the training phase. However, the same problem was found in the QTA transfer task which provided appropriate theories. Despite a measurable improvement in QTA, theory application still seems difficult for students.

Most likely, these persistent problems even after the training are at least partly due to the short intervention time. Additionally, the expert explanations provided in both experimental conditions might have been too complex for the participants. The distance between the explanations of the learners and the expert solutions was too far and for some learners even insurmountable. Here the concept of the „zone of proximal development“ (Vygotski, 1986) can be conducive for further optimization efforts; the idea being that the students' next level of competence is lower than the one provided in the expert explanations and that therefore, students were not able to grasp all „essential steps“ of a correct explanation (see. 2.4.2, measurement of procedural knowledge). To improve the effect of the explanations, they could either be simplified, which would not fit idea of the „optimal“ solution, or scaffolded by a structural expert explanation that explicates the essential steps.

4.2 Subjective dimensions

Regarding subjective learning success, both experimental groups were in the upper third of the theoretical maximum, while the control group was closer towards the middle of

the scale. With regard to the correlations to the learning outcomes, our findings show that students in all groups were capable of estimating their learning success at least to a certain degree.

The difference between the experimental groups was possibly influenced by their respective performance in the QEI *pre-test*. While the EBLE group trained explicitly for this task, which was certainly new to the students, the PFLE group, and the control group which reported significantly lower subjective learning success, expressed disappointment in a debriefing session, since the pretests had made them aware of their difficulties, but their learning materials did not specifically help them to overcome them.

Regarding acceptance, both experimental groups reported scores in the upper third of the theoretical maximum. This was consistent with oral reports in the debriefing session, in which students expressed overall acceptance of the training.

4.3 Limitations

Providing more post tests would have allowed for more reliable measures of QTA and QEI. However, due to the findings of the pilot study regarding time-on-task for the post tests, and the constraints of a field study in a seminar, we relied on fewer measures. Complete randomization including the control group would have been desirable. In our study, only the experimental groups were completely randomized, however, the students in the control group stem from the same population of student teachers choosing their seminar based on preferences regarding their time table. Additionally, relevant variables in the group comparison were controlled.

Regarding authenticity of the scenarios, we acknowledge that it is hardly possible for a teacher to invest as much time and effort into problems of single pupils as such a thorough analysis would require, nor is there always an educational psychologist at hand to supply the appropriate theories. The point of the training, however, was not to have future teachers analyze any given situation this elaborately, but to foster transferrable competences that can be called upon *if* a situation requires them (see also *tacit knowing view*, Neuweg, 2005). To improve authenticity nonetheless, the methods acquired in the training could be applied to actual case materials from the students' internships and practical experience as an intermediate phase and extension of applying theory to written narratives.

4.4 Conclusion: Theoretical and pedagogical implications

From a theoretical perspective, this study contributed to research on situated learning in teacher education, a field that is often neglected in favour of more well-structured domains. In addition, constraints and affordances of productive failure and learning from advocacy errors were contrasted in the study (see also productive failure vs. „vicarious learning“, Kapur, 2014). With regard to the additional knowledge dimension provided by negative knowledge in the EBLE, this approach warrants further investigation. Admittedly, it cannot cover all possible problems and feedback might more helpful if it relates directly to a learner’s specific error or problem. Learning from advocacy errors, on the other hand, allows teacher educators to address relevant problems directly. Having learners make errors on their own is somewhat idiosyncratic, since not all learners commit the same mistakes or even the same number. Additionally, they might not even realize that they committed an error, which defeats the purpose of learning from errors (Ohlsson, 1996).

Additionally, research on the effects of having teachers analyze and reflect on their own classroom actions versus those of others showed that they were less likely to notice „critical incidents“ when analyzing themselves in classroom videos; i.e. they were less aware of their own errors (Seidel, et al., 2011). Pointing out errors explicitly should therefore foster awareness of and consequently learning from errors. Finally, higher subjective learning success and acceptance reports in the EBLE support the implementation of similar training approaches into the teacher education curriculum.

References

- Bainbridge, A. (2011). Beginning teaching: the theory/practice divide. *Cliopsy*, 6, 25-32.
- Baumert, J. & M. Kunter (2006): Stichwort: professionelle Kompetenz von Lehrkräften. [Key word: Professional competence of teachers]. *Zeitschrift für Erziehungswissenschaften*. 9 (4), 469-520.
- Beck, K. & Krapp, A. (2006). Wissenschaftstheoretische Grundfragen der Pädagogischen Psychologie [Fundamental questions on science theory in pedagogical psychology]. In A. Krapp & B. Weidenmann (Eds.), *Pädagogische Psychologie* [Pedagogical psychology] (p. 33-73). Weinheim: Beltz PVU.
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* [Research methods and evaluation in human and social sciences]. Berlin: Springer.
- Bransford, J., Darling-Hammond, L., & LePage, P. (2005): Introduction. In: Darling-Hammond, L., Bransford, J. (Eds.): *Preparing Teachers for a Changing World*. – San Francisco, S. 1-39.
- Bromme, R. (2008). Lehrerexpertise. [Teacher expertise] In W. Schneider & M. Hasselhorn (Eds.). *Handbuch der Pädagogischen Psychologie* [Handbook of pedagogical psychology] (p. 159- 167). Göttingen: Hogrefe.
- Cognition and Technology Group at Vanderbilt (1993). Designing learning environments that support thinking: The Jasper series as a case study. In T.M. Duffy, J. Lowyck, D. H. Jonassen, & T. M. Welsh (Eds.), *Designing environments for constructive learning* (p. 9-36). Berlin: Springer.
- Cohen, J. (1988): *Statistical Power Analysis for the Behavioral Sciences*, 2nd. Ed., Hillsdale: Lawrence Erlbaum Associates.
- Curry, L. (2004). The effects of self-explanations of correct and incorrect solutions on algebra problem solving performance. In K. Forbus, D. Gentner, and T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, 1548. Mahwah, NJ: Erlbaum.
- Dabbagh, N. & Dass, S. (2013). Case problems for problem-based pedagogical approaches: A comparative analysis. *Computers & Education*, (64), 161-174
- De Jong, T., & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31, 105-113.

- Dochy, F., Segers, M., Van den Bossche, P. M., & Gijbels, D. (2003). Effects of problem-based learning. A meta-analysis. *Learning and Instruction*, 13, 533–568.
- Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction*, 22, 206-214.
- Groeben, N., Wahl, D., Schlee, J., & Scheele, B. (1988). *Das Forschungsprogramm Subjektive Theorien*. [The research program Subjective Theories] Tübingen: Francke.
- Große, C., & Renkl, A. (2004). Learning from worked examples: What happens if errors are included? In P. Gerjets, J. Elen, R. Joiner & P. Kirschner (Eds.), *Instructional design for effective and enjoyable computer supported learning* (p. 356-364). Tübingen: Knowledge Media Research Center.
- Große, C., & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction*, 17, 612-634.
- Huang, T. H., Liu, Y. C., & Shiu, C. Y. (2008). Construction of an online learning system for decimal numbers through the use of cognitive conflict strategy. *Computers & Education*, 50, 61-76
- Ingvarson, L., & Rowe, K. (2008). *Conceptualising and evaluating teacher quality: substantive and methodological issues*. Retrieved from: <http://www.freepatentsonline.com/article/Australian-Journal-Education/181463888.html> [02/08/2013].
- Kapur, M. (2008). Productive Failure. *Cognition and Instruction*, 26, 3, p. 379-424.
- Kapur, M (2010). Productive failure in mathematical problem solving. *Instructional Science*, 38, (6), 523-550.
- Kapur, M. (2014), Productive Failure in Learning Math. *Cognitive Science*. doi: 10.1111/cogs.12107
- Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *The Journal of the Learning Sciences*, 21, 45–83.
- Kirschner, P. A., Sweller, J., & Clark, R.E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational psychologist*, 41, 75-86.
- Klein, M., Wagner, K. & Stark, R. (2012). Theoretisieren für die Praxis - Analyse und Förderung der Anwendung wissenschaftlichen Wissens in schulischen Kontexten. In : Krämer, M., Dutke, S., & Barenberg, J. (Hrsg.). *Psychologiedidaktik und Evaluation IX* (Auflage: 1.). Aachen: Shaker.

- Kopp, V., Möltner, A. & Fischer, M.R. (2006): Key-Feature-Probleme zum Prüfen von prozeduralem Wissen: Ein Praxisleitfaden. [Key-feature problems to test procedural knowledge]. *GMS Z Med Ausbild* 23 (3), Doc 50.
- Krause, U.-M. (2007). *Feedback und kooperatives Lernen*. [Feedback and cooperative learning] Münster: Waxmann.
- Kuhn, D. (2010). Teaching and learning science as argument. *Science Education*, 94(5), 810–824
- Le Cornu, R. & Ewing, R. (2008). Reconceptualising professional experiences in pre service teacher education -reconstructing the past to embrace the future, *Teaching and Teacher Education*, 24 (7), 1799-1812.
- McNeill, K. L., Harris, C. J., Heitzman, M., Lizotte, D. J., Sutherland, L. M., & Krajcik, J. (2004). How can I make new stuff from old stuff? In J. Krajcik & B. J. Reiser (Eds.), *IQWST: Investigating and Questioning - Our World Through Science and Technology*. Ann Arbor, MI: University of Michigan.
- Meier, A. (2006). Theorienutzungskompetenz als Schlüsselqualifikation von Lehrkräften. [Theory application competence as a key skill of teachers] In: Y.M. Nakamura, C. Böckelmann & D. Tröhler (Eds.). *Theorie versus Praxis?* [Theory versus practice?] (p. 89-106) Verlag Pestalozzianum: Zürich.
- Neuweg, G. H. (2005): Der Tacit Knowing View. Konturen eines Forschungsprogramms. [The tacit knowing view. An outline of a research program] *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 101, 557–573.
- Neuweg, G. H. (2007): *Wie grau ist alle Theorie, wie grün des Lebens goldener Baum? LehrerInnenbildung im Spannungsfeld von Theorie und Praxis*. [Grey, dear friend, is all theory, and green the golden tree of life. Teachers' education in the conflict between theory and practice] bwp@, Ausgabe 12. http://www.bwpat.de/ausgabe12/neuweg_bwpat12.pdf [18.10.2014].
- Ohlsson, S. (1992). The cognitive skill of theory articulation: A neglected aspect of science education. *Science & Education*, 1, 181-192.
- Ohlsson, S. (1996). Learning from performance errors. *Psychological Review*, 103, 241 - 262.
- Oser, F. (2007). Aus Fehlern lernen. [Learning from mistakes] In: M. Göhlich, Ch. Wulf & J. Zirfas (Eds.): *Pädagogische Theorien des Lernens* [Pedagogical theories of learning] (p. 203-212). Weinheim und Basel: Beltz Verlag.

- Oser, F., & Spychiger, M. (2005). *Lernen ist schmerzhaft. Zur Theorie des Negativen Wissens und zur Praxis der Fehlerkultur*. [Learning is painful. On the theory of negative knowledge and the practice of error culture]. Weinheim: Beltz
- Pathak, S. A., Kim, B., Jacobson, M. J., & Zhang, B. H. (2011). Learning the physics of electricity: A qualitative analysis of collaborative processes involved in productive failure. *International Journal of Computer-Supported Collaborative Learning*, 6(1), 57–73.
- Reinmann-Rothmeier, G., & Mandl, H. (2006). Unterrichten und Lernumgebungen gestalten. [Designing teaching and learning environments] In: A. Krapp & B. Weidenmann (Eds.), *Pädagogische Psychologie. Ein Lehrbuch* [Pedagogical Psychology. A textbook] (S. 613-658). Weinheim: Beltz.
- Ritchie, J.R. and Wilson, D.E. (2001). *Teacher Narrative as Critical Inquiry: Rewriting the Script*. New York: Teachers College Press.
- Schneewind, K.A. (1977). Zum Verhältnis von Psychologie und Wissenschaftstheorie. [On the relation between psychology and the philosophy of science] In: K.A. Schneewind (Ed.), *Wissenschaftstheoretische Grundlagen der Psychologie* [The theory of science in psychology] (p.11-25). München: Reinhardt.
- Seidel, T. & Prenzel, M. (2008). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen - Erfassung pädagogisch-psychologischer Kompetenzen mit Videosequenzen [How teachers perceive and evaluate lessons - Assessment of pedagogical competences with video sequences]. In M. Prenzel, I. Gogolin & H.-H. Krüger (Eds.), *Kompetenzdiagnostik* [Competence diagnostic] (p. 201-216). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K.(2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others?, *Teaching and Teacher Education*, Volume 27, Issue 2, February 2011, Pages 259-267.
- Sherif, M. (1966). *In Common Predicament: Social Psychology of Intergroup Conflict and Cooperation*. Boston: Houghton Mifflin Company. S. 24–61.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4- 31.
- Siegler, R. S., & Chen, Z. (2008). Differentiation and integration: Guiding principles for analyzing cognitive change. *Developmental Science*, 11, 433-448.

- Smith, K. (2005). Teacher Educators' professional knowledge- How does it differ from teachers' professional knowledge? *Teaching and Teacher Education*, 21, 177-192.
- Spiro, R. J., Collins, B., & Thota, J. (2003). Cognitive flexibility theory: Hypermedia for complex learning, adaptive knowledge application, and experience acceleration. *Educational Technology*, 43, 5-10.
- Stamouli, E. Gruber, H., & Schmid, C. (2010). Expertiseforschung und Lehrer/innen(fort)bildung [Expertise research and teacher training]. In F. Müller, A. Eichenberger, M. Lüders & J. Mayr (Eds.), *Lehrerinnen und Lehrer lernen - Konzepte und Befunde der Lehrerfortbildung* [Teachers learn - concepts and findings of teacher training]. Münster, New York, München, Berlin: Waxmann Verlag.
- Stark, R. (2005). Constructing arguments in educational discourses. In H. Gruber, C. Harteis, R. Mulder & M. Rehl (Eds.), *Bridging individual, organisational, and cultural aspects of professional learning* (p. 64–71). Regensburg: S. Roderer.
- Stark, R., Herzmann, P., & Krause, U.-M. (2010). Effekte integrierter Lernumgebungen – Vergleich problembasierter und instruktionsorientierter Seminarkonzeptionen in der Lehrerbildung. [Effects of integrated learning environments – a comparison between problem-based and instruction-oriented concepts for seminars in the education of teachers] *Zeitschrift für Pädagogik*, 56, 548-563.
- Stark, R., Kopp, V., & Fischer, M. R. (2011). Case-based learning with worked examples in complex domains: Two experimental studies in undergraduate medical education. *Learning and Instruction*, 21, 22-33.
- Swinkels, M.F.J., Koopman, M., & Beijaard, D. (2013). Student teachers' development of learning-focused conceptions. *Teaching and Teacher Education*, 34, 26-37.
- Tyroller, M. (2005). *Effekte metakognitiver Prompts beim computerbasierten Statistikkennen*. [Effects of metacognitive prompts in computer-based learning of statistics] Dissertation, Ludwig-Maximilians-Universität München. Retrieved from: http://edoc.ub.unimuenchen.de/archive/00005514/01/Michael_Tyroller.pdf [2.1.2011].
- VanLehn, K., Siler, S., Murray, R. C., Yamauchi, T., & Baggett, W.B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, 21, 209-249.
- Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Trans.). Cambridge, MA: The MIT Press.
- Wagner, K., Stark, R., Daudbasic, J., Klein, M., Krause, U.-M., & Herzmann, P. (2013). Effektivität integrierter Lernumgebungen in der universitären Lehrerbildung – eine

quasiexperimentelle Feldstudie. [Effectiveness of integrated learning environments in university education of teachers – a quasiexperimental field study] *Journal für Bildungsforschung online*, 5, 115-140.

7. Studie 2: Instruktionale Unterstützung beim Lernen aus fremden Fehlern in integrierten Lernumgebungen

Das folgende Kapitel basiert auf einem Artikel von Wagner, Klein, Klopp und Stark (2014), bei dem ich Zweitautor bin und erweitert diesen um Informationen zur Konzeption der Lernumgebung sowie zusätzliche Analysen und Diskussionspunkte. Es handelt sich dabei um folgenden Artikel:

Wagner, K., Klein, M., Klopp, E. & Stark, R. (2014) Instruktionale Unterstützung beim Lernen aus advokatorischen Fehlern in der Lehramtsausbildung: Effekte auf die Anwendung wissenschaftlichen Wissens. *Psychologie in Erziehung und Unterricht*, 61, 287-301. Ernst Reinhardt Verlag München, Basel.

Die darin beschriebene Studie 2 diente zur Entwicklung und Evaluation einer integrierten Lernumgebung basierend auf dem advokatorischen Lernen aus fremden Fehlern nach Oser (2007; Oser & Spychiger, 2005). Dieser Ansatz wurde aufgrund der Ergebnisse aus Studie 1 gewählt. Damit das folgende Kapitel analog zu den Aufsätzen zu Studie 1 und 3 auch für sich alleinstehend gelesen werden kann, werden die wichtigsten theoretischen Hintergründe kurz zusammengefasst. Eine ausführlichere Darstellung des Lernens aus Fehlern und integrierter Lernumgebungen findet sich in Kapitel 3 und 4.

Zusammenfassung

In den Standards der Lehrerbildung wird als Ansatz zur Förderung anwendbaren bildungswissenschaftlichen Wissen u.a. die „Analyse simulierter, filmisch dargebotener oder tatsächlich beobachteter komplexer Schul- und Unterrichtssituationen und deren methodisch geleitete Interpretation“ (KMK, 2014, S.6) genannt. Die in Studie 2 beschriebene Lernumgebung basiert auf Beschreibungen authentischer, komplexer schulischer Situationen, die anhand ausgearbeiteter Lösungsbeispiele unter Bezugnahme auf bildungswissenschaftliches Wissen erklärt werden. Die Aufgabe der Lernenden besteht dabei im Nachvollzug der Lösung. Die Lernumgebung wird durch instruktionale Unterstützungsmaßnahmen didaktisch angereichert. Zudem wurde die Auswahl der dargestellten Fehler im Vergleich zu Studie 1 durch eine *Fehlermatrix* erweitert und systematisiert. Die Lernwirksamkeit dieser integrierten Lernumgebung wurde in Pflichtseminaren des bildungswissenschaftlichen Lehramtsstudiums im Rahmen einer

experimentellen Feldstudie mit 288 Probanden untersucht. In einem Kontrollgruppendesign mit Messwiederholung wurden konzeptuelles Fehlerwissen, strategisches Fehlervermeidungswissen, prozedurales Fehleridentifikationswissen sowie Wissen zur Erklärung schulischer Situationen erhoben. Dabei wurden zwei instruktionale Unterstützungsmaßnahmen (Prompts zur Präsentation bzw. Rekonstruktion von Fehlervermeidungsstrategien) variiert. Die integrierte Lernumgebung basierend auf dem advokatorischen Lernen aus Fehlern erwies sich als effektiv und effizient, die Kombination der beiden instruktionalen Unterstützungsmaßnahmen erbrachte den höchsten Lernerfolg. Im Nachtest waren die Probanden mit maximaler instruktorischer Unterstützung den Probanden mit geringerer Unterstützung sowie der Kontrollgruppe in allen Variablen signifikant überlegen.

7.1 Problemdarstellung

Wie bereits in Kapitel 2.1 dargelegt, ist die Theorienutzung zur Erklärung komplexer schulischer Situationen anhand wissenschaftlichen Wissens eine wichtige Voraussetzung für effektives Lehrerhandeln im Schulalltag (vgl. Kap.1; Bromme, 2008; Meier, 2006). Zur Förderung dieser Kompetenz eignet sich das schon in Studie 1 zur Operationalisierung herangezogene Konzept der *Theorieartikulation* (Ohlsson, 1992) als die „Anwendung einer Theorie auf eine bestimmte Situation und eine Ausarbeitung dessen, was die Theorie implizit oder explizit über die Situation aussagt.“ (S.182; Übers. d. A.). Beobachtete Phänomene werden in Bezug zu einer Theorie gesetzt, was das Zusammenspiel verschiedener Wissensarten erfordert. Die Erklärung einer schulischen Situation kann damit als *Anwendung* wissenschaftlichen Wissens betrachtet werden (Krause, 2007, s. Kap. 3.2.1). Lehramtsstudierende haben allerdings zahlreiche Schwierigkeiten mit der Anwendung wissenschaftlichen Wissens (Seidel & Prenzel, 2008; Star & Strickland, 2008), was u.a. auf ungünstige Einstellungen zu Theorien zurückgeht (Stark, Herzmann & Krause, 2010; s.a. Kap. 2.3). Daraus resultieren typische Fehler wie z. B. die Erklärung der Situation anhand von Alltagswissen oder die unzulängliche Bezugnahme auf empirische Evidenz (Stark, 2005; Stark & Krause, 2006).

Wie schon in Studie 1 sollten diese Fehler als Lerngelegenheiten genutzt werden, um dieser Problematik zu begegnen. Dazu wurde im Rahmen einer experimentellen Interventionsstudie eine integrierte Lernumgebung (Reinmann & Mandl, 2006) auf Basis des advokatorischen Lernens aus Fehlern (Oser & Spychiger, 2005) entwickelt. Darin werden systematisch instruktionsorientierte und problemorientierte Designprinzipien kombiniert, indem

ausgearbeitete Lösungsbeispiele (hier umgesetzt als *fehlerhafte* Lösungsbeispiele) anhand schulischer Szenarien in ein narratives Format eingebettet wurden und an Musterlösungen kontrastiert werden. Dabei werden verschiedene Maßnahmen zur instruktionalen Unterstützung erprobt.

7.2 Theoretischer Hintergrund

7.2.1 Lernen aus Fehlern

Advokatorisches Lernen aus Fehlern basiert auf dem Erwerb *negativen Wissens* (Oser, Hascher & Spychiger, 1999). Dieses umfasst Wissen darüber, wie etwas *nicht ist* (deklaratives negatives Wissen), wie etwas *nicht funktioniert* (prozedurales negatives Wissen) und welche Strategien nicht zu einer Lösung führen (negatives strategisches Wissen; vgl. Oser, 2007, S. 4). Die Taxonomie von Oser (ebd.) kann mit den Definitionen *positiver* Wissensarten von De Jong und Ferguson-Hessler (1996, s. Kap. 2) kontrastiert werden. Diese unterscheiden *konzeptuelles Wissen* über Fakten und Konzepte, *prozedurales Wissen* über adäquate Handlungen in einer Domäne und *strategisches Wissen* über Handlungspläne, die die einzelnen Handlungsschritte zur Lösung eines Problems in eine zielführende Sequenz bringen. Anwendbarkeit wird nach Krause (2007) als eine *Qualität* von Wissen betrachtet, die das Zusammenwirken dieser drei Wissensarten bei der Lösung von Problemstellungen beschreibt und die Ableitung konkreter Handlungen zur Bearbeitung einer Problemstellung erlaubt. Anhand des Begriffs der Anwendbarkeit kann auch negatives Wissen beschrieben und so z.B. von trägem Wissen unterschieden werden (Gruber & Renkl, 2000; s.a. Kap. 2.2). Negatives Wissen dient zur Abgrenzung falscher von korrekten Konzepten und damit dem Schutz vor Fehlern. Dies erhöht die Wahrscheinlichkeit *richtigen Handelns* (Oser, 2007; Oser et al., 1999; Oser & Spychiger, 2005). Um negatives Wissen zu erwerben, müssen Lernende den Fehler identifizieren, verstehen und anhand der Kontrastierung von falschen und korrekten Konzepten (z.B. Curry, 2004; Oser & Spychiger, 2005) eine Möglichkeit zur Korrektur des Fehlers erwerben (Oser et al., 1999). Fehler werden also nicht „gelöscht“, sondern als *Kontraste* zu positivem Wissen im episodischen Gedächtnis enkodiert (Stark, Kopp & Fischer, 2009). Diese Verknüpfung von negativem und positivem Wissen fördert die Entwicklung von Strategien zur Vermeidung der Fehler (Oser & Spychiger, 2005).

Beim advokatorischen Lernen aus Fehlern kommt als ergänzende Voraussetzungen dazu, dass eine Rollenübernahme durch die Lernenden stattfinden muss. Sie müssen sich mit den Protagonisten, deren Fehler dargestellt werden, identifizieren. Dies erfordert authentische, relevante Problemstellungen.

7.2.2 Lernen in integrierten Lernumgebungen

Diese Voraussetzungen sowie die Wirkmechanismen des Lernens anhand advokatorischer Fehler wurden bei der Entwicklung einer *integrierten Lernumgebung* (Reinmann & Mandl, 2006) berücksichtigt. Die Lernwirksamkeit integrierter Lernumgebungen, die auf der Kombination problemorientierter und instruktionsorientierter Lehr-Lern-Ansätze basieren, wurde bezüglich des Erwerbs anwendbaren Wissens in komplexen Domänen mehrfach gezeigt (Klein, Wagner & Stark, 2012; Krause, Stark & Herzmann, 2011; Stark et al., 2010; Wagner, Klein, Klopp, Puhl & Stark, 2013).

Problemorientierte Designprinzipien präsentieren Problemstellungen häufig in einem authentischen, narrativen Format (z. B. *anchored instruction*; Cogniton and Technology Group at Vanderbilt [CGTV], 1992). Lernen erfolgt *situationsgebunden* (Resnick, 1987), die aktive Auseinandersetzung mit authentischen und relevanten Problemstellungen ist hierbei von zentraler Bedeutung (Gräsel, 1997). Studien im Bereich der Lehrerbildung zeigen, dass solche Lernkonzeptionen insbesondere den Erwerb anwendbaren Wissens (Fölling-Albers, Hartinger & Mörtl-Hafizovic, 2004) sowie motivationale Aspekte (Wagner et al., 2013a) fördern. Ummel (2010) betrachtet problemorientierte Konzeptionen des Fallverstehens als besonders geeignet, um angehende Lehrer im Umgang mit problematischen Situationen in der Schule auszubilden.

Aufgrund der hohen Anforderungen problemorientierten Lernens wird häufig instruktionale Unterstützung angeboten. *Instruktionsorientierte Designprinzipien* eignen sich vor allem zur *systematischen* Wissensvermittlung und zum Erwerb konzeptuellen Wissens (Hmelo-Silver, Duncan & Chinn, 2007; Kirschner, Sweller & Clark, 2006; Wagner et al., 2013b) und werden häufig bei Lernenden mit geringem Vorwissen eingesetzt (Clark & Mayer, 2003; Renkl, 2002). Ein dabei häufig verwendeter Ansatz sind *ausgearbeitete Lösungsbeispiele* (Renkl, 2001).

7.2.3 Ausgearbeitete Lösungsbeispiele

Im Gegensatz zu reinen Problemlöseaufgaben (vgl. Reinmann & Mandl, 2006) werden dem Lernenden beim Lernen anhand ausgearbeiteter Lösungsbeispiele Problemstellungen *und* der dazugehörige Lösungsweg sowie die Lösung selbst möglichst detailliert präsentiert. Aufgabe des Lernenden ist es, die Modelllösungen zu elaborieren (z.B. Stark, Mandel, Gruber & Renkl, 2002). Oft wird dabei eine zusätzliche instruktionale Unterstützung angeboten (z.B. Stark, 1999; 2001). Die Lernwirksamkeit beispielbasierten Lernens ist von der Qualität und Intensität der Beispielelaboration (Schworm & Renkl, 2007; Stark, 1999) bzw. der

Selbsterklärungen der Lernenden (Chi et al., 1989; Renkl, 1997; Rittle-Johnson, 2006) abhängig. Stark (1999) wies einen positiven Zusammenhang von *prinzipienorientierter*, d.h. an den den Beispielen zugrundeliegenden Prinzipien und Konzepten angelehnte Beispielelaboration und dem Lernerfolg nach. Weniger lernwirksam waren Paraphrasierungen und wiederholtes Ablesen der Lösungsbeispiele. Chi et al. (1989) zeigten, dass lernwirksame Selbsterklärungen Aussagen über Prinzipien und Strukturen der Aufgaben sowie Handlungsbedingungen und Bezüge zur Zielsetzung beinhalten.

Die Lernwirksamkeit ausgearbeiteter Lösungsbeispiele ist gut belegt (Atkinson, Derry, Renkl & Wortham, 2000; Hilbert, Renkl, Kessler & Reiss, 2008; Renkl, 2001). Die Überlegenheit des Lernens mit Lösungsbeispielen gegenüber normalen Problemlöseaufgaben konnte mehrfach bestätigt werden (*worked example effect*; Clark & Mayer, 2003). Studien von McLaren, Lim und Koedinger (2008) zeigten, dass Lernende anhand ausgearbeiteter Lösungsbeispiele rund zwanzig Prozent schneller und damit effizienter lernten als mit Problemlöseaufgaben mit Hilfestellungen. Insbesondere beim initialen Lernen in gut strukturierten Domänen (z.B. Renkl & Atkinson, 2002; Stark, 2001) bzw. bei Lernenden mit geringem Vorwissen (Renkl, 1997) wird das Lernen mit Lösungsbeispielen häufig empfohlen, es gibt jedoch auch Belege für die Lernwirksamkeit in komplexeren Domänen (z.B. wissenschaftliches Argumentieren, Schworm & Renkl, 2007).

Die Effekte sind anhand der Cognitive Load-Theorie erklärbar (Sweller, 2005; Sweller, van Merriënboer & Paas., 1998). Bei ausgearbeiteten Lösungsbeispielen werden dem Lernenden sowohl der Zielzustand als auch die notwendigen Zwischenschritte vorgegeben. Dadurch müssen Zwischenschritte oder Teillösungen nicht im Arbeitsgedächtnis behalten werden, was bei geeigneter Gestaltung den *extraneous load*, also die durch ungeeignete Aufgabengestaltung hervorgerufene kognitive Belastung sowie den *intrinsic load*, der auf die Interaktion zwischen den einzelnen Aufgabenelementen zurückgeht, verringert. Die freigewordene Kapazität kann für *germane load*, also für Verstehensprozesse und zum Aufbau geeigneter Schemata genutzt werden (Sweller, 2005).

Ausgearbeitete Lösungsbeispiele eignen sich gut zur Umsetzung des advokatorischen Lernens aus Fehlern nach Oser (2007). Die wesentlichen Schritte einer relevanten, komplexen Problemstellung, einer detaillierten Fehleranalyse zur Sicherung negativen Wissens und die Kontrastierung an einer korrekten Lösung zum Erwerb von Schutzwissen decken sich mit den Designprinzipien ausgearbeiteter Lösungsbeispiele. Die u.a. von Schworm und Renkl (2007) sowie Stark (1999) geforderte Beispielelaboration bezieht sich dabei auf den aktiven Nachvollzug des Fehlers; die für den Kontrast benötigte korrekte Lösung entspricht einer

elaborierten Feedbackmaßnahme. Elaboriertes Feedback wurde in einer Studie von Stark und Fischer (2008) unter Zuhilfenahme von Laut-Denk-Protokollen (Ericsson & Simon, 1993) mit der Tiefe der Beispielerlaboration und damit der Lernwirksamkeit in Verbindung gebracht. Auch in den o.g. Studien von Große und Renkl (2004; 2007) sowie Durkin und Rittle-Johnson (2012) wurde das Lernen aus Fehlern in Form ausgearbeiteter Lösungsbeispiele umgesetzt. Der reine Nachvollzug birgt jedoch die Gefahr einer oberflächlichen Bearbeitung. Diese muss durch geeignete instruktionale Maßnahmen aufgefangen werden, damit die Wirkmechanismen des Lernens aus ausgearbeiteten Lösungsbeispielen zum Tragen kommen können (Stark, 2001). Instruktionale Unterstützung kann als unverzichtbarer Bestandteil integrierter Lernumgebungen (Reinmann & Mandl, 2006) in verschiedenen Formen umgesetzt werden.

7.2.4 Induktion von Selbsterklärungen durch Prompts

Die Qualität und Intensität der Beispielerlaboration (Stark, 1999; 2001) bzw. der Selbsterklärungen der Lernenden (Chi et al., 1989; Renkl, 1997) als Kriterium der Lernwirksamkeit ausgearbeiteter Lösungsbeispiele kann durch indirekte und direkte instruktionale Maßnahmen unterstützt werden. Indirekte Maßnahmen beziehen sich auf das Design der Lösungsbeispiele wie z.B. die Anordnung im Sinne einer *strukturbetonenden Beispielsequenz* (Quilici & Mayer, 2006). Eine solche Aufgabensequenz zeichnet sich dadurch aus, dass deren einzelne Lösungsbeispiele analoge strukturelle Merkmale aufweisen, wie etwa identische Lösungsschritte. Direkte instruktionale Maßnahmen sind z.B. explizite Anleitungen zur Selbsterklärung (Atkinson, Renkl & Merrill, 2003) oder Aufforderungen zur Reflexion über das erworbene Wissen (z. B. *Selbsterklärungsprompts*; Schworm & Renkl, 2007). Beispielsweise werden Lernende aufgefordert, die den ausgearbeiteten Lösungsbeispielen zugrunde liegenden Prinzipien zu elaborieren (Conati & VanLehn, 2000). Derartige Maßnahmen verhindern, dass sich beim reinen Nachvollzug ausgearbeiteter Lösungsbeispiele ein Scheinverständnis einstellt, das zu einer oberflächlichen Rezeption des Lerninhalts führen kann (*Verstehensillusion*, Stark, 2001). Zudem werden Lernende bei ausgearbeiteten Lösungsbeispielen kaum zu Eigenleistung aufgefordert. Sie können ihr neuerworbenes Wissen und ihre neuen Lösungsstrategien nicht selbst auf ihre Wirksamkeit überprüfen, was über einen längeren Zeitraum hinweg die Motivation negativ beeinflussen kann (Clark & Mayer, 2003).

Instruktionale Unterstützung, die Lernende bei der aktiven Wissenskonstruktion unterstützt, ist somit von Vorteil. In der o.g. Studie von Stark, Kopp und Fischer (2011) wurden die Lernenden zur Sicherung des Lernerfolgs durch geeignete Feedbackmaßnahmen unterstützt

und zur *aktiven Selbsterklärung* angeregt (s. a. Curry, 2004; Stark, 2001). Möglichkeiten zur Realisierung solcher Feedbackmaßnahmen bieten *instruktionale Aufforderungen* zum Abruf und zur Rekonstruktion und damit zur Konsolidierung des erworbenen Wissens (*Prompts*; Stark et al., 2011). Diese sind ökonomischer als interaktive und an den individuellen Lernfortschritten orientierte, adaptive Feedbackmaßnahmen. Zudem können sie genutzt werden, um Prozessdaten zu erheben, die weitere Erkenntnisse über die Wirkmechanismen fehlerbasierten Lernens erbringen.

Ein derartiger Ansatz wird u.a. bei Durkin und Rittle-Johnson (2012) verfolgt. Diese induzierten Lernen aus Fehlern im Bereich der Mathematik (Dezimalzahlen) durch sogenannte *reflection prompts*. Dabei wurden Lernende beim Vergleich korrekter und inkorrektter Lösungsbeispiele gezielt unterstützt. Sie wurden u.a. zur Reflexion darüber angeregt, warum ein Beispiel falsch sei, wie sich Beispiele unterschieden und wie sie einem anderen Schüler erklären würden, wo der Fehler lag. Stark, Mandl, Gruber und Renkl (1999) konnten im Bereich der Zinsrechnung im Rahmen der kaufmännischen Erstausbildung positive Effekte von Prompts auf kognitive und metakognitive Prozesse und die kausale Bedeutung von Selbsterklärungen für erfolgreiches Lernen mit Lösungsbeispielen nachweisen. Auch für die komplexere Kompetenz des Argumentierens wurde die Effektivität von Selbsterklärungsprompts nachgewiesen (Schworm & Renkl, 2007).

Dabei sind fokussierte Prompts, die die Aufmerksamkeit der Lernenden auf spezifische Konzepte und Prinzipien des Lernmaterials lenken, besonders lernwirksam und offeneren Formaten überlegen (Berthold & Renkl, 2010). Eine Studie von Berthold, Röder, Knörzer, Kessler und Renkl (2011) zeigte aber auch die Wichtigkeit der Balance fokussierter Prompts bzgl. verschiedener Wissensarten. Auf konzeptuelle Aspekte fokussierte Prompts, die auf das tiefere Verständnis von zugrunde liegenden Konzepten abzielen, förderten hier den Erwerb konzeptuellen Wissens. Dies ging jedoch zu Lasten des anwendbaren Wissens. Prompts sind somit nicht ohne weiteres für alle Wissensdimensionen förderlich und können sich auf einzelne Dimensionen sogar ungünstig auswirken (Heitzmann, 2014). Daher sollten Prompts nicht oder zumindest nicht ausschließlich auf konzeptuelles Wissen fokussieren. Anwendbares Wissen konnte in einer Studie von Berthold, Eysink und Renkl (2009) gefördert werden, in der Lernende theoriebasierte Begründungen für bestimmte Vorgehensweisen formulieren mussten. Basierend auf der in Kap. 3.2.1 dargestellten Wissenstaxonomie von De Jong und Ferguson-Hessler (1996) wurde hier also vor allem strategisches Wissen benötigt. Prompts sollten demzufolge also nicht nur die Elaboration konzeptuellen Wissens anregen, sondern auch handlungsnähere Wissensdimensionen wie prozedurales und strategisches

Wissen berücksichtigen, um im Sinne von Krauses (2007) Konzeption anwendbaren Wissens das Zusammenwirken verschiedener Wissensarten zu unterstützen.

Weitere wichtige Hinweise für die Gestaltung lernwirksamer Prompts ergaben sich in einer Studie von Heitzmann, Fischer und Fischer (2013) zur Förderung der Diagnosekompetenz im Kontext der universitären Lehrerbildung anhand fehlerhafter Lösungsbeispiele. Die dort eingesetzten Fehleranalyse-Prompts fokussierten mehrere Wissensarten und zielten auf die Identifikation und Erklärung instruktionaler Fehler sowie die daraus von den Lernenden zu ziehenden Schlussfolgerungen für weiteres Handeln ab. Die Nutzung dieser Prompts wirkte sich unerwarteterweise negativ auf die Förderung anwendbaren Wissens aus. Daraus lässt sich der Schluss ziehen, dass Prompts auch zu komplex sein können. Prompts sollten zwar die Aufmerksamkeit der Lernenden gezielt auf die Aufgabe, in diesem Fall die Analyse und Erklärung der instruktionalen Fehler, richten, dabei aber nicht im Sinne einer „dual task“-Aufgabe mit spontanen, von den Lernenden generierten Fehlererklärungen interferieren. Lernende dürfen zudem nicht daran gehindert werden, auf ihr Vorwissen zurückzugreifen, d.h. das Material weniger elaborieren als Lernende mit unspezifischen oder ohne Prompts, indem zu viel vorweggenommen wird.

Beim Lernen aus *fehlerhaften* Lösungsbeispielen ist instruktionale Unterstützung besonders angezeigt (Große & Renkl, 2007; Kopp, Stark, Heitzmann & Fischer, 2010). Beim Vergleich inkorrekt und korrekter Lösungsbeispiele, der beim Lernen aus Fehlern unbedingt erforderlich ist, sind Prompts lernwirksam, die eine Reflexion über diesen *Vergleich* anregen (*reflection prompts*; Durkin & Rittle-Johnson, 2012). Hier sind Prompts angezeigt, die sich explizit auf Strategien beziehen, Fehler zu vermeiden (vgl. Oser & Spychiger, 2005). Die Effektivität derartiger Prompts wurde noch nicht untersucht. In diesem Zusammenhang stellt sich insbesondere die Frage, wie intensiv sich Studierende mit Fehlervermeidungsstrategien auseinandersetzen müssen. Angesichts der Komplexität der Problemstellungen beim Lernen aus Fehlern in der Pädagogischen Psychologie ist anzunehmen, dass es nicht ausreicht, Strategien lediglich zu präsentieren. Es erscheint zielführender, Strategien von den Lernenden zusätzlich reproduzieren zu lassen (zur Rekonstruktion von Informationen aus Texten s. Steiner, 2006).

7.3 Ableitung der Fragestellung

Ausgehend von den theoretischen Überlegungen und den Ergebnissen von Studie 1 wurde eine integrierte Lernumgebung basierend auf dem advokatorischen Lernen aus Fehlern entwickelt. Diese soll anhand der Kontrastierung *inkorrekt* an *korrekten* ausgearbeiteten

Lösungsbeispielen in einem authentischen, relevanten Kontext den Erwerb konzeptuellen Wissens über typische Fehler bei der Erklärung schulischer Situationen und der Erwerb von Wissen über Strategien zur Vermeidung dieser Fehler (konzeptuelles Fehlerwissen und strategisches Fehlervermeidungswissen) fördern. Gleichzeitig soll der Erwerb anwendbaren Wissens in Form prozeduralen Wissens zur Identifikation der Fehler (Fehleridentifikationswissen) und der Erwerb von Wissen zur Erstellung korrekter Erklärungen schulischer Situationen (Erklärungswissen) unterstützt werden. Ausgehend von den dargestellten Implikationen instruktionaler Unterstützung wird untersucht, welches Ausmaß instruktionaler Unterstützung nötig ist. Dazu werden zwei Maßnahmen zur instruktionalen Unterstützung (Prompts zur *Präsentation* bzw. *Rekonstruktion* von Vermeidungsstrategien) experimentell variiert. Zudem werden die Experimentalgruppen in Bezug auf die unterschiedenen Wissensarten mit einer Kontrollgruppe verglichen.

7.4 Fragestellung und Hypothesen*

Welchen Einfluss haben a) die Lernumgebung und b) das Ausmaß instruktionaler Unterstützung auf den Erwerb von konzeptuellem Fehlerwissen, strategischem Fehlervermeidungswissen, prozeduralem Fehleridentifikationswissen sowie Erklärungswissen?

1. Es wurde angenommen, dass alle Experimentalgruppen der Kontrollgruppe hinsichtlich dieser Variablen im Nachtest überlegen sind (Haupteffekt Gruppenzugehörigkeit). Die höchste Performanz im Nachtest sollte die Gruppe mit maximaler Unterstützung zeigen, die Gruppen mit mittlerer Unterstützung sollten schlechter abschneiden. Der geringste Performanzunterschied gegenüber der Kontrollgruppe wurde in der Gruppe ohne Unterstützung erwartet (Trendeffekt).
2. Es wurde angenommen, dass hinsichtlich des Erwerbs konzeptuellen Fehlerwissens, strategischen Fehlervermeidungswissens und Erklärungswissens in allen Experimentalgruppen ein Lernfortschritt von Vor- zu Nachtest stattfindet (Haupteffekt Messzeitpunkt).
3. Es wurde zudem erwartet, dass der Lernfortschritt in der Gruppe mit maximaler instruktionaler Unterstützung (Präsentation *und* Rekonstruktion von Vermeidungsstrategien) am größten ist (Interaktionseffekt Gruppenzugehörigkeit und Messzeitpunkt).
4. Bezüglich der subjektiven Dimensionen wurde erwartet, dass der subjektive Lernerfolg sowie die Akzeptanz wie in Studie 1 hoch ausfallen würden. Der höchste subjektive Lernerfolg wurde dabei in der Gruppe mit maximaler Unterstützung erwartet. Auch bezüglich der Akzeptanz wurden Gruppenunterschiede erwartet. Diese sollte in den Gruppen mit höherer instruktionaler Unterstützung höher ausfallen als in der Gruppe ohne Unterstützung, da diese keine Unterstützung bezüglich der in den Wissenstests verlangten Fehlervermeidungsstrategien erhielt.

*Fragestellungen und Hypothesen wurden unverändert aus dem Originalartikel übernommen und lediglich um eine Frage zu den subjektiven Dimensionen erweitert, um die Vergleichbarkeit zu Studie 1 zu erhöhen.

7.5 Methode

7.5.1 Stichprobe und Design

Die Stichprobe umfasste 288 Studierende (187 weiblich; 8 k. A.), die in Pflichtseminaren der bildungswissenschaftlichen Lehramtsausbildung an der Universität des Saarlandes rekrutiert wurden. Die Studierenden befanden sich im ersten Studienabschnitt (Semesterzahl $M = 4.38$; $SD = 2.3$). Das Durchschnittsalter der Teilnehmer lag bei 22.6 ($SD = 5.21$) Jahren. In einem experimentellen Kontrollgruppendesign mit Messwiederholung wurden die Bearbeitung der Lernumgebung und zwei Maßnahmen instruktorischer Unterstützung systematisch variiert. Hierzu wurden die Probanden randomisiert vier Experimentalgruppen (s. Tab. 1) und einer Kontrollgruppe zugewiesen.

Tabelle 1*: Versuchsdesign – experimentelle Variation des Ausmaßes instruktorischer Unterstützung

Instruktionale Unterstützung	Experimentalgruppen			
	EG 1 (MIU)	EG 2 (REK)	EG 3 (PRÄ)	EG 4 (KIU)
	$n = 53$	$n = 63$	$n = 58$	$n = 53$
Prompts zur Präsentation der Vermeidungsstrategien	Ja	Nein	Ja	Nein
Prompts zur Rekonstruktion der Vermeidungsstrategien	Ja	Ja	Nein	Nein

Experimentalgruppe 1 bearbeitete die Lernumgebung mit *maximaler instruktorischer Unterstützung* (= MIU), hier wurden Prompts zu Präsentation *und* Rekonstruktion von Vermeidungsstrategien gegeben ($n = 53$; 40w; 2 k. A. $M_{\text{Alter}} = 22.6$ ($SD = 5.21$) Jahre).

*Tabellen wurden, soweit nicht anders vermerkt, unverändert aus dem Originalartikel übernommen

Die Experimentalgruppe 2 (REK) erhielt nur die Prompts zur *Rekonstruktion* ($n = 63$; 43w; 2 k. A.; $M_{\text{Alter}} = 22.4$ ($SD = 3.44$) Jahre), Experimentalgruppe 3 (PRÄ) nur diejenigen zur Präsentation von Vermeidungsstrategien ($n = 58$; 33w; $M_{\text{Alter}} = 22.6$ ($SD = 3.72$) Jahre). Experimentalgruppe 4 ($n = 53$; 34w; 4 k. A.; $M_{\text{Alter}} = 21.9$ ($SD = 2.20$) Jahre) erhielt *keine instruktionale Unterstützung* (KIU). Die Kontrollgruppe umfasste 61 Studierende (37 w) mit einem Durchschnittsalter von 23.2 ($SD = 4.69$) Jahren. Die Gruppen waren hinsichtlich Alter und Geschlecht vergleichbar ($F(4,275) = 0.76$, $p = .55$; $\chi^2(4) = 7.24$, $p = .12$).

7.5.2 Ablauf der Untersuchung

Das Experiment wurde in den Seminaren von zwei Versuchsleitern durchgeführt, die Standardisierung des Ablaufs erfolgte anhand eines detaillierten Leitfadens für die Versuchsleiter. Alle Vortests wurden am ersten Seminartermin durchgeführt, an den beiden folgenden Seminarterminen bearbeiteten die Studierenden die Trainingsszenarien der Lernumgebung. In der vierten Sitzung erfolgte der Nachtest. Die Seminare fanden wöchentlich statt. Um die Bearbeitungszeit des Trainings zu standardisieren, war die Dauer der Trainingssitzungen auf 90 Minuten (eine volle Seminarsitzung) limitiert. Die Mehrzahl der Studierenden nahm die vollen 90 Minuten in Anspruch. Die Kontrollgruppe bearbeitete schriftliche Aufgaben zu vergleichbaren Inhalten (Lern- und Persönlichkeitstheorien).

7.5.3 Die Fehlermatrix

Zur Systematisierung der in Studie 2 dargestellten Fehler diente eine Fehlermatrix, die auf Grundlage der Ergebnisse von Studie 1 und der von Stark (2005) identifizierten Fehlertypen zwölf relevante Fehler bei der Erklärung schulischer Situationen anhand der Kategorien *Fehlertyp* und *Komplexitätsgrad* definiert.

Basierend auf den Ergebnissen von Stark (ebd.) werden drei Fehlertypen unterschieden, pro Fehlertyp wurden jeweils zwei Fehler aufgenommen. So ergaben sich sechs isolierte Fehler. Fehlertyp 1 (Analyse komplexer schulischer Situationen und Phänomene anhand von Alltagswissen und subjektiven Theorien) umfasst den *fundamentalen Attributionsfehler* (z.B. Ross & Nisbett, 1991) und die *Verwendung einer nicht-wissenschaftlichen Theorie*. Fehlertyp 2 (nicht-optimale Auswahl wissenschaftlicher Theorien, Modelle und Konzepte zur Erklärung) wird anhand der *Verwendung einer ungeeigneten Theorie* (Antezedenzbedingungen der Theorie in der Situation nicht relevant) und die *Verwendung einer nicht-aktuellen Theorie* dargestellt. Fehlertyp 3 (fehlerhafte Bezugnahme auf empirische Evidenz) umfasst die

Zirkelerklärung (s. a. Westermann, 2000) und die *oberflächliche oder unzureichende Interpretation empirischer Befunde*.

Die Kategorie Komplexitätsgrad ordnet die Fehler in drei Abstufungen ansteigender Komplexität an (isolierter Fehler – kombinierter Fehler – komplexer Fehler). Die drei vorgestellten kombinierten Fehler ergeben sich durch sukzessive, nicht-redundante Kombination der Fehlertypen. So wurde z.B. die Kombination des fundamentalen Attributionsfehlers mit einer Zirkelerklärung vorgestellt:

[...] Die wiederholten Angriffe von Schüler X auf Schüler Y deuten auf dessen aggressives Temperament hin. Dieses stark ausgeprägte Persönlichkeitsmerkmal bei X könnte seine wiederholten Übergriffe erklären. [...]

Die drei komplexen Fehler *Premature Closure*, die *monoperspektivische Betrachtung* und eine Kombination dieser beiden bilden die höchste Abstufung der Kategorie Komplexitätsgrad. *Premature Closure* (z.B. Wollman, Eylon & Lawson, 1980) ist das vorschnelle Festlegen auf eine (geeignete oder ungeeignete) Erklärung, woraufhin die Suche nach anderen möglichen Erklärungen eingestellt und Belege für die Erklärung gesucht werden. Gegenteilige Evidenz wird ignoriert. *Premature Closure* ist komplex, da es jedem Fehler machen vorgeschaltet sein kann und so z.B. die Auswahl einer nicht passenden Theorie oder vor die unzulängliche Interpretation empirischer Befunde beeinflusst. Die *monoperspektivische Betrachtung* beinhaltet die Eindimensionalität einer Erklärung. Die zu erklärende Situation wird nur aus einer einzigen Perspektive (z.B. Individuum, Klasse, Elternhaus) betrachtet. Dies reicht jedoch zur umfassenden, erschöpfenden Erklärung eines schulischen Problems oft nicht aus, hier ist die *Triangulation* mehrerer Perspektiven angebracht (z.B. Weinstock, Neuman & Tabak, 2004). Insgesamt wurden den Lernenden damit zwölf relevante Fehler bei der Erklärung schulischer Situationen (inkorrekte Theorieartikulationen) präsentiert.

7.5.4 Die Lernumgebung

Eine ausführliche Beschreibung der Lernumgebung findet sich in Wagner et al. (2014). Die Umsetzung *problemorientierter Designprinzipien* fokussiert insbesondere die Schaffung eines narrativen Rahmens für die Lernumgebung anhand verschiedener Handlungsfiguren. Stellvertretend für den Lernenden macht der junge Lehrer „Roman“ Fehler bei der Erklärung problematischer Schulsituationen. Der Schulpsychologe „Herr Jung“ korrigiert diese anhand einer detaillierten Fehleranalyse. Dieses authentische Lernsetting sollte die Identifikation mit dem Protagonisten erleichtern (vgl. Oser & Spychiger, 2005).

Die Umsetzung *instruktionsorientierter Designprinzipien* erfolgte anhand der Konstruktion der Problemstellungen der Lernumgebung als ausgearbeitete Lösungsbeispiele. Die notwendige Kontrastierung negativen und positiven Wissens erfolgte durch inkorrekte und korrekte Lösungsbeispiele, die als *inkorrekte* und *korrekte Theorieartikulationen* konzipiert wurden. Romans fehlerhafte Erklärung der Situation wird dabei als inkorrekte Theorieartikulation verstanden, während die Erklärung des Schulpsychologen eine korrekte Theorieartikulation darstellt. Die Integration der Designprinzipien wurde durch die Einbettung der ausgearbeiteten Lösungsbeispiele in das narrative Format realisiert. Hierdurch sollte eine intensive Beispielelaboration durch die Lernenden ermöglicht werden (Stark, 2001).

Die Lernumgebung umfasst vier Trainingsszenarien (zu Bullying, Lernmotivation, Gruppenprozessen und Disziplinproblemen).

Die Trainingsszenarien sind im Sinne einer strukturbetonenden Aufgabensequenz (Quilici & Mayer, 1996) jeweils identisch strukturiert und in die beiden Abschnitte *Fehleranalyse* und *Fehlervermeidung* gegliedert (s. Tab. 2).

Tabelle 2: Struktur der Trainingsszenarien

	Präsentation des Problems	Es wird eine schulische Situation geschildert, in der ein Problem auftritt. Inhalte: Bullying, Lernmotivation, Gruppenprozesse und Disziplinprobleme.
Fehleranalyse	Präsentation der fehlerhaften Erklärung	Die fehlerhafte Erklärung von „Roman“ enthält drei Fehler aus der Fehlermatrix; inkorrekte Theorieartikulation
	Fehleranalyse des Schulpsychologen	Hier wird elaboriert, aus welchen Gründen die fehlerhafte Erklärung falsch ist. Die drei Fehler werden einzeln ausführlich behandelt.
	Abruf von konzeptuellem Fehlerwissen	Es werden drei Fragen im Multiple-Choice Format nach der Definition der Fehler gestellt. Je drei Antworten, zwei Distraktoren.

Fehlervermeidung	Präsentation der Musterlösung	Sachverhalt – Theorie – Erklärung; Erklärung der Situation anhand wissenschaftlicher Theorien: Theorieartikulation des Schulpsychologen.				
	Experimentelle Variation	<table border="0"> <tr> <td rowspan="2" style="font-size: 3em; vertical-align: middle;">{</td> <td>Präsentation der Vermeidungsstrategien</td> <td>Pro Fehler wird eine Vermeidungsstrategie vorgestellt.</td> </tr> <tr> <td>Rekonstruktion der Vermeidungsstrategien</td> <td>Pro Strategie wird eine offene Frage zur Rekonstruktion der Strategie gestellt.</td> </tr> </table>	{	Präsentation der Vermeidungsstrategien	Pro Fehler wird eine Vermeidungsstrategie vorgestellt.	Rekonstruktion der Vermeidungsstrategien
{	Präsentation der Vermeidungsstrategien	Pro Fehler wird eine Vermeidungsstrategie vorgestellt.				
	Rekonstruktion der Vermeidungsstrategien	Pro Strategie wird eine offene Frage zur Rekonstruktion der Strategie gestellt.				

Ziel des Abschnitts *Fehleranalyse* ist die Vermittlung konzeptuellen Wissens über Fehler (Fehlerdefinitionen) und prozeduralen Wissens zur Fehleridentifikation. Die Fehleranalyse beginnt mit der Schilderung einer schulischen Situation und deren fehlerhaften Erklärung durch Roman. Die vier Trainingsszenarien präsentieren die zwölf Fehler aus der Fehlermatrix mit ansteigendem Schwierigkeitsgrad (drei pro Szenario). In den ersten beiden Szenarien überwiegen isolierte, in Szenario drei und vier kombinierte und komplexe Fehler. Die in der Fehlermatrix kategorisierten Fehler werden so anhand der Figur „Roman“ operationalisiert. Die inkorrekte Theorieartikulation wird von einem fiktiven Schulpsychologen analysiert. Die detaillierte *Fehleranalyse* des Schulpsychologen beinhaltet die Definition des jeweiligen Fehlers und deren situationsbezogene Konkretisierung anhand der Frage „*Was ist falsch und warum ist das falsch?*“. So wird ein *inhaltlicher* Bezug hergestellt. Zum Abschluss wird das erworbene Wissen mittels dreier Multiple-Choice-Fragen über Fehlerdefinitionen konsolidiert. Als narrativer Anker (CGTV, 1992) fungiert jeweils ein fiktiver vergesslicher Kollege. Im Abschnitt Fehleranalyse wird so vornehmlich negatives deklaratives Wissen vermittelt: Die Lernenden erfahren, was *keine* korrekte Erklärung der Situation ist und welche Fehler zur inkorrekten Theorieartikulation und so zur fehlerhaften Erklärung der Situation geführt haben (Oser, 2007). Da der Erwerb anwendbaren Wissens nicht ohne eine Basis konzeptuellen Wissens auskommt (z. B. Krause, 2007), wurde diese Maßnahme nicht experimentell variiert.

Im Abschnitt *Fehlervermeidung* wird das deklarative negative Wissen über Fehlerdefinitionen anhand des Kontrasts der Fehler an der korrekten Lösung gesichert (Oser & Spychiger, 2005). Das negative Wissen wird mit strategischem Wissen darüber ergänzt, wie fehlerhafte Erklärungen zukünftig vermieden werden können und wird so zu

Schutzwissen (ebd.). Hierzu stellt der Schulpsychologe eine *Musterlösung* zur Erklärung der schulischen Situation vor, die einer *korrekten* Theorieartikulation (Ohlsson, 1992) entspricht und in die Abschnitte Sachverhalt – Theorie – Erklärung gegliedert ist. Auf eine kurze Zusammenfassung der Problemsituation folgt eine Sammlung von Theorien, die das Problem erklären können. In der folgenden korrekten Erklärung der Situation erläutert der Schulpsychologe, welche Erklärungskraft die Theorien jeweils für die Situation haben und stellt Bezüge zwischen den situationalen Fakten und den Theorien her. Die Musterlösung wird als *korrekte* Theorieartikulation verstanden. Das Zusammenwirken konzeptuellen, prozeduralen und strategischen Wissens sollte den Erwerb von *Erklärungswissen* unterstützen, welches zur Erstellung eigener Erklärungen angewendet werden kann.

7.5.5 Experimentelle Variation: Instruktionale Unterstützung

Das Ausmaß instruktionaler Unterstützung wurde experimentell variiert, indem die Lernenden angelehnt an *reflection prompts* (Durkin & Rittle-Johnson, 2012) zwei Arten von Prompts erhielten. Die hier eingesetzten Prompts zur *Präsentation* bzw. *Rekonstruktion* von Strategien zur Fehlervermeidung fokussierten vor dem Hintergrund der Befunde von Heitzmann et al. (2013) strategisches Wissen über die Vermeidung von Fehlern. Dies erfordert auch konzeptuelles Wissen über die behandelten Fehler, wodurch die Balance verschiedener Wissensarten sichergestellt wurde (Berthold et al., 2009; 2011). Durch den Verzicht auf unmittelbare Handlungsableitung wurde im Vergleich zu Heitzmann et al. (2013) die Komplexität der Prompts reduziert, um die Fokussierung auf den von Oser et al. (1999, Oser, 2007) geforderten Kontrast an der korrekten Lösung zu unterstützen.

Die Experimentalgruppe MIU erhielt beide Unterstützungsmaßnahmen. Zunächst wurde nach der Bearbeitung der Musterlösung zu jedem der drei Fehler eines Trainingsszenarios eine Vermeidungsstrategie (Oser & Spychiger, 2005) präsentiert (*Prompts zur Präsentation der Vermeidungsstrategien*). Dieser umfasst eine kurze, prägnante Anleitung, wie der Fehler allgemein (nicht nur im speziellen, vorliegenden Kontext) vermieden werden kann. Dazu werden aus der Fehlerdefinition (identisch zur Fehleranalyse) Schlussfolgerungen zur Vermeidung abgeleitet. Aufgabe der Lernenden war es, die Strategien genau zu studieren. Die Experimentalgruppe PRÄ erhielt ebenfalls die *Prompts zur Präsentation der Vermeidungsstrategien*.

Beispiel Vermeidungsstrategie „Verwendung einer ungeeigneten Theorie“:

Um sicherzustellen, dass eine wissenschaftliche Theorie Erklärungskraft in einer Situation besitzt, muss kritisch überprüft werden, ob die beobachteten Phänomene überhaupt anhand der Theorie erklärt werden können. Dazu ist es nötig, den Wenn – Teil der Theorie mit den Gegebenheiten der Situation zu vergleichen. Stimmen diese überein, ist es je nach Reichweite der Theorie möglich, die Situation anhand der Theorie zu erklären.

Anschließend wurden die Lernenden der Experimentalgruppe MIU aufgefordert, die soeben präsentierten Strategien mit eigenen Worten anhand einer offenen Frage schriftlich zu rekonstruieren (*Prompts zur Rekonstruktion der Vermeidungsstrategien*). Auch diese Sicherungsmaßnahme war narrativ an der Figur des vergesslichen Kollegen verankert. Die Experimentalgruppe REK erhielt ebenfalls die *Prompts zur Rekonstruktion der Vermeidungsstrategien*, da dort vorab aber keine Prompts zur Präsentation gegeben wurden, musste hier keine Rekonstruktions- sondern vielmehr eine *Konstruktionsleistung* bezüglich der Vermeidungsstrategien erbracht werden.

Die Experimentalgruppe KIU bearbeitete nur die Lernumgebung bis zur korrekten Erklärung des Schulpsychologen.

7.5.6 Instrumente

Alter und Geschlecht wurden mittels Fragebögen ermittelt. Die Kontrollvariable *Einstellungen zu Theorien* wurde mittels einer 6-stufigen Skala mit 16 Items (*stimmt genau - stimmt überhaupt nicht*) ermittelt. Inhaltsspezifisches Interesse wurde mittels einer 5-stufigen Skalen basierend auf Items mit inhaltlicher Verankerung erhoben (z.B.: *4. Hohes Interesse: Ich interessiere mich bereits stark für dieses Thema und freue mich über die Gelegenheit, mehr darüber zu erfahren*).

Alle offenen Fragen wurden von zwei geschulten Bewertern ausgewertet. Diese waren nicht darüber informiert, aus welcher Bedingung die auszuwertenden Tests stammten. Abweichungen in den Urteilen wurden im Diskurs erörtert. Die Wissenstests wurden mit Ausnahme des Tests zum prozeduralen Fehleridentifikationswissens (nur Nachtest, s.u.) jeweils im Vor- und Nachtest durchgeführt, zu jedem Messzeitpunkt wurde als Maß der Übereinstimmung der Bewerter die Interrater-Reliabilität (Cohens κ) ermittelt. Zur Überprüfung der internen Validität wurde neben Alter und Geschlecht die Performanz der Probanden in den Vortests (Vorwissen) herangezogen.

Das *konzeptuelle Fehlerwissen* wurde mittels zwölf Multiple-Choice-Fragen mit drei Antwortmöglichkeiten bzgl. der Definitionen der Fehler erhoben. Jede richtige Antwort wurde mit einem Punkt bewertet (z. B. „Wie wird die Zirkelerklärung definiert?“; Cronbachs $\alpha = .80$ (Vortest) bzw. $.91$ (Nachtest); theoret. Maximum: 12 Punkte). Das *strategische Fehlervermeidungswissen* wurde zu jeder der zwölf Fehlerdefinitionen anhand der offenen Frage „Wie können Sie diesen Fehler vermeiden?“ erfasst (Cohens $\kappa = .87$ (Vortest) bzw. $.85$ (Nachtest); theoret. Maximum: 24 Punkte).

Prozedurales Fehleridentifikationswissen wurde nur im Nachtest erfasst, da dieses in der Pilotstudie (Klein et al., 2012) ohne explizites Training nicht nachgewiesen werden konnte. Hierzu wurde eine inkorrekte Theorieartikulation präsentiert; diese beinhaltete vier Fehler (zwei isolierte, einen kombinierten, einen komplexen Fehler). Die Studierenden wurden aufgefordert, die Fehler zu identifizieren. In einem offenen Antwortformat waren maximal 32 Punkte zu erreichen (Cohens $\kappa = .87$).

Das *Erklärungswissen* stellt in Anlehnung an Krause (2007) ein Maß für die Anwendbarkeit des erworbenen Wissens dar und wurde anhand der Qualität der von den Probanden verfassten Theorieartikulationen erfasst. Hierzu mussten die Studierenden mit Hilfe kurzer, vorgegebener Zusammenfassungen verschiedener Theorien eine schulische Situation selbständig erklären. Dazu waren Bezüge zwischen den beschriebenen Fakten der Situation und den Theorien herzustellen. Die von den Studierenden verfassten Theorieartikulationen wurden mit einer Musterlösung verglichen. Für jede Übereinstimmung mit der Musterlösung wurden Punkte vergeben (Cohens $\kappa = .84$ (Vortest) bzw. $.88$ (Nachtest); theoret. Maximum 16 Punkte). Darüber hinaus war es möglich, Punkte für eine sinnvolle wissenschaftliche Argumentation zu erreichen, die nicht in der Musterlösung enthalten war. So sollte eine zu starke Normativität bei der Bewertung vermieden werden.

Subjektiver Lernerfolg und *Akzeptanz* der didaktischen Gestaltung der Lernumgebung (subjektive Dimensionen) wurden in den Experimentalgruppen mittels 6-fach gestufter Skalen mit 15 bzw. 5 Items erfasst. Cronbach's Alpha für die Skalen betrug $.88$ bzw. $.64$. Die Akzeptanz bildet die Bereitschaft der Lernenden zur Nutzung des Angebots ab (Bürg, 2005).

Die Wissensdimensionen korrelierten zu t2 signifikant positiv in mittlerer Höhe (s. Tab. 3).

Tabelle 3: Korrelationsmatrix der Wissenstests zu t2

	strat. Fehlervermeidungs- wissen	Erklärungs- wissen	Fehleridentifikat ionswissen
konz. Fehlerwissen	.55*	.28*	.39*
strat. Fehlervermeidungs- wissen	-	.41*	.63*
Erklärungswissen	-	-	.55*

Anm.: * $p < .01$

7.5.7 Statistische Analysen*

Die Haupt- und Interaktionseffekte von Gruppenzugehörigkeit und Messzeitpunkt wurden mit zweifaktoriellen Varianzanalysen mit Messwiederholung berechnet. Der Effekt der Gruppenzugehörigkeit, also des Ausmaßes instruktorischer Unterstützung auf die Performanz hinsichtlich der einzelnen Nachtestvariablen sowie der subjektiven Dimensionen wurde mittels polynomialer Trendtests in Kombination mit paarweisen geplanten Kontrastanalysen überprüft. Trendtests inspizieren den Verlauf von Faktorstufenmittelwerten und eignen sich dazu, Mittelwerte verschiedener Gruppen anhand von Polynomen höherer Ordnung zu beschreiben. Die anschließenden Kontrastanalysen erlauben auch bei nicht signifikantem Trend Unterschiede zwischen mindestens zwei der untersuchten Gruppen zu identifizieren (Bühner & Ziegler, 2009) und damit eine ökonomischere Auswertung als die Darstellung der Einzelkontraste. Als Maß der Effektstärke wurde das partielle Eta-Quadrat genutzt. Das Signifikanzniveau der Analysen zur Sicherung interner Validität wurde aufgrund der Gleichheitshypothesen auf .20 adjustiert. Fehlende Werte in den Tests wurden listenweise ausgeschlossen. Es wurden nur Probanden mit in die Stichprobe aufgenommen, die das Training in vollem Umfang absolvierten, d.h. an allen vier Sitzungen teilgenommen hatten. Unterschiedliche Gruppengrößen gingen gewichtet mit in die Varianzanalysen ein. Wiesen die Levene-Tests auf Varianzheterogenität hin, wurden die Freiheitsgrade entsprechend adjustiert.

*Die Anmerkungen zu den statistischen Analysen wurden bis auf die Anmerkung zu Einzelkontrasten unverändert aus dem Originalartikel übernommen

7.6 Ergebnisse*

7.6.1 Analysen zur internen Validität

Die Gruppen waren hinsichtlich des Vorwissens vergleichbar (*konzeptuelles Fehlerwissen* $F(4,278) = 2.10, p = .29$; *strategisches Fehlervermeidungswissen* $F(4,273) = 1.06, p = .38$; *Erklärungswissen* $F(4,266) = 0.10, p = .98$). Auch hinsichtlich der Kontrollvariablen *Einstellungen zu Theorien* $F(4,188) = 1.30, p = .24$) und *Inhaltliches Interesse* $F(4,188) = 1.30, p = .39$) zeigten sich keine Unterschiede.

7.6.2 Effekte von Lernumgebung und instruktionaler Unterstützung

Konzeptuelles Fehlerwissen

Tabelle 4 sind die deskriptiven Statistiken zu entnehmen. Alle Experimentalgruppen zeigten Nachttestleistungen im oberen Drittel der maximal zu erreichenden Punkte, dabei erzielte Experimentalgruppe MIU den höchsten Wert. Hinsichtlich des Erwerbs konzeptuellen Fehlerwissens trat wie erwartet ein signifikanter und großer Effekt der Gruppenzugehörigkeit auf ($F(4,242) = 14.3, p < .001, \eta_p = .90$). Die Anordnung der Gruppenmittelwerte zu t2 kann durch einen kubischen Trend beschrieben werden ($F(1,249) = 10.42, p < .001, \eta_p = .03$), es lagen also signifikante Unterschiede zwischen mindestens zwei Gruppen vor. Der Effekt war klein. Die Vergleiche der Gruppen REK und PRÄ ($t(116) = -0.30; p = .77$) und der Gruppen REK und KIU ($t(112) = 0.38; p = .71$) waren nicht signifikant. Auch der Vergleich von PRÄ und KIU war nicht signifikant ($t(108) = 0.72; p = .47$). Alle anderen geplanten Kontraste waren signifikant (alle $p < .01$). Damit waren die Experimentalgruppe MIU allen Gruppen, die Experimentalgruppen REK, PRÄ und KIU jedoch nur jeweils der Kontrollgruppe signifikant überlegen.

Ebenso zeigte sich wie erwartet ein signifikanter und großer Effekt des Messzeitpunktes ($F(1,242) = 331, p < .001, \eta_p = .58$), alle Experimentalgruppen verbesserten sich deutlich von t1 nach t2. Auch die Interaktion von Gruppenzugehörigkeit und Messzeitpunkt war erwartungsgemäß signifikant, der Effekt war groß ($F(4,242) = 14.9, p < .001, \eta_p = .20$).

*Die Ergebnisse wurden, soweit nicht anders vermerkt, unverändert aus dem Originalartikel übernommen und lediglich um Analysen zu den subjektiven Dimensionen erweitert.

Tabelle 4: *konzeptuelles Fehlerwissen* -- Mittelwerte und Standardabweichungen

Messzeit- punkt	MIU <i>n</i> = 49 <i>M</i> (<i>SD</i>)	REK <i>n</i> = 61 <i>M</i> (<i>SD</i>)	PRÄ <i>n</i> = 57 <i>M</i> (<i>SD</i>)	KIU <i>n</i> = 52 <i>M</i> (<i>SD</i>)	KG <i>n</i> = 28 <i>M</i> (<i>SD</i>)
t1	5.14 (3.01)	4.11 (2.84)	5.63 (3.51)	4.50 (3.15)	4.36 (2.95)
t2	11.0 (1.12)	9.97 (2.96)	10.1 (2.33)	9.62 (2.62)	4.43 (3.52)

Anm.: theoret. Maximum: 12 Punkte

Strategisches Fehlervermeidungswissen

Die Gruppenmittelwerte waren unterschiedlich weit vom theoretischen Maximum entfernt, die höchste Punktzahl erreichte die Experimentalgruppe MIU (s. Tab. 5). Hinsichtlich des Erwerbs strategischen Fehlervermeidungswissens wurde wie erwartet ein signifikanter und großer Effekt der Gruppenzugehörigkeit festgestellt ($F(4,233) = 26.6, p < .001, \eta_p = .70$). Die Anordnung der Gruppenmittelwerte zu t2 kann durch einen linearen Trend beschrieben werden ($F(1,243) = 160, p < .001, \eta_p = .40$), auch hier gab es signifikante Unterschiede zwischen mindestens zwei Gruppen. Der Effekt war groß. Bis auf den Vergleich der Experimentalgruppen REK und PRÄ ($t(112) = 1.31; p = .19$) waren alle geplanten paarweisen Kontraste signifikant (alle $p < .02$). Der Effekt des Messzeitpunktes war erwartungsgemäß signifikant und groß ($F(1,233) = 428, p < .001, \eta_p = .65$), alle Experimentalgruppen verbesserten sich von t1 nach t2. Erwartungsgemäß zeigte sich hinsichtlich der Interaktion von Gruppenzugehörigkeit und Messzeitpunkt ein signifikanter und großer Effekt, auch hier erzielte die Experimentalgruppe MIU den größten Lernfortschritt ($F(4,233) = 45.2, p < .001, \eta_p = .44$).

Tabelle 5: *strategisches Fehlervermeidungswissen* -- Mittelwerte und Standardabweichungen

Messzeit- punkt	MIU <i>n</i> = 49 <i>M</i> (<i>SD</i>)	REK <i>n</i> = 57 <i>M</i> (<i>SD</i>)	PRÄ <i>n</i> = 55 <i>M</i> (<i>SD</i>)	KIU <i>n</i> = 49 <i>M</i> (<i>SD</i>)	KG <i>n</i> = 28 <i>M</i> (<i>SD</i>)
t1	1.67 (2.99)	1.32 (2.02)	1.89 (2.60)	1.00 (1.61)	1.75 (1.94)
t2	13.9 (5.50)	9.53 (4.65)	8.15 (5.01)	5.44 (4.45)	0.96 (1.42)

Anm.: theoret. Maximum: 24 Punkte

Prozedurales Fehleridentifikationswissen

Die maximal erreichte Punktzahl lag deutlich über der Hälfte des theoretischen Maximums und wurde auch hier von der Experimentalgruppe MIU erzielt (s. Tab. 6). Auch beim prozeduralen Fehleridentifikationswissen trat erwartungsgemäß ein signifikanter und großer Effekt der Gruppenzugehörigkeit auf ($F(4,262) = 49.7, p < .001, \eta_p = .43$). Die Anordnung der Gruppenmittelwerte zu t2 kann auch hier durch einen linearen Trend beschrieben werden ($F(1,266) = 192, p < .001, \eta_p = .42$), es lagen also signifikante Unterschiede zwischen mindestens zwei Gruppen vor. Auch dieser Effekt war groß. Bis auf den Vergleich von Experimentalgruppe REK mit Experimentalgruppe PRÄ ($t(114) = 0.84; p = .40$) waren alle geplanten paarweisen Kontraste signifikant (alle $p < .02$).

Tabelle 6: *prozedurales Fehleridentifikationswissen* -- Mittelwerte und Standardabweichungen

	MIU	REK	PRÄ	KIU	KG
Messzeit-	$n = 53$	$n = 63$	$n = 58$	$n = 52$	$n = 41$
punkt	$M (SD)$				
t2	19.2 (5.64)	15.0 (5.73)	14.1 (6.52)	11.5 (4.48)	4.12 (2.92)

Anm.: theoret. Maximum: 32 Punkte

Erklärungswissen

Tabelle 7 sind die deskriptiven Statistiken zu entnehmen, auch hier erzielte die Experimentalgruppe MIU den höchsten Wert. Hinsichtlich des Erklärungswissens zeigte sich erwartungsgemäß ein signifikanter und großer Effekt der Gruppenzugehörigkeit ($F(4,245) = 22.3, p < .001, \eta_p = .27$). Auch bei dieser Variable kann die Anordnung der Gruppenmittelwerte zu t2 durch einen linearen Trend beschrieben werden ($F(1,266) = 187, p < .001, \eta_p = .40$), signifikante Unterschiede zwischen mindestens zwei Gruppen lagen vor. Der Effekt war ebenfalls groß. Alle geplanten paarweisen Kontraste waren signifikant (alle $p < .03$).

Erwartungsgemäß war der Effekt des Messzeitpunktes signifikant und groß ($F(1,245) = 228, p < .001, \eta_p = .48$), alle Experimentalgruppen verbesserten sich von t1 nach t2. Auch der Interaktionseffekt von Gruppenzugehörigkeit und Messzeitpunkt war wie erwartet signifikant und groß, die Experimentalgruppe MIU erzielte hier ebenso den größten Lernfortschritt ($F(4,245) = 38.7, p < .001, \eta_p = .39$).

Tabelle 7: *Erklärungswissen* -- Mittelwerte und Standardabweichungen

Messzeit- punkt	MIU <i>n</i> = 52 <i>M</i> (<i>SD</i>)	REK <i>n</i> = 63 <i>M</i> (<i>SD</i>)	PRÄ <i>n</i> = 57 <i>M</i> (<i>SD</i>)	KIU <i>n</i> = 50 <i>M</i> (<i>SD</i>)	KG <i>n</i> = 28 <i>M</i> (<i>SD</i>)
t1	4.04 (2.21)	3.82 (1.67)	3.87 (1.53)	3.90 (1.91)	4.11 (2.36)
t2	10.8 (3.31)	7.79 (3.22)	6.61 (2.79)	5.41 (1.67)	3.41 (2.15)

Anm.: theoret. Maximum: 16 Punkte

7.6.3 Subjektive Dimensionen

Subjektiver Lernerfolg

Tabelle 8 zeigt die deskriptiven Ergebnisse. Die Unterschiede zwischen den Experimentalgruppen waren minimal, der Trendtest erreichte keine statistische Signifikanz ($p = .32$). Der erwartete Vorteil der Experimentalgruppe MIU blieb aus. Alle Gruppen schätzten aber alle ihren Lernerfolg im oberen Drittel der Skala (im Schnitt 4 von 6 Punkten) ein.

Tabelle 8: *Subjektiver Lernerfolg* - Mittelwerte und Standardabweichungen

MIU <i>n</i> = 51 <i>M</i> (<i>SD</i>)	REK <i>n</i> = 61 <i>M</i> (<i>SD</i>)	PRÄ <i>n</i> = 52 <i>M</i> (<i>SD</i>)	KIU <i>n</i> = 51 <i>M</i> (<i>SD</i>)
4.29 (.56)	4.15 (.51)	4.28 (.53)	4.22 (.52)

Anm.: theoret. Maximum: 6 Punkte. Erhebung zu t2

Akzeptanz

Die deskriptiven Ergebnisse finden sich in Tabelle 9. Auch hier fallen die deskriptiven Unterschiede gering aus, es zeigte sich erwartungsgemäß ein signifikanter kubischer Trend für die Anordnung der Gruppenmittelwerte ($F(1,210) = 7.92$; $p = .004$; $\eta_p^2 = .06$). Es lagen also signifikante Unterschiede zwischen mindestens zwei Gruppen vor. Der Effekt war klein. Die Vergleiche der Gruppen MIU und REK ($t(210) = .530$; $p = .60$) und der Gruppen MIU und KIU ($t(210) = -1.92$; $p = .06$) waren nicht signifikant, ebenso der Vergleich von PRÄ und KIU ($t(210) = .690$; $p = .49$). Alle anderen geplanten Kontraste waren signifikant (alle $p < .02$).

Tabelle 9: Akzeptanz - Mittelwerte und Standardabweichungen

MIU	REK	PRÄ	KIU
$n = 51$	$n = 59$	$n = 53$	$n = 51$
$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$
4.36 (.64)	4.3 (.55)	4.66 (.53)	4.58 (.58)

Anm.: theoret. Maximum: 6. Erhebung zu t2

Korrelationen

Die Korrelationen zwischen der subjektiven Einschätzung des Lernerfolgs sowie der Akzeptanz und den Ergebnissen der Wissenstests zu t2 findet sich in Tabelle 10. Subjektiver Lernerfolg und Akzeptanz korrelierten signifikant in mittlerer Höhe. Der subjektive Lernerfolg korrelierte außerdem signifikant mit konzeptuellem Fehlerwissen und strategischem Fehlervermeidungswissen, während die Akzeptanz nur mit ersterem korrelierte.

Tabelle 10: Korrelationsmatrix zu t2

	Akzeptanz	konz. Fehlerwissen	strat. Fehlervermei- dungswissen	Erklärungs- wissen	Fehleridentifi- - kationswissen
Subjektiver Lernerfolg	.42**	.21**	.29**	.01	.12
Akzeptanz	-	.17*	.09	-.09	-.05

**= $p < .001$; *= $p < .05$

7.7 Diskussion

7.7.1 Interne Validität

Aufgrund der Randomisierung, des standardisierten Vorgehens und der vergleichbaren Lernzeit sowie der Vergleichbarkeit der Gruppen im Vorwissen kann die interne Validität der Studie als gesichert angesehen werden. Die Ergebnisse der Kontrollvariablen Alter, Geschlecht, Einstellungen zu Theorien und Inhaltliches Interesse stützen diese Annahme. Testeffekte zwischen Vor- und Nachtest können weitgehend ausgeschlossen werden, da zwischen den Messzeitpunkten eine Zeitspanne von vier Wochen lag.

7.7.2 Effekte von Lernumgebung und instruktionaler Unterstützung

Die Lernwirksamkeit der Lernumgebung konnte bezüglich aller getesteten Wissensarten nachgewiesen werden. Alle Experimentalgruppen erzielten signifikante Lernfortschritte von Vor- zu Nachtest und waren der Kontrollgruppe im Nachtest überlegen. Ebenso wurden die Hypothesen zum Ausmaß der instruktionalen Unterstützung bestätigt, die Anordnung der Gruppenmittelwerte in den einzelnen Nachtests entsprach den Erwartungen. Die Ergebnisse belegen, dass *beide* Maßnahmen instruktionaler Unterstützung lernwirksam sind und ihre *Kombination* den größten Lerneffekt bewirkt. Die Gruppe mit maximaler instruktionaler Unterstützung erzielte den größten Lernfortschritt und war nach dem Training besser in der Lage, schulische Problemsituationen zu erklären, als die weniger intensiv unterstützten Gruppen. Angesichts der durchweg großen Effektstärken kann hier von substanziellen, praktisch bedeutsamen Unterschieden ausgegangen werden.

Im *konzeptuellen Fehlerwissen* zeigten sich keine signifikanten Unterschiede zwischen den Gruppen, die als instruktionale Unterstützung nur eine oder keine der Maßnahmen erhielten. Bezüglich des Erwerbs konzeptuellen Fehlerwissens erfordert eine deutliche Zunahme der Lernwirksamkeit offenbar die Kombination beider Maßnahmen. Bei der Interpretation dieser Befunde muss jedoch berücksichtigt werden, dass *alle* Gruppen die Maßnahme zum Abruf konzeptuellen Wissens während des Trainings erhielten. So sollte eine Wissensbasis zum Erwerb anwendbaren Wissens gesichert werden (vgl. Kapitel 3.2.2.; Krause, 2007).

Diese unter theoretisch-didaktischen Gesichtspunkten notwendige Maßnahme konfundierte aber unter Umständen die Erfassung des konzeptuellen Fehlerwissens im Nachtest. Die erreichte Punktzahl im Nachtest war in allen Gruppen mit Ausnahme der Kontrollgruppe sehr hoch, was die Varianz in dieser Wissensdimension eingeschränkt. Um eine größere Varianz

zwischen den Gruppen zu erzeugen, könnte in zukünftigen Untersuchungen die Schwierigkeit dieses Tests erhöht werden.

Betrachtet man den Lernfortschritt von t1 nach t2, fällt auf, dass die Gruppe mit maximaler Unterstützung und die Gruppe mit Rekonstruktion der Vermeidungsstrategien numerisch den gleichen Lernfortschritt zeigen. Allerdings fällt dieser in der Gruppe mit maximaler Unterstützung homogener aus, die Werte streuen zu t2 in geringerem Maß um den – im Vergleich signifikant höchsten – Mittelwert. Die Kombination der Unterstützungsmaßnahmen bewirkte also eine größere Leistungsdichte. Erklärbar wird dieses Ergebnis unter der Annahme, dass durch die zweimalige Beschäftigung mit den auf den Fehlerdefinitionen basierenden Fehlervermeidungsstrategien bei intensiver Unterstützung auch in Bezug auf das konzeptuelle Fehlerwissen eine stärkere Beispielelaboration (Stark, 2001) angeregt wurde.

Interessanterweise unterschieden sich die Gruppen, die jeweils nur eine Unterstützungsmaßnahme erhielten, beim Erwerb von *strategischem Fehlervermeidungswissen* und *prozeduralem Fehleridentifikationswissen* nicht signifikant. Es macht demnach bei beiden Wissensdimensionen keinen Unterschied, ob zur instruktionalen Unterstützung Prompts zum Abruf *oder* zur Rekonstruktion der Vermeidungsstrategien angeboten werden, beide scheinen hier gleich lernwirksam zu sein. Dies ist ein Beleg für einfach zu realisierende Unterstützungsmaßnahmen beim Textverständnis (s. z.B. Steiner, 2006), steht aber im Widerspruch zur Annahme, dass die Kombination der instruktionalen Maßnahmen, ähnlich wie beim konzeptuellen Fehlerwissen, die Beispielelaboration stärker fördert als eine einzelne Maßnahme.

Der Erwerb von *Erklärungswissen* wurde anhand der Qualität der von den Studierenden verfassten Theorieartikulationen erfasst. Die Experimentalgruppe, die nur die Prompts zur Rekonstruktion der Vermeidungsstrategien erhielt, erzielte signifikant bessere Ergebnisse als die Gruppe, der nur die Strategien präsentiert wurden. Erstere musste die Strategien selbst konstruieren. Aus konstruktivistischer Perspektive spricht dieser Befund für die Lernwirksamkeit expliziter Konstruktionsleistungen, die effektiver ist als die reine Rezeption von Informationen (vgl. Reinmann & Mandl, 2006). Zudem ist anzunehmen, dass durch diese Maßnahme qualitativ hochwertigere und damit lernwirksamere Selbsterklärungen induziert wurden, da keine Lösung vorgegeben war und damit ein Scheinverständnis (Stark, 1999) vermieden wurde. Selbsterklärungen sind ebenso wie Beispielelaboration erwiesenermaßen äußerst lernwirksame Mechanismen beim Lernen mit Lösungsbeispielen (Schworm & Renkl, 2007; Stark, 2001).

Insgesamt zeigen die Ergebnisse dieser Studie, dass die Kontrastierung inkorrektur an korrekten Konzepten den Erwerb der untersuchten Wissensarten auch in der Domäne der Pädagogischen Psychologie unterstützt. Durch die explizite Unterstützung von Fehlervermeidungsstrategien kann das Erstellen *korrekter Erklärungen schulischer Situationen* auf ökonomische Weise substantiell gefördert werden. Die Befunde sind konform mit der Forschung zum Lernen aus Fehlern (Durkin & Rittle-Johnson, 2012; Oser & Spychiger, 2005) und mit Studien zur Effektivität integrierter Lernumgebungen (Stark et al., 2010; Krause et al., 2011) und liefern einen weiteren Beleg für die Notwendigkeit adäquater instruktionaler Unterstützung beim Lernen in komplexen Domänen.

7.7.3 Subjektive Dimensionen

Bei der subjektiven Einschätzung des Lernerfolgs zeigte sich lediglich ein Zusammenhang mit den deklarativen Wissensarten. Offenbar hatten die Lernenden Schwierigkeiten damit, ihren Lernerfolg beim prozeduralen Fehleridentifikationswissen und in der Anwendungsdimension (Erklärungswissen) richtig einzuschätzen. Ein Grund dafür dürfte sein, dass die Studierenden in allen Experimentalgruppen bezüglich der Fehlerdefinitionen mehrfach die korrekten Lösungen erhielten und auf dieser Grundlage ihr Abschneiden in den Tests zum konzeptuellen Fehlerwissen besser beurteilen konnten. Zumindest für die Experimentalgruppen mit einer oder der Kombination der Unterstützungsmaßnahmen gilt das auch für die Fehlervermeidungsstrategien. Fehlervermeidungsstrategien und damit das strategische Fehlervermeidungswissen. Ein derartiges Feedback, das 1:1 auf die Wissenstests übertragbar ist, existierte nicht für das prozedurale Fehleridentifikationswissen und das Erklärungswissen.

Die Akzeptanz des didaktischen Ansatzes erscheint dagegen relativ unabhängig vom Abschneiden in den Wissenstests. Offenbar war der subjektive Lernerfolg, der ja auch mit der Akzeptanz korrelierte, in jedem Fall hoch genug, dass der Ansatz durch die Lernenden gut angenommen wurde. Erwartungsgemäß zeigten sich hier jedoch geringe, aber signifikante Unterschiede zwischen den Experimentalgruppen. Insgesamt berichtete die Gruppe, die die Strategien zur Fehlervermeidung selbst konstruieren musste, die höchste Akzeptanz. Die deskriptiv und im inferenzstatistischen Vergleich geringste Akzeptanz zeigte sich dagegen wider Erwarten nicht in der Gruppe, die keine Unterstützung erhielt, sondern in der Gruppe, in der die Fehlervermeidungsstrategien präsentiert wurden. Hier lässt sich wiederum spekulieren, ob die aktive Konstruktion von Wissen, auch ohne spezifische Aufforderung, auch in Hinsicht auf motivationale Eigenschaften der Lernumgebung wirksamer ist als der

reine Nachvollzug. Diese These wird unter anderem von Clark und Mayer (2003) gestützt, die auf die üblicherweise geringe Eigenleistung Lernender bei der Arbeit mit ausgearbeiteten Lösungsbeispielen hinweisen, welche die Motivation negativ beeinflussen kann.

7.7.4 Pädagogische Konsequenzen und weitere Forschung

Die Ergebnisse sprechen für eine systematische Implementation der Lernumgebung in die Lehramtsausbildung. Mit dem hier verwendeten Ansatz können mit vergleichsweise geringem Aufwand gute Lernergebnisse besonders im Hinblick auf den Erwerb anwendbaren Wissens erzielt werden. Die Befunde müssen allerdings in weiteren Studien repliziert werden, dabei ist insbesondere die Untersuchung der Stabilität der erzielten Effekte (z. B. durch Follow-Up-Tests am Semesterende) von Interesse. Gezieltere Analysen *motivationaler Bedingungen* der Lernwirksamkeit der Unterstützungsmaßnahmen (ATI-Effekte) und potenzieller *einstellungsbezogener Effekte* der Arbeit mit der Lernumgebung (z. B. positive Veränderung der Einstellung gegenüber wissenschaftlichen Theorien) sollten dabei das Untersuchungsdesign ergänzen. Ebenso müssen die der Lernumgebung zugrunde liegenden Wirkmechanismen systematisch untersucht werden. Eine vielversprechende Analyseperspektive bieten hier Laut-Denk-Protokolle (Ericsson & Simon, 1993) und eine Triangulation qualitativer und quantitativer Verfahren. Mit Blick auf die beobachteten Effekte in der Gruppe mit aktiver Wissenskonstruktion bietet sich zudem eine Untersuchung von Effekten der selbständigen Konstruktion von Trainingsszenarien nach dem Muster der Lernumgebung durch die Lernenden an. Möglicherweise könnte dadurch auch in den Anwendungsdimensionen der Lernerfolg noch gesteigert werden. Um die Lernwirksamkeit der Lernumgebung selbst weiter zu steigern, sind verschiedene Optimierungsmaßnahmen vorgesehen. In einer Folgestudie soll eine kollaborative Bearbeitung der Lernumgebung untersucht werden. Zudem ist die Integration weiterer Themen aus dem Bereich Analyse von Unterricht (z. B. Prüfungs- und Leistungsangst) vorgesehen. Eine Limitierung der Studie stellt neben dem Erreichen der Testdecke beim konzeptuellen Wissen z.B. die Übertragung der Befunde auf die tatsächliche Schulpraxis dar. Hier wäre in zukünftigen Studien anhand einer Stichprobe von Lehrern zu untersuchen, inwiefern sich das Trainingsprogramm auf deren Erklärungen schulischer Probleme auswirkt und ob eine derartige Förderung der Anwendung wissenschaftlichen Wissens tatsächlich die in der Literatur (z.B. Bromme, 2008; Meier, 2006; Neuweg, 2007) beschriebenen Effekte auf die Unterrichtsqualität hat.

Literatur*

- Atkinson, R. K., Derry, S. J., Renkl, A. & Wortham, D. (2000). Learning from Examples: Instructional Principles from the Worked Examples Research. *Review of Educational Research, 70* (2), 181-214. doi:10.3102/00346543070002181.
- Atkinson, R., Renkl, A. & Merrill, M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology, 95*(4), 774-783.
- Berthold, K., Eysink, T. H. S. & Renkl, A. (2009). Assisting Self-Explanation Prompts Are More Effective than Open Prompts when Learning with Multiple Representations. *Instructional Science: An International Journal of the Learning Sciences, 37*(4), 345-363.
- Berthold, K. & Renkl, A. (2010). How to Foster Active Processing of Explanations in Instructional Communication. *Educational Psychology Review, 22*(1), 25-40.
- Berthold, K., Röder, H., Knörzer, D., Kessler, W. & Renkl, A. (2011). The double-edged effects of explanation prompts. *Computers in Human Behavior, 27*(1), 69-75.
- Bromme, R. (2008). Lehrerexpertise. In W. Schneider & M. Hasselhorn (Hrsg.). *Handbuch der Pädagogischen Psychologie* (S. 159-167). Göttingen: Hogrefe.
- Bühner, M. & Ziegler, M. (2009). *Statistik für Sozialwissenschaftler und Psychologen*. München: Addison Wesley.
- Bürg, O. (2005). Akzeptanz von E-Learning in Unternehmen. Die Bedeutung von institutionellen Rahmenbedingungen, Merkmalen des Individuums und Merkmalen der Lernumgebung für die Akzeptanz von E-Learning. Berlin: Logos.
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P. & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.
- Clark, R. C. & Mayer, R. E. (2003). *e-Learning and the science of instruction*. San Francisco: Pfeiffer.
- Cognition and Technology Group at Vanderbilt (1992). The jasper series as an example of anchored instruction: Theory, program, description, and assessment data. *Educational Psychologist, 27*, 291-315. doi:10.1207/s15326985ep2703_3.

*Aus Gründen der Übersichtlichkeit wird die Literatur für dieses Kapitel, analog zu den Aufsätzen, in einem eigenen Verzeichnis aufgeführt.

- Conati, C. & VanLehn, K. (2000). Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation. *International Journal of Artificial Intelligence in Education*, 11, 398-415.
- Curry, L. 2004. The effects of self-explanations of correct and incorrect solutions on algebra problem solving performance. In K. Forbus, D. Gentner, and T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, p. 1548. Mahwah, NJ: Erlbaum.
- De Jong, T. & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31, 105-113. doi:10.1207/s15326985ep3102_2.
- Durkin, K. & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction* 22, 206-214. doi:10.1016/j.learninstruc.2011.11.001.
- Ericson, A. & Simon, H. A. (1993). *Protocol Analysis. Verbal reports as data*. Cambridge: MIT Press.
- Fölling-Albers, M., Hartinger, A. & Mörtl-Hafizović, D. (2004). Situiertes Lernen in der Lehrerbildung. *Zeitschrift für Pädagogik*, 50, 727-747.
- Gräsel, C. (1997). *Problemorientiertes Lernen*. Göttingen: Hogrefe.
- Große, C. S. & Renkl, A. (2004). Learning from worked examples: What happens if errors are included? In P. Gerjets, J. Elen, R. Joiner, & P. Kirschner (Eds.), *Instructional design for effective and enjoyable computer-supported learning* (s. 356-364). Tübingen: Knowledge Media Research Center.
- Große, C. S., & Renkl, A. (2007). Finding and fixing errors in worked examples: can this foster learning outcomes? *Learning and Instruction*, 17, 612-634. doi:10.1016/j.learninstruc.2007.09.008.
- Gruber, H. & Renkl, A. (2000). Die Kluft zwischen Wissen und Handeln: Das Problem des trägen Wissens. In G. H. Neuweg (Hrsg.), *Wissen-Können-Reflexion. Ausgewählte Verhältnisbestimmungen* (S. 155-174). Innsbruck: Studienverlag.
- Heitzmann, N. (2014). *Fostering diagnostic competence in different domains*. Dissertation, Ludwig-Maximilians-University, Munich, Germany.
- Heitzmann, N., Fischer, F. & Fischer, M. (2013). Teachers' diagnostic competence: Validation of a model and evaluation of instructional support. 15th Biennial Conference of the European Association for Research in Learning and Instruction (EARLI). München, 27.08.-31.08.2013.

- Hilbert, T. S., Renkl, A., Kessler, S. & Reiss, K. (2008). Learning to prove in geometry: Learning from heuristic examples and how it can be supported. *Learning and Instruction, 18* (1), 54-65. doi:10.1016/j.learninstruc.2006.10.008.
- Hmelo-Silver, C. E., Duncan, R. & Chinn, C. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42* (2), 99-107. doi:10.1080/00461520701263368.
- Huang, T. H., Liu, Y. C. & Shiu, C. Y. (2008). Construction of an online learning system for decimal numbers through the use of cognitive conflict strategy. *Computers & Education, 50*, 61-76 doi:10.1016/j.compedu.2006.03.007.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational psychologist, 41*, 75-86.
- Klein, M., Wagner, K. & Stark, R. (2012). Theoretisieren für die Praxis - Analyse und Förderung der Anwendung wissenschaftlichen Wissens in schulischen Kontexten. *Tagungsband zur 9. Fachtagung Psychologiedidaktik und Evaluation 2012*.
- Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2004). Standards für die Lehrerbildung: Bildungswissenschaften (Beschluss der Kultusministerkonferenz vom 16.12.2004). Bonn: KMK, verfügbar unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf.
- Kopp, V., Stark, R., Heitzmann, N. & Fischer, M. R. (2010). Self-regulated learning with case-based worked examples: effects of errors. *Evaluation & Research in Education, 22* (2-4), 107-119.
- Krause, U.-M. (2007). *Feedback und kooperatives Lernen*. Münster: Waxmann.
- Krause, U.-M., Stark, R. & Herzmann, P. (2011): Förderung anwendbaren Theoriewissens in der Lehrerbildung: Vergleich problembasierter und instruktionsorientierter Lernens. *Psychologie in Erziehung und Unterricht, 58* (2), 106-115. doi:10.2378/peu2011.art27d.
- McLaren, B. M., Lim, S.-J., & Koedinger, K. R. (2008). When is assistance helpful to learning? Results in combining worked examples and intelligent tutoring. In Proceedings of the 9th International Conference on Intelligent Tutoring Systems (pp. 677-680).

- Meier, A. (2006). Theorienutzungskompetenz als Schlüsselqualifikation von Lehrkräften. Überlegungen zu einem Grundauftrag von Pädagogischen Hochschulen. In: Y.M. Nakamura, C. Böckelmann & D. Tröhler (Hrsg.). *Theorie versus Praxis?* (S. 89-106). Zürich: Verlag Pestalozzianum.
- Neuweg, G. H. (2007): *Wie grau ist alle Theorie, wie grün des Lebens goldener Baum? LehrerInnenbildung im Spannungsfeld von Theorie und Praxis*. bwp@, Ausgabe 12. http://www.bwpat.de/ausgabe12/neuweg_bwpat12.pdf [18.10.2014].
- Ohlsson, S. (1992). The cognitive skill of theory articulation: A neglected aspect of science education. *Science & Education*, 1, 181-192. doi:10.1007/BF00572838.
- Oser, F. (2007). Aus Fehlern lernen. In: M. Göhlich, Ch. Wulf & J. Zirfas (Hrsg.): *Pädagogische Theorien des Lernens* (S. 203-212). Weinheim und Basel: Beltz Verlag.
- Oser, F., Hascher, T. & Spychiger, M. (1999). Lernen aus Fehlern. Zur Psychologie des 'negativen' Wissens. In: W. Althof (Hrsg.): *Fehlerwelten* (S. 11-41). Opladen: Leske + Budrich.
- Oser, F. & Spychiger, M. (2005). *Lernen ist schmerzhaft. Zur Theorie des Negativen Wissens und zur Praxis der Fehlerkultur*. Weinheim: Beltz.
- Quilici, J. L. & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology*, 88 (1), 144-161. doi:10.1037//0022-0663.88.1.144.
- Reinmann, G. & Mandl, H. (2006). Unterrichten und Lernumgebungen gestalten. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 613-658). Weinheim: Beltz.
- Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science: A Multidisciplinary Journal*, 21(1), 1-29.
- Renkl, A. (2001). Lernen aus Lösungsbeispielen: Eine Einführung. *Unterrichtswissenschaft*, 29 (1), 2-4.
- Renkl, A. (2002). Worked-out examples: instructional explanations support learning by self-explanations. *Learning and Instruction*, 12 (5), 529-556. doi:10.1016/S0959-4752(01)00030-5.
- Renkl, A. & Atkinson, R. K. (2002). Learning from examples: Fostering self-explanations in computer-based learning environments. *Interactive Learning Environments*, 10 (2), 105-119. doi:10.1076/ilee.10.2.105.7441.

- Resnick L. B. (1987). Learning in School and Out. *Educational Researcher*, 16 (9), 13-20. doi:10.3102/0013189X016009013.
- Rittle-Johnson, B. (2006). Promoting Transfer: Effects of self-explanation and direct instruction. *Child Development*, 77, 1-15. doi:10.1111/j.1467-8624.2006.00852.x.
- Ross, L. & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York, NY: McGraw-Hill.
- Schworm, S. & Renkl, A. (2007). Learning argumentations skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology*, 99, 285-296. doi:10.1037/0022-0663.99.2.285.
- Seidel, T. & Prenzel, M. (2008). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen - Erfassung pädagogisch-psychologischer Kompetenzen mit Videosequenzen. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 201-216). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-90865-6_12
- Star, J. R., & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice teachers' ability to notice. *Journal of Mathematics Teacher Education*, 11, 107-125. doi:10.1007/s10857-007-9063-7.
- Stark, R. (1999). *Lernen mit Lösungsbeispielen*. Göttingen: Hogrefe-Verlag.
- Stark, R. (2001). *Analyse und Förderung beispielbasierten Lernens: Anwendung eines integrativen Forschungsparadigmas*. Unveröffentlichte Habilitation, Ludwig-Maximilians-Universität München.
- Stark, R. (2005). Constructing arguments in educational discourses. In H. Gruber, C. Harteis, R. Mulder & M. Rehl (Eds.), *Bridging individual, organisational, and cultural aspects of professional learning* (pp. 64-71). Regensburg: S. Roderer.
- Stark, R. & Fischer, M. (2008). Förderung diagnostischer Kompetenz durch beispielbasiertes Lernen: Einfluss von Prozessorientierung und Fehler. Unveröffentlichter Arbeitsbericht über die Projektarbeit im Förderzeitraum 1 (FI 720/2-1 und STA 569/3-1).
- Stark, R., Herzmann, P. & Krause, U.-M. (2010). Effekte integrierter Lernumgebungen - Vergleich problembasierter und instruktionsorientierter Seminarkonzeptionen in der Lehrerbildung. *Zeitschrift für Pädagogik*, 56 (4), 548-563.
- Stark, R., Kopp, V. & Fischer, M. R. (2009). Förderung der Diagnosekompetenz bei Studierenden der Medizin durch fallbasierte Lösungsbeispiele: der Einfluss von Fehlern und Feedback. *Psychologie in Erziehung und Unterricht*, 56 (2), 137-149.

- Stark, R., Kopp, V. & Fischer, M. R. (2011). Case-based learning with worked examples in complex domains: Two experimental studies in undergraduate medical education. *Learning and Instruction*, 21, 22-33.
- Stark, R. & Krause, U.-M. (2006). Konzeption einer computerbasierten Lernumgebung zur Förderung von Kompetenzen zum wissenschaftlichen Argumentieren. In G. Krampen & H. Zayer (Hrsg.), *Didaktik und Evaluation in der Psychologie* (S. 218-230). Göttingen: Hogrefe.
- Stark, R., Mandl, H., Gruber, H. & Renkl, A. (1999). Instructional means to overcome transfer problems in the domain of economics: Empirical studies. *International Journal of Educational Research*, 31(7), 591-609.
- Stark, R., Mandl, H., Gruber, H. & Renkl, A. (2002). Conditions and effects of example elaboration. *Learning & Instruction*, 12 (1), 39-60. doi:10.1016/S0959-4752(01)00015-9.
- Steiner, G. (2006). Lernen und Wissenserwerb. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 137-203). Weinheim: Beltz.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E Mayer (Ed.), *The Cambridge handbook of multimedia learning*, (p. 19-30). Cambridge: University Press.
- Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10 (3), 251-296.
- Ummel, H. (2010). Fallverstehen als Gegenstück zum Training von Routinen. *Journal für Lehrerinnen- und Lehrerbildung* 10 (2), 42-47.
- Wagner, K., Klein, M., Klopp, E., Puhl, T. & Stark, R. (2013a). Effects of Instruction-Supported Learning with Worked Examples in Quantitative Method Training. *Higher Education Studies*, 3 (3), 1-12. doi:10.5539/hes.v3n3p1.
- Wagner, K., Klein, M., Klopp, E. & Stark, R. (2014). Theoretisieren für die Praxis: Förderung anwendbaren pädagogischen Wissens anhand advokatorischer Fehler. *Journal für Lehrerinnen- und Lehrerbildung*, im Druck.
- Wagner, K., Stark, R., Daudbasic, J., Klein, M., Krause, U.-M. & Herzmann, P. (2013b). Effektivität integrierter Lernumgebungen in der universitären Lehrerbildung – eine quasiexperimentelle Feldstudie. *Journal für Bildungsforschung online*, 5 (1). 115-140.

- Weinstock, M., Neuman, Y. & Tabak, I. (2004). Missing the point or missing the norms? Epistemological norms as predictors of student's ability to identify fallacious arguments. *Contemporary Educational Psychology* 29, 77-94.
- Westermann, R. (2000). *Wissenschaftstheorie und Experimentalmethodik*. Göttingen: Hogrefe.
- Wollman, W., Eylon, B.-S. & Lawson, A. E. (1980). An analysis of premature closure in science and developmental stages. *Journal of Research in Science Teaching*, 17, 105-114. doi:10.1002/tea.3660170204.

8. Studie 3: Implementation des Lernens aus Fehlern in Seminaren anhand von Blended learning und Fading

Klein, M., Wagner, K., Klopp, E. & Stark, R. (2015, angenommen). Förderung anwendbaren bildungswissenschaftlichen Wissens bei Lehramtsstudierenden anhand fehlerbasierten kollaborativen Lernens: Eine Studie zur Replikation bisheriger Befunde sowie zur Nachhaltigkeit und Erweiterung der Lernumgebung. *Unterrichtswissenschaft*.

Studie 3 beschreibt die Implementation der Lernumgebung in ein Seminar der Lehrerbildung. Gleichzeitig sollte in der Studie neben einer Replikation der Befunde zur Lernumgebung aus Vorgängerstudien (Studie 2, s.a. Wagner, Klein, Klopp & Stark, 2015, angenommen) auch die Nachhaltigkeit der Lerneffekte getestet werden. Dazu wurde eine reguläre Seminarkonzeption mit einer erweiterten Seminarkonzeption verglichen. Diese stellt systematisch Bezüge zur Lernumgebung her, indem Studierende eigene Problemszenarien nach dem Muster der in der Lernumgebung genutzten Trainingsszenarien erstellten. Hier wurde der Bezug zu Fehlern dadurch realisiert, dass Studierende eigene fehlerhafte Erklärungen als Beispiele für Fehler in der Theorieanwendung sowie eine korrekte Erklärung formulieren mussten. Da zuvor die in Studie 2 beschriebene Lernumgebung bearbeitet wurde, wurde ein aktiver Umgang mit dem in der Lernumgebung erworbenen Wissen gefördert.

Damit sollte gleichzeitig die Möglichkeit gegeben werden, eigene (problematische) Unterrichtserfahrungen auf Grundlage bildungswissenschaftlichen Wissens zu reflektieren. Dies entspricht einem Ansatz aus den Standards zur Lehrerbildung, die zur Förderung der beschriebenen Kompetenzen (s. Kap. 2.1) die „Analyse und Reflexion der eigenen biographischen Lernerfahrungen mit Hilfe der theoretischen Konzepte“ vorsehen (KMK, 2004; 2014, S. 6). Motivationale Variablen wie das themenspezifische Interesse und die Selbstwirksamkeit der Lernenden (Bandura, 1997), die bereits in früheren Studien deutliche Effekte zeigten (z. B. Stark, 2004; Stark & Mandl, 2005), können durch die Verknüpfung mit eigenen Erfahrungen positiv beeinflusst werden.

Die erweiterte Seminarkonzeption wurde anhand von *Blended Learning* (Mandl & Kopp, 2006; Steffens & Reiß, 2009) als Wechsel zwischen Präsenzphasen mit Instruktion durch die Dozenten und selbstgesteuerten Lernphasen strukturiert. Gleichzeitig wurde die steigende Kompetenz der Lernenden im Sinne des Einflusses von Lernermerkmalen wie dem

Vorwissen auf die Wirksamkeit instruktionaler Unterstützung durch *Fading der instruktionalen Unterstützung* (Wecker, 2012) berücksichtigt.

8.1. Stabilität der Lerneffekte

In Bezug auf die Effektivität fehlerbasierten Lernens stellt sich auch die Frage, inwieweit sich Effekte fehlerbasierten Lernens in den unterschiedlichen Lernbedingungen unmittelbar nach der Trainingsphase oder erst bei einer späteren Testung zeigen. Verzögerte Effekte sind bei instruktionalen Maßnahmen, die auf Reflexion abzielen, keine Seltenheit (Mamede et al., 2012). Aus einer pädagogischen Perspektive stellt sich auch die Frage nach der Stabilität instruktionaler Effekte. In den Vorgängerstudien konnten insofern erste Hinweise für die Stabilität der Effekte der integrierten Lernumgebung gewonnen werden (Studie 1 & 2; Wagner et al., 2015, angenommen), als dass die Testung der Lerneffekte eine Woche nach der letzten Lernsitzung stattfand. Längerfristige Effekte und Effekte unterschiedlicher Formen der Einbettung der Lernumgebung in ein Seminar (Bezug auf die Lernumgebung vs. reguläres Seminar ohne direkten Bezug) sind das Thema der dritten Studie.

8.2 Strukturierung der Seminarkonzeption anhand von Blended Learning

Die Stabilität instruktionaler Effekte ist verbunden mit der Frage nach externer bzw. ökologischer Validität der laborexperimentell erzielten Effekte und damit auch der Frage der Verwertbarkeit neuer Erkenntnisse in der pädagogischen Praxis (Stark & Mandl, 2007, Trempler et al., 2014). Studie 3 untersucht, inwiefern das Lernen aus Fehlern bei der Anwendung wissenschaftlichen Wissens zur Reflexion eigener Erfahrungen lernwirksam ist, ob also ein konzeptuelles Verständnis von Fehlern auch genutzt werden kann, um eigene Fehler oder Fehler Dritter bei der Erklärung komplexer schulischer Situationen in der Rückschau zu identifizieren und ob auf Basis des erworbenen Schutzwissens eine korrekte Erklärung unter Vermeidung ähnlicher Fehler selbst formuliert werden kann.

Die Lernenden arbeiteten dabei größtenteils eigenständig. Um dennoch eine Strukturierung der Wechsel zwischen Präsenzlehre und selbstgesteuerten Lernphasen an den einzelnen Seminarterminen herzustellen, wurde ein Ansatz aus der Forschung zum *Blended Learning* herangezogen. Der Begriff „Blended learning“ bezeichnet eine Verknüpfung traditioneller Lehrmethoden in Seminaren oder Vorlesungen mit unterschiedlichen Arten bzw. Anteilen von Online-Lernangeboten (Driscoll, 2002; House, 2002; Rossett, 2002). In einem weiteren Sinne umfasst er jedoch alle Lernansätze, die eine Kombination aus selbstgesteuertem, orts-

und zeitunabhängigem Lernen und instruktionaler Unterstützung bei der Präsenzlehre bieten und die zu vermittelnden Lerninhalte auf verschiedene Medien und Methoden verteilt werden (Arnold, Kilian, Thillosen & Zimmer, 2004; eine Übersicht über mögliche Kombinationen bieten z.B. Oliver und Trigwell, 2005). Wichtig ist dabei die gezielte didaktische Verzahnung von Präsenzlehre und selbstgesteuertem Lernen (Steffens & Reiß, 2009). Die einzelnen Komponenten beim Blended Learning stehen nicht nebeneinander, sondern werden integriert und in ein soziales Umfeld eingebettet (Mandl & Kopp, 2006). Dies soll die Vorteile der Präsenzlehre (u.a. direkte Rückfragen an die Lehrperson, soziale Interaktion) mit denen selbstgesteuerten Lernens (flexibel, orts- und zeitunabhängig, wiederholbar; z.B. Reinmann & Mandl, 2006) verbinden.

Blended Learning setzt voraus, dass Lernende über ein gewisses Maß an Selbststeuerungs-, Medien- und Kooperationskompetenzen verfügen, die zugleich durch die Gestaltung der Lernumgebung gefördert werden sollen. Lernende müssen wissen, wie sie Informationen recherchieren, sie verarbeiten und für Problemlösungen nutzen können, um anwendbares Wissen zu erwerben. Neben diesen kognitiven Strategien sind metakognitive Strategien der Planung, Überwachung und Steuerung des Lernprozesses ebenso wichtig wie motivationale und volitionale Strategien, mit deren Hilfe der Lernprozess auch bei Ablenkung aufrechterhalten wird (Mandl & Kopp, 2006; Weinert, 1996, zur Volition s. Heckhausen, Gollwitzer & Weinert, 1987). Medienkompetenz stellt eine weitere wichtige Voraussetzung für Blended Learning-Ansätze dar, in Studie 3 insbesondere die Fähigkeit, Medien gezielt zur Informationssuche einzusetzen. Dazu gehört auch, (bildungswissenschaftliches) Wissen und Evidenzen hinsichtlich ihrer Adäquatheit und Nützlichkeit zu bewerten, darüber zu reflektieren und für die weitere Arbeit zu nutzen (Bauer et al., 2010; Terhart, Lohmann & Seidel, 2010; Wenglein, Bauer & Prenzel, 2015).

Studien zu Effekten von Blended Learning auf den Erwerb anwendbaren Wissens wurden u.a. von Nistor, Schnurer und Mandl (2005; s.a. Schnurer, 2005) durchgeführt. Dabei zeigten sich positive Effekte auf die Akzeptanz des Ansatzes, sowie die Einschätzung des Lernprozesses und des Lernerfolgs durch die Lernenden. Die Akzeptanz bildete die Bereitschaft der Lernenden zur Nutzung des Angebots und die tatsächliche Nutzung ab (Bürg, 2005). Die Einschätzung des Lernprozesses umfasste Angaben zur eigenen Motivation, der didaktischen Aufbereitung des Lehr-Lern-Angebots und der sozialen Einbettung des Lernens. Der Lernerfolg wurde anhand von Tests zur Medienkompetenz sowie Fakten-, Konzept- und Anwendungswissen ermittelt. Der individuelle Lernerfolg wurde von den Lernenden in allen Dimensionen (Fakten-, Konzept-, Anwendungswissen sowie Medienkompetenz) subjektiv

hoch eingeschätzt. In den Wissenstests wurden vor allem beim Anwendungswissen hohe Werte erzielt, konzeptuelles Wissen konnte nur in geringerem Maß erworben werden. Auch eine Studie von Kopp, Ertl und Mandl (2006), die untersuchte, inwiefern Lernende durch eine inhaltliche und soziale Strukturierung der Aufgabe bei der kooperativen Wissensanwendung in Videokonferenzen unterstützt werden können, erbrachte positive Ergebnisse hinsichtlich der Wissensanwendung. Eine neuere Studie von Lin, Wen, Jou & Wu (2013) zeigte positive Effekte von Blended learning auf kritisches Denken und Motivation.

In Studie 3 wurde die Verknüpfung von Präsenzlehre und selbstgesteuertem Lernen anhand der standardisierten Vorgabe von Arbeitsaufträgen und Beispielen sowie regelmäßigen Feedbackterminen realisiert. Beim selbstgesteuerten Lernen in Form der Bearbeitung schriftlich vorliegender Arbeitsaufträge und Hilfsmaterialien konnten Lernende zudem jederzeit Rückfragen stellen. Die Anwesenheit der Studierenden im Seminar war dabei nach der Vorgabe der Arbeitsaufträge für die jeweilige Sitzung freiwillig. Durch die Betreuung fand außerdem eine kontinuierliche Kontrolle des Lernfortschritts durch die Lehrenden statt. Die Überwachung des Lernprozesses (*monitoring*) beeinflusste in Studien von Wecker (2012) und Wecker und Fischer (2011) die Wirksamkeit des Fadings (s. Kap. 8.3). Diese hängt davon ab, dass die Lernenden tatsächlich konzentriert an der Aufgabe arbeiten.

8.3 Fading instruktionaler Unterstützung

Verschiedene Lernermerkmale wirken sich auf den Lernprozess und den Lernerfolg aus. Insbesondere das Vorwissen hat dabei im Sinne von Aptitude-Treatment-Interaktionen (Cronbach & Snow, 1977) einen starken Einfluss (Hasselhorn & Gold, 2013). Im Allgemeinen wird das Vorwissen, gerade auch im Umgang mit instruktionalen Fehlern, dabei als förderlich für den Lernprozess betrachtet (Große & Renkl, 2004; 2007; Stark et al., 2011). Gerade im Zusammenhang mit Maßnahmen zur instruktionalen Unterstützung sind jedoch auch gegenteilige Effekte in der Art mathematantischer Effekte oder des expertise-reversal-Effekts denkbar, der in Studien zum beispielbasierten Lernen wiederholt auftrat (Kalyuga & Renkl, 2010). Eine Studie von Stark und Fischer (2008) zur Förderung der Diagnosekompetenz in der Medizin zeigte, dass mit Fortschreiten des Lernprozesses und damit einhergehendem Anstieg der Kompetenz der Lernenden eine instruktionale Unterstützungsmaßnahme in Form elaborierten Feedbacks an Wirkung verlor. Instruktionale Unterstützung ist insbesondere beim initialen Lernen von hoher Bedeutung (Reinmann & Mandl, 2006; Renkl, 2005), kann im Laufe des Wissenserwerbs jedoch reduziert werden. Die

graduelle Reduzierung der instruktionalen Unterstützung ist unter dem Begriff „Fading“ auch Bestandteil des cognitive apprenticeship-Ansatzes (Collins et al., 1989, s. Kap. 4) und berücksichtigt, dass Lernende zu Anfang intensive instruktionale Unterstützung benötigen, diese jedoch im Verlauf des Lernprozesses zunehmend weniger Wirkung zeigt (Eiriksdottir & Catrambone, 2011; Salden, Aleven, Schwonke & Renkl, 2010; Sweller, Ayres & Kalyuga, 2011; Wecker, 2012). Fading dient also zur Anpassung des Grades der instruktionalen Unterstützung an den Lernfortschritt der Lernenden (Renkl & Atkinson, 2010). Analog zur möglichen lernhinderlichen Wirkung von Prompts (Berthold et al., 2010; Heitzmann et al., 2013, s. Studie 2) können auch andere Formen instruktionaler Unterstützung, wie etwa tutorielle Unterstützung oder der Vortrag eines *didactic leader* (Krause et al., 2011) mit spontan ablaufenden Selbsterklärungen Lernender interferieren.

Analog zu der Unterscheidung zwischen konstruktions- und instruktionsorientierten Ansätzen kann zwischen zwei Arten von Fading unterschieden werden. Beim *informationsbezogenem Fading* werden z.B. zu Beginn ausgearbeitete Lösungsbeispiele vorgegeben, die während des Lernprozesses unvollständiger werden, bis hin zu komplett selbständig zu lösenden Lernaufgaben. Beim *vorgehensbezogenem Fading* kann z.B. durch die graduell geringer werdende Häufigkeit von Prompts, die das *Lernverhalten* beeinflussen (zu Prompts s. Studie 2), eine Steuerung vorgenommen werden. Die Unterstützung bei beiden Formen kann dabei mit steigendem Wissens- und Fertigungsgrad der Lernenden im Verlauf des Lernprozesses reduziert werden.

Belege für die Lernwirksamkeit einer solchen *adaptiven instruktionalen Unterstützung* durch Fading finden sich neben den theoretischen Konzeptionen und Studien im Bereich des Cognitive Apprenticeship-Ansatzes (Collins et al., 1989) u.a. im Bereich der Argumentationskompetenz bei Wecker (2012) sowie Wecker und Fischer (2011) und Salden et al. (2010). Verschiedene Formen der Operationalisierung werden bei Eiriksdottir und Catrambone (2011) diskutiert, die u.a. zwischen prozeduralen (aufgabenorientierten) *Instruktionen* und *Beispielen*, die ein Modell zur Durchführung einer Aufgabe bieten, unterscheiden. Sowohl bei Instruktionen als auch Beispielen wird initiales Lernen durch eine detaillierte, aufgabennahe Darstellung gefördert. Reflektion und Transfer erfordern jedoch einen höheren Abstraktionsgrad. Daher wurden detaillierte aufgabennahe Instruktionen und Lösungsbeispiele in der vorliegenden Studie mit steigender Kompetenz der Lernenden durch allgemeinere, transferierbare Instruktionen ersetzt.

8.4 Förderung anwendbaren bildungswissenschaftlichen Wissens bei Lehramtsstudierenden anhand fehlerbasierten kollaborativen Lernens: Eine Studie zur Replikation und Stabilität bisheriger Befunde und zur Erweiterung der Lernumgebung

Fostering of applicable pedagogical knowledge in student teachers by collaborative error-based learning: A study on the replicability and stability of previous results and an expansion of the learning environment

Martin Klein, Kai Wagner, Eric Klopp und Robin Stark
Universität des Saarlandes, Saarbrücken

Zusammenfassung

Die Anwendung bildungswissenschaftlichen Wissens zur Erklärung schulischer Situationen ist eine Grundlage effektiven Lehrerhandelns im Schulalltag. Eine solche *Theorienutzungskompetenz* im Sinne wissenschaftlich reflektierten Handelns wird als Schlüsselqualifikation von Lehrkräften betrachtet. Studierenden unterlaufen dabei jedoch häufig Fehler. Um diese instruktional nutzbar zu machen, wurde eine auf *advokatorischen* Fehlern basierende Lernumgebung entwickelt, die in Seminaren der Lehramtsausbildung kollaborativ bearbeitet wurde. Ziel war die Förderung der Kompetenz zur Erklärung schulischer Situationen auf Basis bildungswissenschaftlicher Theorien. Die vorliegende Studie untersucht zusätzlich zur Replikation bisheriger Ergebnisse die Stabilität der erzielten Lerneffekte unter verschiedenen Bedingungen. Experimentell variiert wurde die Konzeption einer auf die Bearbeitung der Lernumgebung folgenden Seminarphase. Eine systematisch auf der Lernumgebung aufbauende erweiterte Seminarkonzeption auf Basis von *Blended Learning* und *Fading instruktionaler Unterstützung* wurde mit einer regulären Seminarkonzeption verglichen. Als abhängige Variablen wurden konzeptuelles, strategisches und prozedurales Fehlerwissen sowie Erklärungswissen erhoben. Die Effekte früherer Studien konnten repliziert werden und waren am Seminarende noch nachweisbar. Die erweiterte Seminarkonzeption erbrachte keinen zusätzlichen Lernfortschritt, verbesserte aber die Stabilität der Lerneffekte in zwei Wissensarten gegenüber der Kontrollbedingung.

Schlüsselwörter: Advokatorische Fehler, Theorieanwendung, Blended learning, Fading instruktionaler Unterstützung

Abstract

Theory application competences for the explanation of complex school situations are a key skill for teachers. However, student teachers often commit typical errors in the process. To utilize these errors instructionally, an integrated learning environment based on advocacy errors was implemented in regular teacher education seminars. Student teachers worked collaboratively on authentic case scenarios to acquire competences in the explanation of complex school situations based on scientific pedagogical knowledge. The present study investigates the replicability of previous results and effects of different conditions on the stability of the learning outcomes. An expanded seminar conception based on *blended learning* and *fading of instructional support* was compared to a regular seminar. Conceptual error-knowledge, strategic error-avoidance- knowledge, procedural error-identification-knowledge and knowledge needed for the explanation of school situations were assessed. Learning outcomes of earlier studies could be replicated and were stable until the end of the semester. The second seminar phase did not lead to additional learning; however, stability of two knowledge types was improved.

Keywords: Advocatory errors, theory application, blended learning, fading of instructional support

1. Problemstellung

Die Forderung nach einer stärkeren Evidenzbasierung professionellen Entscheidens und Handelns hat in den vergangenen Jahren im Bildungsbereich an Bedeutung gewonnen (Fischer et al., 2014). Bereits 2004 wurde in den *Standards der Lehrerbildung* festgelegt, dass zukünftige Lehrkräfte dazu in der Lage sein sollen, (bildungs-)wissenschaftliche Forschungsergebnisse für die Unterrichtspraxis zu nutzen (KMK, 2004; 2014). Empirische Befunde werden jedoch erst anhand kritischer Interpretation vor dem Hintergrund theoretischer Modelle und Fragestellungen zu *Evidenzen* (Bromme, Prenzel & Jäger, 2014). Wichtig ist daher neben der Förderung von Kompetenzen angehender Lehrkräfte im Umgang mit empirischen Befunden auch die Förderung eines *kompetenten Umgangs mit wissenschaftlichen Theorien*. Ohlssons (1992) Konzept der *Theorieartikulation* umschreibt diesen als die „[...] Anwendung einer Theorie auf eine bestimmte Situation und eine Ableitung dessen, was die Theorie implizit oder explizit über die Situation aussagt“ (S.182; Übers. d. A.). Dies umfasst auch die *Erklärung* komplexer schulischer Situationen anhand bildungswissenschaftlichen Wissens, die als eine der Grundlagen effektiven Lehrerhandelns im Schulalltag gilt (Bromme, 2008). Da diese komplexe Tätigkeit hohe Ansprüche an die Studierenden stellt, kommt es häufig zu Fehlern. Zur Erklärung werden z. B. Alltagswissen oder ungeeignete Theorien herangezogen oder empirische Evidenz wird nicht adäquat verwendet (Stark, 2005).

Um diese Fehler als Lerngelegenheiten zu nutzen, wurde eine auf advokatorischen Fehlern (Oser, 2007) basierende integrierte Lernumgebung (Reinmann & Mandl, 2006) entwickelt, die an der Universität des Saarlandes in Seminaren des Lehramtsstudiums eingesetzt wird.

Die Effektivität dieser Lernumgebung wurde in verschiedenen Studien unserer Arbeitsgruppe nachgewiesen (Wagner, Klein, Klopp & Stark, 2014b; Wagner, Klein, Klopp & Stark, 2015 angenommen). Die vorliegende Studie untersucht die Replizierbarkeit der bisherigen Befunde sowie die Stabilität der Lerneffekte. Dazu wurde die Lernwirksamkeit einer der Bearbeitung der Lernumgebung (erste Seminarphase) nachgeordneten, zweiten Seminarphase untersucht. Eine *erweiterte* Seminarkonzeption im *Blended-Learning*-Format (z.B. Steffens & Reiß, 2009) inkl. *Fading instruktionaler Unterstützung* (Wecker, 2012) wurde hierbei mit einer *regulären* Seminarkonzeption mit standardisierten Arbeitsaufträgen verglichen. Beide Bedingungen fokussierten die Anwendung bildungswissenschaftlicher Theorien, jedoch griff nur die erweiterte Seminarkonzeption systematisch Struktur und Konzeption der Lernumgebung auf.

2. Theoretischer Hintergrund

2.1 Replikation der Ergebnisse zur Lernwirksamkeit der Lernumgebung

Die in der ersten Seminarphase eingesetzte Lernumgebung war bereits Gegenstand früherer Studien unserer Arbeitsgruppe (z.B. Wagner et al., 2014b; 2015, angenommen). Sie basiert auf *Advokatorischem Lernen aus Fehlern*, bei dem *negatives Wissen* darüber erworben wird, wie etwas *nicht ist* oder *nicht gemacht wird* (Oser, 2007). Um zu verdeutlichen, was aus welchem Grund falsch ist und wie die korrekte Lösung lautet, werden typische Fehler in einer Domäne präsentiert und mit korrekten Konzepten kontrastiert (z.B. Durkin & Rittle-Johnson, 2012).

Die Umsetzung des Lernens aus Fehlern erfolgte in einer *integrierten Lernumgebung*, die problem- und instruktionsorientierte Designprinzipien verbindet (Reinmann & Mandl, 2006). Problemorientierte Ansätze basieren auf relevanten, authentischen Problemstellung in einem narrativen Format (z.B. *anchored instruction*; CTGV, 1992), die Lernende mit hoher Eigenaktivität bearbeiten. Instruktionsorientierte Ansätze setzen dagegen auf systematische Wissensvermittlung, etwa durch ausgearbeitete Lösungsbeispiele, bei denen Lernende Modelllösungen komplexer Probleme nachvollziehen (Renkl & Atkinson, 2010). Dabei können sie durch Selbsterklärungs- oder Reflexionsprompts (z.B. Schworm & Renkl, 2007) unterstützt werden.

Vor dem Hintergrund dieser Designprinzipien werden in der Lernumgebung fehlerhafte und korrekte ausgearbeitete Lösungsbeispiele kontrastiert, die mittels authentischer Problemszenarien in einen narrativen Rahmen eingebettet sind. Es werden problematische Situationen in der Klasse eines jungen Lehrers beschrieben, dieser formuliert (stellvertretend für die Lernenden) eine fehlerhafte Erklärung. Ein Schulpsychologe weist ihn auf die Fehler hin und präsentiert eine wissenschaftlich korrekte Erklärung als korrekte Lösung (s. Abb 1 und 2). Die enthaltenen Fehler wurden anhand einer *Fehlermatrix* (Wagner, Klein, Klopp & Stark, 2014a) systematisiert, die auf den Ergebnissen von Stark (2005) basiert, d.h. in diesem Kontext festgestellten Fehlern. Die dargestellten Fehler waren nicht auffällig falsch, hatten eine hohe Auftretenswahrscheinlichkeit und Plausibilität sowie nachteilige Folgen.

Abbildung 1: Ausschnitt aus einer fehlerhaften Erklärung

2. Romans Erklärung der Situation

Roman ist sehr beunruhigt über das Ausmaß an Aggression, dass er in der Situation beobachtet hat. Um sich auf die Konferenz vorzubereiten, auf der er seine Kollegen und den Schulpsychologen um Rat fragen will, hat er ein Handout erstellt, auf dem seine Erklärung der Situation zu lesen ist. Roman legt Wert darauf, dass seine Erklärung wissenschaftlichen Standards genügt, allerdings ist er sich dabei nicht ganz sicher.

Bitte lesen Sie das Handout aufmerksam!

Erklärung des Vorfalls zwischen Markus (Täter) und Peter (Opfer):

- Da Markus Peter schon mehrmals angegriffen hat, scheint er ein eher aggressives Temperament zu haben. Es kann also gut sein, dass dieses Persönlichkeitsmerkmal bei Markus sehr ausgeprägt ist. Das Persönlichkeitsmerkmal „Aggressivität“ könnte demzufolge auch die wiederholten Übergriffe gegenüber Peter erklären.
- Frustrations-Aggressions-Hypothese (Dollard et al, 1941): Aggression wird durch Frustration verursacht. Eine Frustration ist eine Störung einer zielgerichteten Aktivität (Hindernisfrustration). Bei wiederholten oder starken Frustrationen kann es zu aggressivem Verhalten kommen. Peter scheint Markus also ständig in irgendeiner Weise daran zu hindern, ein Ziel zu erreichen. Damit frustriert er Markus und es kommt zu aggressiven Handlungen.

Abbildung 2: Beispiel zur Korrektur eines Fehlers

3. Die Fehleranalyse des Schulpsychologen (1)

Romans Fehler:
Zirkelerklärung

Eine Zirkelerklärung ist ein Fehler, bei dem ein Phänomen mit sich selbst erklärt wird.

Romans
Erklärung

➤ *Da Markus Peter schon mehrmals angegriffen hat, scheint er ein eher aggressives Temperament zu haben. Es kann also gut sein, dass dieses Persönlichkeitsmerkmal bei Markus sehr ausgeprägt ist. Das Persönlichkeitsmerkmal „Aggressivität“ könnte demzufolge auch die wiederholten Übergriffe gegenüber Peter erklären.*

Feedback des
Schulpsychologen

1. Fehler: Was ist falsch und warum?

Eine Zirkelerklärung kann eine Situation nicht erklären, sondern benutzt lediglich das beobachtete Phänomen, um sich selbst zu erklären. Tatsächlich wird nicht geklärt, warum Markus aggressiv handelt.

2. Inhalt der verwendeten Theorie (Inhaltliche Anwendung der Fehlerdefinition):

Roman erklärt das aggressive Verhalten von Markus mit dessen Temperament und das Temperament anhand des Verhaltens, notwendig wäre es das Temperament unabhängig von Markus Verhalten zu erklären:

In der beschriebenen Situation verhält sich Markus aggressiv, weil er aggressiv ist. Und weil Markus aggressiv ist, handelt er aggressiv.

Anhand dieses zirkulären Schlusses ist Roman tatsächlich nicht in der Lage, die Situation zu erklären.

Zur instruktionalen Unterstützung der Lernenden wurden *reflection prompts* (Durkin & Rittle-Johnson, 2012) eingesetzt. In einer Studie zur Lernwirksamkeit der beschriebenen

Lernumgebung wurde das Ausmaß an instruktionaler Unterstützung in Form solcher Lernaufforderungen experimentell variiert (Wagner et al., 2014b). Die Prompts beinhalteten die Präsentation sowie Rekonstruktion von Vermeidungsstrategien bezüglich der in der Lernumgebung enthaltenen Fehler. Dabei zeigte sich ein Vorteil *umfassender* instruktionaler Unterstützung hinsichtlich des Erwerbs verschiedener Wissensarten, die anhand der Konzeptionen von De Jong und Ferguson-Hessler (1996) sowie Krause (2007) operationalisiert wurden. Im Einzelnen wurde konzeptuelles Wissen über Fehlerdefinitionen, strategisches Wissen darüber, wie diese Fehler vermieden werden können sowie prozedurales Wissen zur Identifikation von Fehlern in einer fehlerhaften Erklärung gefördert. Auch die Wissensanwendung zur Erstellung von Erklärungen (Theorieartikulation, s. o.), die das Zusammenwirken dieser Wissensarten erfordert (Krause, 2007), verbesserte sich substantiell.

2.2 Befunde zur kollaborativen Bearbeitung der Lernumgebung

Wagner et al. (2015, angenommen) untersuchten die Lernwirksamkeit der oben beschriebenen Lernumgebung bei *kollaborativer* Bearbeitung in Dyaden. Kollaboratives Lernen basiert auf flexiblen Skripts interaktionsbezogenen Wissens, sogenannte *internalen Kollaborationsskripts* (IKS; *script theory of guidance*, Fischer, Kollar, Stegmann & Wecker, 2013). Sind diese Skripts defizitär, kann produktive Kollaboration behindert werden. In diesem Fall können Lernende durch *externale Kollaborationsskripts* (EKS) bei der Zusammenarbeit angeleitet und so fehlende oder unangemessene IKS kompensiert werden (Fischer et al., 2013). EKS sind sequenzielle Anordnungen kollaborativer Aufgaben unterschiedlicher Komplexität, die den Diskurs der Lernenden zur kollaborativen Wissenskonstruktion fördern (Weinberger, Stegmann, Fischer & Mandl, 2007). EKS haben sich in verschiedenen Studien als lernwirksam erwiesen (z.B. Fischer et al., 2013). Ein hoch strukturiertes EKS verbesserte auch in der o.g. Studie das Fehleridentifikationswissen und das Erklärungswissen im Vergleich zu einer Bedingung mit individueller Bearbeitung (Wagner et al., 2015, angenommen).

2.3 Überlegungen zur Stabilität der Lerneffekte

Angesichts der nachgewiesenen Lernwirksamkeit der Lernumgebung stellt sich in der vorliegenden Untersuchung neben der Replikation der Befunde die Frage nach deren *Stabilität*. Eine gewisse Stabilität des in der Lernumgebung erworbenen Wissens kann auf Basis der bisherigen Ergebnisse zwar angenommen werden (in allen früheren Studien lag eine

Woche zwischen letzter Bearbeitung der Lernumgebung und Test; Wagner et al., 2014b; 2015, angenommen), Nachweise längerfristiger Retention und Untersuchungen zur Implementation in Seminaren stehen jedoch aus. Daher wurde die *erste* Seminarphase, die wie in den bisherigen Studien aus der Bearbeitung der o.g. Lernumgebung bestand, um eine systematisch auf dieser Lernumgebung aufbauende, *zweite* Seminarphase ergänzt. In dieser Phase wurden anhand der Konstruktion eigener Problemszenarien durch die Studierenden die erworbenen Fähigkeiten in der Theorieanwendung mit deren eigenen Erfahrungen verbunden. Durch diese Reflektion biografischer Lernerfahrungen (vgl. KMK, 2014, S.) sollte das in der Lernumgebung Gelernte zusätzlich elaboriert und in der Wissensbasis der Lernenden verankert werden, wodurch auch positive Effekte auf die Stabilität des Gelernten erzielt werden sollten.

Zur didaktischen Gestaltung wurden Konzepte zum *Blended Learning* (Mandl & Kopp, 2006; Steffens & Reiß, 2009) sowie zum *Fading instruktionaler Unterstützung* (Wecker, 2012) herangezogen. Blended Learning diente zur Strukturierung der Anteile von Präsenzlehre und selbstgesteuertem Lernen, Fading zur Anpassung der instruktionalen Unterstützung von Seiten der Dozenten an die steigende Kompetenz der Studierenden.

3. Didaktische Designprinzipien der erweiterten Seminarkonzeption

3.1 Blended Learning

Blended Learning wird hier lediglich als Überbegriff für eine Kombination von Präsenzlehre und selbstgesteuertem Lernen verwendet, bei der verschiedene methodische und mediale Zugänge zum Lerngegenstand die größtmögliche Qualität und Effizienz des Lehr-Lern-Angebots sichern (vgl. Kerres & Jechle, 2002). Mandl und Kopp (2006) betonen die Fokussierung auf die Perspektive der Lernenden als Kriterium für die Verwendung des Begriffs „Blended Learning“. Demzufolge ist Lernen dann „*blended*“, wenn Lernenden die Variationen zwischen verschiedenen Herangehensweisen an den Lerngegenstand deutlich werden. Relevante Aspekte des Lerngegenstands werden durch die verschiedenen Zugangsweisen hervorgehoben. Wichtig ist dabei die gezielte didaktische Verzahnung von Präsenzlehre und selbstgesteuertem Lernen (Steffens & Reiß, 2009). Damit sollen die Vorteile der Präsenzlehre (u.a. direkte Rückfragen an die Lehrperson, soziale Interaktion) mit denen selbstgesteuerten Lernens (flexibel, orts- und zeitunabhängig, wiederholbar; z.B. Reinmann und Mandl, 2006) kombiniert werden.

3.2 Fading instruktionaler Unterstützung

Lernende benötigen beim selbstgesteuerten Lernen mit hoher Eigenaktivität für den effizienten Wissenserwerb *instruktionale Unterstützung* (z.B. beim problemorientierten Lernen; Reinmann & Mandl, 2006; Schmidt, Loyens, Van Gog & Paas, 2007). Instruktionale Unterstützung ist insbesondere beim initialen Lernen von hoher Bedeutung und kann im Laufe des Wissenserwerbs immer weiter reduziert werden. Dieses sogenannte *Fading* (z.B. im *cognitive apprenticeship*-Ansatz; Collins, Brown & Newman, 1989; s.a. Wecker, 2012) berücksichtigt, dass Lernende anfangs intensive instruktionale Unterstützung benötigen, diese jedoch im Verlauf des Lernprozesses zunehmend weniger Wirkung (Eiriksdottir & Catrambone, 2011; Salden, Alevén, Schwonke & Renkl, 2010) oder sogar negative Effekte zeigt (*expertise-reversal*-Effekt, z.B. Kalyuga, Rikers & Paas, 2012). *Fading* dient also zur Anpassung des Grades der instruktionalen Unterstützung an den Lernfortschritt der Lernenden (Renkl & Atkinson, 2010).

Belege für die Lernwirksamkeit von Fading finden sich u.a. bei Wecker (2012) sowie Wecker und Fischer (2011), eine Übersicht über Formen der Umsetzung bieten Eiriksdottir und Catrambone (2011). Diese unterscheiden u.a. zwischen prozeduralen (aufgabenorientierten) *Instruktionen* und *Beispielen*, die ein Modell zur Durchführung einer Aufgabe bieten. In beiden Fällen fördert eine detaillierte, aufgabennahe Darstellung initiales Lernen, Reflektion und Transfer erfordern jedoch einen höheren Abstraktionsgrad. Analog dazu wurden in der vorliegenden Studie detaillierte, aufgabennahe Instruktionen und Lösungsbeispiele zunehmend durch allgemeinere, transferierbare Instruktionen ersetzt.

4. Fragestellung und Hypothesen

Ausgehend von den dargestellten theoretischen Überlegungen wird in dieser Studie untersucht, inwieweit die Ergebnisse früherer Studien unserer Arbeitsgruppe repliziert werden können. Außerdem wird getestet, inwieweit eine erweiterte Seminarkonzeption auf Basis von *Blended Learning* in Kombination mit *Fading instruktionaler Unterstützung* am Ende der Intervention zu höheren Lerneffekten hinsichtlich der Anwendung bildungswissenschaftlichen Wissens zur Erklärung schulischer Situationen führt, als eine reguläre Seminarkonzeption. Es wurden folgende Fragestellungen formuliert:

1. Inwieweit können die Ergebnisse der vorangegangenen Untersuchungen hinsichtlich der Performanz der Lernenden in den Tests zu konzeptuellem Fehlerwissen (konz. FW),

strategischem Fehlervermeidungswissen (strat. FW), prozeduralem Fehleridentifikationswissen (proz. FW) sowie Erklärungswissen (ErklW) repliziert werden?

Es wurde angenommen, dass alle Versuchspersonen in den Variablen *konz. FW*, *strat. FW* und *ErklW* im ersten Nachtest (zweiter Messzeitpunkt = t2) nach der Bearbeitung der Lernumgebung höhere Werte als im Vortest (erster Messzeitpunkt = t1) zeigen. Der Lernerfolg sollte deskriptiv in ähnlicher Höhe wie in der vorausgegangenen Untersuchung ausfallen.

2. Welchen Einfluss hat die erweiterte Seminarkonzeption auf die Performanz in den Tests am Seminarende (dritter Messzeitpunkt = t3)? Verbessert sich die Stabilität der Lerneffekte bzw. ist in Abhängigkeit von der Seminarkonzeption ein zusätzlicher Lernfortschritt nachweisbar?

Bezüglich des Effekts der Seminarkonzeption wurde angenommen, dass die Experimentalgruppe mit erweiterter Seminarkonzeption der Kontrollgruppe mit regulärer Seminarkonzeption hinsichtlich der Variablen *konz. FW*, *strat. FW*, *proz. FW* und *ErklW* zum dritten Messzeitpunkt überlegen ist. Entsprechend wird eine Interaktion zwischen dem Messzeitpunkt und der Gruppenzugehörigkeit erwartet. Die Lerneffekte der Lernumgebung sollten in der Experimentalgruppe stabilisiert werden und deshalb zum dritten Messzeitpunkt größer sein als in der Kontrollgruppe bzw. es sollte sich in der Experimentalgruppe in allen Variablen ein Lernfortschritt vom zweiten zum dritten Messzeitpunkt nachweisen lassen.

5. Methode

5.1 Stichprobe und Design

Die Stichprobe umfasste 143 Studierende (85w; 15 k. A.) aus Pflichtseminaren der bildungswissenschaftlichen Lehrerbildung. Alle Probanden studierten in Lehramtsstudiengängen für weiterführende Schulen (94 [65.7 %] LA Gymnasien u. Gemeinschaftsschulen, 33 [23.1%] LA Real- u. Gemeinschaftsschulen, 13 [9.1%] LA berufliche Schulen, 3 k.A.). Das Durchschnittsalter der Teilnehmer lag bei 22.4 Jahren ($SD = 3.75$). In einem quasiexperimentellen Design mit zwei Messwiederholungen wurde die Teilnahme an der regulären und erweiterten Seminarkonzeption variiert. Die Zuweisung der Seminare zur jeweiligen Seminarkonzeption erfolgte zufallsbasiert. So ergab sich eine Experimentalgruppe mit $n = 68$ Probanden (Pbn; 42 weiblich, 5 k. A.; $M_{\text{Alter}} = 23.4$ [$SD = 3.44$] Jahre) und eine Kontrollgruppe mit $n = 75$ Pbn (43 weiblich, 10 k. A.; $M_{\text{Alter}} = 22.8$ [$SD = 3.81$] Jahre). Die Gruppen waren hinsichtlich Alter ($F(2,128) = 1.78, p = .56$), Geschlecht ($\chi^2(2) = 0.004, p = .95$) und Studiengang (Schulart; $\chi^2(3) = 1.34, p = .72$) vergleichbar.

5.2 Ablauf der Untersuchung

Nach den Vortests (t1) bearbeiteten *beide* Versuchsgruppen in Phase 1 die o.g. Lernumgebung (Wagner et al., 2014a; 2014b) kollaborativ in Dyaden. In Phase 2 wurde die Experimentalgruppe bis zum Seminarende anhand der erweiterten und die Kontrollgruppe anhand der regulären Seminarkonzeption unterrichtet (s. Tab. 1).

Zur Standardisierung von Phase 1 (vier Termine) wurden zwei erfahrene Dozenten mit Hilfe eines Leitfadens instruiert. Nach den Vortests (t1) am ersten Seminartermin bearbeiteten die Pbn an den beiden folgenden Terminen die Lernumgebung in Dyaden unter Supervision der Dozenten. Beide Lernpartner waren jeweils präsent. Um die Bearbeitungszeit zu standardisieren, war die Dauer der Bearbeitung auf 90 Minuten limitiert. Diese wurde von der Mehrzahl der Studierenden vollständig in Anspruch genommen. Zum vierten Termin erfolgte der erste Nachtest (t2). Dieselben Dozenten führten die anschließende *Blended Learning*-Phase (Phase 2, sieben Termine) durch und unterrichteten auch die Kontrollgruppe, die in Phase 2 standardisierte Arbeitsaufträge zu pädagogischen Problemstellungen bearbeitete. Auch in Phase 2 war die Dauer eines Seminartermins auf 90 Minuten begrenzt. Ebenso waren die Testbatterien so konzipiert, dass sie in 90 Minuten zu bewältigen waren. Der zweite Nachtest (t3) fand am letzten Termin am Seminarende statt. Im Vordergrund stand die Frage, inwieweit die bisherigen Befunde zur kollaborativen Bearbeitung der Lernumgebung repliziert werden können und inwieweit diese Effekte durch eine erweiterte Seminarkonzeption stabilisiert oder sogar gesteigert werden können. Da in der zweiten Phase kooperatives Lernen nicht im Zentrum stand, arbeiteten die Pbn dort individuell.

Tabelle1

Versuchsablauf in der Experimental- und Kontrollgruppe

Phase 1: Bearbeitung der Lernumgebung, 4 Wochen		Phase 2: Selbstgesteuertes Lernen mit Unterstützung durch die Dozenten, 7 Wochen		
Termin 1 (90 min)	Termin 2-3 (je 90 min)	Termin 4 (90 min)	Termin 5-12 (je 90 min)	Termin 13 (90 min)
<i>Vorwissenstest:</i> konz.FW strat. FW Erkl.W <i>Soziodemografische Daten</i> und <i>Kontrollvariablen</i>	<i>Bearbeitung der Lernumgebung:</i> Schulische Problemszenarien zu vier Themengebieten: Bullying Gruppenphänomene Leistungsmotivation Disziplinprobleme	<i>Nachtest 1</i> (= Vorwissenstests) + proz. FW	<i>Erweiterte Seminarkonzeption:</i> Schrittweise Konstruktion eines schulischen Problemszenarios Unterstützung durch Dozenten und regelmäßige Feedback-termine (s. Tabelle 2)	<i>Nachtest 2</i> (= Nachtest 1)
			<i>Reguläre Seminarkonzeption:</i> Bearbeitung standardisierter Arbeitsaufträge zur Theorieanwendung auf reguläre Seminarinhalte	

5.3 Phase 1: Bearbeitung der Lernumgebung

Alle Probanden bearbeiteten in der Lernumgebung kollaborativ schulische Problemszenarien in Form fallbasierter ausgearbeiteter Lösungsbeispiele. Dieses Training diente zum Aufbau der o.g. Wissensarten anhand des Vergleichs korrekter mit inkorrekten Theorieartikulationen (Erklärung einer komplexen schulischen Situation). Alle Wissensinhalte waren vorgegeben, die Aufgabe der Lernenden bestand im Nachvollzug der Lösungsbeispiele (zur Lernwirksamkeit Wagner et al., 2015, angenommen).

5.4 Phase 2: Experimentelle Variation – erweiterte und reguläre Seminarkonzeption

Erweiterte Seminarkonzeption. In der *erweiterten Seminarkonzeption* konstruierten die Studierenden eigenständig schulische Problemszenarien nach dem Muster der Lösungsbeispiele aus der Lernumgebung. Dadurch wurden die erworbenen Kenntnisse in der Theorieanwendung mit eigenen Erfahrungen der Studierenden aus Praktika oder anderen schulischen Erfahrungen verbunden und Reflexion, Elaboration und Transfer ermöglicht (vgl. *rückschauendes Begreifen*, Beck & Krapp, 2006). Der systematische Bezug zu Phase 1 wurde dadurch hergestellt, dass Studierende eigene fehlerhafte Erklärungen als Beispiele für Fehler in der Theorieanwendung sowie korrekte Erklärungen formulieren mussten.

Blended Learning wurde durch den Wechsel zwischen Präsenzlehre in der Instruktionsphase der jeweiligen Sitzungen sowie bei den Präsenzterminen und selbstgesteuertem Lernen mit kontinuierlicher Beratung durch die Dozenten realisiert. In der Instruktionsphase wurden detaillierte Arbeitsanweisungen und alle für die selbstgesteuerte Bearbeitung notwendigen Materialien bereitgestellt. Dadurch wurde die notwendige Verzahnung von Präsenzlehre und selbstgesteuertem Lernen hergestellt (Steffens & Reiß, 2009). Die Materialien beinhalteten Informationen zur sukzessiven Erstellung eines *komplexen schulischen Problemszenarios*, z.B. Informationen zu Fehlern bei der Erklärung des Problems und Hinweise zur Erstellung einer Theoriesammlung zur korrekten Erklärung. Die Dozenten erläuterten zu Beginn jeden Seminartermins die genaue Aufgabenbeschreibung sowie ein Beispiel. Für die Pbn war mit Ausnahme zweier Präsenztermine (s. Tab. 2) nur die Anwesenheit in der Einführungsphase des jeweiligen Seminartermins verpflichtend (ca. 15 min), die restliche Zeit konnte selbstgesteuert zur Aufgabebearbeitung (z.B. Theorierecherche) genutzt werden. Die Dozenten standen jedoch für die gesamte verbleibende Zeit zur Beratung zur Verfügung. Dieses Angebot wurde von allen Pbn häufig genutzt. Die Arbeitsergebnisse waren bis zum folgenden Termin bei den Dozenten per Email einzureichen. Diese gaben ebenfalls per Email vor dem nächsten Termin elaborierte Rückmeldungen zur Qualität und Vorschläge zur Verbesserung.

Fading der instruktionalen Unterstützung wurde über die Gesamtdauer von Phase 2 umgesetzt. An den ersten drei Terminen von Phase 2 gaben die Dozenten nach der Einführungsphase auf Nachfrage noch stärker direkte inhaltliche Unterstützung, da z.B. der *Entwurf eines schulischen Problems* und die *Erstellung einer fehlerhaften Erklärung* in der Lernumgebung aus Phase 1 vorgegeben gewesen waren. Die eigenständige Konstruktion dieser Abschnitte war also nicht Teil der Bearbeitung der Lernumgebung und für die Studierenden neu. Gegen Ende von Phase 2 fokussierte die Unterstützung stärker die den Aufgaben zugrunde liegenden Strukturen (s. Tab. 2). Im Sinne von Eiriksdottir und Catrambone (2011) wurde so die Progression von einer detaillierten, aufgabennahen Darstellung zu einem höheren Abstraktionsgrad als Form des *Fadings instruktionaler Unterstützung* zur schrittweisen Bearbeitung von Aufgaben (vgl. Wecker, 2012) umgesetzt.

Nach Abschluss von drei bzw. sechs Aufgaben folgten zwei Präsenztermine, die Studierenden stellten ihre Lösungen beide Male im Plenum vor. Zunächst hatten die Kommilitonen Gelegenheit, die Lösungen zu kommentieren, dann folgte das Feedback der Dozenten. Dabei wurden Fehler der Studierenden bei der Bearbeitung der Arbeitsanweisungen aufgegriffen und zusammen mit einer entsprechenden Vermeidungsstrategie erklärt. Beispielsweise durfte

keine der *Erklärungen* auf Informationen zurückgreifen, die nicht in der Situationsbeschreibung enthalten waren. Die Entwicklung der Szenarien hatte somit einen deutlichen Bezug zur Struktur der Lernumgebung in Phase 1.

Tabelle 2

Seminartermine Phase 2 - Blended Learning

Termin	Dozent	Studierende
5	Hilfestellung zum Aufbau der Problemdarstellung	Entwurf eines schulischen Problemfalls
6	Hilfestellung zum Aufbau einer fehlerhaften Lösung	Erstellung einer fehlerhaften Erklärung
7	Informationen zur Quellenanalyse und Unterscheidung zw. wiss. und Alltagswissen	Theoriesammlung zur Erklärung der Situation
8	Feedback zu bisher fertiggestellten Teilen des Problemszenarios	Vortrag und Diskussion im Plenum
Präsenztermin		
9	Vorgabe der Analysestruktur analog zur Lernumgebung	Analyse der Fehler
10	Vorgabe der Struktur von Vermeidungsstrategien analog zur Lernumgebung	Formulierung von Fehlervermeidungsstrategien mit konkretem Situationsbezug
11	Vorgabe der Struktur einer Musterlösung analog zur Lernumgebung	Formulierung einer korrekten theoriebasierten Erklärung
12	Feedback zum fertigen Problemszenario	Vortrag und Diskussion im Plenum
Präsenztermin		

Reguläre Seminarkonzeption. Die Kontrollgruppe mit regulärer Seminarkonzeption bearbeitete in Phase 2 standardisierte Arbeitsaufträge zur Anwendung von Theorien in der konkreten Unterrichtsplanung und dabei auftretenden Problemen. Analog zur Experimentalbedingung konnten die Studierenden Problemstellungen bearbeiten, die auf eigenen Erfahrungen basierten. Beispielsweise wurden Disziplinprobleme innerhalb einer Klasse anhand sozialpsychologischer Theorien zur Gruppenbildung und zur Entstehung von Konflikten erklärt. Die Studierenden wurden somit ebenfalls zur Theorieanwendung angehalten, jedoch fehlten die systematischen Bezüge zum Lernen aus Fehlern in Phase 1. Die Studierenden erhielten zudem kein elaboriertes Feedback zu den Arbeitsaufträgen,

sondern jeweils eine Musterlösung. Auch in der Kontrollgruppe bestand nur Anwesenheitspflicht während der Vergabe der Arbeitsaufträge.

5.5 Instrumente

Die Instrumente dieser Studie waren mit den vorhergehenden Untersuchungen identisch (Wagner et al., 2014b; 2015, angenommen). Neben Alter und Geschlecht wurden potentiell den Lernerfolg beeinflussende Kontrollvariablen erhoben. *Einstellungen zum kooperativen Lernen* wurden mit fünf Items (z. B. „In Gruppen zu lernen ist besser als allein zu lernen.“; Cronbachs $\alpha = .78$), sowie *Einstellungen zum individuellen Lernen* mit drei Items erfasst (z. B. „Ich lerne am besten, wenn ich allein lerne.“; Cronbachs $\alpha = .79$). Die den Studierenden zur Beantwortung vorgelegten Ratingskalen waren sechsfach gestuft (von 1 = *stimmt überhaupt nicht* bis 6 = *stimmt genau*).

Bis auf den Test zum *proz. FW* wurden die Wissenstests jeweils an allen drei Messzeitpunkt durchgeführt. *Proz. FW* wurde nur in den beiden Nachtests erfasst, da es in einer früheren Studie ohne explizites Training nicht nachgewiesen werden konnte (Klein, Wagner & Stark, 2012).

Konz. FW wurde mittels zwölf Multiple-Choice-Fragen bzgl. Definitionen möglicher Fehler bei der Erklärung schulischer Situationen erhoben (drei Antwortmöglichkeiten, zwei Distraktoren). Jede richtige Antwort wurde mit einem Punkt bewertet (z. B. „Wie wird die Zirkelerklärung definiert?“; Cronbachs $\alpha = .88$ [Vortest]; $.75$ [Nachtest 1]; $.78$ [Nachtest 2]; theoret. Maximum: 12 Punkte).

Strat. FW wurde im Anschluss auf die Fragen zu den zwölf Fehlerdefinitionen jeweils anhand der offenen Frage „Wie können Sie diesen Fehler vermeiden?“ erfasst (Cohens $\kappa = .86$ [Vortest] bzw. $.83$ [Nachtest 1]; $.81$ [Nachtest 2]; theoret. Maximum: 24 Punkte).

Zur Erfassung des *proz. FW* wurde eine inkorrekte Theorieartikulation mit vier typischen Fehlern präsentiert. Die Pbn mussten hier die Fehler identifizieren (offenes Antwortformat, max. 32 Punkte; Cohens $\kappa = .87$ [Nachtest 1]; $.82$ [Nachtest 2] vgl. Abb. 3).

Abbildung 3: Test zum Prozeduralen Fehleridentifikationswissen

3. Ihre Fehleranalyse

Herrn Müllers
Fehler:

Kurze Fehlerdefinition:

Herrn Müllers
Erklärung

➤ Geben Sie kurz an, auf welche Zeilen Sie sich beziehen:

Ihr Feedback:

1. Fehler: Was ist falsch und warum?

2. Inhalt der verwendeten Theorie (Inhaltliche Anwendung der Fehlerdefinition):

ErklW wurde anhand von den Pbn verfassten Theorieartikulationen erfasst. Mit Hilfe kurzer Zusammenfassungen verschiedener Theorien musste eine schulische Situation eigenständig erklärt werden. Dabei waren Bezüge zwischen der Situation und den Theorien herzustellen (offenes Antwortformat, max. 16 Punkte; Cohens $\kappa = .86$ [Vortest] bzw. $.85$ [Nachttest 1]; $.88$ [Nachttest 2] vgl. Abb. 4).

Abbildung 4: Test zum Erklärungswissen

Ihre eigene Erklärung

Ihre Aufgabe

Nachdem Sie Herrn Müller auf die Fehler in seiner Erklärung aufmerksam gemacht haben, bittet Sie dieser, ihm bei der Erstellung einer wissenschaftlich korrekten Erklärung zu helfen. Zu diesem Zweck gibt er Ihnen eine Sammlung von Theorien, die er vom Schulpsychologen Herrn Jung erhalten hat.

1. **Ihre Erklärung:** Formulieren Sie eine Erklärung der Situation anhand der theoretischen Konzepte aus Herrn Jungs Theoriesammlung.
2. **Ihre Strategien:** Geben Sie dabei bitte jeweils an, welche der Theorien oder Phänomene Sie zur Grundlage Ihrer Argumentation machen, und begründen Sie, warum Sie diesen Ansatz gewählt haben.

Die offenen Fragen der jeweiligen Tests wurden von zwei geschulten Bewertern ausgewertet. Diese waren nicht informiert, aus welcher Bedingung die auszuwertenden Tests stammten. Testleistungen wurden mit Musterlösungen abgeglichen, Abweichungen in den Urteilen wurden im Diskurs erörtert. Bis auf die Korrelation von *konz. FW* und *ErklW* zu t1 waren alle Korrelationen der einzelnen Wissenstests zu den drei Messzeitpunkten signifikant (s. Tab. 3).

Tabelle 3

Korrelationen der Wissenstests zu allen Messzeitpunkten

	Strat. FW			ErklW			Proz. FW		
	t1	t2	t3	t1	t2	t3	t1	t2	t3
Konz.FW	.57**	.49**	.39**	.06	.34**	.27**	-	.44**	.41**
Strat.FW	-	-	-	.18*	.34**	.37**	-	.67**	.47**
ErklW	-	-	-	-	-	-	-	.43**	.33**

Anmerkung. * $p < .05$; ** $p < .01$

5.6 Statistische Analysen

Um die Replikation und Stabilität der Lerneffekte der Lernumgebung sowie die Effekte der erweiterten Seminarkonzeption zu untersuchen, wurden für die einzelnen Variablen zweifaktorielle ANOVAs mit dem dreifach gestuften Faktor „Messzeitpunkt“ und dem zweifach gestuften Faktor „Gruppenzugehörigkeit“ durchgeführt. Die Replikation wurde aufgrund der gerichteten Hypothesen zusätzlich anhand paarweiser geplanter

Kontrastanalysen zwischen einzelnen Faktorstufen des Faktors „Messzeitpunkt“ (t1 und t2) überprüft.

Bei den Fragestellungen, die sich auf Mittelwertsunterschiede zwischen Experimental- und Kontrollgruppe in Phase 1 (Bearbeitung der Lernumgebung in Dyaden) und damit *zwischen* den Dyaden beziehen, wurde wie folgt vorgegangen: Nach einer globalen Überprüfung des Haupteffektes Gruppenzugehörigkeit wurde aufgrund der wechselseitigen Abhängigkeit der Daten der Lernpartner und der damit hierarchisch geschachtelten Datenstruktur in den Dyaden die Intraklassenkorrelation (ICC) zu t2 und t3 berechnet. So wurde der Varianzanteil bestimmt, der durch Interaktionsprozesse innerhalb der Dyaden erklärt werden kann. Bei signifikanter ICC wurden die Effekte der Gruppenzugehörigkeit anhand von Mehrebenenanalysen auf Dyadenebene zu t3 modelliert (vgl. Kenny, 2008). Unterschieden sich die Versuchsgruppen im Vorwissen hinsichtlich *konz. FW*, *strat. FW* bzw. *ErklW* signifikant, wurden die Effekte unter Einbezug dieser Variablen als Kovariaten untersucht. Ausgehend von einem Modell mit festen Effekten (Nullmodell) wurden nacheinander Modelle mit *random intercept*, *random slope* sowie *random intercept und random slope* berechnet (vgl. Field, 2009). Die Modelle wurde mittels χ^2 -Differenztests (-2LL) verglichen. Auch die Interaktionseffekte (als Differenz von t2 und t3 in den einzelnen Variablen) wurden auf Dyadenebene unter Einbezug des Vorwissens als Kovariate modelliert.

Das Signifikanzniveau der Analysen zur Überprüfung der internen Validität wurde aufgrund der Gleichheitshypothesen auf .20 adjustiert. Die Stichprobe umfasste nur Pbn, die an allen Terminen zur Bearbeitung der Lernumgebung in Phase 1 anwesend waren. Fehlende Werte in den Wissenstests wurden listenweise ausgeschlossen.

6. Ergebnisse

6.1 Vorbereitende Analysen

Bezüglich *Einstellungen zum kooperativen Lernen* und den *Einstellungen zum individuellen Lernen* unterschieden sich die beiden Versuchsgruppen nicht ($F(2,128) = 1.40, p = .24$; $F(2,127) = 0.82, p = .37$).

Die beiden Gruppen unterschieden sich signifikant im Vorwissen hinsichtlich *konz. FW* ($F(2,130) = 42.9, p < .001$) und *strat. FW* ($F(2,129) = 8.78, p = .003$). Im *ErklW* erübrigte sich eine Signifikanztestung. Deskriptive Statistiken finden sich in Tabelle 4.

Tabelle 4

Vorwissen; Mittelwerte und Standardabweichungen

	Experimentalgruppe <i>M (SD)</i>	Kontrollgruppe <i>M (SD)</i>
Konz. FW	8.03 (3.48)	4.19 (3.20)
Strat. FW	0.73 (1.79)	1.86 (2.48)
ErklW	3.86 (2.36)	3.86 (2.07)

Anmerkung. EG= Experimentalgruppe, KG= Kontrollgruppe

Zum zweiten Messzeitpunkt war die Intraklassenkorrelation in den Dyaden bzgl. *konz. FW*, *strat. FW* und *ErklW* signifikant, d.h. hier lag eine systematische wechselseitige Abhängigkeit der Testwerte vor. Dies war zum dritten Messzeitpunkt bei den Variablen *konz. FW* und *strat. FW* der Fall (vgl. Tab. 3).

Tabelle 5

Intraklassenkorrelation

	<i>ICC</i>	<i>p</i>	
t2	Konz. Fw	.51	.004
	Strat. Fw	.57	.001
	Proz. Fw	.12	.32
	Erklw	.43	.01
t3	Konz. FW	.50	.008
	Strat. FW	.47	.02
	Proz. FW	.32	.10
	ErklW	.31	.11

6.2 Hypothese 1: Replikation

Die deskriptiven Statistiken werden in Tabelle 6 dargestellt. Es zeigte sich ein signifikanter und großer Haupteffekt des Faktors Messzeitpunkt auf das *konz. FW* ($F(2,108) = 215, p < .001, \eta_p = .67$), *strat. FW* ($F(2,97) = 351, p < .001, \eta_p = .78$) und *ErklW* ($F(2,99) = 273, p < .001, \eta_p = .73$).

Die Kontrastanalysen zur *Replikation* waren beim *konz. FW* ($F(1,108) = 233, p < .001, \eta_p = .68$), *strat. FW* ($F(1,97) = 807, p < .001, \eta_p = .89$) und *ErklW* ($F(1,99) = 631, p < .001, \eta_p = .87$) signifikant. Die Effekte waren groß. Beide Gruppen erzielten in allen Variablen einen signifikanten Wissenszuwachs vom ersten zum zweiten Messzeitpunkt.

Tabelle 6

Mittelwerte und Standardabweichungen gesamt zu zwei Messzeitpunkten

	t1: <i>M (SD)</i>	t2: <i>M (SD)</i>
Konz. Fw ($n = 110$)	5.91 (3.86)	10.7 (1.45)
Strat. Fw ($n = 99$)	1.28 (2.24)	14.9 (4.74)
Erklw ($n = 101$)	3.83 (2.13)	12.2 (2.99)

6.3 Hypothese 2: Effekte der Seminarakzeptanz und Stabilität des Lernerfolgs

Haupteffekt Gruppenzugehörigkeit. Die deskriptiven Statistiken finden sich in Tabelle 7. Die Experimentalgruppe erzielte in allen Wissensarten deskriptiv höhere Werte als die Kontrollgruppe.

Der globale Haupteffekt des Faktors Gruppenzugehörigkeit war hinsichtlich der Variable *konz. FW* ($F(1,108) = 54.1, p < .001, \eta_p = .96$) signifikant und groß, hinsichtlich *strat. FW* ($F(1,97) = 4.23, p = .04, \eta_p = .04$) und *ErklW* ($F(1,99) = 4.86, p = .03, \eta_p = .05$) zwar signifikant, aber klein. Im *proz. FW* ($F(1,109) = 3.54, p = .06$) unterschieden sich die Gruppen nicht signifikant.

Aufgrund der signifikanten Vorwissensunterschiede und der signifikanten ICC im *konz. FW* und *strat. FW* wurde der Haupteffekt der Gruppenzugehörigkeit zum dritten Messzeitpunkt unter Einbezug der Kovariate Vorwissen (Leistung in den Variablen zu t1) auf Dyadenebene modelliert.

Beim *konz. FW* konnte weder das Modell mit *random intercept* ($\Delta\chi^2(1) = 3.50, n.s.$) noch das Modell mit *random slope* ($\Delta\chi^2(1) = 2.42, n.s.$) die Daten besser beschreiben als das Nullmodell. Eine Berechnung mit *random intercept und random slope* erübrigte sich somit. Der Effekt der Gruppenzugehörigkeit war unter Einbezug des Vorwissens signifikant, aber

klein ($F(1,110) = 11.3, p = .001, \eta_p = .09$). Der Haupteffekt der Kovariate war nicht signifikant ($F(1,110) = 3.10, p = .08$). Der gefundene globale Haupteffekt der Gruppenzugehörigkeit geht demnach auf die Gruppenunterschiede zum dritten Messzeitpunkt zurück.

Beim *strat. FW* beschrieb das Modell mit *random intercept* gegenüber dem Nullmodell die Daten besser ($\Delta\chi^2(1) = 14.7, p < .01$). Die Modelle mit *random intercept und random slope* unterschieden sich hiervon nicht signifikant ($\Delta\chi^2(1) = 1.59, n.s.$). Wurde der Effekt der Gruppenzugehörigkeit auf Dyadenebene untersucht, ergab sich unter Einbezug des Vorwissens jedoch *kein* signifikanter Unterschied zwischen den Gruppen zum dritten Messzeitpunkt ($F(1,68.6) = 0.64, p = .43$). Der Haupteffekt der Kovariate war nicht signifikant ($F(1,91.9) = 1.84, p = .18$). Der oben berichtete globale Haupteffekt der Gruppenzugehörigkeit geht demnach auf *nicht* auf Gruppenunterschiede zum ersten bzw. dritten MZP, sondern auf den systematischen Einfluss von Interaktionsprozessen innerhalb der Dyaden zurück.

Tabelle 7

Mittelwerte und Standardabweichungen der Versuchsgruppen zu t3

t3	EG <i>M (SD)</i>	KG <i>M (SD)</i>
Konz. Fw ($n = 110$)	11.5 (0.87)	10.2 (1.59)
Strat. Fw ($n = 99$)	13.8 (6.63)	12.6 (5.54)
Erklw ($n = 101$)	9.74 (3.34)	8.22 (3.20)
Proz. Fw ($n = 111$)	19.31 (4.71)	17.4 (4.64)

Interaktionseffekt Messzeitpunkt und Gruppenzugehörigkeit. Tabelle 8 sind die deskriptiven Statistiken zu entnehmen. Der Interaktionseffekt von Messzeitpunkt und Gruppenzugehörigkeit war hinsichtlich *konz. FW* signifikant, aber klein ($F(2,108) = 21.8, p < .001, \eta_p = .17$). Beim *strat. FW* ($F(2,97) = 0.10, p = .90$), *proz. FW* ($F(2,109) = 0.47, p = .49$) sowie beim *Erklw* ($F(2,99) = 1.71, p = .18$) war der Interaktionseffekt nicht signifikant. Auch der Interaktionseffekt (als Differenz zwischen den Ergebnissen zum zweiten und dritten Messzeitpunkt) wurde aufgrund der signifikanten Vorwissensunterschiede und der signifikanten ICC im *konz. FW* und *strat. FW* unter Einbezug des Vorwissens als Kovariate auf Dyadenebene modelliert. Da die Gruppen zum ersten Messzeitpunkt im *Erklw* vergleichbar waren, erübrigte sich eine kovarianzanalytische Auswertung. Auch im *Erklw* wurden die Daten aufgrund der signifikanten ICC auf Dyadenebene modelliert.

Beim *konz. FW* konnte das Modell mit *random intercept* gegenüber dem Nullmodell mit festen Effekten die Daten besser beschreiben ($\Delta\chi^2(1) = 3.94, p < .05$). Der Vergleich mit dem Modell mit *random intercept und random slope* war nicht signifikant ($\Delta\chi^2(1) = 0, n.s.$). Der Interaktionseffekt war auf Dyadenebene unter Einbezug des Vorwissens *nicht* signifikant ($F(1,110) = 2.69, p = .10$), der Haupteffekt der Kovariate war ebenfalls nicht signifikant ($F(1,110) = 1.26, p = .26$). Auch hier geht also der global gefundene Effekt auf eine systematische Interaktion innerhalb der Dyaden zurück. Das Vorwissen hat auch hier keinen Einfluss.

Auch beim *strat. FW* beschrieb das Modell mit *random intercept* gegenüber dem Nullmodell die Daten besser ($\Delta\chi^2(1) = 18.68, p < .01$). Das Modell mit *random intercept und random slope* unterschied sich hiervon nicht signifikant ($\Delta\chi^2(1) = 0.03, n.s.$). Der Interaktionseffekt war auch auf Dyadenebene unter Einbezug des Vorwissens nicht signifikant ($F(1,71.2) = 0, p = .99$), der Haupteffekt der Kovariate war ebenfalls nicht signifikant ($F(1,88.6) = 0.06, p = .82$).

Beim *Erklw* konnte ebenfalls das Modell mit *random intercept* gegenüber dem Nullmodell die Daten besser beschreiben ($\Delta\chi^2(1) = 7.84, p < .01$). Der Vergleich mit dem Modell mit *random intercept und random slope* war auch hier nicht signifikant ($\Delta\chi^2(1) = 0.65, n.s.$). Auch auf Dyadenebene war der Interaktionseffekt nicht signifikant ($F(1,71.2) = 0, p = .99$).

Tabelle 8

Mittelwerte und Standardabweichungen der Versuchsgruppen zu t2 und t3

	t2		t3	
	EG	KG	EG	EG
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Konz. Fw	11.2 (1.55)	10.2 (1.55)	11.48 (0.87)	10.2 (1.60)
Strat. Fw	15.7 (4.61)	14.1 (4.77)	13.8 (6.63)	12.6 (5.54)
Erklw	12.6 (3.07)	11.8 (2.89)	9.74 (3.43)	8.22 (3.20)
Proz. Fw	21.4 (5.21)	20.2 (5.82)	19.31 (4.71)	17.42 (4.64)

7. Diskussion

7.1 Interne Validität

Die Versuchsgruppen unterschieden sich hinsichtlich der im Vortest erhobenen Kontrollvariablen nicht. Kognitive Variablen, in denen sich die Gruppen im Vorwissen unterschieden, wurden als Kovariaten in die Berechnungen aufgenommen, diese waren statistisch nicht signifikant. Die interne Validität der Studie konnte somit in Bezug auf potenziell relevante Störvariablen abgesichert werden.

7.2 Replikation und Stabilität der Lerneffekte

Die Hypothese zur *Replikation* der Studie von Wagner et al. (2015, angenommen) wurde bestätigt. Die Studierenden erzielten in allen Nachtests substanziell höhere Werte als in den Vortests, die Lernwirksamkeit der fehlerbasierten Lernumgebung wurde somit erneut nachgewiesen. Die deskriptiven Ergebnisse in den Nachtests unmittelbar nach dem Training in Phase 1 sind mit denjenigen der Vorgängerstudien vergleichbar (vgl. Wagner et al., 2015, angenommen).

Bezüglich des *Haupteffekts der erweiterten Seminarkonzeption* sind die Ergebnisse weniger eindeutig. Ein Vorteil der erweiterten Seminarkonzeption am Ende des Seminars (dritter Messzeitpunkt) zeigte sich unter Einbezug der Dyadenebene in ein Mehrebenenmodell nur für *konz. FW* und *ErklW*. Zwar erzielten die Lernenden in der erweiterten Seminarkonzeption beim *proz. FW* und *strat. FW* ebenfalls einen höheren Lernerfolg, der Unterschied war aber nicht signifikant.

Entsprechend fielen die Ergebnisse zur Stabilität der Lerneffekte als Interaktion zwischen Seminarkonzeption und Messzeitpunkt aus. Im Vergleich zwischen den Ergebnissen der Nachtests unmittelbar nach dem Training und am Ende des Seminars zeigte sich zwar nur eine geringfügige Verschlechterung, so dass vorsichtig von einem Nachweis der Stabilität der Lerneffekte gesprochen werden kann. Der angenommene Lernfortschritt in der erweiterten Seminarkonzeption konnte aber nur für das *konz. FW* nachgewiesen werden. In den übrigen Wissensarten fiel zumindest die ohnehin geringe Verschlechterung zum Ende des Seminars geringer aus als in der Kontrollgruppe.

Aus einer *statistischen* Perspektive sind die vorhandenen signifikanten Vorteile der erweiterten Seminarkonzeption am Seminarende erklärbar, wenn man die deskriptiv besseren Ergebnisse der Lernenden in der erweiterten Seminarkonzeption beim Nachtest berücksichtigt. Der zu diesem Zeitpunkt nicht signifikante Vorsprung dieser Gruppe blieb bestehen und

fürte am Seminarende unter Einbezug der geringeren Verschlechterung zu einem signifikanten Vorteil gegenüber der Kontrollgruppe mit regulärer Seminarkonzeption. Damit war der Unterschied zwischen den Gruppen zwar groß genug, um Vorteile am Seminarende statistisch nachweisen zu können, jedoch zu klein, um innerhalb der Interaktion signifikant zu werden und damit deutliche Lernvorteile gegenüber der Kontrollgruppe belegen zu können.

Mögliche *inhaltliche* Gründe für die hypothesenwidrig geringe Lernwirksamkeit der erweiterten Seminarkonzeption könnten bei der Operationalisierung des *Blended Learning* oder des *Fadings instruktionaler Unterstützung* liegen. Die selbstgesteuerte Lernphase war möglicherweise für die Bedürfnisse der Studierenden zu unstrukturiert. Der Wechsel der Herangehensweisen könnte zu abrupt gewesen sein, zudem nahmen die Studierenden während der selbständigen Problembearbeitung u.U. zu wenig Unterstützung in Anspruch. Auch könnten die bereitgestellten Arbeitsmaterialien nicht detailliert genug oder die Anweisungen unklar gewesen sein. Zudem war möglicherweise das *Fading* der instruktionalen Unterstützung zu abrupt. Angelehnt an Salden et al. (2010) entsprach das Fading durch die Dozenten eher einem sogenannte *fixed fading*. Dabei wird das Fading fest vorgegeben und nicht an den tatsächlichen Lernfortschritt der Pbn angepasst. Zur genaueren Anpassung im Sinne des von den o.g. Autoren beschriebenen *adaptiven Fadings* an den Lernfortschritt der Studierenden wären beispielsweise Zwischenüberprüfungen des Lernstands oder Selbsteinschätzungen, z.B. in Form metakognitiver Prompts (Bannert, Hildebrand & Mengelkamp, 2009) denkbar. Auch hätte man Zwischenstufen in die Progression von den vollständig ausgearbeiteten Lösungsbeispielen der Lernumgebung hin zur selbständigen Konstruktion von Szenarien einfügen können, in denen beispielsweise nur Teile eines Problemszenarios hätten ergänzt werden müssen.

Effekte von *Blended Learning* und *Fading* konnten leider nicht separat getestet werden, da diese beiden Maßnahmen in unserem Ansatz integrativ miteinander verwoben waren. Die enge Verknüpfung könnte jedoch dazu geführt haben, dass sich die Effekte der beiden Maßnahmen teilweise gegenseitig neutralisierten. Blended Learning erfordert zur Sicherung des Lernerfolgs in den selbstgesteuerten Lernphasen instruktionaler Unterstützung durch klare Arbeitsanweisungen und den Zugriff auf weitere Ressourcen (z.B. Präsenztreffen, unterstützte Transferphasen, Winkler & Mandl, 2003). Die diesbezügliche Unterstützung durch die Dozenten in Phase 2 wurde möglicherweise zu schnell reduziert.

Eine positivere Interpretation ist angesichts der Ergebnisse zum dritten Messzeitpunkt ebenso schlüssig. Die Lernwirksamkeit des Trainings in Phase 1 war analog zu früheren Studien (Wagner et al, 2014a; Wagner et al., 2014b) bereits sehr hoch (im Schnitt 75% der

Maximalpunkte in den einzelnen Wissensarten), so dass durch zusätzliche Interventionen keine starken Verbesserungen zu erwarten waren. Im Zusammenhang damit muss auch der Einsatz identischer Tests am Seminarende hinterfragt werden. Dieser ermöglichte zwar Messwiederholungen, ist aus motivationaler und methodischer Sicht jedoch fraglich. Bei drei identischen Testbatterien war nicht mehr zu erwarten, dass die Studierenden insbesondere in der Wissensanwendung beim ErklW *noch* umfassendere neue Ansätze oder Ideen generieren würden. Denkbar wäre hier die Entwicklung neuer Testszenarien, deren Komplexität durch die Auswahl oder Verknüpfung mehrerer unterschiedlicher Theorieansätze zur Erklärung einer Situation gesteigert wird. Diese könnten die angenommen höheren Fähigkeiten der Pbn nach einer Vertiefungsphase besser abbilden.

7.3 Pädagogische Konsequenzen und Ausblick

Der Einsatz der fehlerbasierten Lernumgebung in Seminaren der Lehramtsausbildung erscheint nach wie vor vielversprechend. Der Umgang mit wissenschaftlichem Wissen wie Evidenzen und Theorien und damit die wissenschaftliche Qualität der Erklärungen konnte bei angehenden Lehrkräften gefördert werden. Jedoch besteht Optimierungsbedarf hinsichtlich der Umsetzung einer zweiten, systematisch auf die Lernumgebung abgestimmten Seminarphase. Insbesondere muss beim verwendeten *Blended Learning*-Konzept von Dozentenseite darauf geachtet werden, dass die Arbeitsanweisungen verstanden werden und die Hilfestellung durch die Dozenten von den Studierenden eingefordert wird.

Zusätzlich zu einer kleinschrittigeren Stufung des Fadings beschreiben Wecker und Fischer (2011) die Möglichkeit von *peer monitoring* beim Fading, also der externen Überwachung des Lernprozesses durch einen Lernpartner. Hier bietet sich analog zur dyadischen Bearbeitung des Trainings in der ersten Seminarphase eine kollaborative Bearbeitung der Arbeitsanweisungen in der zweiten Phase an.

8. Literatur

- Bannert, M., Hildebrand, M. & Mengelkamp, C. (2009). Effects of a metacognitive support device in learning environments. *Computers in Human Behavior*, 25 (4), 829-835
- Beck, K. & Krapp, A. (2006). In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 31-71). Weinheim: Beltz.
- Bromme, R. (2008). Lehrerexpertise. In W. Schneider & M. Hasselhorn (Hrsg.). *Handbuch der Pädagogischen Psychologie* (S. 159-167). Göttingen: Hogrefe.
- Bromme, R., Prenzel, M. & Jäger, M. (2014). Empirische Bildungsforschung und evidenzbasierte Bildungspolitik. Eine Analyse von Anforderungen an die Darstellung, Interpretation und Rezeption empirischer Befunde. *Zeitschrift für Erziehungswissenschaft* (S. 3 - 54). Wiesbaden: VS Springer.
- Cogniton and Technology Group at Vanderbilt (1992). The jasper series as an example of anchored instruction: Theory, program, description, and assessment data. *Educational Psychologist*, 27, 291-315.
- Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale: Lawrence Erlbaum Associates.
- De Jong, T. & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31, 105-113.
- Durkin, K. & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction*, 22, 206-214.
- Eiriksdottir, E., & Catrambone, R. (in press). Procedural instructions, principles, and examples: How to structure instructions for procedural tasks to enhance performance, learning, and transfer. *Human Factors*. Online verfügbar unter <http://hfs.sagepub.com/content/53/6/749.full.pdf+html> [08.10.2014]
- Field, A. P. (2009). *Discovering statistics using SPSS: and sex and drugs and rock 'n' roll*. London: Sage publications.
- Fischer, F. Kollar, I., Stegmann, K. & Wecker, C. (2013) Toward a Script Theory of Guidance in Computer-Supported Collaborative Learning. *Educational Psychologist*, 48 (1), 56-66.
- Fischer, F., Kollar, I., Ufer, S., Sodian, B. Hussmann, H., Pekrun, R., Neuhaus, B., Dorner, B. Pankofer, S., Fischer, M., Strijbos, J., Heene, M., Eberle, J. (2014). Scientific Reasoning

- and Argumentation: Advancing an Interdisciplinary Research Agenda in Education. *Frontline Learning Research* (3), 1-18.
- Kalyuga, S., Rikers, R., Paas, F. (2012). Educational implications of expertise reversal effect in learning and performance of complex cognitive and sensorimotor skills. *Educational Psychology Review*, 24, 313-337
- Kenny, D.A. (2008). *Dyadic Analyses*. verfügbar unter <http://davidakenny.net/dyad.htm> [10.10.2014].
- Kerres, M. & Jechle, T. (2002). Didaktische Konzeption des Telelernens. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia und Internet. Lehrbuch für Studium und Praxis* (S. 266-281). Weinheim: PVU.
- Klein, M., Wagner, K. & Stark, R. (2012). Theoretisieren für die Praxis - Analyse und Förderung der Anwendung wissenschaftlichen Wissens in schulischen Kontexten. In : Krämer, M., Dutke, S., & Barenberg, J. (Hrsg.). *Psychologiedidaktik und Evaluation IX* (Auflage: 1.). Aachen: Shaker.
- Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004). Standards für die Lehrerbildung: Bildungswissenschaften (Beschluss der Kultusministerkonferenz vom 16.12.2004). Bonn: KMK, verfügbar unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf
- Krause, U.-M. (2007). *Feedback und kooperatives Lernen*. Münster: Waxmann.
- Mandl, H. & Kopp, B. (2006): *Blended Learning: Forschungsfragen und Perspektiven*. (Forschungsbericht Nr. 182). LMU München: Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie, Internet, ISSN 1614-6336
- Ohlsson, S. (1992). The cognitive skill of theory articulation: A neglected aspect of science education. *Science & Education*, 1, 181-192.
- Oser, F. (2007). Aus Fehlern lernen. In: M. Göhlich, Ch. Wulf & J. Zirfas (Hrsg.): *Pädagogische Theorien des Lernens* (S. 203-212). Weinheim und Basel: Beltz Verlag.
- Reinmann, G. & Mandl, H. (2006). Unterrichten und Lernumgebungen gestalten. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 613-658). Weinheim: Beltz.
- Renkl, A. & Atkinson, R. K. (2010). Learning from worked-out examples and problem solving. In J. L. Plass, R. Moreno & R. Brünken (Eds.), *Cognitive load theory* (pp. 91-108). Cambridge: Cambridge University Press

- Salden, R. J. C. M., Aleven, V., Schwonke, R. & Renkl, A. (2010). The expertise reversal effect and worked examples in tutored problem solving. *Instructional Science*, 38(3), 289-307.
- Schmidt, H.G, Loyens, S.M.M, Van Gog, T., & Paas, F. (2007). Problem-based learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark. *Educational Psychologist* 42 (2), 91-97
- Schworm, S. & Renkl, A. (2007). Learning argumentations skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology*, 99, 285-296.
- Stark, R. (2005). Constructing arguments in educational discourses. In H. Gruber, C. Harteis, R. Mulder & M. Rehl (Eds.), *Bridging individual, organisational, and cultural aspects of professional learning* (pp. 64-71). Regensburg: S. Roderer.
- Steffens, D. & Reiß, M. (2009): Blended Learning in der Hochschullehre. Vom Nebeneinander der Präsenzlehre und des E-Learning zum integrierten Blended Learning-Konzept. *HSW – das Hochschulwesen*, 57 (4), 115-123.
- Wagner, K., Klein, M., Klopp, E. & Stark, R. (2014a). Theoretisieren für die Praxis: Förderung anwendbaren pädagogischen Wissens anhand advokatorischer Fehler. *Journal für Lehrerinnen- und Lehrerbildung*. 5 (1). 115-140.
- Wagner, K., Klein, M., Klopp, E. & Stark, R. (2014b). Instruktionale Unterstützung beim Lernen aus advokatorischen Fehlern in der Lehramtsausbildung: Effekte auf die Anwendung wissenschaftlichen Wissens. *Psychologie in Erziehung und Unterricht*.
- Wagner, K., Klein, M., Klopp, E. & Stark, R. (2015, angenommen). Förderung anwendbaren bildungswissenschaftlichen Wissens anhand kollaborativem Lernen aus Fehlern. *Unterrichtswissenschaft*. 61, 287 - 301.
- Wecker, C. (2012). *Vom Sollen zum Können. Fading instruktionaler Skripts zur Förderung von Argumentationskompetenz*. Münster: Waxmann.
- Wecker, C. & Fischer, F. (2011). From guided to self-regulated performance of domain-general skills: The role of peer monitoring during the fading of instructional scripts. *Learning and Instruction*, 21 (6), 746-756.
- Weinberger, A., Stegmann, K., Fischer, F., & Mandl, H. (2007). Scripting argumentative knowledge construction in computer-supported learning environments. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Scripting computer-supported collaborative learning: Cognitive, computational, and educational perspectives* (pp. 191–211). New York: Springer.

9. Gesamtdiskussion

Die Lernwirksamkeit des Lernens aus Fehlern zur Förderung der Theorieanwendung konnte im Rahmen der vorliegenden Arbeit für die Domäne der Lehramtsausbildung gezeigt werden. Auch die Implementation in ein Seminar der Lehrerbildung verlief erfolgreich, wenn auch nicht ganz im erwarteten Umfang. Damit eignet sich dieser Ansatz in Form der in Studie 1 und 2 beschriebenen integrierten Lernumgebungen sowohl als „standalone“- oder Zusatzangebot, in Form einer integrierten Seminarkonzeption (Studie 3) aber auch zum Einsatz im Rahmen des regulären Curriculums. Im Zuge dessen kann er zur Lösung zahlreicher im Bereich der Lehramtsausbildung häufig beklagter Probleme beitragen. Dazu zählen die Theorie-Praxis-Kluft (Gordon & O'Brien, 2007; Le Cornu & Ewing, 2008; Neuweg, 2007), Probleme bei der Anwendung wissenschaftlichen Wissens (z.B. Star & Strickland, 2008; Stark, 2005; Stark et al., 2010) sowie ungünstige Einstellungen und die daraus resultierende Ablehnung oder Vernachlässigung der wissenschaftlichen Wissensbasis im Beruf (Bainbridge, 2011; Ritchie & Wilson, 2001; Terhart, 2006).

Damit wird ein Beitrag zur systematischen Förderung der Kompetenzen zur Bewertung und Nutzung bildungswissenschaftlicher Theorien (Neuweg, 2007; Meier, 2006; Stark, Herzmann & Krause, 2010) und Evidenz (Bauer et al., 2010; Terhart, Lohmann & Seidel, 2010; Weber & Achtenhagen, 2009) geleistet. Ebenso können zahlreiche Forderungen aus dem Kompetenzkatalog der Standards der Lehrerbildung (KMK, 2004; 2014) umgesetzt werden. Dort heißt es:

„Ausgehend von dem Schwerpunkt Theorie erschließt die erste Phase die pädagogische Praxis, während in der zweiten Phase diese Praxis und deren theoriegeleitete Reflexion im Zentrum stehen. Das Verhältnis zwischen universitärer und stärker berufspraktisch ausgerichteter Ausbildung ist so zu koordinieren, dass insgesamt ein systematischer kumulativer Erfahrungs- und Kompetenzaufbau erreicht wird.“

Das hier vorgestellte Forschungsprojekt stellt also eine Brücke zwischen diesen beiden Phasen her. Wissenschaftliches Wissen, das in der ersten Phase erworben wird, konnte mit Kompetenzen in der *Anwendung* wissenschaftlichen Wissens angereichert werden, die insbesondere in der zweiten Phase von Bedeutung sind.

9.1 Übergreifende Diskussion aller Studien mit Bezug zu den theoretischen Grundlagen

Wie bereits in Kapitel 4 dargestellt, gibt es verschiedene Ansätze zum Lernen aus Fehlern. Insbesondere liegen teilweise inkonsistente Befunde dazu vor, ob das Lernen aus eigenen oder fremden Fehlern lernwirksamer ist (z.B. Durkin & Rittle-Johnson, 2012; Ivancic & Hesketh, 2000; Kapur, 2014; Maxwell et al., 2010; Stark et al., 2011). Analog dazu, dass verschiedene Handlungen situationsbedingt als richtig oder falsch gelten (vgl. dazu auch Ohlsson, 1996, zur Domänenspezifität von Fehlern), kann auf Grundlage dieser Befunde aus verschiedenen Domänen auch eine gewisse Domänenspezifität bezüglich der geeigneten Methode des Lernens aus Fehlern angenommen werden.

In Studie 1 wurde daher im Bereich der Lehramtsausbildung das Lernen aus eigenen Fehlern (*Productive Failure*, Kapur & Bielaczyc, 2012) mit dem Lernen aus fremden Fehlern (*advokatorisches Lernen aus Fehlern*; Oser, 2007) verglichen. Ziel war, wie in allen vorgestellten Studien, die Förderung der Theorieanwendung zur Erklärung komplexer schulischer Situationen und Phänomene. Hier zeigte sich, dass das advokatorische Lernen aus Fehlern durch die explizite Analyse und Erklärung vorgegebener Fehler gegenüber dem Lernen aus eigenen Fehlern in besonderem Maße den Erwerb *negativen* Wissens (vgl. Kap. 4.3) fördert. Die Befunde sind damit konsistent zu Maxwell et al. (2010), speziell in Bezug auf den Vergleich zwischen Lernen aus eigenen und fremden Fehlern. Auch andere Studien aus dem Bereich der Mathematik, Lehrerbildung und Medizin (z.B. Durkin & Rittle-Johnson, 2012; Klein et al., 2012; Stark et al., 2011) fanden positive Effekte des Lernens aus fremden Fehlern. Die Ergebnisse stehen im Gegensatz zu den Ergebnissen von Kapur (2014) oder Ivancic und Hesketh (2000), bei denen das Lernen aus eigenen Fehlern im Bereich der Mathematik und bei einem Fahrtraining überlegen war. Kapur (ebd.) operationalisierte jedoch das Lernen aus fremden Fehlern nicht unter Berücksichtigung der Bedingungen, die bei Oser und Spsychiger (2005) genannt werden. Ähnlich war es auch bei der Studie von Ivancic und Hesketh (2010). Hier kann zusätzlich die Emotionalität des Lernens aus eigenen Fehlern zur Erklärung herangezogen werden. Im Kontext eines Fahrtrainings, in dem Fehler als Schritte im Lernprozess ausdrücklich erlaubt sind, wird durch eigene Fehler eher eine „produktive Beschämung“ im Sinne von Oser und Spsychiger (2005) ausgelöst. Demzufolge sind die negativen Emotionen, belegt durch die Verringerung des Vertrauens in die eigenen fahrerischen Fähigkeiten, hier durchaus lernförderlich. In der Bedingung mit fremden Fehlern zeigte sich kein derartiger Effekt, allerdings kann es als zweifelhaft angesehen werden, dass

durch die Vorführung eines Videos eine Rollenübernahme durch die Lernenden stattfand (vgl. Spychiger, 2004). Entsprechend wurden hier keine negativen Emotionen ausgelöst und damit der Lerneffekt stark verringert. In Studie 1 wurde dagegen die Rollenübernahme durch verschiedene Maßnahmen explizit unterstützt (vgl. Kap. 6.3).

Auf Grundlage der Befunde aus Studie 1 wurde das advokatorische Lernen als geeignetere Methode zur Förderung der Theorieanwendung ausgewählt. Es zeigte sich jedoch auch Optimierungsbedarf, insbesondere bei der Erfassung der verschiedenen Wissensarten. Der Einsatz szenariobasierter Tests auf der Grundlage des Key-Feature-Ansatzes (Kopp, Möltner & Fischer, 2006), der eine hochinferente Auswertung erforderte, ermöglicht nur Analysen zur *Wissensanwendung*, das zugrundeliegende konzeptuelle Wissen wurde nicht dezidiert erfasst. Im Sinne des von Krause (2007; vgl. Kap. 3.2.3) beschriebenen notwendigen Zusammenspiels verschiedener Wissensarten bei der Wissensanwendung konnte dieses zwar angenommen, aber nicht nachgewiesen werden.

Dieser Mangel sollte in der Folgestudie behoben werden, gleichzeitig sollte eine umfassendere Unterstützung der Bedingungen des advokatorischen Lernens aus Fehlern erfolgen. Diese umfassen neben Verständnis der Fehler und den Aufbau korrekten Wissens sowie von Strategien zur Fehlervermeidung auch den Kontrast an der korrekten Lösung (Curry, 2004; Oser & Spychiger, 2005). Der dazu notwendige aktive Nachvollzug der Fehler durch „umgekehrtes Modelllernen“ (Spychiger, 2004, S. 3) wird durch eine Rollenübernahme durch die Lernenden sichergestellt. Dabei kann zusätzliche instruktionale Unterstützung die Lernwirksamkeit erhöhen.

Studie 2 untersuchte die Umsetzung dieser Maßnahmen in Form einer integrierten Lernumgebung auf Basis ausgearbeiteter Lösungsbeispiele. Untersucht wurde dabei, wie viel instruktionale Unterstützung in Form verschiedener Prompts notwendig ist, um insbesondere den Erwerb von Fehlervermeidungsstrategien zu fördern. Die Inhalte der Lernumgebung entsprachen dabei größtenteils Studie 1. In der logischen Weiterentwicklung von Studie 1 wurden auch die dort verwendeten Fehler in Form einer Fehlermatrix systematisiert und erweitert (vgl. Kap. 7).

Um das Zusammenwirken der unterschiedlichen Wissensarten bei der Wissensanwendung (Krause, 2007, s. Kap. 3.2.3) überprüfen zu können, wurden auf Grundlage der Taxonomie von DeJong und Ferguson-Hessler (1996) neue Testverfahren entwickelt, um insbesondere konzeptuelles Wissen über Fehler und strategisches Wissen über deren Vermeidung getrennt von deren Anwendung zu erfassen. Gleichzeitig wurden die Tests zur Wissensanwendung

überarbeitet und die Anzahl der szenariobasierten Tests aus Gründen der Testökonomie reduziert.

In Verbindung mit der relativ großen Stichprobe konnte überzeugend nachgewiesen werden, dass die Lernumgebung in Bezug auf die o.g. Kompetenzen in hohem Maße lernwirksam ist. Der Erwerb konzeptuellen Wissens über Fehler und Strategien zu deren Vermeidung, aber auch die Anwendung in der Identifikation von Fehlern und die Vermeidung von Fehlern bei der Formulierung einer eigenen Erklärung konnten gefördert werden. Dabei erzielte hypothesenkonform die Bedingung mit der maximalen instruktionalen Unterstützung die höchste Lernwirksamkeit. Diese Befunde sind konsistent mit den Ergebnissen von Durkin und Rittle-Johnson (2012) und Stark et al. (2011).

In Studie 2 ergaben sich bei der Auswertung Probleme in Form von Deckeneffekten bzw. geringer Varianz in den Tests zum konzeptuellen Fehlerwissen, zudem wurde von Studierenden rückgemeldet, dass eine soziale Komponente bei der Arbeit mit der Lernumgebung vermisst wurde. Auf Grundlage dieser Ergebnisse wurde zunächst in einer weiteren Studie von Wagner, Klein, Klopp und Stark (2015; angenommen) eine kollaborative Variante der Arbeit mit der Lernumgebung untersucht. Diese erbrachte insbesondere bei der Fehleridentifikation sowie in der Anwendungsdimension Erklärungswissen zusätzliche, wenn auch geringe, Wissenszuwächse. Um die bei zahlreichen Autoren (z.B. Lind, 2001; Spychiger, 2004) geforderte Authentizität und Relevanz durch einen stärkeren Bezug zur Lebenswelt zu verstärken und die theoriegeleitete Reflexion *eigener* Erfahrungen (KMK, 2004; 2014) zu fördern, wurde eine umfassende integrierte Seminarkonzeption entwickelt. Darin wurde die kollaborative Bearbeitung der Lernumgebung durch eine zweite Seminarphase ergänzt, die systematisch Bezug auf das in der Lernumgebung erworbene Wissen nahm.

Ziel war neben einer Replikation früherer Befunde der Lernumgebung auch die Überprüfung der Stabilität und ggf. Verbesserung der Lerneffekte in Abhängigkeit von der Seminarkonzeption. In Studie 3 wurde eine erweiterte Seminarkonzeption auf Basis von Blended Learning, kombiniert mit Fading instruktionaler Unterstützung, mit einer regulären Seminarkonzeption verglichen. Die Studierenden konstruierten dabei in Einzelarbeit Problemszenarien nach dem Muster der Lernumgebung. Inhaltlich wurde die Auswahl der Problemstellungen von den Studierenden vorgenommen. So konnten in der Lernumgebung erworbene Kompetenzen in der Erklärung schulischer Situationen im Sinne der Reflexion eigener biografischer Lernerfahrungen (vgl. KMK, 2004; 2014, S. 6) angewendet werden.

Eine soziale Komponente wurde zusätzlich zur kollaborativen Bearbeitung der Lernumgebung in Phase 1 durch die enge Zusammenarbeit mit den Dozenten (analog zur

Zusammenarbeit zwischen Lehrperson/Meister und Schüler im Cognitive Apprenticeship-Ansatz, Collins et al., 1989) sowie durch das Feedback der anderen Studierenden bei Besprechungen und Vorstellungen der eigenen Problemszenarien im Rahmen dafür vorgesehener Präsenztermine realisiert.

Die Konstruktion von Problemszenarien in Einzelarbeit, nachdem zuvor die Lernumgebung kollaborativ bearbeitet worden war, geschah dabei in erster Linie aus experimentalpsychologischen und methodologischen Überlegungen. Die Auswertung der Daten erforderte bereits auf Grundlage von Phase 1 eine Überprüfung der Abhängigkeit der individuellen Lernergebnisse von der Zugehörigkeit zu einer Lerndyade und damit die Berechnung eines Mehrebenenmodells. Gruppenarbeit in Phase 2 hätte eine weitere Ebene in diesem Modell notwendig gemacht, was die Anzahl auswertbarer Einzelergebnisse dramatisch verringert hätte.

Möglicherweise hätte aber in Phase 2 die Arbeit in Gruppen im Sinne von *peer monitoring* (Wecker & Fischer, 2011), also der gegenseitigen Überwachung des Lernfortschritts und der Fokussierung auf die Aufgabe, die Ergebnisse verbessert. Hier zeigte sich nämlich wider Erwarten ein Effekt der erweiterten Seminarkonzeption nur insofern, als die Experimentalgruppe in den Tests am Semesterende minimal, aber statistisch signifikant, geringere Wissensverluste aufwies als die Kontrollgruppe. Zwar waren diese in beiden Gruppen sehr klein, so dass die Lerneffekte durchaus als stabil gelten können, der erwartete Lernfortschritt in der erweiterten Seminarkonzeption blieb jedoch aus.

Dies weist neben Schwächen in der Umsetzung, insbesondere des Fadings in Phase 2, auch auf methodische Probleme des Einsatzes identischer Tests an drei Messzeitpunkten hin. Mögliche Zusatzeffekte der Reflexion an eigenen, teilweise sehr komplexen Erfahrungen, konnten eventuell anhand der Auswertung der vorliegenden Testbatterie nicht erfasst werden. Eine zusätzliche qualitative Auswertung könnte zeigen, ob es Veränderungen im Aufbau der Argumentation der Lernenden (bei gleichbleibender Punktzahl) gab, ob also beispielsweise Aspekte oder Ansätze herangezogen wurden, die in den Tests unmittelbar nach dem Training nicht angesprochen wurden. Auf Grundlage der Studie zur Kollaboration von Wagner et al. (2015, angenommen), sowie Kopp, Ertl und Mandl (2006), kann angenommen werden, dass sich derartige Effekte insbesondere in den Anwendungsdimensionen zeigen würden.

Anwendung erfordert einen gewissen Grad an Automatisierung (Stark, 2001, s.a. Kap. 3.2.1). Dieser bezieht sich mit Blick auf die Qualitäten von Wissen im Sinne von DeJong und Ferguson-Hessler (1996) vor allem auf den Abruf von Wissen - aktiv bei deklarativem, d.h. unverbundenem, unstrukturiertem und wenig elaboriertem Wissen; automatisch, umfassend

und intuitiv bei kompiliertem Wissen (s. Kap. 3.2.1). Es ist anzunehmen, dass die zweite Seminarphase die Automatisierung förderte, da die eingehende Beschäftigung mit den behandelten Konzepten über das Training hinaus Struktur und Verarbeitungstiefe des Wissens verbessert haben sollte und die Qualitäten im Modell von DeJong und Ferguson-Hessler (ebd.) verbunden sind (ein höheres Niveau bzgl. einer Qualität von Wissen bedeutet ein höheres Niveau der anderen). Die vorhandenen Tests stoßen aber in Hinsicht auf die Erfassung der Qualitäten als Einflussfaktoren bei der Wissensanwendung an ihre Grenzen. Hier kann nur post-hoc aus den Ergebnissen der Tests zum Fehleridentifikation- und Erklärungswissen geschlossen werden, ob Wissen strukturierter und elaborierter ist und sich die Automatisierung verbessert hat. Es müssen spezifische Tests entwickelt werden, um Veränderungen in Struktur, Verarbeitungstiefe und Automatisierungsgrad direkt erfassen zu können. Angesichts der bereits sehr guten Ergebnisse bei der Messung unmittelbar nach dem Training wäre zudem möglicherweise ein höherer Schwierigkeitsgrad für die Tests am Seminarende angebracht gewesen, um eine höhere Varianz hervorzubringen.

Hier ist also die Entwicklung komplexerer Testverfahren angezeigt. In Anlehnung an die bewährten Verfahren aus den drei Studien bietet sich hier die Konstruktion szenariobasierter Tests nach dem Key-Feature-Ansatz (Kopp et al., 2006) an. Struktur und Verarbeitungstiefe von Wissen könnten durch Prompts zur Rekonstruktion von Wissen über die behandelten Konzepte (z.B. Zusammenfassung oder Auflistungen relevanter Theorien, Rekonstruktion der Fehlermatrix durch die Lernenden) abgebildet werden (vgl. DeJong & Ferguson-Hessler, 1995). Zur Erfassung des Automatisierungsgrades empfehlen DeJong und Ferguson-Hessler (ebd.) die Messung der Bearbeitungsgeschwindigkeit (*speed tests*). Die Komplexität der Tests könnte durch die zusätzliche Auswahl aus geeigneten und ungeeigneten Theorien gesteigert werden, inklusive einer expliziten Aufforderung an die Lernenden, ihre Auswahl zu begründen (*external material*, vgl. DeJong & Ferguson-Hessler, ebd). Dadurch würden gleichzeitig Kompetenzen in der Bewertung von wissenschaftlichen Theorien und Evidenzen (Bauer et al., 2010; Terhart, Lohmann & Seidel, 2010; Weber & Achtenhagen, 2009) erfasst. Deren Erwerb durch die Bearbeitung der Lernumgebung und die eigenständige, durch die Dozenten unterstützte Recherche in Phase 2 wurde zwar angenommen, aber durch die vorhandenen Tests nicht überprüft. Derartige Testverfahren wären auch in Bezug auf echte Nachhaltigkeit, also die „Dauerhaftigkeit der Lernresultate im Hinblick auf ihre Bedeutung für zukünftig zu bewältigende Handlungsproblematiken“ (Holzkamp 1993, S. 183; vgl. auch Transfer; Kap. 3.2.2) von Interesse.

9.3 Limitationen

Die Limitationen des Forschungsprojekts zeigen sich aus einer erkenntnisorientierten Perspektive vor allem auf der Prozessebene. Offen ist etwa die Frage, ob die Lernwirksamkeit des Lernens aus Fehlern tatsächlich auf die von Stark (2001) postulierte Beispielelaboration und damit verbundene Selbsterklärungen (VanLehn, 1999) zurückgeht und ob diese durch den von Oser und Spychiger (2005; s.a. Oser, 1999) geforderten Vergleich zwischen Fehler und korrekter Lösung initiiert wurden. Zukünftige Untersuchungen könnten hier etwa anhand von Prompts (s. Kap. 7) oder Laut-Denk-Protokollen (Ericsson & Simon, 1993) Prozessdaten bezüglich der Lernmechanismen erheben. Eine weitere Möglichkeit wäre die Auswertung der Kommunikation zwischen Dozenten und Studierenden in Studie 3. Die schriftliche Dokumentation durch den verpflichtenden E-Mail-Verkehr zur Bewertung und Rückmeldung der studentischen Problemszenarien in Studie 3 könnte hinsichtlich der prozessualen Entwicklung der verschiedenen Kompetenzen wie etwa der Identifikation und Analyse von fehlerhaften Erklärungsansätzen ausgewertet werden. Dies würde eine qualitative Auswertung erfordern, die gleichzeitig dazu genutzt werden könnte, den Anschluss an die in Kapitel 4.4 skizzierte Conceptual Change-Forschung stärker herauszuarbeiten und damit eine weitere theoretische Perspektive auf das Lernen aus Fehlern zu ermöglichen.

Weitere Limitationen des Forschungsprojekts aus einer praxisorientierten Perspektive liegen u.a. bei der Überprüfung der Nachhaltigkeit (Holzkamp, 1993; s.o.) vor. Hier bietet sich, wie bereits oben erwähnt, die Entwicklung weiterer Testverfahren an. Auch Kompetenzen in der Beurteilung der Qualität wissenschaftlichen Wissens, seien es Theorien in Lehrbüchern oder aktuelle empirische Studien, sind weder Bestandteil der Lernumgebung noch der Tests. Eine informelle Förderung findet zwar in der zweiten Phase der integrierten Seminarkonzeption statt, in der Studierende selbständig nach wissenschaftlich fundierten Erklärungsansätzen für ihre Problemszenarien suchen, diese müsste allerdings systematisiert und ihre Effekte überprüft werden. Möglichkeiten hierzu finden sich etwa in Studien von Wenglein et al. (2015) oder Trempler et al. (2015).

Nichtsdestotrotz ist die gezielte Förderung der Anwendung wissenschaftlichen Wissens bereits während des Studiums wichtig. Subjektive Theorien erfahrener Praktiker unterscheiden sich möglicherweise auf den von Bortz und Döring (2006, s. Kap. 2) genannten Dimensionen aufgrund ständiger Überprüfung und Bewährung im Berufsalltag in Bezug auf Systematik und Struktur kaum noch von wissenschaftlichem Wissen und sind auf einer Handlungsdimension möglicherweise in eng begrenzten Bereichen sogar zielführender. Wissenschaftliches Wissen erhebt aber den Anspruch, für weit größere Bereiche gültig zu

sein. Gerade diese Eigenschaft kann es Alltagswissen überlegen machen, wenn wenig eigene Erfahrung vorhanden ist. In diesem Fall ermöglicht die Struktur wissenschaftlicher Theorien durch ihre Formulierung in Wenn-Dann-Hypothesen die Prognose und Steuerung von Handlungen zumindest mit einer etwas größeren Sicherheit als der Bezug auf eigenes vorwissenschaftliches Wissen. Da sich Studierende zu Beginn der zweiten Phase der Lehramtsausbildung in genau dieser Lage befinden, ist der Rückgriff auf wissenschaftliches Wissen notwendig und sinnvoll. Es erscheint außerdem annehmbar, dass die Studierenden durch das Training anhand der Lernumgebung und die systematische Reflexion eigener Erfahrungen lernen, flexibler und systematischer zu denken. Hier kann spekuliert werden, dass sich das auch positiv auf Handlungen in der Praxis auswirkt. Diese Frage müsste aber in eigenen Studien untersucht werden.

Im Zusammenhang damit bleibt auch die Frage offen, ob die Anwendung wissenschaftlichen Wissens und die Förderung diesbezüglicher Kompetenzen positive Effekte bei Lehrern hervorrufen, die bereits jahrlange Berufserfahrung aufweisen. Eine entsprechende Annahme ist in der Literatur sehr weit verbreitet (Bauer & Prenzel, 2012; Bainbridge, 2011; Beck & Krapp, 2006; Bromme, 2008; Gordon & O'Brien, 2007; KMK, 2004; 2014; Meier, 2006; Neuweg, 2007; Ritchie & Wilson, 2001; Smith, 2005; Stark et al., 2010; Terhart, 2006; Terhart et al., 2010; Weber & Achtenhagen, 2009). Empirische Untersuchungen dazu, ob der systematische Bezug auf wissenschaftliches Wissen im Vergleich zu Alltagswissen wie gut entwickelten subjektiven Theorien erfahrener Lehrer tatsächlich besseren Unterricht fördert oder ob erfahrene Lehrer überhaupt auf wissenschaftliches Wissen zurückgreifen, sind selten (z.B. Hetmanek, 2014a; 2014b).

Diese Überlegung führt zu der Frage, inwieweit wissenschaftliches Wissen und darauf basierende Erklärungen und Lösungsansätze tatsächlich handlungsleitend werden. In der vorliegenden Form bleibt die Lernumgebung auf der Ebene der Erklärung, hier bietet sich eine Weiterentwicklung im Hinblick auf eine stärkere Handlungsorientierung an. Analog zum bereits bewährten Ansatz könnte diese auf dem Lernen aus Fehlern aufbauen. Zur Absicherung praxisrelevanter Fehler sollten diese jedoch in Zusammenarbeit mit Praktikern wie erfahrenen Lehrern oder Lehrerbildnern erarbeitet werden. Dazu sind etwa strukturierte Leitfadeninterviews (Meuser & Nagel, 1991) geeignet.

9.4 Pädagogische Konsequenzen

Aufgrund der positiven Resultate der Arbeit mit der Lernumgebung und auch der erfolgreichen Einbettung in ein Seminar erscheint der in Rahmen des vorliegenden

Forschungsprojekts eingesetzte Ansatz des Lernens aus Fehlern zur Förderung der Theorieanwendung trotz aller Einschränkungen aus einer pragmatischen Perspektive gut geeignet. Wünschenswert wäre nun die systematische Einbindung dieses oder ähnlicher Ansätze in das Lehramtsstudium. Dabei könnte der vorliegende Ansatz mit anderen Ansätzen im Bereich der Förderung von Kompetenzen in der Anwendung von Theorien oder Evidenzbasierung (z.B. KOMPARE, Hetmanek, 2014a; 2014b; Trempler et. al, 2015; Wenglein et al., 2015) ergänzt werden. Idealerweise könnte ein Baustein zur Vermittlung derartiger Kompetenzen etwa in das Kerncurriculum Erziehungswissenschaft (DGfE, 2004) aufgenommen werden. Um der etwa bei Terhart (2006, s. auch Ritchie & Wilson, 2001) beklagten Vernachlässigung wissenschaftlichen Wissens im Referendariat und im späteren Berufsleben entgegenzuwirken, könnte ein derartiges Training auch im Sinne der KMK (2004) in die zweite Phase der Lehramtsausbildung hinein erweitert und dort zur theoriegeleiteten Reflexion tatsächlicher Unterrichtserfahrungen genutzt werden.

10. Literatur Synopse

- Albanese, M. A. & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine*, 68, 52-81.
- Althof, W. (Hrsg.)(1999). *Fehlerwelten*. Opladen: Leske + Budrich.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge: Harvard University Press.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111, 1036-1060.
- Anderson, J. R., Reder, L. M. & Simon, H. A. (1996). Situated learning and education. *Educational Researcher*, 25(4), 5-11.
- Anderson, J.R. & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson (Ed.), *Handbook of reading research* (S. 255-291). New York: Longman.
- Arnold, P., Kilian, L., Thillosen, A. & Zimmer, G. (2004). *E-Learning - Handbuch für Hochschulen und Bildungszentren: Didaktik, Organisation, Qualität*. Nürnberg: BW Bildung und Wissen.
- Ausubel, D. (1974). *Psychologie des Unterrichts*. (Band 1 und 2). Stuttgart: Klett-Cotta.
- Bainbridge, A. (2011). Beginning teaching: the theory/practice divide. *Cliopsy*, 6, 25-32.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bärenfänger, O. (2002): Automatisierung der mündlichen L2-Produktion: Methodische Überlegungen. In W. Börner & K. Vogel (Hrsg.)(2002): *Grammatik im Fremdsprachenunterricht. Kognitive, psycholinguistische und erwerbstheoretische Perspektiven* (119-140). Tübingen: Narr
- Bartlett, F. C. (1932). *Remembering: An experimental and social study*. Cambridge: Cambridge University Press.
- Bauer, J. & Mulder, R. H. (2008). Conceptualisation of learning through errors at work - a literature review. In S. Billett, C. Harteis & A. Eteläpelto (Eds.), *Emerging perspectives of workplace learning* (S. 115-128). Rotterdam: Sense Publishers.
- Bauer, J. & Prenzel, M. (2012, August). Scientific and professional journals as resource for teachers' professional learning and evidence-based practice. Paper presented at the

- EARLI SIG 14 Learning and Professional Development Conference, Antwerp (Belgium)
- Bauer, J., Drechsel, B., Retelsdorf, J., Sporer, T., Rösler, L., Prenzel, M. & Möller, J. (2010). Panel zum Lehramtsstudium - PaLea: Entwicklungsverläufe zukünftiger Lehrkräfte im Kontext der Reform der Lehrerbildung. *Beiträge zur Hochschulforschung*, 32, 34-55.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J. & Weiß, M. (Hrsg.). (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich
- Beck, K. & Krapp, A. (2006). Wissenschaftstheoretische Grundfragen der Pädagogischen Psychologie. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 33-73). Weinheim: Beltz PVU.
- Bendorf, M. (2013). *Bedingungen und Mechanismen des Wissenstransfers: Lehr- und Lern-Arrangements für die Kundenberatung in Banken* (Google eBook). Springer
- Bereiter, C. (1995). A dispositional view of transfer. In A. McKeough, J. L. Lupart, & A. Marini (Eds.), *Teaching for transfer: Fostering generalization in learning* (S. 21-34). Hillsdale, NJ: Erlbaum
- Berkson, L. (1993). Problem-based learning: Have the expectations been met? *Academic Medicine*, 68 (October supplement), 79-88.
- Berliner, D. C. (1992). The nature of expertise in teaching. In F. K. Oser, A. Dick & J.-L. Patry (Eds.), *Effective and responsible teaching: The new synthesis*. San Francisco, CA: Jossey Bass.
- Berliner, D. C. and Carter, K. J. (1989). Differences in processing classroom information by expert and novice teachers. In J. Lowyck and C. Clark (Eds.), *Teacher thinking and professional action*. Leuven, Belgium: Leuven University Press.
- Berthold, K., Röder, H., Knörzer, D., Kessler, W. & Renkl, A. (2011). The double-edged effects of explanation prompts. *Computers in Human Behavior*, 27(1), 69-75.
- Billett, S. (2004). Workplace participatory practices. Conceptualising workplaces as learning environments. *Journal of Workplace Learning*, 16 (6), 312–324.
- BMJ Case Reports (2015). <http://casereports.bmj.com/> [24.4.2015].
- Bohnsack, R. (2000): *Rekonstruktive Sozialforschung. Einführung in Methodologie und Praxis qualitativer Forschung*. Opladen

- Booth, J. L., Lange, K. E., Koedinger, K. R., & Newton, K. J. (2013). Using example problems to improve student learning in algebra: Differentiating between correct and incorrect examples. *Learning and Instruction, 25*, 24-34.
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* [Research methods and evaluation in human and social sciences]. Berlin: Springer.
- Boshuizen, H. & Schmidt, H. (2008). The development of clinical reasoning expertise. In J. Higgs, M. A. Jones, S. Loftus, & N. Christensen (Hrsg.), *Clinical reasoning in the health professions*. Butterworth-Heinemann.
- Boshuizen, H. P. A., Schmidt, H. G., Custers, E. J. F. M., & van de Wiel, M. W. (1995). Knowledge development and restructuring in the domain of medicine: The role of theory and practice. *Learning and Instruction, 5*(4), 269–284.
- Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (Vol. 24, S. 61–100). Washington, DC: American Educational Research Association.
- Bromme, R. (1992). *Der Lehrer als Experte: Zur Psychologie des professionellen Wissens*. Bern: Huber.
- Bromme, R. (2008). Lehrerexpertise. In W. Schneider & M. Hasselhorn (Hrsg.). *Handbuch der Pädagogischen Psychologie* (S. 159-167). Göttingen: Hogrefe.
- Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review, 31*, 21-32
- Bürg, O. (2005). *Akzeptanz von E-Learning in Unternehmen. Die Bedeutung von institutionellen Rahmenbedingungen, Merkmalen des Individuums und Merkmalen der Lernumgebung für die Akzeptanz von E-Learning*. Berlin: Logos.
- Çalik, M., Okur, M., & Taylor, N. (2010). A Comparison of Different Conceptual Change Pedagogies Employed within the Topic of “Sound Propagation”. *Journal of Science Education and Technology, 20* (6), 729-742,
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P. & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.
- Chi, M.T.H. & Roscoe, R.D. (2002). The processes and challenges of conceptual change. In M. Limon & L. Mason (Eds.), *Reconsidering Conceptual Change: Issues in Theory and Practice* (S. 3-27). The Netherlands: Kluwer Academic Publishers.

- Clark, R. C. & Mayer, R. E. (2003). *e-Learning and the science of instruction*. San Francisco: Pfeiffer.
- Clifford, M. M. (1984). Thoughts on a Theory of Constructive Failure. *Educational Psychologist*, 19 (2), 1984
- Cognition and Technology Group at Vanderbilt (1993). Designing learning environments that support thinking: The Jasper series as a case study. In T.M. Duffy, J. Lowyck, D. H. Jonassen, & T. M. Welsh (Eds.), *Designing environments for constructive learning* (p. 9-36). Berlin: Springer.
- Cognition and Technology Group at Vanderbilt (1997). *The Jasper project: Lessons in curriculum, instruction, assessment, and professional development*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Cognition and Technology Group at Vanderbilt (1992). The jasper series as an example of anchored instruction: Theory, program, description, and assessment data. *Educational Psychologist*, 27, 291-315. doi:10.1207/s15326985ep2703_3
- Cohen, I. B. (1952). The education of the public in science. *Impact of Science on Society*, 3, 67–101.
- Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (S. 453-494). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and Instructional Methods a Handbook for Research on Interactions*. Irvington Pub.
- Curry, L. (2004). The effects of self-explanations of correct and incorrect solutions on algebra problem solving performance. In K. Forbus, D. Gentner & T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (S. 15-48). Mahwah, NJ: Erlbaum.
- Custers, E. J. F. M., Boshuizen, H. P. A., & Schmidt, H. G. (1999). The role of illness scripts in the development of medical expertise: Results from an interview study. *Cognition and Instruction*, 16(4), 367-398.)
- De Jong, T., & Ferguson-Hessler, M. G. (1995). *Types and qualities of knowledge. An application to physics problem solving*. IST Internal Memo, 95-05. Enschede (The Netherlands): University of Twente, Faculty of Educational Science and Technology.

- De Jong, T., & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational Psychologist, 31*, 105-113.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik, 39*, 223-238.
- Demme, Silke. (1993). Fehleranalyse und Fehlerkorrektur. Die Anwendung fehleranalytischer Erkenntnisse in der didaktischen Ausbildung von Fremdsprachenlehrern (Deutsch als Fremdsprache). *Fremdsprachen lehren und lernen, 22* (1993), 161-174.
- Deutsche Gesellschaft für Erziehungswissenschaft (2004). Kerncurriculum für das Hauptfachstudium Erziehungswissenschaft
http://www.dgfe.de/fileadmin/OrdnerRedakteure/Stellungnahmen/2004_01_KC_HF_EW.pdf [24.04.2015]
- Dochy, F., Segers, M., Van den Bossche, P. M., & Gijbels, D. (2003). Effects of problem-based learning. A meta-analysis. *Learning and Instruction, 13*, 533-568.
- Driscoll, M. (2002). Blended learning: Let's get beyond the hype. *E-Learning, 3* (3), p. 54.
- Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*(3), 287-312.
- Duffy, T. M. & Jonassen, D. H. (1991). Constructivism: New implications for instructional technology? *Educational Technology, 31*(5), 7-11.
- Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction, 22*, 206-214.
- Eiriksdottir, E. & Catrambone, R. (in press). Procedural instructions, principles, and examples: How to structure instructions for procedural tasks to enhance performance, learning, and transfer. *Human Factors*. Online verfügbar unter <http://hfs.sagepub.com/content/53/6/749.full.pdf+html> [08.10.2014]
- Ericsson, A. & Simon, H. A. (1993). *Protocol Analysis. Verbal reports as data*. Cambridge: MIT Press.
- Fenstermacher, G. D. (1994). The knower and the known: The nature of knowledge in research on teaching. In L. Darling-Hammond (Ed.), *Review of Research in Education*, Vol. 20 (S. 3-56). Washington: American Educational Research Association

- Fölling-Albers, M., Hartinger, A. & Mörtl-Hafizović, D. (2004). Situiertes Lernen in der Lehrerbildung. *Zeitschrift für Pädagogik*, 50, 727-747.
- Frankel, A., Knoblock, K. & Berwick, D.(2015). Learning from Medical Errors (Part 2) <http://www.ihl.org/education/ihopenschool/resources/Pages/Activities/PerspectiveTheMistakePart2.aspx> [24.04.2015]
- Frese, M. & Zapf, D. (1994). Action as the core of work psychology: A German approach. In H. C. Triandis, M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology*, Volume 4 (S. 271-340). Palo Alto: Consulting Psychologists Press.
- Gagné, R. (1973). *Die Bedingungen des menschlichen Lernens* (3. Aufl.) Hannover: Schrödel.
- Gordon, M. & O'Brien, T. (2007). Introduction. In: M. Gordon, T.V. O'Brien & E.R. Smith (Hrsg.), *Bridging Theory and Practice in Teacher Education. Education Faculty Book and Media Gallery*. (xi-xvi). Book 30. <http://digitalcommons.fairfield.edu/education-books/30> [24.04.2015]
- Gräsel, C. & Mandl, H. (1999). Problemorientiertes Lernen in der Methodenausbildung des Pädagogikstudiums. *Empirische Pädagogik*, 13(4), 371-391.
- Gräsel, C. (1997). *Problemorientiertes Lernen*. Göttingen: Hogrefe
- Groeben, N., Wahl, D., Schlee, J., & Scheele, B. (1988). *Das Forschungsprogramm Subjektive Theorien*. [The research program Subjective Theories] Tübingen: Francke.
- Große, C., & Renkl, A. (2004). Learning from worked examples: What happens if errors are included? In P. Gerjets, J. Elen, R. Joiner & P. Kirschner (Eds.), *Instructional design for effective and enjoyable computer supported learning* (S. 356-364). Tübingen: Knowledge Media Research Center.
- Große, C., & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction*, 17, 612-634.
- Gruber, H. & Renkl, A. (2000). Die Kluft zwischen Wissen und Handeln: Das Problem des trägen Wissens. In G. H. Neuweg (Hrsg.), *Wissen-Können-Reflexion. Ausgewählte Verhältnisbestimmungen* (S. 155-174). Innsbruck: Studienverlag.
- Gruber, H. (1999). *Erfahrung als Grundlage kompetenten Handelns*. Bern: Huber.
- Gruber, H., Mandl, H. & Renkl, A. (2000). Was lernen wir in Schule und Hochschule: Träges Wissen? In H. Mandl & J. Gerstenmaier (Hrsg.), *Die Kluft zwischen Wissen und Handeln. Empirische und theoretische Lösungsansätze* (S. 139-156). Göttingen: Hogrefe.

- Gschwendtner, T., Geißel, B., Nickolaus, R. (2007): Förderung und Entwicklung der Fehleranalysefähigkeit in der Grundstufe der elektrotechnischen Ausbildung. In: *bwp@*, Ausgabe 13 (http://www.bwpat.de/ausgabe13/gschwendtner_etal_bwpat_13.pdf) [24.04.2015]
- Guldimann, T. & Zutavern, M. (1999): „Das passiert uns nicht noch einmal!“. Schülerinnen und Schüler lernen gemeinsam den bewussten Umgang mit Fehlern. In W. Althof (Hrsg.): *Fehlerwelten. Vom Fehlermachen und Lernen aus Fehlern* (233-258). Opladen: Leske + Budrich.
- Harteis, C., Bauer, J. & Gruber, H. (2008). The culture of learning from mistakes: How employees handle mistakes in everyday work. *International Journal of Educational Research*, 47 (4), 223–231.
- Hasselhorn M. & Gold, A. (2013). *Pädagogische Psychologie. Erfolgreiches Lernen und Lehren* (3. Auflage). Stuttgart: Kohlhammer.
- Heckhausen, H., Gollwitzer, P. M. & Weinert, F. E. (Hrsg.). (1987). *Jenseits des Rubikon. Der Wille in den Humanwissenschaften*. Berlin: Springer.
- Heitzmann, N. (2014). *Fostering diagnostic competence in different domains*. Dissertation, Ludwig-Maximilians-University, Munich, Germany.
- Heitzmann, N., Fischer, F. & Fischer, M. (2013, August). Teachers' diagnostic competence: Validation of a model and evaluation of instructional support. Vortrag auf der 15th Biennial Conference of the European Association for Research in Learning and Instruction (EARLI). München, 27.08.-31.08.2013.
- Hempel, C. G. & Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science*, 15, 567-579.
- Hetmanek, A., Wecker, C., Kiesewetter, J., Trempler, K., Hake, C., Gräsel, C., Fischer, M. & Fischer, F. (2014, September). *Wozu nutzen Lehrkräfte welche Ressourcen? Eine Interviewstudie zur Schnittstelle zwischen bildungswissenschaftlicher Forschung und professionellem Handeln*. Vortrag auf dem 49. Kongress der Deutschen Gesellschaft für Psychologie (DGPs) „Die Vielfalt der Psychologie“, Bochum, 21. bis 25. September 2014.
- Hetmanek, A., Wecker, C., Kiesewetter, J., Trempler, K., Hake, C., Gräsel, C., Fischer, M.R., & Fischer, F. (2014). *Lowering the bar for evidence-based practice in teaching – what structured abstracts may contribute*. Paper presented at the EARLI Special Interest Group „Research in Teaching and Teacher Education (SIG 11)“ Conference, Frauenchiemsee, Germany, Juni 16-18, 2014.

- Hmelo, C. E. & Lin, X. (2000). Becoming self-directed learners: Strategy development in problem-based learning. In D. H. Evensen, & C. E. Hmelo (Eds.), *Problem-based learning: A research perspective on learning interactions* (S. 227-250). Lawrence Erlbaum Associates, Publishers: Mahwah, NJ.
- Hmelo, C. E. (1998). Problem-based learning. Effects on the early acquisition of cognitive skills in medicine. *Journal of the Learning Sciences*, 7, 173-208.
- Hmelo-Silver, C. E., Duncan, R. & Chinn, C. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42 (2), 99-107. doi:10.1080/00461520701263368
- Holzkamp, K.(1993). *Lernen - Subjektwissenschaftliche Grundlegung*. Campus, Frankfurt
- House, R. (2002, January 8). Clocking in column. *Spokesman-Review*.
- Huang, T. H., Liu, Y. C., & Shiu, C. Y. (2008). Construction of an online learning system for decimal numbers through the use of cognitive conflict strategy. *Computers & Education*, 50, 61-76.
- Ickler, T. (1994): Geborgter Reichtum – ehrliche Armut. Psychologische Sprache als semiotisches Problem zwischen Mentalismus und Behaviorismus. *Sprache & Kognition*, 13 (2), 103–112.
- Ingvarson, L., & Rowe, K. (2008). *Conceptualising and evaluating teacher quality: substantive and methodological issues*. Verfügbar: <http://www.freepatentsonline.com/article/Australian-Journal-Education/181463888.html> [02/08/2013].
- Ivancic, K. & Hesketh, B. (2000) Learning from errors in a driving simulation: effects on driving skill and self-confidence. *Ergonomics*, 43(12), 2000
- Jacobs, B. & Strittmatter, P. (1979). *Der schulängstliche Schüler*. München, Wien, Baltimore; Urban & Schwarzenberg.
- Jacobs, B. (2010). *Testfragen selbst beantworten oder Musterlösungen studieren?*. Verfügbar unter: <http://psydok.sulb.uni-saarland.de/volltexte/2010/2693>. [24.04.2015].
- Jones, B. F. (1992). Cognitive Designs in instruction. In M. C. Alkin (Hrsg.), *Encyclopedia of Educational Research*, Bd. 1 (S. 166-177). New York: McMillan.
- Kagan, D. M. (1992). Implications of research on teacher belief. *Educational Psychologist*, 27(1), 65-90.
- Kalyuga, S., & Renkl, A. (2010). Expertise reversal effect and its instructional implications: Introduction to the special issue. *Instructional Science*, 38, 209-215

- Kalyuga, S., Rikers, R., Pass, F. (2012). Educational implications of expertise reversal effect in learning and performance of complex cognitive and sensorimotor skills. *Educational Psychology Review*, 24, 313-337
- Kapur, M (2010). Productive failure in mathematical problem solving. *Instructional Science*, 38 (6), 523-550.
- Kapur, M. & Bielaczyc, K. (2012). Designing for productive failure. *The Journal of the Learning Sciences*, 21, 45–83.
- Kapur, M. & Rummel, N. (2009). The assistance dilemma in CSCL. In A. Dimitracopoulou, C. O'Malley, D. Suthers, & P. Reimann (Eds.), *Computer Supported Collaborative Learning Practices- CSCL2009 Community Events Proceedings*, Vol. 2 (S. 37-42). International Society of the Learning Sciences.
- Kapur, M. (2008). Productive Failure. *Cognition and Instruction*, 26 (3), 379-424.
- Kapur, M. (2014), Productive Failure in Learning Math. *Cognitive Science*. doi: 10.1111/cogs.12107
- Kießling, A. (1925). *Die Bedingungen der Fehlsamkeit*. Leipzig: Klinkhardt.
- Kirschner, P. A., Sweller, J., & Clark, R.E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational psychologist*, 41, 75-86.
- Klein, M., Park, B., Seufert, T. & Brünken, R. (2009). „Seductive Details“ in multimedialen Lehr-Lernsystemen: Effekte unterschiedlicher Präsentationsformen von Seductive Details auf den Lernerfolg und die Kognitive Belastung. 12. Fachtagung der Fachgruppe Pädagogische Psychologie, Saarbrücken.
- Klein, M., Wagner, K. & Stark, R. (2012). Theoretisieren für die Praxis - Analyse und Förderung der Anwendung wissenschaftlichen Wissens in schulischen Kontexten. In : Krämer, M., Dutke, S., & Barenberg, J. (Hrsg.). *Psychologiedidaktik und Evaluation IX*, 1.Auflage (Seitenzahlen). Aachen: Shaker.
- Klein, T. A., Neumann, J., Reuter, M., Hennig, J., von Cramon, D. Y. & Ullsperger, M. (2007). Genetically Determined differences in Learning from Errors. *Science*, 318 (5856), 1642–1645. doi:10.1126/science.1145044
- Klopp, E., Stark, R., Kopp, V. & Fischer, M. R. (2013). Psychological Factors Affecting Medical Students' Learning with Erroneous Worked Examples. *Journal of Education and Learning*, 2(1), 158-170.
- Kobi, E.E. (1994). Fehler. *Die neue Schulpraxis*, 64 (2), 5-10.

- Koedinger, K. R. & Alevan, V. (2007). Exploring the assistance dilemma in experiments with Cognitive Tutors. *Educational Psychology Review*, 19 (3), 239-264.
- Koedinger, K. R., Pavlik Jr., P. I., McLaren, B. M., & Alevan, V. (2008). Is it better to give than to receive? The assistance dilemma as a fundamental unsolved problem in the cognitive science of learning and instruction. In B.C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (S. 2155-2160). Austin, TX: Cognitive Science Society.
- Kollar, I. & Fischer, F. (2013). Orchestration is nothing without conducting – but arranging ties the two together!. *Computers & Education*, 69, 507-509.
- Köller, O. (2008). Lehr-Lern-Forschung. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch Pädagogische Psychologie* (S. 210 – 222). Göttingen: Hogrefe.
- Kolodner, J. L. (1983). Towards an understanding of the role of experience in the evolution from novice to expert. *International Journal of Man-Machine Studies*, 19(5), 497-518.
- Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004). Standards für die Lehrerbildung: Bildungswissenschaften (Beschluss der Kultusministerkonferenz vom 16.12.2004). Bonn: KMK, verfügbar unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf [24.04.2015].
- Kopp, B. & Mandl, H. (2005). Wissensschemata. (Forschungsbericht Nr. 177). LMU München: Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie. Verfügbar unter: <http://epub.ub.uni-muenchen.de/744/1/FB-177.pdf> [24.4.2015]
- Kopp, B., Ertl., B. & Mandl, H. (2006). Wissensschema und Skript. Förderung der Anwendung von Theoriewissen auf die Aufgabenbearbeitung in Videokonferenzen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 38 (3), 132-138.
- Kopp, V., Möltner, A. & Fischer, M.R. (2006): Key-Feature-Probleme zum Prüfen von prozeduralem Wissen: Ein Praxisleitfaden. [Key-feature problems to test procedural knowledge]. *GMS Z Med Ausbild*, 23 (3), Doc 50.
- Kopp, V., Stark, R. & Fischer, M. R. (2008). Fostering diagnostic knowledge through computer-supported, case-based worked examples: effects of erroneous examples and feedback. *Medical Education*, 42(8), 823-829.

- Kopp, V., Stark, R., Heitzmann, N. & Fischer, M. R. (2009). Self-regulated learning with case-based worked examples: effects of errors. *Evaluation & Research in Education*, 22 (2), 107-119.
- Kopp, V., Stark, R., Kühne-Eversmann, L. & Fischer, M. R. (2009). Do worked examples foster medical students' diagnostic knowledge of hyperthyroidism? *Medical Education*, 43 (12), 1210-1217.
- Krause, U.-M. & Stark, R. (2009). Förderung anwendbaren Wissens über kooperatives Lernen. *Journal für Lehrerinnen- und Lehrerbildung*, 9 (2), 44-48.
- Krause, U.-M. (2007). *Feedback und kooperatives Lernen*. Münster: Waxmann.
- Krause, U.-M., Stark, R. & Herzmann, P. (2011): Förderung anwendbaren Theoriewissens in der Lehrerbildung: Vergleich problembasierter und instruktionsorientierter Lernens. *Psychologie in Erziehung und Unterricht*, 58 (2), 106-115. doi:10.2378/peu2011.art27d
- Krohne, H. W. (1996). *Angst und Angstbewältigung*. Stuttgart: Kohlhammer
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96, 674-689.
- Kuhn, L., & Reiser, B. J. (March, 2005). *Students constructing and defending evidence based scientific explanations*. Paper presented at NARST 2005. Dallas, TX.
- Le Cornu, R. & Ewing, R. (2008). Reconceptualising professional experiences in pre service teacher education -reconstructing the past to embrace the future, *Teaching and Teacher Education*, 24 (7), 1799-1812.
- Lee, H. S. & Anderson, J. R. (2013). Student learning: What has instruction got to do with it? *Annual Review of Psychology*, 64, 445-469.
- Lehner, P. N (1979). *Handbook of ethological methods*. New York: Cambridge University Press.
- Leinhardt, G. (1993). On teaching. In R. Glaser (Eds.), *Advances in instructional psychology* (S. 1-54). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Leppink, J., Paas, F., Van Gog, T., Van der Vleuten, C.P.M. & Van Merriënboer, J.J.G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, 30, 32-43. doi:10.1016/j.learninstruc.2013.12.001.
- Levin, A. (2010) Effekte strukturierter Situationsanalysen auf Wissensvertiefung und Entwicklung von Reflexionsvermögen. Vortrag auf der 74. Tagung der AEPF in Jena vom 13.–15. September 2010

- Limón, M. (2006). The domain generality-specificity of epistemological beliefs: A theoretical problem, a methodological problem or both? *International Journal of Educational Research*, 45, 7–27.
- Lind, G. (2001). Von der Praxis zur Theorie. Eine Neubestimmung der Funktion der Praxis für die Lehrerbildung. [From Practice to Theory. Turning Teacher Education on its Feet]. Manuscript, University of Konstanz.
- Loke, Y.K., Price, D., Derry, S. & Aronson, K.A. (2006). Case reports of suspected adverse drug reactions—systematic literature survey of follow-up. *BMJ* 2006, 332-335
- Lowyck, J. (1991). The field of instructional design. In J. Lowyck, P. DePotter & J. Elen (Hrsg.), *Instructional Design: Implementation Issues* (S.1-30). Proceedings of the I.B.M./V.U. Leuven Conference, La Hulpe, Dec. 17-19, 1991.
- Mähler, C. & Hasselhorn, M. (2001). Transfer. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 721-730). Weinheim: Beltz PVU.
- Mahler, J. F. (1995). Erwachsene erinnern sich an negative Schulerfahrungen: Eine retrospektive Untersuchung hinsichtlich spezifischer Bedingungen, Bewertungs- und Bewältigungsversuche zur Erklärung subjektiv erlebter Schattenseiten während der Schulzeit : ein aktueller Beitrag zu: „Schatten über der Schule“ von W. Schohaus (1930). Freiburg [Schw.: s.n.]
- Mamede, S., Van Gog, T., Moura, A. S., De Faria, R. M. D., Peixoto, J. M., Rikers et al. (2012). Reflection as a strategy to foster medical students' acquisition of diagnostic competence. *Medical Education*, 46 (5), 464-472.
- Mandl, H. & Kopp, B. (2006): Blended Learning: Forschungsfragen und Perspektiven. (Forschungsbericht Nr. 182). LMU München: Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie, Internet, ISSN 1614-6336
- Mandl, H., Friedrich, H. F. & Hron, A. (1988). Theoretische Ansätze zum Wissenserwerb. In H. Mandl & H. Spada (Hrsg.), *Wissenspsychologie* (S. 123-160). München: Psychologie Verlags Union.
- Martinez-Legaz, J.E. & Soubeyran, A.(2003). *Learning from Errors*. UFAE and IAE Working Papers 557.03, Unitat de Fonaments de l'Anàlisi Econòmica (UAB) and Institut d'Anàlisi Econòmica (CSIC).
- Mathan, S. A. & Koedinger, K. R. (2005) Fostering the Intelligent Novice: Learning from errors with metacognitive tutoring. *Educational Psychologist*. 40 (4), 257-265.

- Maxwell, J. P., Masters, R. S., Kerr, E., & Weedon, E. (2001). The implicit benefit of learning without errors. *The Quarterly Journal of Experimental Psychology*, 54(4), 1049-1068.
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R.E. Mayer (Eds.), *The Cambridge Handbook of Multimedia Learning* (S. 31-48). New York: Cambridge University Press.
- McLaren, B. M., Lim, S.-J., & Koedinger, K. R. (2008). When is assistance helpful to learning? Results in combining worked examples and intelligent tutoring. In *Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (S. 677-680).
- McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, 55 (1), 51–62.
- Mehl, K. (1993). *Über einen funktionalen Aspekt von Handlungsfehlern: Was lernt man wie aus Fehlern?* Münster: LIT
- Meier, A. (2006). Theorienutzungs-kompetenz als Schlüsselqualifikation von Lehrkräften. In: Y.M. Nakamura, C. Böckelmann & D. Tröhler (Eds.). *Theorie versus Praxis?* (S. 89-106). Zürich: Verlag Pestalozzianum.
- Merenluoto, K. & Lehtinen, E. (2004). Number concept and conceptual change: towards a systemic model of the processes of change. *Learning and Instruction*, 14, 519–534
- Meurier C.E., Vincent, C.A. & Parmar, D.G. (1997). Learning from errors in nursing practice. *Journal of Advanced Nursing*. Wiley Online Library; 26:111–9
- Meuser, M. & Nagel, U. (1991): ExpertInneninterviews - vielfach erprobt, wenig bedacht. Ein Beitrag zur qualitativen Methodendiskussion. In D. Gartz & K. Kraimer (Hrsg.), *Qualitativ-empirische Sozialforschung* (S.441-471). Opladen: Westdeutscher Verlag.
- Mindnich, A., Wuttke, E. & Seifried, J. (2008). Aus Fehlern wird man klug? Eine Pilotstudie zur Typisierung von Fehlern und Fehlersituationen. In E.-M. Lankes (Hrsg.), *Pädagogische Professionalität als Gegenstand Empirischer Forschung* (S. 153-164). Münster: Waxmann.
- Minnameier, G. (2008). Zur empirischen Analyse des Umgangs mit Fehlern im wirtschaftskundlichen Unterricht. *Modernisierung der Berufsbildung - Neue Forschungserträge und Perspektiven der Berufs- und Wirtschaftspädagogik*
- Minsky, M. (1994). Negative expertise. *International Journal of Expert Systems*, 7 (1), 13–19.

- Montada, L (1968). *Über die Funktion der Mobilität in der kognitiven Entwicklung*. Stuttgart: Klett.
- Neuweg, G. H. (2005). Der Tacit Knowing View. Konturen eines Forschungsprogramms. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 101, 557–573.
- Neuweg, G. H. (2007). Wie grau ist alle Theorie, wie grün des Lebens goldener Baum? LehrerInnenbildung im Spannungsfeld von Theorie und Praxis. *bwp@*, Ausgabe 12. http://www.bwpat.de/ausgabe12/neuweg_bwpat12.pdf [18.10.2014].
- Neuweg, G.H. (2011). Distanz und Einlassung. Skeptische Anmerkungen zum Ideal einer „TheoriePraxis-Integration“ in der LehrerInnenbildung. *Erziehungswissenschaft*, 22 (43), 33-45.
- Nickel, G. (19772). *Fehlerkunde. Beiträge zur Fehleranalyse, Fehlerbewertung und Fehlertherapie*. Berlin, Cornelsen.
- Nistor, N., Schnurer, K. & Mandl, H. (2005). Akzeptanz, Lernprozess und Lernerfolg in virtuellen Seminaren. Wirkungsanalyse eines problemorientierten Seminarkonzepts. *Medienpädagogik*, 1-20. Verfügbar unter: www.medienpaed.com05-2/nistor1.pdf.
- Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, 88 (1), 1-15.
- Ohlsson, S. (1992). The cognitive skill of theory articulation: A neglected aspect of science education. *Science & Education*, 1, 181-192.
- Ohlsson, S. (1996). Learning from errors and the design of task environments. *International Journal of Educational Research*, 25(5), 419-448. Oxford: Elsevier
- Oliver, M. & Trigwell, K. (2005). Can ‚blended learning‘ be redeemed? *ELearning*, 2 (1), 17-26.
- Oser, F. & Spsychiger, M. (2005). *Lernen ist schmerzhaft. Zur Theorie des Negativen Wissens und zur Praxis der Fehlerkultur*. Weinheim: Beltz
- Oser, F. (2007). Aus Fehlern lernen. In M. Göhlich, Ch. Wulf & J. Zirfas (Hrsg.): *Pädagogische Theorien des Lernens* (S. 203-212). Weinheim und Basel: Beltz Verlag.
- Oser, F., Hascher, T. & Spsychiger, M. (1999). Lernen aus Fehlern. Zur Psychologie des ‚negativen‘ Wissens. In W. Althof (Hrsg.): *Fehlerwelten* (S. 11-41). Opladen: Leske + Budrich.
- Pajares, M. F. (1992). Teachers’ beliefs and educational research: cleaning up a messy construct. *Review of Educational Research*, 62 (3), 307-333.
- Park, B., Flowerday, T. & Brünken, R. (2015). Cognitive and affective effects of seductive details in multimedia learning. *Computers in Human Behavior*, 44, 267-278.

- Pathak, S. A., Kim, B., Jacobson, M. J., & Zhang, B. H. (2011). Learning the physics of electricity: A qualitative analysis of collaborative processes involved in productive failure. *International Journal of Computer-Supported Collaborative Learning*, 6 (1), 57–73.
- Piaget, J. (Ed.). (1968). *Six psychological studies*. New York: Random House.
- Pintrich, P. R., Marx, R. W. & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, 63, 167-199.
- Popper, K. R. (1994). *Logik der Forschung* (10. Aufl.). Tübingen: Mohr.
- Posner, G. J., Strike, K. A., Hewson, P. W. & Gertzog, W. A. (1982). Accommodation of a Scientific Conception: Toward a Theory of Conceptual Change. *Science Education*, 66 (2), 211-227.
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rolff, H.-G., Rost, J. & Schiefele, U. (Hrsg.).(2004). *PISA-Konsortium Deutschland PISA 2003 Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- Rach, S., Ufer, S., Heinze, A. (2013). Learning from errors: effects of teachers' training on students' attitudes towards and their individual use of errors. *PNA*, 8 (1), 21-30.
- Rasmussen, J. (1982). Human errors: taxonomy for describing human malfunction in industrial installations. *Journal of Occupational Accidents*, 4, 311-333.
- Rasmussen, J. (1987) The definition of human error and a taxonomy for technical system design. In J. Rasmussen, K. Duncan & J. Leplat (Eds.), *New Technology and Human Error* (S. 23-30). New York, NY: John Wiley & Sons.
- Reason, J. T. (1990). *Human error*. Cambridge, England: Cambridge University Press.
- Reigeluth, C. M. & Stein, F. S. (1983). The Elaboration Theory of Instruction. In C. M. Reigeluth (Ed.), *Instructional Design Theories and Models: An Overview of their Current States*, S. 338-381 Hillsdale, NJ: Lawrence Erlbaum.
- Reigeluth, C. M. (1979). In search of a better way to organize instruction: The elaboration theory. *Journal of Instructional Development*, 2 (3), 8-15.
- Reinmann, G. & Mandl, H. (2006). Unterrichten und Lernumgebungen gestalten. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 613-658). Weinheim: Beltz.

- Renkl, A. & Atkinson, R. K. (2002). Learning from examples: Fostering self-explanations in computer-based learning environments. *Interactive Learning Environments*, 10 (2), 105-119. doi:10.1076/ilee.10.2.105.7441
- Renkl, A. & Atkinson, R. K. (2010). Learning from worked-out examples and problem solving. In J. L. Plass, R. Moreno & R. Brünken (Eds.), *Cognitive load theory* (S. 91-108). Cambridge: Cambridge University Press
- Renkl, A. (1996). Träges Wissen: Wenn Erlerntes nicht genutzt wird. *Psychologische Rundschau*, 47, 78-92.
- Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science: A Multidisciplinary Journal*, 21 (1), 1-29.
- Renkl, A. (2001). Lernen aus Lösungsbeispielen: Eine Einführung. *Unterrichtswissenschaft*, 29 (1), 2-4.
- Renkl, A. (2005). The worked-out examples principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (S. 229-245). Cambridge: Cambridge University Press.
- Renkl, A. (2006). Träges Wissen. In: Detlev H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 778-781). Weinheim: Beltz.
- Renkl, A. (September 1994): Träges Wissen: Die „unerklärliche“ Kluft zwischen Wissen und Handeln. (Forschungsbericht Nr. 41). LMU München: Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie, Internet, ISSN 1614-6336
- Resnick L.B. (1987). Learning in School and Out. *Educational Researcher*, 16 (9), 13-20. doi:10.3102/0013189X016009013
- Ritchie, J.R. and Wilson, D.E. (2001). *Teacher Narrative as Critical Inquiry: Rewriting the Script*. New York: Teachers College Press.
- Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93, 346-362. doi: 10.1037//0022-0663.93.2.346
- Rohrbach, D. (1999). *Lernen aus Fehlern im Laufe des Lebens. Erkenntnisse aus einer empirischen Untersuchung*. Freiburg/Schweiz: Universität Freiburg, Departement Erziehungswissenschaften.
- Rollet, B. (1999) Auf dem Weg zu einer Fehlerkultur. Anmerkungen zur Fehlertheorie von Fritz Oser. In W. Althof (Hrsg.), *Fehlerwelten* (S. 71-87). Opladen: Leske + Budrich.

- Rook, M., Frey, D. & Irle, M. (2001). Wissenschaftstheoretische Grundlagen sozialpsychologischer Theorien. In D. Frey & M. Irle (Hrsg.), *Kognitive Theorien der Sozialpsychologie*. Bern: Huber.
- Rossett, A. (2002). *The ASTD e-learning handbook*. New York: McGraw-Hill.
- Rost, D.H. & Schermer, F.J. (2006): Leistungsängstlichkeit. In D.H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz PsychologieVerlagsUnion.
- Rybowiak, V., Garst, H., Frese, M., & Batinic, B. (1999). Error orientation questionnaire (EOQ): reliability, validity, and different language equivalence. *Journal of Organizational Behavior*, 20 (4), 527–547.
- Salden, R. J. C. M., Aleven, V., Schwonke, R. & Renkl, A. (2010). The expertise reversal effect and worked examples in tutored problem solving. *Instructional Science*, 38 (3), 289-307.
- Schank, R. C. & Abelson, R. P. (1977). *Scripts, Plans, Goals and Understanding. An Inquiry into Human Knowledge Structures*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Schank, R. C. (1986). *Explanation Patterns: Understanding Mechanically and Creatively*. Erlbaum, Hillsdale, NJ.
- Schmassmann, Margret (1992). Der Zugang von körperbehinderten Kindern zur Mathematik. In: *SVFK Bulletin*, 92, S. 12-54.
- Schmidt, H. G., Rotgans, J. I., Yew, E. (2011). The Process of Problem-Based Learning: What Works And Why. *Medical Education*, 45 (8), 792-806.
- Schmidt, H., Loyens, S., Van Gog, T. & Paas, F. (2007). Problem-based learning is compatible with human cognitive architecture. Commentary on Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42 (2), 91-97.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, 3, 207-217. Reprinted, 1993, *Effective School Practices*, 12, 36-48; Reprinted, D.A. Robin, K. Yorkston, and D.R. Beukelman (Eds.). (1995). *Disorders of motor speech: Recent advances in assessment, treatment, and clinical characterization*. Baltimore: Brookes Publishing Co.
- Schneewind, K.A. (1977): Zum Verhältnis von Psychologie und Wissenschaftstheorie. In K.A. Schneewind (Hrsg.), *Wissenschaftstheoretische Grundlagen der Psychologie* (S.11-25). München: Reinhardt.

- Schnotz, W., Vosniadou, S. & Carretero, M. (Eds.). (1999). *New perspectives on conceptual change*. Oxford: Pergamon Press.
- Schnurer, K. (2005). *Kooperatives Lernen in virtuell-asynchronen Hochschulseminaren. Eine Prozess-Produkt-Analyse des virtuellen Seminars „Einführung in das Wissensmanagement“ auf der Basis von Felddaten*. Berlin: Logos.
- Schohaus, W. (1933). *Schatten über der Schule. Eine kritische Betrachtung. Mit einer Betrachtung von 82 Schulbekenntnissen*. Zürich: Schweizer Spiegel Verlag.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82, 498-504
- Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for learning: The hidden efficiency of original student production in statistics instruction. *Cognition and Instruction*, 22, 129–184.
- Seidel, T. & Prenzel, M. (2008). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen - Erfassung pädagogisch-psychologischer Kompetenzen mit Videosequenzen. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 201-216). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-90865-6_12
- Seufert, T., Schütze, M., & Brünken, R. (2009). Memory characteristics and modality in multimedia learning: An aptitude–treatment–interaction study. *Learning and Instruction*, 19 (1), 28–42. doi:10.1016/j.learninstruc.2008.01.002
- Sherin, M.G. (2007). The development of teachers' professional vision in video clubs. In R. Goldman, R. Pea, B. Barron & S. Derry (Eds.), *Video research in the learning sciences* (S. 383-395). Hillsdale, NJ: Lawrence Erlbaum.
- Sherin, M.G., & van Es, E.A. (2009). Effects of Video Club Participation on Teachers' Professional Vision. *Journal of Teacher Education*, 60, 20-37.
- Shuell, T. J. (1993). Toward an integrated theory of teaching and learning. *Educational Psychologist*, 28, 291-311.
- Sinatra, G. M., & Mason, L. (2008). Beyond knowledge: Learner characteristics influencing conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (560-582). New York: Routledge.
- Skinner, B. F. (1971). *Beyond Freedom and Dignity*. New York: Knopf.
- Skinner, B.F. (1958). Teaching machines. *Science*, 128 (1958), 969–977.

- Smith, J. P., diSessa, A. A. & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, 3, 115-163.
- Smith, K. (2005). Teacher Educators' professional knowledge- How does it differ from teachers' professional knowledge? *Teaching and Teacher Education*, 21, 177-192.
- Snow, R. E. (1989). Toward assessment of cognitive and conative structures in learning. *Educational Researcher*, 18 (9), 8-14.
- Spiro, R. & Jehng, J. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. Spiro (Hrsg.), *Cognition, education and multimedia: Exploring ideas in high technology* (S. 163-205). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Spiro, R. J., Collins, B. & Thota, J. (2003). Cognitive flexibility theory: Hypermedia for complex learning, adaptive knowledge application, and experience acceleration. *Educational Technology*, 43, 5-10.
- Spychiger, M. (2004). Bitter, edel oder leicht? Ausführungen zum biografischen Lernen aus Fehlern. Schriftliche Version des Beitrags zum Kongress für Bildungsforschung der DGfE, SGBF und ÖFEB. http://www.paed-kongress04.unizh.ch/downloads/publikationen/AG43_Spychiger.pdf
- Stahl, E. & Bromme, R. (2007). The CAEB: An instrument for measuring connotative aspects of epistemological beliefs. *Learning and Instruction*, 17 (6), 773-785.
- Star, J. R. & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice teachers' ability to notice. *Journal of Mathematics Teacher Education*, 11, 107-125. doi:10.1007/s10857-007-9063-7
- Stark, R. & Fischer, M. (2008). Förderung diagnostischer Kompetenz durch beispielbasiertes Lernen: Einfluss von Prozessorientierung und Fehler. unveröffentlichter Arbeitsbericht über die Projektarbeit im Förderzeitraum 1 (FI 720/2-1 und STA 569/3-1).
- Stark, R. & Koch, B. (2007). WALe – A problem-oriented computer-based learning environment to foster quality of interpretation and argumentation. Proceedings of the *Sixth IASTED International Conference on Web-based Education WBE 2007*. V. Uskov (Eds.)Chamonix, France: ACTA Press.
- Stark, R. & Mandl, H. (2005). Lernen mit einer netzbasierten Lernumgebung im Bereich empirischer Forschungsmethoden. Effekte zusätzlich implementierter Maßnahmen und Bedeutung von Lernvoraussetzungen. *Unterrichtswissenschaft*, 33(1), 3-29.

- Stark, R. & Mandl, H. (2007). Bridging the gap between basic and applied research by an integrative research approach. *Educational Research and Evaluation*, 13(3), 249-261.
- Stark, R. (1999). *Lernen mit Lösungsbeispielen*. Göttingen: Hogrefe-Verlag.
- Stark, R. (2001). *Analyse und Förderung beispielbasierten Lernens: Anwendung eines integrativen Forschungsparadigmas*. Unveröffentlichte Habilitationsschrift, Ludwig-Maximilians-Universität München.
- Stark, R. (2003). Conceptual Change: kognitiv oder situiert? *Zeitschrift für Pädagogische Psychologie*, 17 (2), 133-144.
- Stark, R. (2004). *Lehren und Lernen mit ausgearbeiteten Lösungsbeispielen – zentrale Befunde, Optimierungsmaßnahmen und Anwendung im Kontext der Methodenlehre*. Heidelberg: Universität Heidelberg.
- Stark, R. (2005). Constructing arguments in educational discourses. In H. Gruber, C. Harteis, R. Mulder & M. Rehl (Eds.), *Bridging individual, organisational, and cultural aspects of professional learning* (S. 64-71). Regensburg: S. Roderer.
- Stark, R., Herzmann, P. & Krause, U.-M. (2010). Effekte integrierter Lernumgebungen -- Vergleich problembasierter und instruktionsorientierter Seminarkonzeptionen in der Lehrerbildung. *Zeitschrift für Pädagogik*, 56 (4), 548-563.
- Stark, R., Kopp, V. & Fischer, M. R. (2009). Förderung der Diagnosekompetenz bei Studierenden der Medizin durch fallbasierte Lösungsbeispiele: der Einfluss von Fehlern und Feedback. *Psychologie in Erziehung und Unterricht*, 56 (2), 137-149.
- Stark, R., Kopp, V., & Fischer, M. R. (2011). Case-based learning with worked examples in complex domains: Two experimental studies in undergraduate medical education. *Learning and Instruction*, 21, 22-33.
- Stark, R., Mandl, H., Gruber, H. & Renkl, A. (1999). Instructional means to overcome transfer problems in the domain of economics: Empirical studies. *International Journal of Educational Research*, 31(7), 591-609.
- Stark, R., Mandl, H., Gruber, H. & Renkl, A. (2002). Conditions and effects of example elaboration. *Learning & Instruction*, 12 (1), 39-60. doi:10.1016/S0959-4752(01)00015-9
- Stark, R., Puhl, T. & Krause, U.-M. (2009). Improving scientific argumentation skills by a problem-based learning environment: effects of an elaboration tool and relevance of student characteristics. *Evaluation & Research in Education*, 22 (1), S.51-68.

- Stark, R., Tyroller, M., Krause, U.-M. & Mandl, H. (2008). Effekte einer metakognitiven Promptingmaßnahme beim situierten, beispielbasierten Lernen im Bereich Korrelationsrechnung. *Zeitschrift für Pädagogische Psychologie*, 22 (1), 59-71.
- Stefanko, G., Lancashire, B., Coombes, J.S. & Fassett, R.G. (2009). Pulmonary oedema and hyponatraemia after an ironman triathlon. *BMJ Case Reports*, online journal. doi:10.1136/bcr.04.2009.1764
- Steffe, L. & Gale (1995). *Constructivism in education*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Steffens, D. & Reiß, M. (2009): Blended Learning in der Hochschullehre. Vom Nebeneinander der Präsenzlehre und des E-Learning zum integrierten Blended Learning-Konzept. *HSW – das Hochschulwesen*, 57 (4), 115-123.
- Strobel, J. & van Barnefeld, A. (2009). When is PBL More Effective? A Meta-synthesis of Meta-analyses Comparing PBL to Conventional Classrooms. *Interdisciplinary Journal of Problem-Based Learning*, 3(1), 44-58. DOI: <http://dx.doi.org/10.7771/1541-5015.1046>.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E Mayer (Ed.), *The Cambridge handbook of multimedia learning*, (p. 19-30). Cambridge: University Press.
- Sweller, J., Ayres, P., & Kalyuga S. (2011). *Cognitive load theory*. New York: Springer.
- Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10 (3), 251-296.
- Swinkels, M.F.J., Koopman, M. & Beijaard, D. (2013). Student teachers' development of learning-focused conceptions. *Teaching and Teacher Education*, 34, 26-37.
- Terhart, E. (2006). Was wissen wir über gute Lehrer? *Pädagogik*, 58.Jg. (5), 42-47.
- Terhart, E., Czerwenka, K., Ehrich, K., Jordan, F. & Schmidt, H. J. (1994). *Berufsbiografien von Lehrern und Lehrerinnen*. Frankfurt/Main: Peter Lang.
- Terhart, E., Lohmann, V. & Seidel, V. (2010). *Die bildungswissenschaftlichen Studien in der universitären Lehrerbildung. Eine Analyse aktueller Studienordnungen und Modellhandbücher an Universitäten in Nordrhein-Westfalen. Abschlussbericht an das Ministerium für Schule und Weiterbildung NRW*. Münster: Westfälische Wilhelms-Universität Münster.
- Trempler, K., Hasselkuß, M., Heckersbruch, C. M., Gräsel, C., Baedeker, C. & Schneidewind, U. (2014). Implementation von Bildungsinnovationen in Netzwerken - Analyse

- von Schul-Unternehmens-Kooperationen [Supplement]. *Zeitschrift Für Erziehungswissenschaft*, 17(5), 79-95.
- Trempler, K., Hetmanek, A., Wecker, C., Kiesewetter, J., Wermelt, M., Fischer, M., Fischer, F. & Gräsel, C. (2015, März). Nutzung von Evidenz im Bildungsbereich – Validierung eines Instruments zur Erfassung von Kompetenzen der Informationsauswahl und Bewertung von Studien. Vortrag auf der 3. Tagung der Gesellschaft für empirische Bildungsforschung (GEBF). Bochum
- Tversky, A. & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases, *Science*, 185, 1124-1131.
- Ummel, H. (2010). Fallverstehen als Gegenstück zum Training von Routinen. *Journal für Lehrerinnen- und Lehrerbildung*, 10 (2), 42--47.
- Universität des Saarlandes (2007). *Studienordnung der Universität des Saarlandes für die Studiengänge Lehramt an beruflichen Schulen (LAB), Lehramt an Gymnasien und Gesamtschulen (Klassenstufen 5-13) (LAG), Lehramt an Hauptschulen und Gesamtschulen (LAH) und Lehramt an Realschulen und Gesamtschulen(LAR) vom 26. April 2007. Dienstblatt Nr. 30. http://www.uni-saarland.de/fileadmin/user_upload/Campus/Service/Recht_und_Datenschutz/Recht_der_Universitaet/Ausbildungs-Pruefungs-Studienordnungen/Lehramtsstudiengaenge_modularisiert_DB08-541.pdf* [20.3.2014]
- Van den Broek, P. & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: the role of co-activation in confronting misconceptions. *Applied Cognitive Psychology*, 22, 335-351. doi:10.1002/acp.1418
- Van Es, E.A., & Sherin, M.G. (2002): Learning to notice: scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, 10(4), 571–596.
- van Gog, T., Paas, F., & van Merriënboer, J. (2004). Recommendations for research on task formats that model expert approaches to problem solving (Paper presentation): Educational Technology Expertise Center, Open University of the Netherlands
- VanLehn, K. (1999). Rule-learning events in the acquisition of a complex skill: An evaluation of cascade. *Journal of the Learning Sciences*, 8 (1), 71-125.
- VanLehn, K., Siler, S., Murray, R. C., Yamauchi, T., & Baggett, W.B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, 21, 209-249.

- Vosniadou, S., Ioannides, C., Dimitrakopoulou, A. & Papademetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction, 11*, 381-419.
- Wagner, K., Klein, M., Klopp, E. & Stark, R. (2014). Theoretisieren für die Praxis: Förderung anwendbaren pädagogischen Wissens anhand advokatorischer Fehler. *Journal für Lehrerinnen- und Lehrerbildung*, im Druck.
- Wagner, K., Klein, M., Klopp, E. & Stark, R. (2015, angenommen). Förderung anwendbaren bildungswissenschaftlichen Wissens anhand kollaborativem Lernen aus Fehlern. *Unterrichtswissenschaft, 61*, 287 - 301.
- Wagner, K., Klein, M., Klopp, E., Puhl, T. & Stark, R. (2013). Effects of Instruction - Supported Learning with Worked Examples in Quantitative Method Training. *Higher Education Studies, 3* (3), 1-12. doi:10.5539/hes.v3n3p1
- Wagner, K., Stark, R., Daudbasic, J., Klein, M., Krause, U.-M. & Herzmann, P. (2013a). Effektivität integrierter Lernumgebungen in der universitären Lehrerbildung – eine quasiexperimentelle Feldstudie. *Journal für Bildungsforschung online, 5* (1). 115-140.
- Wagner, K., Stark, R., Krause, U.-M., Gutenberg, N. & Klein, M. (2013b). Evaluation eines Trainingsprogramms zur Förderung der sprachlichen Kompetenzen von Schülerinnen und Schülern im letzten Hauptschuljahr. *Zeitschrift für Angewandte Linguistik, 147-169*. doi: 10.1515/zfal-2013-0015.
- Walker, A., & Leary, H. (2009). A Problem Based Learning Meta Analysis: Differences Across Problem Types, Implementation Types, Disciplines, and Assessment Levels. *Interdisciplinary Journal of Problem-Based Learning, 3* (1).
- Weber, S., & Achtenhagen, F. (2009). Forschungs- und evidenzbasierte Lehrerbildung. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus, R. Mulder (Hrsg.), *Lehrprofessionalität - Bedingungen, Genese, Wirkungen und ihre Messung*. (S. 477-487). Weinheim: Beltz (ISBN 978-3-407-32103-9).
- Wecker, C. & Fischer, F. (2011). From guided to self-regulated performance of domain-general skills: The role of peer monitoring during the fading of instructional scripts. *Learning and Instruction, 21* (6), 746-756.
- Wecker, C. (2012). *Vom Sollen zum Können. Fading instruktionaler Skripts zur Förderung von Argumentationskompetenz*. Münster: Waxmann.

- Weibell, C. J. (2011). Principles of learning: 7 principles to guide personalized, student-centered learning in the technology-enhanced, blended learning environment. Retrieved July 4, 2011 from [<https://principlesoflearning.wordpress.com>].
- Weimer, H. (1925). *Psychologie der Fehler*. Leipzig: Klinkhardt.
- Weimer, H. (1930). Fehlerbekämpfung. In H. Nohl & L. Pallat (Hrsg.): *Handbuch der Pädagogik, 3. Band. Allgemeine Didaktik und Erziehungslehre* (S. 119-128). Berlin: Beltz
- Weimer, H. (1926). Der Kampf gegen die objektiven Bedingungen der Fehlsamkeit. *Zeitschrift für pädagogische Psychologie, Experimentelle Pädagogik und jugendliche Forschung*, 27, 220-230.
- Weinert, F. E. (1996). Lerntheorien und Instruktionsmodelle. In F. E. Weinert (Hrsg.), *Psychologie des Lernens und der Instruktion. Pädagogische Psychologie* (S. 1-48). Göttingen: Hogrefe.
- Weinert, F. E. (2001): Vergleichende Leistungsmessung in Schulen – Eine umstrittene Selbstverständlichkeit. In F. E. Weinert, (Hrsg.), *Leistungsmessungen in Schulen*. Weinheim: Beltz.
- Weingardt, M. (2004): *Fehler zeichnen uns aus. Transdisziplinäre Grundlagen zur Theorie und Produktivität des Fehlers in Schule und Arbeitswelt*. Bad Heilbrunn. Klinkhardt.
- Weinstock, M., Neuman, Y. & Tabak, I.(2004)..Missing the point or missing the norms? Epistemological norms as predictors of students' ability to identify fallacious arguments. *Contemporary Educational Psychology*, 29 (1),77-94.
- Wenglein, S. (2015). *E4teach – Evidence for teachers. Training der Kompetenz angehender Lehrkräfte zum Umgang mit Evidenz*. Vortrag auf der 3. Tagung der Gesellschaft für empirische Bildungsforschung (GEBF). Bochum
- Westermann, R. (2000). *Wissenschaftstheorie und Experimentalmethodik: Ein Lehrbuch zur Psychologischen Methodenlehre*. Göttingen: Hogrefe.
- Whitehead, A. N. (1929). *The aims of education*. New York: Macmillan.
- Wieland, G. (Hrsg.) (1991). *Mathematik Forum. Fehler! - Fehler?*. Bericht zum XIII Schweizerischen Mathematikforum in Weinfelden, 19. - 21.11.1990. Bern: EDK
- Wills, A.J., Lavric A., Croft G. & Hodgson T.L. (2007). Predictive learning, prediction errors and attention: Evidence from event-related potentials and eye-tracking. *Journal of Cognitive Neuroscience*, 19, 843–854

- Winn, W. D. (1993). An account for how people search for information in diagrams. *Contemporary Educational Psychology*, 18, 162-185.
- Wollman, W., Eylon, B. & Lawson, A.E. (1980). An analysis of premature closure in science and developmental stages. *Journal of Research in Science Teaching*, 17 (2), 105-114.
- Wollring, Bernd (2009): Zur Kennzeichnung von Lernumgebungen für den Mathematikunterricht in der Grundschule. In A. Peter-Koop, G. Lilitakis & B. Spindeler (Hrsg.): *Lernumgebungen – Ein Weg zum kompetenzorientierten Mathematik in der Grundschule* (S. 9-23). Offenburg: Mildenerger Verlag.
- Woolfolk, A. (2008). *Pädagogische Psychologie*. München: Pearson.
- World Health Organisation (2008). *World Alliance for Patient Safety. Forward Programme 2008-2009*.
http://www.who.int/patientsafety/information_centre/reports/Alliance_Forward_Programme_2008.pdf
- Yen-Ting Lin, Ming-Lee Wen, Min Jou, & Din-Wu Wu. 2014. A cloud-based learning environment for developing student reflection abilities. *Computers in Human Behaviour*, 32 (March 2014), 244-252. DOI=10.1016/j.chb.2013.12.014
<http://dx.doi.org/10.1016/j.chb.2013.12.014>
- Zink, S. (2002). *Warum Menschen aus Fehlern nicht lernen. Aufbau von negativem Wissen*. Seminararbeit, Universität Freiburg, Departement Erziehungswissenschaften.
