

Die Förderung kognitiver Prozesse beim Schreiben mit Prompts

I N A U G U R A L D I S S E R T A T I O N

zur Erlangung des Grades eines Doktors der Philosophie
im Fachbereich Erziehungswissenschaften
der Johann Wolfgang Goethe-Universität
zu Frankfurt am Main

vorgelegt von

Dipl.-Päd. Christin Picard

2 0 1 1

GUTACHTER: PROF. DR. UDO RAUIN
PROF. DR. MARGARETE IMHOF

Inhaltsverzeichnis

1	Einleitung	1
2	Theoretischer Hintergrund	5
2.1	Definition des Schreibbegriffes	5
2.2	Theoretische Perspektiven auf Schreiben als Lernmedium	6
2.2.1	Zusammenfassung und Ausblick	29
2.3	Studien mit Lerntagebüchern und Lernprotokollen	31
2.3.1	Prompts zum gezielten Aktivieren von Lernstrategien beim Schreiben	37
2.3.2	Forschungsbedarf	42
3	Die Konzeption der Lernumgebung	50
3.1	Die Gestaltung der Prompts	50
3.1.1	Vorwissen und cognitive load	51
3.1.2	Strategien, die sich als effektiv erwiesen haben	53
3.1.3	Darbietungszeitpunkt	55
3.1.4	Zusammenfassung und Ableitung der Prompts	56
3.2	Konzeptueller und methodischer Hintergrund des Feedbacks	56
3.2.1	Funktionen des Feedbacks	58
3.2.2	Feedbackinhalte und Konstruktion des Feedbacks	62
3.3	Das Lösungsbeispiel	64
4	Fragestellungen	66
4.1	Studie 1: Evaluation der Prompts und des Analyse-Instrumentes	67
4.1.1	Fragestellungen	67
4.2	Studie 2: Effekte verschiedener Feedbackformen auf die Verarbeitungstiefe beim Schreiben und auf die Motivation	68
4.2.1	Fragestellungen	69

5	Eigene empirische Untersuchungen	72
5.1	Beschreibung und Analyse der Lernergebnisse	73
5.2	Studie 1: Evaluation der Prompts und des Textanalyseinstrumentes	84
5.2.1	Methode	84
5.2.2	Ergebnisse	87
5.2.3	Zusammenfassung und Konsequenzen	94
5.3	Studie 2: Effekte verschiedener Feedbackformen auf die Verarbeitungstiefe beim Schreiben und auf die Motivation	95
5.3.1	Methode	96
5.3.2	Ergebnisse	105
6	Diskussion	120
6.1	Zur Kategorisierung der Texte mit der SOLO-Taxonomie	122
6.2	Können Prompts die Verarbeitungstiefe beim Schreiben beeinflussen?	125
6.3	Feedback und Motivation	126
6.4	Feedback und Verarbeitungstiefe	129
6.5	Implikationen für die instruktionale Praxis	131
6.6	Grenzen der Untersuchung und Perspektiven für zukünftige Forschung	133
	Anhang	137
.1	Prompts der zweiten Studie	138
.2	Selbstfeedbackfragen	140
.3	Fragen aus eHElp	143
.4	Kodierleitfaden	145
.5	Musterbeispiel	147
.6	Vorwissenstest	149
.7	FAM (Fragebogen zur aktuellen Motivation)	154
.8	Austausch mit John Biggs bezüglich des Skalenniveaus der SOLO-Taxonomie	156
	Literatur	158

Abbildungsverzeichnis

1	Verteilung der Lernjournaltexte über die Taxonomiestufen	89
2	Prozentuale Verteilung der Texte in der Kategorie „Alltagsbezug“	91
3	Prozentuale Verteilung der Texte in der Kategorie „Beruflicher Bezug“	91
4	Prozentuale Verteilung der Texte in der Kategorie „Bewertung“ .	92
5	Prozentuale Verteilung der Texte in der Kategorie „Offene Fragen“	92
6	Prozentuale Verteilung der Texte in der Kategorie „Beispiele“ . .	93
7	Prozentuale Verteilung der Texte in der Kategorie „Kontextübergreifende Bezüge“	93
8	Verteilungen der Texte auf der SOLO-Taxonomie in der ersten Studie (links) und in der zweiten Studie (rechts)	109
9	Verlauf der Motivationskomponente <i>Herausforderung</i> über die drei Messzeitpunkte	114
10	Verlauf der Motivationskomponente <i>Erfolgserwartung</i> über die drei Messzeitpunkte	114
11	Verlauf der Motivationskomponente <i>Misserfolgsbefürchtung</i> über die drei Messzeitpunkte	114
12	Verlauf der Motivationskomponente <i>Interesse</i> über die drei Messzeitpunkte	115
13	Vorgehensweise bei der Untersuchung der Feedbackeffekte	130

Tabellenverzeichnis

1	Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)	76
1	Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)	77
1	Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)	78
1	Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)	79
1	Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)	80
1	Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)	81
2	Merkmalsausprägungen der Prompts	86
2	Merkmalsausprägungen der Prompts	87
3	Chronologie der Untersuchung in der zweiten Studie	97
4	Friedman-Messwiederholungsplan	98
5	Wilcoxon Messwiederholungsplan	100
6	Die Prompts in der zweiten Studie	102
6	Die Prompts in der zweiten Studie	103
7	Feedbackreihenfolge in den Gruppen	104
8	Reliabilitäten (Cronbachs α) des FAM	106
9	Itemschwierigkeiten P und Standardabweichungen der Items aus dem Vorwissenstest (n=19) und dem Nachtest	107
10	Mittelwerte und Standardabweichung (in Klammern) der Gruppen A, B und C zu drei Messzeitpunkten mit dem FAM	111
11	Mittelwerte und Standardabweichungen (in Klammern) der Ein- stufungen der Texte mit der SOLO-Taxonomie (P= Prompt) . . .	116

12	Vorgehen bei der Untersuchung der Feedbackeffekte auf die Textqualität	116
13	Korrelationen zwischen den Subskalen des FAM zu Z1 und der Verarbeitungstiefe	118
14	Feedbackreihenfolge in den Gruppen	121
15	Taxonomie-Niveaus und Indikatoren	145
15	Taxonomie-Niveaus und Indikatoren	146

Danksagung

Ich danke Herrn Professor Udo Rauin für die Zustimmung zu meinem Promotionsthema. Er ermöglichte mir, diese Arbeit am Fachbereich Erziehungswissenschaften durchzuführen und unterstützte mich vor allem in methodischen Fragen durch wertvolle Anregungen und Ratschläge. Des Weiteren danke ich den wissenschaftlichen Hilfskräften Lena Schares, Konstanze Jenderek und Michael Weigand, die mir unglaublich gewissenhaft und zuverlässig bei der Datenanalyse und Datenauswertung geholfen haben. Judith Junk gilt ebenfalls großer Dank, da sie die Arbeit trotz eigener Auslastung komplett korrektur gelesen hat. Sie stellte mir überdies uneingeschränkt ihre Wohnung zum Schreiben zur Verfügung, wenn ich einen Umgebungswechsel brauchte. Dr. Susanne Frühauf erklärte sich bereit, die erste Untersuchung in zwei ihrer Seminare durchzuführen. Dafür danke ich ihr herzlich. Ich danke außerdem meiner Mutter, die mich in der Endphase der Dissertation sehr ermutigt und gestärkt hat und mit Vitaminpräparaten und einem einzigartigen Cateringservice für meine Durchhaltekraft gesorgt hat. Mein ganz besonderer Dank gilt Frau Professor Margarete Imhof, die den Charakter der Arbeit mit geprägt hat und mit der ich mich jederzeit darüber austauschen durfte. Sie hat mich in allen Phasen unterstützt und dadurch einen großen Beitrag zum Gelingen der Arbeit geleistet.

Zusammenfassung

Im Mittelpunkt des Dissertationsvorhabens steht das Schreiben als Lerninstrument im Rahmen von Lernjournalen. Ausgehend von der Annahme, dass durch Schreiben höherrangige kognitive Prozesse angeregt werden können, ist das Ziel der vorliegenden Arbeit die Untersuchung der Effekte von Schreibinstruktionen (Prompts) auf die Qualität kognitiver Prozesse bei der Textproduktion. Im Rahmen zweier Untersuchungen werden dazu Studierende angeleitet, mit den Aufgaben Texte zu verschiedenen Themengebieten zu verfassen. Die Texte werden mithilfe einer Taxonomie, mit der die Komplexität von Lernergebnissen abgebildet wird, analysiert und ausgewertet. Im Zentrum der ersten Untersuchung steht die Frage, inwiefern sich verschiedene Anforderungsspezifika der Prompts auf die Komplexität der produzierten Texte auswirken. Die zweite Untersuchung geht der Frage nach, wie Schreibprozesse durch verschiedene Formen von Feedback beeinflusst werden und befasst sich mit den Wechselwirkungen des Feedbacks, der Motivation der Lerner und der aus den Texten erkennbaren Verarbeitungstiefe. Die Ergebnisse zeigen zum einen, dass Lerner mit Prompts bereits zu Beginn eines Lernzyklus dazu angeregt werden können, Elaborationsstrategien einzusetzen und komplexe Wissensstrukturen beim Schreiben aufzubauen. Zum anderen deuten sich Effekte des Feedbacks auf die Motivation der Lerner an, da sich in Abhängigkeit der Feedbackform unterschiedliche Motivationsverläufe nachweisen ließen. Die Befunde weisen also darauf hin, dass Feedback den Lernprozess auf der motivationalen Ebene beeinflusst.

Insgesamt werden in der Arbeit Faktoren untersucht, die bei der Gestaltung einer Lernumgebung mit Lernjournalen zu beachten sind. Die Befunde geben Hinweise darauf, wie Lernjournale in eine Seminarstruktur eingebettet werden können, um die Komplexität der Schreibprozesse erwachsener Lerner zu erhöhen.

1 Einleitung

In der universitären Lehrerausbildung werden seit den achtziger Jahren verschiedene Formen des Schreibens (z.B. Portfolios und Lernjournale) eingesetzt, um Lernprozesse zu dokumentieren und zu unterstützen. Innerhalb dieser didaktischen Maßnahmen nimmt das Schreiben einen zentralen Stellenwert ein, weil Lehrpersonen von einer epistemischen Funktion des Textproduzierens ausgehen. Lerner, so wird angenommen, verstehen Lerninhalte besser und können diese auf neue Kontexte anwenden, nachdem sie schriftlich darüber reflektiert haben.

Betrachtet man die aktuellen kognitiven Modelle des Schreibprozesses (vgl. Bereiter & Scardamalia, 1987; Hayes, Flower, Schriver, Stratman & Carey, 1987), werden starke Ähnlichkeiten zu Lernprozessmodellen (vgl. Boekaerts, 1997; Zimmerman, 1989) deutlich: Beide werden als rekursive Prozesse konzipiert, die von informationsverarbeitenden Strategien (kognitiven Strategien) und Strategien zur Überwachung (metakognitiven Strategien) des Prozesses geprägt sind. Lernstrategien stehen also auch im Zentrum des Schreibens und kennzeichnen es als Lerninstrument.

Ähnlich wie Lernprozesse ziehen jedoch auch Schreibprozesse nicht per se akademischen Erfolg nach sich oder führen zu einer tieferen Verarbeitung der Lerninhalte, wie die Analyse von Lerntagebüchern zeigt (Borsch & Imhof, 2007; Nückles & Fries, 2004). Diese Studien deuten darauf hin, dass Lerner eher oberflächliche Texte verfassen, die keine intensive Auseinandersetzung mit dem Lernstoff widerspiegeln. Um dem entgegenzuwirken, werden Schreibinstruktionen (sogenannte Prompts) eingesetzt, die anspruchsvolle Strategien anregen sollen. Die bisherigen Untersuchungen mit Prompts legen nahe, dass Lerner von dieser expliziten Anleitung der Lernstrategien in der Schreibinstruktion im Rahmen von Portfolio und Lerntagebüchern profitieren.

Welche Effekte haben Prompts jedoch auf die Wissensstruktur der Textproduzenten und inwiefern beeinflussen sie den Lernprozess? Gelingt es Lernern durch das Schreiben mit Prompts, komplexe Wissensstrukturen aufzubauen? Welche Art des Lernens wird durch Prompts gefördert?

Bislang wurde der Erfolg von Schreib-Interventionen mit Prompts an den Effekten auf den Wissenszuwachs festgemacht, der zumeist mittels Multiple-Choice-Test erfasst wird. Das Ziel von Schreibinstruktionen ist es jedoch, Lerner zu einer Verarbeitung anzuregen, die ein tiefgreifendes Verstehen der Lerninhalte nach sich zieht. Ob dies mit Prompts gelingt, ist mit der Methodik von Multiple-Choice-Tests nicht zu beantworten.

Eine weitere, bislang nicht geklärte Frage betrifft die Voraussetzungen auf Seiten der Lerner und wie diese den Schreibprozess beeinflussen: Wirkt sich z.B. die Motivation der Textproduzenten auf die Qualität des Schreibprozesses aus? Und inwiefern wirkt das Schreiben auf die Motivation zurück?

Diese Fragen sind Gegenstand dieser Arbeit. Vor dem beschriebenen Hintergrund ist das Ziel die experimentelle Untersuchung der Effekte einer Prompts-Intervention auf die Qualität von Schreibprozessen bei Studierenden. Im Zentrum steht dabei die Förderung anspruchsvoller Lernstrategien. Um zu ergründen, inwiefern sich die Intervention auf die Wissensstruktur auswirkt, werden die mit den Prompts produzierten Texte auf ihre Verarbeitungstiefe hin analysiert und Zusammenhänge zwischen Merkmalen der Prompts und der in den Texten abgebildeten Wissensstrukturen hergestellt.

Zusätzlich soll im Rahmen der Untersuchung geklärt werden, ob Informationen über das Schreibprodukt in Form von Feedback die Optimierung des Schreibprozesses unterstützen. Zur Untersuchung dieses Themenfeldes werden zwei verschiedene Formen des Feedbacks erprobt.

Die vorliegende Arbeit setzt sich aus sechs Kapiteln zusammen, deren wesentliche Aspekte im Folgenden skizziert werden.

Kapitel 2: Den Ausgangspunkt der Arbeit bildet die Darstellung der theoretischen und empirischen Grundlagen für das Schreiben als Lerninstrument. Es werden unterschiedliche theoretische Perspektiven beschrieben, die sich mit der Frage befassen, welche Mechanismen und Faktoren ausschlaggebend für die Lernwirksamkeit des Schreibens sind. Im Mittelpunkt stehen dabei ausschließlich solche Theorien und Studien, die kognitive Prozesse als Bedingung für das Lernen durch Schreiben in den Blick nehmen. Schließlich werden Forschungsergebnisse zum Zusammenhang zwischen Schreibinstruktionen und dem Lernerfolg beschrie-

ben, wobei Studien zu Lerntagebüchern und Prompts einen Teilbereich bilden. Aus den dargestellten Studien wird sodann ein Forschungsbedarf erarbeitet, der zu den Fragestellungen der Arbeit führt.

Kapitel 3: In Kapitel 3 werden auf der Basis der empirischen Studien Bedingungen für eine effektive Schreibinstruktion mit Prompts herausgearbeitet. Das Ziel der Instruktion ist die Nachbearbeitung von Seminarinhalten, die über ein halbes Semester lang durch das Schreiben eines Lerntagebuchs mit Prompts dokumentiert werden sollen. Durch das Schreiben sollen die Seminarteilnehmer zu einer tiefgreifenden Verarbeitung der Lerninhalte angeregt werden. Dazu wird eine Lernumgebung aus den Komponenten Feedback und Prompts entwickelt, deren Wirksamkeit im empirischen Teil der Arbeit geprüft wird.

Kapitel 4: Die Fragestellungen, denen die Arbeit nachgeht, bilden Kapitel 4. Zur Erprobung der Lernumgebung werden zwei Studien durchgeführt, für die aufeinander aufbauende Fragekomplexe entwickelt werden. In der ersten Studie geht es zunächst um die Erprobung inhaltlich verschieden ausgestalteter Prompts. Im Fokus steht die Frage, welche Merkmale von Prompts sich besonders eignen, um anspruchsvolle kognitive Strategien anzuregen. Mit einem Analyse-Instrument (SOLO-Taxonomie), dessen Eignung ebenfalls im Rahmen der ersten Untersuchung überprüft wird, werden dazu die Texte im Hinblick auf ihre Komplexität eingestuft. Daraus werden Rückschlüsse auf die Verarbeitungstiefe beim Schreiben gezogen.

Aufbauend auf den Erkenntnissen der ersten Studie werden in der zweiten Studie zusätzlich Lernervariablen und zwei Feedbackvarianten in den Blick genommen. Der zweite Fragekomplex befasst sich mit den Wechselwirkungen von Feedback, Verarbeitungstiefe und Motivation der Lerner.

Kapitel 5: In Kapitel 5 werden zwei empirische Studien vorgestellt, die die Effekte verschiedener Prompts auf die Textqualität untersuchen. In der ersten Studie geht es um die Erprobung der Prompts und eines Analyse-Instrumentes, mit dessen Hilfe die Textqualität eingestuft wird. Die zweite Untersuchung integriert zusätzlich Lernervariablen (Vorwissen und Motivation) und zwei Formen des Feedbacks (Selbstfeedback und Dozentenfeedback).

Kapitel 6: Den Abschluss dieser Arbeit bildet die zusammenfassende Diskussion der Erkenntnisse aus den theoretischen Überlegungen und den empirischen Befunden zur Prompts-Maßnahme. Im Anschluss daran werden Grenzen der Un-

tersuchung aufgezeigt und Perspektiven für zukünftige Untersuchungen entwickelt.

2 Theoretischer Hintergrund

Die Untersuchung des Schreibens als Lerninstrument ist eine relativ junge Forschungsrichtung. War die universitäre Lehre lange Zeit mündlich geprägt, stellt das Verfassen von Texten zum Zweck des Wissenserwerbs heutzutage ein Kernelement schulischen und universitären Lernens dar (vgl. Kruse, Jakobs & Ruhmann, 1999). So schreiben Studierende und Schüler Forschungsartikel, Hausarbeiten, verfassen Berichte über wissenschaftliche Experimente und führen Lernjournale oder Portfolios. Betrachtet man die Vielfalt der Textarten und Textgenres in bildenden Kontexten, so wird deutlich, dass Schreiben nicht nur der Kommunikation von Wissen dient, sondern dass es von Lehrenden als intellektueller Prozess verstanden wird.

Schreiben dient der Organisation von Wissen (z.B. Abschlussarbeiten, Lernjournale), der Erkundung von Zusammenhängen zwischen Ideen (z.B. Essays) und der Entwicklung neuer Ideen (z.B. Kreatives Schreiben). Seit den siebziger Jahren des vergangenen Jahrhunderts werden die Zusammenhänge zwischen Schreiben und Lernen bzw. Schreiben und Denken untersucht. Die Arbeiten waren über die letzten Jahrzehnte von wechselnden Auffassungen und wechselnden Forschungstraditionen geprägt. Im Zentrum der Untersuchungen standen dabei Fragen wie: Trägt Schreiben zum Lernen bei? Welche Formen des Lernens werden durch Schreiben gefördert? Welche Prozesse konstituieren den Schreibprozess? Welche Variablen beeinflussen ihn? Das folgende Kapitel beginnt mit einer Definition des Schreibbegriffes, der dieser Arbeit zugrunde liegt. Im Anschluss daran werden theoretische Perspektiven auf das Schreiben als Lerninstrument und ihr empirischer Gehalt dargestellt.

2.1 Definition des Schreibbegriffes

Schreiben kann aus vielen verschiedenen Perspektiven betrachtet werden und entsprechend facettenreich ist die Forschung zu dem Gegenstand. So kann Schreiben z.B. als grafisch-motorische Tätigkeit aufgefasst werden und das Forschungs-

teresse sich auf das Erlernen der motorischen Fertigkeit richten (Eigler, 1996; Molitor-Lübbert, 1996). In der vorliegenden Arbeit jedoch steht das Produzieren von Texten im Mittelpunkt. Dem hier zugrundeliegenden Begriffsverständnis nach ist Schreiben eine mentale Aktivität, die durch kognitive Teilprozesse geprägt ist (z.B. die Planung, Redaktion und Überarbeitung von Texten). Nach Molitor-Lübbert (1996) wird für diese Auffassung von Schreiben der Begriff „Textproduzieren“ oder „Textproduktion“ verwendet. Aus Gründen der Sprachgefälligkeit wird jedoch im weiteren Verlauf der Arbeit der Begriff „Schreiben“ synonym zu dem Begriff „Textproduktion“ gebraucht.

Die Besonderheit dieses Ansatzes liegt im Forschungsfokus, der sich dadurch auszeichnet, dass er den Vorgang der Textproduktion in den Blick nimmt. Im Zentrum steht der Autor, der, ausgerichtet an Textsorte, Intention des Textes, Adressat und in Abhängigkeit seiner kognitiven Voraussetzungen, einen Text produziert. Texte werden dementsprechend als Produkte einer Vielzahl mentaler Prozesse betrachtet und auf kommunikative, argumentative und rhetorische Strukturen hin analysiert (vgl. Kruse, 1997).

In der vorliegenden Arbeit wird Schreiben jedoch ausschließlich als Medium des Lernens untersucht. Diese Forschungsperspektive geht auf eine von England und den USA ausgehende pädagogische Initiative ab 1970 zurück, die unter der Bezeichnung „*Writing to learn*“ oder „*Writing across the curriculum*“ bekannt geworden ist. Der daraus entstandene Forschungszweig befasst sich mit der Frage, inwiefern Schreiben nicht nur ein Kommunikationsmittel, sondern ein intellektueller Prozess ist, der die Strukturierung und Organisation von Wissensbeständen fördert und letztlich Problemlösungen und effektive Wissensnutzung begünstigt.

2.2 Theoretische Perspektiven auf Schreiben als Lernmedium

Im folgenden Abschnitt wird beschrieben, wie sich das Interesse an Schreiben als Lernmedium entwickelt hat und welche theoretischen Annahmen ihm zugrunde liegen. Der Zusammenhang zwischen Lernen und Schreiben unterlag in den letzten vierzig Jahren wechselnden Auffassungen und Forschungsparadigmen. Bis in die siebziger Jahre hinein wurde Schreiben ausschließlich als Kommunikationsmittel aufgefasst und das Forschungsinteresse fokussierte vor allem auf die Frage, welche Merkmale kommunikationswirksame Texte konstituieren und wie sich Tex-

te anhand dieser optimieren ließen (Feilke, 1996; Schnotz, 2001). Dazu wurden Texte als Produkte des Schreibens nach syntaktischen Kategorien analysiert. Da Schreiben nicht als Instrument des Wissenserwerbs verstanden wurde, fand in den Bildungsinstitutionen eine didaktische Trennung von Schreiben als Kommunikationsinstrument und dem Erlernen fachlicher Inhalte statt (Quinn, 1995). Unter dem Einfluss der „kognitiven Wende“ ab circa 1970 veränderte sich das wissenschaftliche Verständnis von Lernen jedoch und dies beeinflusste die Schreibforschung (Tynjälä, Mason & Lonka, 2001) und veranlasste Erzieher und Forscher, Hypothesen über die Beziehung zwischen Lernen und Schreiben anzustellen. Mit dem kognitiven Paradigma rückten Denk- und Verarbeitungsprozesse und deren Koordination und Interaktivität in den Mittelpunkt der Lernforschung und ebenso der Textproduktionsforschung. Kognitive Lerntheorien fassen Lernen als einen Informationsverarbeitungsprozess auf, für den die Eigenaktivität des Lerners zentral ist (vgl. Edelmann, 2000).

Die kognitive Wende beeinflusste die Textproduktionsforschung in zweierlei Hinsicht. Zum einen fanden kognitive Prozesse Eingang in die Schreibmodelle (Flower & Hayes, 1981) und Schreiben wurde nicht länger als eine lineare und chronologisch angeordnete Abfolge von Produktionsstufen (z.B. Beaugrande, 1982) gesehen, sondern als ein sich aus rekursiv verlaufenden Teilprozessen zusammensetzender Prozess aufgefasst. Zum anderen führte die kognitive Wende dazu, dass einige Erzieher Schreiben mit Lernen und Denken gleichsetzten, weil sie der Auffassung waren, alle drei Tätigkeiten seien durch ähnliche kognitive Prozesse geprägt. Besondere Beachtung in der Schreibforschung hat in dieser Hinsicht das Essay „Writing as a mode of learning“ von Emig (1977) gefunden. Die Autorin vergleicht darin die Eigenschaften des Schreibprozesses mit denen „besonders wirkungsvoller“ Lernstrategien. Aufgrund der Ähnlichkeiten, die sie in Bezug auf die Attribute beider Tätigkeiten vorfindet, kommt sie zu dem Schluss, dass Schreiben als Lernform verstanden werden muss: „Writing serves learning uniquely because writing as process-and-product possesses a cluster of attributes that correspond uniquely to certain powerful learning strategies“ (Emig, 1977, S. 122). Nach Emig (1977) sind es vor allem vier Merkmale, die beide Tätigkeiten teilen und elf Merkmale insgesamt, welche Schreiben als Lernform auszeichnen:

1. Der Schreiber / Lerner bestimmt das Tempo selbst.

2. Beide Tätigkeiten beinhalten die aktive Organisation individueller Wissensbestände.
3. Der Lerner oder Schreiber benötigt und sucht sich Feedback; durch das Schreiben werden Inhalte zugänglich für Selbstreflexion.
4. Schreiben wie auch Lernen aktivieren beide Gehirnhälften.

Emig (1977) bezieht sich im vierten Punkt auf die Hemisphärentheorie von Roger Wolcott Sperry. Man nahm damals an, dass die Aktivitäten der linken und rechten Gehirnhälfte durch Schreiben integriert würden, was wiederum für Lernprozesse förderlich wäre. Diese Annahme ist jedoch nach dem aktuellen Stand der Hirnforschung stark umstritten. Dennoch wurde die Bedeutung des Schreibens für das Lernen weiter untersucht. Neben Emig begannen eine Vielzahl weiterer Erziehungswissenschaftler, den Wert des Schreibens für das Lernen und Denken zu betonen und zu untersuchen.

Die Frage, welche Mechanismen des Schreibens lernförderlich oder erkenntnisförderlich sind, wird von verschiedenen theoretischen Perspektiven unterschiedlich beantwortet. Klein (1999) ordnet diese Perspektiven im Rahmen eines Forschungsüberblicks über die Textproduktionsforschung vier Kategorien zu, die er als Hypothesen versteht. Die Unterscheidungsmerkmale der Hypothesen sind dabei die Handlungen und Prozesse, die beim Schreiben angenommen werden. Klein (1999) postuliert, dass die Hypothesen insofern kompatibel miteinander sind, als dass sie ganz unterschiedliche Prozesse und Handlungen adressieren. Als verbindendes Merkmal der unterschiedlichen Ansätze ist die Lernkonzeption festzuhalten, die ihnen jeweils zugrunde liegt. So wird nach Klein (1999) insbesondere tiefenorientiertes Lernen als Indikator für den Erfolg von Schreiben erhoben. Im Kontrast zu oberflächenorientiertem Lernen (z.B. Wiederholen einzelner Aspekte des Lernstoffs) wird tiefenorientiertes Lernen über Aufgaben gemessen, bei denen geprüft wird, inwiefern der Lerner die entsprechenden Inhalte auf konzeptueller Ebene verstanden hat. Weitere Hinweise auf die Lernkonzeption geben die Kategorien, mittels derer die Effekte und Mechanismen des Schreibens beschrieben werden. Begriffe wie „Wissenstransformation“, „*conceptual change*“ (Bereiter & Scardamalia, 1987; Fellows, 1994) und Sinnkonstruktion (*construction of meanings*) (Spivey, 1990) zeigen, dass Lernen als ein Prozess der aktiven Auseinandersetzung mit Lerninhalten verstanden wird, der zu Veränderungen des bereits

vorhandenen Wissens führt. Das Verständnis von Lernen wird innerhalb der *Writing to learn*-Bewegung folglich klar eingegrenzt.

Die erste Hypothese (*hypotheses related to the spontaneous idea generation*) wird durch verschiedene Varianten der Annahme geprägt, dass Schreiber während des Schreibens spontan Wissen generieren. Ein prominenter Vertreter ist Britton (1980), der den Begriff „*shaping at the point of utterance*“ einführte. Britton (1980) nahm an, dass durch freies und spontanes Schreiben implizites Wissen, auf das Personen sonst keinen Zugriff haben, explizit und damit der Reflexion zugänglich gemacht wird: „Shaping at the point of utterance, according to Britton, converts tacit knowledge into explicit knowledge“ (Klein, 1999, S. 212). Darüber hinaus betont der Ansatz die Rolle des Schreibens für kreative Prozesse, denn nach Britton entwickeln sich auch neue Ideen beim Schreiben. Diese Grundannahme spiegelt sich in dem in diesem Zusammenhang häufig zitierten Ausspruch von E. M. Forster: „How can I know what I think until I see what I say?“ wider (Applebee, 1984). Expressives Schreiben stellt nach dieser Perspektive die lernförderlichste Art des Schreibens dar, weil der Schreiber seinen Gedanken freien Lauf lassen kann und nicht durch die Anforderungen einer Aufgabe eingeschränkt wird. Wenngleich durch spontanes Schreiben kreative Prozesse ausgelöst werden und neuen Ideen zustande kommen, bleibt dennoch zu bedenken, dass diese Variante des Schreibens im Falle von Fehlkonzepten nicht ohne den Zusatz von Feedback oder die Zuhilfenahme von Texten lernförderlich sein dürfte, da es unwahrscheinlich ist, dass Lerner sich durch schlichtes Versprachlichen ihrer Fehlkonzepte bewusst werden. Lernen im Sinne von Konzeptwechsel (*conceptual change*), also der Neuinterpretation von Wissen und dem Beheben von Fehlkonzepten, bedarf eher feedbackorientierten Formen des Lernens: „In contrast, conceptual change appears to depend on students taking their prior beliefs as objects for talk, thinking *about* them, rather than *with* them“ (Klein, 1999, S. 219).

Die zweite Hypothese (*forward search hypothesis*) enthält Ansätze, die den Schritt des Überarbeitens (*review*) beim Schreiben betonen. Schreiben ermöglicht demzufolge eine Umwandlung und qualitative Verbesserung des Wissens, weil der Lernende, indem er Text produziert und somit seine Gedanken externalisiert, die Möglichkeit hat, die Skizze zu überarbeiten, Widersprüche zu entdecken, Schlüsse zu ziehen und Verbindungen zwischen verschiedenen Aspekten des

Textes herzustellen. Zahlreiche Studien haben die Effekte von Überarbeitungsprozessen überprüft (z.B. Benton, Kiewra, Whitfill & Dennison, 1993; Bereiter & Scardamalia, 1987; Hayes et al., 1987; Wallace & Hayes, 1991). Dabei zeigte sich, dass erfahrene Schreiber ihre Texte zum Zwecke der Qualität und Kohärenz überarbeiten. Hayes et al. (1987) konnten zudem nachweisen, dass die Fähigkeit, die Probleme eines Textes richtig zu diagnostizieren und die Art der Überarbeitungsprozesse einen Effekt auf Textkohärenz haben. Um diesen Effekt zu überprüfen, ließen sie professionelle Textproduzenten (Journalisten) und Studierende (Novizen) einen fehlerhaften Text redigieren und zeichneten dabei Protokolle Lauten Denkens (Personen verbalisieren ihre Gedanken handlungsbegleitend) auf. Im Anschluss daran befragten sie die Probanden nach den Veränderungen, die diese am Text vorgenommen hatten. Hayes et al. (1987) kamen zu dem Ergebnis, dass die Experten die Probleme des Textes besser diagnostizieren und daraus adäquate Maßnahmen ableiten konnten: „The experts diagnosed problems, rather than simply detecting them, and these diagnoses pointed to appropriate revisions“ (Klein, 1999, S. 228). Die Vorgehensweise der Experten und der Studierenden unterschied sich des Weiteren darin, dass die Studierenden eher einzelne Sätze überarbeiteten, während die Experten den gesamten Text redigierten.

Während diese Studie für die Bedeutung des Überarbeitungsschrittes spricht, deuten andere Befunde darauf hin, dass dieser bei kurzen Aufgaben nicht notwendig ist (z.B. Bereiter & Scardamalia, 1987; Gould & Boies, 1978). Gould und Boies (1978) trainierten acht geübte Diktierer und acht Novizen (Collegeabsolventen im Alter von 25-45) einen Tag lang im Diktieren von Geschäftsbriefen. Im Anschluss daran verfassten die Probanden 16 Briefe mittels unterschiedlicher Methoden (Diktieren, unsichtbares Schreiben, Schreiben, Sprechen). Beim unsichtbaren Schreiben schrieben die Teilnehmer der Trainingsstudie mit einem Holzstift auf ein Blatt, unter dem Durchschlagpapier lag. Die Briefe wurden sodann auf ihre Qualität hin eingestuft. Dabei zeigte sich, dass die Briefe sich über die verschiedenen Gruppen hinweg im Mittel hinsichtlich der Qualität nicht unterschieden, wenngleich das Überarbeiten beim Diktieren oder Sprechen schwerer als beim Schreiben ist und beim unsichtbaren Schreiben sogar unmöglich. Bei kurzen Aufgaben ist also ein Überarbeitungsschritt nicht unbedingt notwendig. Dies bestätigen auch die Untersuchungen von Bereiter und Scardamalia (1987), die Probanden ebenfalls unsichtbar schreiben ließen. Auch hier zeigten sich keine

Unterschiede zwischen der originalen und der redigierten Fassung.

Mehrere von Benton et al. (1993) durchgeführte Experimente mit unterschiedlichen Notizen-Formaten lassen außerdem vermuten, dass auch andere Faktoren als Überarbeitungsprozesse die Textkohärenz beeinflussen, wie die Enkodierspezifität. Das Prinzip der Enkodierspezifität nach Tulving und Thomson (1973) geht von einem Zusammenhang zwischen der Speicherung neuer Informationen und ihrem Abruf aus. Es besagt, dass beim Lernen von Informationen immer auch Aspekte der Lernumgebung (z.B. die Art der Aufgabe) und des Kontextes (z.B. ein bestimmter Raum, die emotionale Befindlichkeit) gespeichert werden. Die Wahrscheinlichkeit, sich an Informationen zu erinnern, ist diesem Prinzip zufolge umso höher, je mehr sich die Bedingung des Abrufes und die der Speicherung (Enkodierung) ähneln.

Benton et al. (1993) führten mit Studierenden der Psychologie vier Experimente durch, in denen sie die Wirkung von Notizen unterschiedlicher Formate auf die Lernleistung (in Form eines vergleichenden Essays erfasst) überprüften. In den ersten zwei Experimenten präsentierten sie den Studierenden eine Video-Vorlesung zu Kreativitätsformen und wiesen die Probanden an, sich Notizen zu machen oder nur der Vorlesung zu folgen. Das Notizen-Format variierte von *freien, individuell gestaltbaren Notizen*, von den Dozenten vorgegliederten Notizen (*Outline*), für die die Forscher Überschriften vorgaben, und tabellarisch aufgebauten Notizen (*Matrix*). Im ersten Experiment schrieben die Probanden direkt im Anschluss an die Vorlesung ein Essay, für das jeweils eine Gruppe aus jedem Notizen-Format ihre Aufzeichnungen zur Hilfe nehmen durfte, während die andere Gruppe aus dem Gedächtnis schrieb. Das zweite Experiment stellte eine Wiederholung der ersten Untersuchung dar mit dem Unterschied, dass das Essay mit einer Woche Verzögerung geschrieben wurde. Die Essays der Probanden wurden sodann auf zwei Maße der Organisation (Kohärenz und Kohäsion) und auf die Anzahl der Wörter und der Texteinheiten hin kategorisiert.

Im Ergebnis zeigte sich, dass in beiden Experimenten die Studenten, die das Essay unter Zuhilfenahme ihrer Notizen schrieben, längere und besser organisierte Essays verfassten als die Vergleichsgruppe, die ihre Notizen nicht verwenden durfte. Beim zweiten Experiment mit verzögertem Schreiben zeigte sich das gleiche Resultat, das allerdings noch prägnanter ausfiel.

Da die ersten zwei Experimente die Wirkung der Verfügbarkeit von Notizen

nicht isoliert vom Notizenmachen untersucht hatten, führten Benton et al. (1993) zwei weitere Experimente mit derselben Video-Vorlesung durch, in denen die Studierenden keine eigenen Notizen machten, sondern fremde Notizen zur Verfügung gestellt bekamen. Die Studierenden durften während der Vorlesung keine Notizen machen, sondern erhielten nach dem Vortrag fremde, bereits fertige Notizen derselben Formate wie in den ersten beiden Experimenten. Die Kontroll-Gruppe erhielt keine Notizen. Sodann wurden die Probanden angewiesen, unter Zuhilfenahme der fremden Notizen ein Essay zu schreiben. In Experiment drei schrieben die Studierenden direkt im Anschluss an den Vortrag, in Experiment vier wiederum verzögert um eine Woche.

Im dritten Experiment zeigten sich keine Unterschiede zwischen der Qualität der Essays der Probanden mit und deren Essays ohne Notizen. Die Verfügbarkeit von Notizen alleine ist also nicht hilfreich in Bezug auf die Gliederung und Kohärenz eines Textes. Die Autoren stellten die Vermutung an, dass die Gruppen mit Notizen im dritten Experiment unter Berücksichtigung des Prinzips der Enkodierspezifität in den Essays besser abgeschnitten hätten als die Gruppe ohne Notizen: „The principle of encoding specificity (Tulving & Osler, 1968) may offer an explanation for why no external-storage effect was found in this experiment. [...] Similarly, in this experiment, providing students with someone else’s notes after they have viewed the lecture fails to enhance writing processes. According to the principle of encoding specificity, the notes must be available when listening to the lecture in order to serve as an effective retrieval cue at the time of writing“ (Benton et al., 1993, S. 276). Für die Rolle der Enkodierspezifität könnte jedoch sprechen, dass nur die selbstverfassten Notizen in Kombination mit deren Verfügbarkeit während des Essayschreibens zu signifikanten Effekten hinsichtlich der Kohärenz und der Organisation der Texte führten. Hier lagen also während der Informationsaufnahme ähnliche Kontextbedingungen vor wie beim Abruf.

Zusammenfassend lässt sich beim Vergleich der verschiedenen Untersuchungen festhalten, dass die bisherigen Studien keine eindeutigen Zusammenhänge zwischen Überarbeitungsprozessen und Lernerfolg aufdecken. So wurden zwar direkte Effekte auf die Textkohärenz untersucht, nicht jedoch, inwiefern die Revision auch das Wissen verändert, z.B. indem dieses danach besser strukturiert oder tiefer verankert ist. Es besteht in dieser Hinsicht noch Forschungsbedarf und Klein (1999) fasst zusammen: „If these revisions can be shown to affect writers’

knowledge this will provide further support for the forward search hypothesis“ (S. 230).

Gemäß der dritten und am intensivsten untersuchten *Genre-Hypothese* ist das Wissen des Schreibers über Genre-Strukturen ausschlaggebend für die Effekte des Schreibens auf das Lernen. Das Genre eines Textes beschreibt dessen rhetorische Funktion, die durch ein Zusammenspiel bestimmter Textelemente (z.B. Vergleiche anstellen, Kontrastieren, Analogien bilden, Belege anführen etc.) repräsentiert wird: „Generally, a genre is distinguished by a rhetorical intention, expressed through discourse elements that form particular relationships with one another“ (Klein, 1999, S. 230). Der Genre-Hypothese zufolge erfordern verschiedene Text-Genres unterschiedliche kognitive Strategien vom Schreiber: “For example, analytic writing requires writers to link claims, warrants, and evidence, but analogies and metaphors require writers to link new domains to other domains with which they are already familiar“ (Klein, 1999, S. 231).

In den unter diesem Stichwort referierten Studien werden kognitive Strategien (z.B. Elaboration und Organisation) als Wirkfaktor für den Lernerfolg angesehen, wobei es zwei unterschiedliche Erklärungsansätze für die Rolle des Genres gibt. So setzen einige Studien nach Klein (1999) implizit voraus, dass das Schreibgenre als Stimulus funktioniert, der spezifische kognitive Prozesse anregt (*response*), die wiederum Lernen zur Folge haben (Reiz-Reaktions-Modell). Der Autor zählt dazu solche Studien, die die kognitiven Strategien, die zu einem bestimmten Textprodukt führen, vernachlässigen und stattdessen Beziehungen zwischen Aufgaben aus einem bestimmten Genre und dem Lernergebnis untersuchen. Diese Perspektive greift insofern zu kurz, als dass kognitive Strategien an individuelle Ziele des Textproduzenten gebunden sind, wie die Lernforschung zeigt (vgl. Klein, 1999).

Im Kontrast dazu wird Schreiben innerhalb des zweiten Erklärungsansatzes als Prozess aufgefasst, innerhalb dessen der Textproduzent ein Genre als Ziel adaptiert und sodann nach geeigneten kognitiven Strategien sucht, um dieses Ziel zu erreichen. Die Strategien beinhalten Handlungen am Text (z.B. Beweisführung, Kontrastieren) und gedankliche Aktivitäten (z.B. den Inhalt planen, Gründe untersuchen). Diese Perspektive setzt zum einen auf Seiten des Textproduzenten voraus, dass dieser die notwendigen „Genre-Bausteine“ kennt und sie flexibel einsetzen kann. Zum anderen muss der Textproduzent, wenn der Text innerhalb eines Lernkontextes verfasst wird, die Genre-Aufgabe auf die vom Lehrenden intendier-

te Weise interpretieren (*task interpretation*). Hier deuten sich Herausforderungen der Genre-Perspektive an, auf die im Weiteren eingegangen wird.

Die innerhalb der Genre-Hypothese dargestellten Studien gehören zu einem umfangreichen Untersuchungsfeld, das für die vorliegende Arbeit essentielle Informationen über die Bedingungen der Gestaltung einer Lernumgebung, innerhalb der Schreibaufgaben eingesetzt werden sollen, bietet. Die Befunde von Klein (1999) werden daher im Folgenden in einer eigenen Gliederung dargestellt. Es wird erörtert, welche kognitiven und metakognitiven Prozesse durch Schreiben angeregt werden, welche Formen des Lernens dadurch evoziert werden und welche Bedingungen seitens der Aufgabe und des Lerners zu bedenken sind.

Kognitive Strategien als Wirkfaktoren beim Schreiben

Während Britton (1980) davon ausgeht, dass Lerner durch die kreativen Gedankenprozesse, die im Rahmen expressiven Schreibens ausgelöst werden, lernen, zeigt die Schreibforschung, dass beim Erlernen fachspezifischer Inhalte vor allem analytisches Schreiben Lerneffekte nach sich zieht (Bangert-Drowns, Hurley & Wilkinson, 2004; Langer & Applebee, 1987; Wiley & Voss, 1996). Analytische Schreibformen sind das Essay, Lernfragen und Texte, die vom Textproduzenten eine kohärente und fokussierte Darstellung eines bestimmten Inhaltes unter Einhaltung einer formellen Sprache verlangen. Das gemeinsame bestimmende Merkmal analytischer Schreibformen ist eine kritische Auseinandersetzung mit spezifischen Inhalten, wobei der Textproduzent z.B. eigene Hypothesen formuliert, Aspekte des Lernstoffs kontrastiert oder sich Analogien ausdenkt.

Hinsichtlich der Effektivität analytischer Schreibformen zeigt sich z.B. beim Vergleich der Lernergebnisse von Lernern, die sich einen Text mithilfe einer Schreibaufgabe aneignen sollen, und Lernern, die sich ohne Schreiben die Inhalte aneignen, dass die Textproduzenten ein besseres und tiefergreifendes Verständnis für Lernkonzepte oder Kernaussagen eines Textes entwickeln (Langer & Applebee, 1987; Wiley & Voss, 1996, 1999). Weitere Befunde zeigen, dass das Schreiben zu einer aktiveren Auseinandersetzung mit den Lerninhalten führt und dass Textproduzenten während des Schreibens neue Informationen in ihr Vorwissen integrieren (Hayes, 1987; Newell, 1984).

Eine Gemeinsamkeit der zahlreichen Befunde zum Lernen durch analytisches Schreiben ist die zentrale Bedeutung, die kognitiven Strategien beigemessen wird.

Mandl und Friedrich (2006) fassen darunter solche Handlungen zusammen, die der Informationsverarbeitung dienen. Dazu zählen z.B. Elaborieren (eigene Beispiele ausdenken), Organisieren (in eigenen Worten zusammenfassen) und Memorieren. Diese sogenannten Lernstrategien können auch ohne zu schreiben angewandt werden und nicht jede Form des Schreibens schließt Lernstrategien ein (z.B. Tagebuchschreiben). Soll das Schreiben allerdings dem Lernen dienen, so sind es diese Handlungen, die für den Lernerfolg ausschlaggebend sind.

Aus Analysen von Protokollen Lauten Denkens, die während der Bearbeitung analytischer Schreibaufgaben aufgezeichnet wurden, geht hervor, dass verschiedenartige Aufgaben unterschiedliche kognitive Strategien nach sich ziehen, die sowohl in ihrer Quantität als auch in der Qualität variieren (z.B. Langer & Applebee, 1987; Wiley & Voss, 1996, 1999). Bangert-Drowns et al. (2004) fassen die Rolle des Schreibens für das Lernen deswegen auch als die eines Strategieaktivators auf: „One way to conceptualize the potential benefits of writing for learning is to consider its impact on discrete learning processes“ (S. 32).

Zahlreiche Studien sind der Frage nachgegangen, welche kognitiven Prozesse durch analytische Schreibaufgaben evoziert werden und inwiefern dies Auswirkungen auf den Wissenszuwachs hat.

Eine besonders detaillierte Untersuchung nahmen Langer und Applebee (1987) vor. In insgesamt drei Studien analysierten sie die Effekte verschiedener Schreibaufgaben (Notizen machen, Essay schreiben, Verständnisfragen beantworten) auf das Behalten und Verstehen fachspezifischer Lerninhalte und versuchten, in Abhängigkeit von der Schreibaufgabe angewandte Strategien der Textproduzenten zu identifizieren. Methodisch griffen sie zum einen auf Protokolle des Lauten Denkens und zum anderen auf Textanalysen von Essays zurück, die die Lerner nach der Bearbeitung der eigentlichen Aufgabe verfassten. In zwei der Studien werden die Schreibaufgaben zudem mit dem Erlernen eines Textes ohne Schreiben verglichen (*read and study*). Als Stichprobe dienten Schüler der neunten und elften Klassenstufe, denen die Autoren zunächst Textpassagen präsentierten. Im Anschluss daran bearbeiteten die Gruppen verschiedene Schreibaufgaben.

Die erste Studie diente vor allem der Identifikation kognitiver Strategien in Zusammenhang mit unterschiedlichen Schreibformen (Notizen machen, Lernfragen, Essay). Langer und Applebee (1987) ließen dazu sechs Highschool-Schüler jeweils einmal jede Schreibaufgabe bearbeiten und analysierten anschließend die Proto-

kolle Lauten Denkens. Die darin vorgefundenen kognitiven Strategien fassten sie zu folgenden Kategorien zusammen:

1. *Questioning* (relativ offene Suche nach einer Antwort)
2. *Hypothesizing* (Vorhersage über Konzepte aus dem Text)
3. *Using schemata* (Kommentare über den Inhalt)
4. *Schemata evaluated* (Evaluieren eigener Ideen)
5. *Making metacomments* (die eigenen Versuche kommentieren)
6. *Evidence and validation* (Auswerten der eigenen Interpretationen)

Im Ergebnis zeigte sich, dass das Schreiben von Essays die größte Bandbreite und die höchste Anzahl kognitiver Strategien nach sich zog. Besonders häufig kommentierten die Lerner während des Schreibens die Textinhalte (*using schemata*) und stellten Hypothesen über Konzepte aus dem Text an (*hypothesizing*). Die Autoren beobachteten hier, dass die Lerner sich beim Schreiben nicht so stark auf den Text rückbezogen, sondern die Informationen eigenständiger weiterverarbeiteten als es bei den anderen beiden der Fall war. Bei den Lernfragen enthielten die Protokolle Lauten Denkens die geringste Menge an kognitiven Strategien. Die Schüler konzentrierten sich bei dieser Aufgabe am häufigsten auf das Auffinden von Informationen im Text (*questioning*), ohne diese jedoch tiefer zu verarbeiten (z.B. kritisch zu hinterfragen oder sich eigene Beispiele auszudenken): „The students paid little attention to what they thought they knew or had learned“ (Langer & Applebee, 1987, S. 97). Beim Notizenmachen lag der Schwerpunkt der Handlungen hingegen auf dem Kommentieren der Inhalte (*using schemata*). Die Lerner entnahmen hier, ähnlich wie beim Beantworten der Lernfragen, in erster Linie Informationen aus dem Text, um diese in eigenen Worten zusammenzufassen. Dabei zeigte sich wiederum eine geringe Intensität hinsichtlich der Verarbeitung der Inhalte: „They spent little time considering how the ideas related to each other or to other things they knew. While engaged in the note-taking activity, the students tended to read the passage in small segments and to use the temporal structure of the passage to structure their notes. They did not stop to integrate the information into larger units that might have then been used to

structure their notes. Instead, they tended to use the text's paragraphs as their organizational frame" (Langer & Applebee, 1987, S. 97).

Die zwei daran anschließenden Untersuchungen von Langer und Applebee (1987) wurden an einer größeren Stichprobe vorgenommen (Studie 2: 208 Schüler der neunten und elften Klasse, Studie 3: 112 Schüler der neunten und elften Klasse) und fokussierten stärker auf die differenzierten Wirkungen der Schreibaufgaben auf das Lernen. In der zweiten Studie bearbeiteten 208 Schüler aus der neunten und elften Klasse drei verschiedene Schreibaufgaben (Notizen machen, Essay schreiben, Verständnisfragen beantworten), nachdem sie einen Text gelesen hatten. Außerdem lernte in jeder Klassenstufe eine Gruppe zur Kontrolle den Text ohne Schreibaufgaben. Als Lernmaße dienten die Menge des Themenwissens, das Inhaltsverständnis und das Anwenden der neuen Information. Im Ergebnis der Studie schnitten die Probanden, die das Essay verfasst hatten, bei der Menge des Themenwissens, das direkt im Anschluss an die Aufgabe erhoben wurde, am schwächsten ab. Bei der Anwendung der neuen Inhalte, die vier Wochen später mithilfe eines Essays getestet wurde, erzielten die Schüler aus der Essaygruppe jedoch die besten Ergebnisse. Die Essays wurden auf ihre Kohärenz und ihre Struktur hin analysiert. Textkohärenz wird in der Schreibforschung allgemein als Hauptfaktor für das Verstehen von Lerninhalten gedeutet (vgl. McNamara, Kintsch, Songer & Kintsch, 1996). Die Analyse der Essays im Post-Test zeigte, dass die Probanden aus der Essay-Gruppe komplexere Verbindungen zwischen einzelnen Konzepten herstellten und insgesamt kohärentere Texte verfassten.

Als ausschlaggebend für den Kontrast zwischen dem geringeren Themenwissen und dem besseren Verständnis im Post-Test sehen Langer und Applebee (1987, S. 110) den Aufmerksamkeitsfokus, der durch die unterschiedlichen Schreibaufgaben variierte: „The scores at the immediate post-test suggest that the essay task focused students' attention on a narrower range of information in the passage, thus providing them with fewer specific associations for the passage knowledge measure and leading to lower passage-specific knowledge scores“. Der engere Fokus der Essayaufgabe, so folgern die Autoren, führt zwar insgesamt dazu, dass sich die Essayschreiber weniger Wissen aneignen. Dafür erfordert das Essay jedoch eine intensivere Auseinandersetzung mit einem kleineren Ausschnitt des Textes und fördert dadurch ein tieferes Verständnis.

Dieser Befund wird in der dritten Studie von Langer und Applebee (1987)

bestätigt, in der die Autoren Schüler der neunten und elften Klasse vier Untersuchungskonditionen zuordneten. Während die Kontrollgruppe einen Text nur lesend lernen sollte, repräsentierten die drei Schreibaufgaben nach Langer und Applebee (1987) zunehmend komplexere Formen des Schreibens:

1. Verständnisfragen
2. Zusammenfassung des Textes
3. Essay

Zusammengefasst führten die Studien von Langer und Applebee (1987) zu drei Hauptergebnissen.

1. Beim Vergleich der schreibenden Gruppen mit einer nur lesenden Kontrollgruppe schnitten die Probanden mit Schreibaufgabe besser ab. Die Autoren führen dies anhand der Protokolle Lauten Denkens darauf zurück, dass die schreibenden Probanden das Lernmaterial flexibler interpretierten (also mehr kognitive Strategien anwendeten) als die Personen in der Untersuchungskondition ohne Schreiben. Lernstoff wird demnach umso besser verstanden und länger behalten, je mehr kognitive Operationen der Lerner damit vornimmt. Dies bestätigt sich auch beim Vergleich der drei Schreibformen (Notizen machen, Essay schreiben, Verständnisfragen beantworten) miteinander, wobei die Befunde hier komplexer ausfallen (siehe Punkt drei).

2. Die Effekte des Schreibens sind auf die im ersten Durchgang aktiv verarbeitete Information begrenzt. Es werden nur Informationen oder Ideen erinnert und reproduziert, über die Lerner tatsächlich auch geschrieben hatten. Darüber hinaus lässt sich keine lernförderliche Wirkung des Schreibens nachweisen.

3. Verschiedene Schreibgenres führen zu einem breiteren oder engeren Fokus und sind ausschlaggebend dafür, wie viel und wie tief neue Information verarbeitet wird. Bei der Menge des Themenwissens, das direkt im Anschluss an die Aufgabe erhoben wurde, schnitten die Probanden, die die Essay-Aufgabe bearbeitet hatten, am schwächsten ab. Bei der Anwendung der neuen Inhalte, die vier Wochen später erneut mithilfe eines Essays getestet wurde, erzielten die Schüler aus der Essaygruppe jedoch die besten Ergebnisse. So stellten diese Probanden im Post-Test komplexere Verbindungen zwischen einzelnen Konzepten her, verfassten kohärentere Texte und ihre Essays deuteten insgesamt auf ein besseres Verständnis des Stoffes hin.

Die Befunde von Langer und Applebee (1987) zeigen, dass Textproduzenten in Abhängigkeit von der Form des Schreibens unterschiedliche Strategien einsetzen, die entweder zu einem eher oberflächlichem Verstehen des Textes führen, das sich durch eine geringe Nachhaltigkeit des Wissens auszeichnet, oder zu einem tiefergreifenden, dafür aber nicht so breiten Verstehen der Lerninhalte, das allerdings auch mehr Anstrengung vom Textproduzenten erfordert. Dies zeigt sich zum einen in der Art und zum anderen in der Menge der kognitiven Prozesse.

Dass Essays eine vergleichsweise intensive Informationsverarbeitung nach sich ziehen, bestätigt auch die Untersuchung von Newell (1984). In einer Feldstudie ließ Newell (1984) acht Schüler der elften Klasse über sechs Sitzungen nach Lesen eines Textes mit historischem oder naturwissenschaftlichem Thema jeweils eine Schreibaufgabe (Essay, Notizenmachen, Verständnisfragen beantworten) bearbeiten. Da kein fester Zeitrahmen vorgegeben wurde, konnten die Schüler solange schreiben, wie es für die Fertigstellung der Aufgabe erforderlich war. Es zeigte sich, dass sie wesentlich mehr Zeit für das Essay benötigten als für die restlichen Schreibaufgaben.

Unterschiedliche Schreibaufgaben sind demnach für verschiedene Formen des Lernens geeignet. Es zeigte sich aus den Untersuchungen, dass die Bedingungen, unter denen sich Lerner am besten Fakten aneignen, andere sind als die optimalen Bedingungen für Textverständnis. Dies bestätigt sich auch in weiteren Studien (vgl. Mayer, 1999; Newell, 1984; Wiley & Voss, 1999). Daraus lässt sich festhalten, dass tiefgreifendes Lernen eher durch eine Schreibinstruktion angeregt wird, die eine intensive, dafür aber möglicherweise inhaltlich auf einen engeren Ausschnitt begrenzte Beschäftigung mit dem Lerngegenstand fördert. Lernfragen, die hingegen mehrere Aspekte des Lernstoffes abdecken, tragen dafür zum kurzfristigen Behalten von mehr Information bei. Der Lernerfolg beim Schreiben hängt folglich auch von der Verarbeitungsintensität ab. Es ist auf der Basis der Befunde von Langer und Applebee (1987) anzunehmen, dass eine Verarbeitung, wie sie beim Essay angeregt wird, dazu führt, dass Lerner die neue Information in ihr Vorwissen integrieren und diese daher nachhaltiger lernen. Die Effekte des Schreibens innerhalb der Genre-Hypothese können folglich nicht auf das Genre allein zurückgeführt werden, sondern implizieren zumindest den Faktor der Verarbeitungsintensität. Es stellt sich daher die Frage, ob nicht grundsätzlich die Verarbeitungsintensität entscheidend für den Lernerfolg ist und nicht das Genre-

wissen eines Lernalers.

Aus den bisherigen Studien zum Lernen durch Schreiben lassen sich Schnittmengen zum Lernen mit Texten erkennen, da als Maß für Lernerfolg innerhalb der *Writing-to-Learn*-Studien überprüft wird, wie gut ein Lerner einen Text verstanden hat. Die Studien aus dem Bereich der Schreibforschung bleiben allerdings im Hinblick auf das Verstehensniveau insgesamt sehr vage. Um differenzierte Aussagen über die Wirkungen von Schreibinterventionen zu machen, wäre eine genauere Einstufung jedoch notwendig. Mehr Aufschluss über Abstufungen des Wissensniveaus geben die Arbeiten zum Textverständnis aus dem Bereich der Psycholinguistik, wie z.B. das Textverarbeitungsmodell von Van Dijk und Kintsch (1983), das verschiedene, aufeinander aufbauende Ebenen des Textverständnisses postuliert. Textverständnis wird als Konstruktion abstrakter Wissenseinheiten (*Propositionen*) aufgefasst. Diese entstehen in Abhängigkeit der Verarbeitungsschritte des Lesers und spiegeln sich in der Struktur der mentalen Repräsentation des gelesenen Textes wieder, die mehr oder weniger komplex ausfallen kann. Um Verstehensebenen zu unterscheiden, nehmen Van Dijk und Kintsch (1983) eine *wörtliche Repräsentationsebene* des Textes an, auf der der Leser den Wortlaut eines Textes verstanden hat, eine *propositionale Repräsentation*, die die Textbedeutung einschließt und die *situative Repräsentation*, die neben dem Verstehen des Sachverhaltes auch die Integration eigener Wissensinhalte des Lesers einschließt und für das tiefste Verständnis steht. In Zusammenhang mit dem Schreibgenre „Essay“ wäre es interessant zu prüfen, inwiefern dabei solche Prozesse angeregt werden, die zum Aufbau einer *situativen Repräsentation* beitragen.

Einige empirische Studien deuten daraufhin, dass sich qualitative Unterschiede in den Verstehensleistungen tatsächlich auf den Repräsentationsebenen von Van Dijk und Kintsch (1983) abbilden lassen (Perrig & Kintsch, 1985; Schmalhofer, Boschert & Kühn, 1990). Diese scheinen außerdem in Zusammenhang mit vorgegebenen Lernzielen zu stehen.

So konnten Schmalhofer und Glavanov (1986) in drei Experimenten zeigen, dass sich die Veränderung vorgegebener Lernziele beim Lernen aus einem Text auf das Verstehensniveau der Lerner auswirkt. Man kann diese Lernziele auch analog als Schreibgenre auffassen. Beide geben dem Lerner einen mehr oder weniger spezifischen Fokus beim Bearbeiten der Aufgabe vor, der offenbar zu verschiedenen Bearbeitungsqualitäten führt. Die Stichprobe in dem Experiment von Schmal-

hofer und Glavanov (1986) bestand aus 64 Studierenden, die eine Anleitung für ein Programm mit unterschiedlichen Lernzielen lasen. Die Studierenden wurden instruiert, den Text entweder mit dem Ziel zu erarbeiten, eine Zusammenfassung des Textes zu schreiben oder sich auf der Basis der darin enthaltenen Information Kenntnisse über das Programm anzueignen. Im Anschluss daran wurde überprüft, welche Art der Textrepräsentation die Lerner konstruiert hatten. Zu diesem Zweck präsentierten Schmalhofer und Glavanov (1986) den Lernern einen Test mit Sätzen, die vier Kategorien zugeordnet waren: Originalsätze aus dem Text, paraphrasierte Sätze, Sätze mit verändertem Wortlaut, aber korrektem Bedeutungsgehalt (*situational correctness*) und zuletzt Sätze mit verändertem Wortlaut und falschem Bedeutungsgehalt. Die Aufgabe bestand darin zu entscheiden, ob die Sätze aus dem Originaltext stammen. Im Ergebnis zeigte sich, dass die Lerner tatsächlich verschiedene Repräsentationen des Textes konstruiert hatten: „As a consequence of these differential encoding processes, subjects who studied for text summarization remembered more propositional information while subjects with a knowledge acquisition goal remembered more situational information. It was found that the situational encoding and retrieval processes proceeded faster than the respective propositional processes“ (Schmalhofer & Glavanov, 1986, S. 279).

Die Befunde aus diesem Experiment zeigen, dass die Qualität der Verarbeitungsprozesse von Lernern einen substanziellen Einfluss auf die Wissenskonstruktion hat. Voraussetzung für die Wirksamkeit vorgegebener Lernziele ist jedoch, dass der Lerner diese auch als solche adaptiert. Wie sich in der im nächsten Abschnitt vorgestellten Studie zeigt, stellt die Aufgabeninterpretation eine mögliche Problemquelle analytischen Schreibens dar.

Die Interpretation der Aufgabe beeinflusst die Herangehensweise an das Schreiben

Penrose (1992) ist der Frage nachgegangen, wie die Interpretation der Aufgabe die Schreibqualität beeinflusst. Um die Effekte der Aufgabeninterpretation auf die Herangehensweise an das Schreiben zu prüfen, verglich Penrose (1992) 40 Studierende im ersten Studienjahr, die zwei Lernbedingungen zugewiesen wurden. Die Aufgabe der Probanden bestand in beiden Gruppen darin, zwei Texte zu lesen und im Anschluss daran jeweils eine Lernaufgabe zu bearbeiten, wobei die Reihenfolge der Lernaufgaben zwischen den Gruppen variierte. Eine Aufgabe

bestand darin, den Text zu lesen und zu lernen („Lesen und Lernen“). Dabei wurde die Wahl der Strategien den Probanden überlassen. Ein Essay stellte die zweite Aufgabe dar. Alle Personen schrieben somit ein Mal ein Essay und lernten einen Text ohne Schreibaufgabe. Personen mit inhaltsspezifischem Vorwissen wurden im Vorfeld von der Untersuchung ausgeschlossen. Um Informationen über die kognitiven Prozesse zu gewinnen, zeichnete die Forscherin Protokolle Lauten Denkens auf. Zudem führte sie direkt nach der Bearbeitung der Aufgabe einen Verstehenstest mit vier verschiedenen Qualitäten des Verstehens durch (*simple recall, complex recall, macrostructure, application*).

Die Ergebnisse für den Verstehenstest fielen überraschend aus. So schnitten die Probanden ohne Schreibaufgabe sowohl beim Abrufen der verarbeiteten Information als auch bei deren Anwendung in einem neuen Kontext besser ab als die schreibenden Probanden.

Aufschluss über dieses erwartungswidrige Ergebnis verschaffte eine Analyse der Notizen der Studierenden, die diese während der „Lesen und Lernen“-Kondition erstellt hatten. Hierbei zeigten sich deutliche qualitative Unterschiede der Aufzeichnungen, die auf zwei grundsätzlich verschiedene Herangehensweisen an die Lernaufgaben schließen ließen. So hatten einige Lerner aus der Gruppe trotz der Anweisung, nur zu lesen und zu lernen, sehr elaborierte Notizen angefertigt. Dementsprechend fanden sich auch bei der Schreibaufgabe sehr heterogene Textprodukte. Zur Systematisierung dieser Unterschiede kategorisierte Penrose (1992) die Notizen in *generative* oder *nongenerative notes*. Zu den *nongenerative notes* zählten solche, die keine Hinweise für eine tiefgehende Verarbeitung des Lernmaterials enthielten: „Students whose activities were classified under this heading evidenced little selection and no manipulation of the source text material. They either jotted down brief facts and figures from the source text, copied virtually everything from the source text, or took no notes at all“ (Penrose, 1992, S. 475). Der Kategorie „*generative notes*“ hingegen wurden Notizen zugeordnet, die Hinweise auf elaborative Strategien enthielten: „Students whose study strategies were classified as generative wrote idea-based notes, extracting key points and unifying ideas from the text, sometimes adding connections or reordering text material“ (Penrose, 1992, S. 475). Um die Qualität der Notizen aus der „Lesen und Lernen“-Untersuchungsbedingung mit den Texten der Schreibaufgabe zu vergleichen, wurden letztere ebenfalls hinsichtlich ihrer Qualität eingestuft.

Trennscharfe Kategorien sind bei qualitativen Untersuchungen eher schwierig zu bewerkstelligen und da es Penrose (1992) um das Kontrastieren der Herangehensweisen an das Schreiben und Lernen ging, verwendete sie nur zwei Kategorien für die Texte: *paraphrase* (ein eng am Ursprungstext orientiertes Produkt, aus dem keine selbstständige Weiterverarbeitung des Textes sichtbar wird) und *constructive summary* (elaborierte Verarbeitung des Textes mit eigenen Beispielen, eigenständiger Anordnung der Themen und Schlussfolgerungen).

Die Datenanalyse mit den kategorisierten Daten zielte zum einen auf die Effekte der Bearbeitungsqualität sowie auf das Ausmaß und die Qualität kognitiver Strategien und zum anderen auf die Effekte auf verschiedene Verstehensmaße (*simple recall, complex recall, macrostructure, application*). Hinsichtlich der kognitiven Strategien wiesen die *generative notes* bei dem ersten Lerntext auf signifikant mehr Elaborationsstrategien hin (dazu zählen hier Paraphrasieren, Evaluieren und zusätzliche Informationen generieren) als die Texte der anderen drei Gruppen (*nongenerative notes, paraphrase, constructive summary*). Bei dem zweiten Text ließ sich das Ergebnis allerdings nicht replizieren und verkehrte sich sogar ins Gegenteil. Hier zeigten die *nongenerative notes* die meisten Elaborationsstrategien.

In Bezug auf das Verstehen der Lerninhalte zeigten sich folgende Befunde:

1. Im Gesamtvergleich der vier Gruppen schneiden die Probanden, deren Texte der Kategorie *constructive summary* zugeordnet wurden, in allen Maßen am schwächsten ab, obwohl von dieser Lernform die intensivste Auseinandersetzung erwartet wurde.

2. Die *generative notes*-Gruppe erzielt vor allem bei den komplexeren Verstehensmaßen sehr gute Ergebnisse.

3. Die Gruppe der *nongenerative notes* schneidet beim Faktenerinnern am besten von den vier Schreibgruppen (*constructive summary, paraphrase, generative notes, nongenerative notes*) ab.

Die Studie von Penrose (1992) bestätigt zunächst einmal den Befund von Langer und Applebee (1987), wonach analytisches Schreiben nicht die optimale Methode für das Erlernen von Fakten ist (*simple recall*). Hier erlernten Probanden, die nur lasen und Notizen machten, mit geringerem Zeitaufwand mehr. Es zeigte sich allerdings auch, dass die Herangehensweise der Lerner nicht nur durch die Form der Aufgabe bestimmt wurde, sondern auch durch ihre Interpretation

derselben. So machten einige Lerner aus der „Lesen und Lernen“-Gruppe elaborierte Notizen (obwohl sie „nur“ lernen sollten), wohingegen Probanden aus der Schreibgruppe, die eigentlich analytische Essays schreiben sollten, oberflächliche Zusammenfassungen schrieben.

Die Autorin führt am Rande der Studie den Kommentar einer Probandin an, deren Essay sich durch ein hohes Maß an persönlichen Sichtweisen auszeichnete. Aus der Äußerung wird deutlich, dass die Lernerin den Sinn der Aufgabe zwar verstanden, diese aber dennoch auf ihre individuelle Weise umgesetzt hat. Zwar hat diese Ausführung nur anekdotischen Charakter, doch konnte in zahlreichen anderen Studien aus dem Bereich der Lernforschung gezeigt werden, dass die Art, wie Lerner Lernaufgaben bearbeiten (z.B. welche Strategien sie einsetzen), stark von ihrer Haltung zur Aufgabe abhängt (z.B. Biggs, 1988; Entwistle, 1995; Marton & Säljö, 1984).

Insbesondere in den Forschungsarbeiten von Marton und Säljö (1984) werden zwei grundsätzlich verschiedene Herangehensweisen postuliert, die die Art der Strategien, die ein Lerner zur Bewältigung der Aufgabe einsetzt, determinieren. Diese Ansätze bezeichnen die Autoren als oberflächenorientierten *surface approach* und als tiefenorientierten Ansatz *deep level approach*. Sie beschreiben die intentionale Haltung, die der Lerner zum Lernmaterial einnimmt und die beeinflusst, welche Lernstrategien er einsetzt. Der *surface approach* ist charakterisiert durch eine passive Haltung des Lerners gegenüber dem Lernmaterial. Säljö (1979) konnte mithilfe von Interviews zeigen, dass Lerner sich bei diesem Ansatz Wissen als extern in separaten Einheiten existierend vorstellen. Dies hat zur Folge, dass vornehmlich Oberflächenstrategien (z.B. auswendig lernen) eingesetzt werden und keine eigenständige Verarbeitung der Informationen stattfindet. Das Lernen zielt auf die Reproduktion des Lernstoffes ab. Demgegenüber steht der *deep level approach*, bei dem der Lerner Wissen als etwas von ihm Konstruiertes auffasst und das Lernen infolgedessen aktiv gestaltet. Damit in Verbindung stehen Tiefenstrategien des Lernens (z.B. die Vernetzung von Informationen, Herstellen interner Verknüpfungen).

Die Untersuchung von Penrose (1992) verdeutlicht die Zusammenhänge zwischen Aufgabeninterpretation und Strategieeinsatz. Analytisches Schreiben kann demzufolge fehlschlagen, wenn der Lerner die Aufgabe anders interpretiert als vom Lehrenden intendiert, und infolgedessen Strategien einsetzt, die zu einem

anderen als dem bezweckten Lernergebnis führen.

Insgesamt zeigen die bis jetzt beschriebenen Studien, dass die Zusammenhänge zwischen Schreiben und Lernen komplex sind und durch ein Zusammenspiel verschiedener Faktoren, wie der Aufgabenart, der Aufgabeninterpretation und der Qualität der kognitiven Prozesse, bestimmt werden.

Analytisches Schreiben, wie es im Zusammenhang mit der Genre-Hypothese dargestellt wurde, zieht Lernerfolg nach sich, wenn entsprechend anspruchsvolle Strategien dabei zum Einsatz kommen. In den Studien von Langer und Applebee (1987) zeigte sich diesbezüglich, dass Lerner innerhalb unterschiedlicher Genres (wie z.B. Notizenmachen, Essay schreiben, Lernfragen beantworten) verschiedene Arten von Textproduktionsstrategien (z.B. eigene Ideen evaluieren, nach Informationen im Text suchen, eigene Schreibversuche kommentieren) einsetzen und dass die Qualität des Lernens in Abhängigkeit der Strategien hinsichtlich der Menge an Informationen, deren Nachhaltigkeit und der Komplexität des Verstehens variiert. In den Schreibstudien werden in dieser Hinsicht typischerweise das Lernen von Fakten und tiefenorientiertes Lernen miteinander kontrastiert.

Es wurde deutlich, dass das Faktenlernen durch Aufgaben gefördert wird, die keine weitreichende Auseinandersetzung mit dem Stoff erfordern, sondern für deren Bewältigung das Auffinden von Information aus einem Text genügt (z.B. das Beantworten kurzer Lernfragen). Analytisches Schreiben, wie die Schreibform des Essays, eignet sich hingegen eher, um das Verstehen von Zusammenhängen zu fördern. Dies ist darauf zurückzuführen, dass die Lerner beim Schreiben eines Essays anspruchsvollere Verarbeitungsschritte durchführen müssen. So müssen sie die Informationen in eine eigene Struktur bringen und die Lerninhalte hinsichtlich ihrer Relevanz beurteilen. Tynjälä (1998) fasst zusammen: „A general conclusion that may be drawn from studies of different writing tasks is that the more a writing assignment involves active manipulation of the information to be studied, the better the learning outcomes are likely to be“ (Tynjälä, 1998, S. 214).

Die Studie von Penrose (1992) deutet darauf hin, dass die Aufgabeninterpretation sich als Problemquelle analytischen Schreibens erweisen kann, da sie von außen schwierig zu beeinflussen ist. Als Folge einer falschen Aufgabeninterpretation kann es passieren, dass Lerner nicht das angestrebte Lernergebnis (z.B. Zusammenhänge erkennen) erreichen, sondern einen Lernweg beschreiten, der zu einem nicht vorgesehenen Ergebnis führt. Eine Möglichkeit, Lerner mittels kon-

kreter Anweisungen auf den intendierten Lernweg zu bringen, stellen Schreibinstruktionen in Form von Prompts dar. Im nächsten Abschnitt wird zunächst die letzte Hypothese von Klein (1999) dargestellt. Im Anschluss daran werden die für den empirischen Teil dieser Arbeit relevanten Ergebnisse im Hinblick auf die Gestaltung der Lernumgebung zusammengefasst und ausgewertet.

Die vierte Hypothese (*backward search*) von Klein (1999) fasst Theorien und Hypothesen zum Schreiben zusammen, die davon ausgehen, dass der Schreiber Wissen konstruiert, indem er sich rhetorische Ziele setzt, Inhalte zu diesen Zielen generiert und schließlich den Text im Hinblick auf das Schreibziel nochmals überprüft, um den Inhalt anzupassen. Beispiele für diese Hypothese sind die Schreibmodelle von Hayes und Flower (1980) und Bereiter und Scardamalia (1987).

Das Schreibmodell von Hayes und Flower (1980) stellte Schreiben zum ersten Mal als Problemlöseprozess dar und übte großen Einfluss auf die weitere Modellbildung aus, in der Schreiben fortan aus der Problemlöseperspektive betrachtet wurde. Nach Schnotz (2001) besteht das vom Textproduzenten zu lösende Problem darin, „einen Text über einen bestimmten Sachverhalt für bestimmte Adressaten mit dem Ziel zu verfassen, ihnen Wissen über den Sachverhalt zu vermitteln, ihnen eine bestimmte Sichtweise nahezulegen, sie eine bestimmte Argumentation nachvollziehen zu lassen usw.“ (Schnotz, 2001, S. 154). Bezeichnend für diese Konzeption sind zum einen die Art der Prozesse, von denen angenommen wird, dass sie zur Problemlösung notwendig sind (z.B. Zielsetzung und Zielüberwachung), und zum anderen der rekursive und interaktive Charakter des Schreibprozesses, der aus Teilprozessen besteht.

Der Problemlöse-Ansatz ersetzte die bis dahin vorherrschenden Stufenmodelle des Schreibens (z.B. Beaugrande, 1982), die den Schreibprozess als eine feststehende Abfolge von Produktionsschritten auffassen, bei dem „pragmatische, dann semantische, syntaktische und lexikalische Entscheidungen getroffen werden, die schließlich mittels Buchstabenketten graphisch umgesetzt werden“ (Molitor-Lübbert, 1996, S. 1005) und konzipierte das Schreiben als einen rekursiv verlaufenden Prozess, der sich aus sich wechselseitig beeinflussenden Teilprozessen zusammensetzt.

Auf der Basis von Protokollen Lauten Denkens und orientiert an den gängigen Kategorien der Problemlöse-Literatur identifizierten Hayes und Flower (1980) Prozesse wie Zielsetzung, Zielüberwachung und Suchprozesse nach adäquaten

Strategien und Bewertungs- und Analyseschritte zur Überwachung des gesamten Prozesses. Ihr Schreibmodell setzt sich aus den drei Hauptkomponenten Schreibprozess, Aufgabenumgebung und Langzeitgedächtnis zusammen, die miteinander interagieren und in ihrer Gesamtheit den Schreibprozess konstituieren (vgl. Abbildung). Die Aufgabenumgebung (*task environment*) besteht aus der Schreibaufgabe, die das Thema und die Adressaten des Textes vorgibt, und dem bereits produzierten Text, der immer wieder überprüft und überarbeitet wird und insofern als Ziel und Referenz dient. Die zweite Hauptkomponente ist das *Langzeitgedächtnis*, aus dem der Textproduzent inhaltliches Wissen über das Textthema, die Zielgruppe des Textes und Textschemata abrufen. Der *Schreibprozess* ist als dritte Hauptkomponente eingebettet in die Aufgabenumgebung und das Langzeitgedächtnis. Er gliedert sich in die Subprozesse *Planen*, *Übertragen in schriftliche Form* und *Prüfen*, die ihrerseits aus verschiedenen Komponenten bestehen. Während des Planungsprozesses werden aus dem Langzeitgedächtnis relevante Informationen abgerufen, die organisiert und strukturiert werden. Während der Planungsphase findet zudem die Zielsetzung statt, die für die Regulation des gesamten Prozesses entscheidend ist. Diese Unterkomponenten bilden den Schreibplan, nach dem der Textproduzent vorgeht. Eine metakognitive Kontrollinstanz (*Monitor*) überwacht den Planungsprozess anhand der Aufgabenorientierung. Zum Schreibprozess gehören außerdem der Übersetzungsprozess, während dem das Wissen in geschriebene Sprache übersetzt wird, und verschiedene Überarbeitungsprozesse.

Das Modell von Hayes und Flower (1980) konzipiert Schreiben als zielgerichteten Prozess und identifiziert verschiedene Teilprozesse, deren Abfolge rekursiv ausgelegt wird. Diese Auffassung setzte sich in weiteren Schreibmodellen durch und ist insofern als Meilenstein für die Textproduktionsforschung zu verstehen.

Als Kritikpunkt wird häufig die Ausrichtung des Modells an einem kompetenten Schreiber und das Fehlen von Hinweisen auf Entwicklungsbedingungen von Novizen genannt (Eigler, 1990). Schnotz (2001) gibt zu bedenken, dass es unrealistisch ist anzunehmen, dass Schreiben isoliert von externen Wissensquellen stattfindet und der Textproduzent ausschließlich auf Wissensinhalte aus dem eigenen Gedächtnis zurückgreift: „Der Prozess der Textproduktion erfolgt im allgemeinen vielmehr durch ein Zusammenspiel von internen Informationen, die aus dem Gedächtnis abgerufen werden, und externen Informationen, die in Form von Dokumenten vorliegen“ (Schnotz, 2001, S. 155). Darüber hinaus zeigen Bereiter und

Scardamalia (1987), dass Schreibprozesse qualitativen Unterschieden unterliegen und nicht immer wie in der in Hayes und Flowers (1980) Modell abgebildeten „Idealform“ ablaufen. Der Textproduzent kann, in Abhängigkeit seiner Schreibfertigkeiten und Ziele, Teilkomponenten des Schreibprozesses überspringen oder routinemäßig durchführen, wie Bereiter und Scardamalias (1987) Modell zeigt.

Insgesamt ist das Modell jedoch noch immer von großem heuristischen Wert für nachfolgende Untersuchungen, weil es die verschiedenen Komponenten des Schreibprozesses miteinander in Verbindung setzt und die Bedeutung verschiedener Wissensgebiete für das Schreiben herausarbeitet. Eigler (1990) hält daher fest: „Wissen ist eine Bedingung von Textproduzieren“ (S. 10).

In Erweiterung dazu differenziert das ebenfalls kognitiv fundierte Modell von Bereiter und Scardamalia (1987) zwischen qualitativen Unterschieden der Schreibstrategien erfahrener und weniger erfahrener Textproduzenten. Im Zentrum der beiden Schreibstrategien stehen die während des Schreibens ablaufenden mentalen Prozesse. Erfahrene Schreiber, so nehmen die Autoren an, bewegen sich zwischen zwei Problemräumen: Einem inhaltlichen, in dem die Frage „Was will ich sagen?“ im Vordergrund steht, und einem rhetorischen, in dem der Schreiber das Problem löst, *wie* er das Intendierte sagen will. Diese dialektische Bewegung aus dem Abgleich zwischen rhetorischen Zielen und den zu explizierenden Inhalten ermöglicht nach Bereiter und Scardamalia (1987) Lernen. Wissens Elemente werden dabei reorganisiert oder transformiert zu höherrangigem Wissen und der Schreiber erlangt ein tieferes Verständnis des Stoffes. Wenn es einer Person gelingt, auf diese Weise zu schreiben, sprechen die Autoren von *knowledge transforming*. Im Kontrast dazu nennen sie das Schreiben von Novizen *knowledge telling*. Dabei fehlt im Unterschied zum *knowledge transforming* die Bewegung zwischen den beiden Problemräumen. Man könnte auch sagen, es fehlt das reflexive Element beim Schreiben. Der Textproduzent ruft bei dieser Schreibstrategie vielmehr bereits vorhandenes Themenwissen aus dem Gedächtnis ab und präsentiert dieses zwar in einer kohärenten Weise, orientiert sich dabei jedoch nicht an dem zu adressierenden Publikum oder an rhetorischen Zielen, wie z.B. Klarheit und Logik, sondern schreibt nieder, was er zu einem Thema weiß. Dass dennoch ein stimmiger Text entsteht, wird nach Bereiter und Scardamalia (1987) durch die Orientierung an Textschemata möglich, auf die die Schreibaufgabe Hinweise gibt. Schreiben erfordert auf diese Weise keinen größeren Planungsaufwand,

was allerdings auch zur Folge hat, dass die epistemische Funktion des Schreibens ausbleibt.

Bereiter und Scardamalia (1987) entwickeln in ihrem Modell zwei grundsätzlich verschiedene Herangehensweisen an das Schreiben, die sie als zwei Extreme des Entwicklungsstandes der Schreibfertigkeiten verstehen. Auswirkungen auf den Wissenserwerb hat jedoch nur die *knowledge transforming*-Strategie, bei der der Textproduzent im Wechsel von wissensbezogenen und textbezogenen Überlegungen seinen Text verbessert und mithilfe komplexer Verarbeitungsprozesse einen gut organisierten Text herstellt, mit dem er sein Wissen erweitert (Eigler, 1990). Im Unterschied zu Flower und Hayes (1981) gehen Bereiter und Scardamalia (1987) davon aus, dass Textproduzenten beim Übergang von der *knowledge telling*-Schreibstrategie zur *knowledge transforming*-Schreibstrategie unterstützt werden können. Ansätze dazu sehen sie in Koordinationshilfen bei den verschiedenen Teilprozessen, die ein Schreiber bewältigen muss: „Diese Unterstützung kann z.B. die Konkretisierung von Zielen betreffen oder in der Empfehlung geeigneter Mittel einer sogenannten prozeduralen Erleichterung bestehen. Außerdem gilt es, Lernende dafür zu sensibilisieren, wann welche Teilaktivitäten des Schreibprozesses zu vollziehen sind, da die Qualität von Texten weniger vom globalen Ausmaß verschiedener Teilaktivitäten, sondern vor allem vom Zeitpunkt abhängt, zu dem diese Aktivitäten durchgeführt werden“ (Schnotz, 2001, S. 156).

In der aktuellen Forschung, die sich mit Schreiben als Lerninstrument befasst, zeichnen sich vielfach Bezüge zu dem Schreibmodell von Bereiter und Scardamalia (1987) ab. So werden im Zusammenhang mit Lerntagebüchern oft prozedurale Hilfestellungen (z.B. metakognitive Fragen zur Unterstützung der Zielüberwachung oder des Verstehens) zur Förderung des Schreibens eingesetzt. Studien aus diesem Bereich sind Gegenstand von Abschnitt 2.3.

2.2.1 Zusammenfassung und Ausblick

Klein (1999) klassifiziert in seinem Forschungsüberblick die theoretische und empirische Literatur zum Schreiben als Lerninstrument bis 1999 in vier Kategorien, die er als Hypothesen versteht. Unter jeder der vier Hypothesen werden jeweils ähnliche Annahmen oder Theorien zusammengefasst, die der Frage „How do writers select operations that allow them to learn during writing?“ nachgehen (Klein, 1999, S. 211). Im Zentrum der zusammengefassten Studien stehen die kognitiven

Prozesse, die beim Schreiben eine Rolle spielen. Kleins Studie beinhaltet ausschließlich Literatur, die sich theoretisch fundiert mit diesen Prozessen befasst.

Klein (1999) versteht die vier vorgestellten Hypothesen als miteinander kompatibel. Sein Ziel ist es, einen Überblick über die Forschungslandschaft zu geben. Was die Validität der einzelnen Hypothesen angeht, resümiert er, dass bislang nur die Genre-Hypothese genügend empirische Belege für Bedingungen liefert, unter denen Lerner vom Schreiben profitieren können. Unter der Genre-Hypothese werden Studien subsumiert, die von Zusammenhängen zwischen dem Schreibgenre und den kognitiven Prozessen des Textproduzenten ausgehen. Dabei wird angenommen, dass durch vorgegebene Textstrukturen (z.B. Essay) bestimmte kognitive Strategien angeregt werden, die ihrerseits zu unterschiedlichen Verarbeitungstiefen und Lernqualitäten (z.B. Faktenlernen, Zusammenhänge verstehen) führen. Stark vereinfacht kann man aus den Studien die Erkenntnis ableiten, dass das Schreibgenre sich dann als lernwirksam herausstellt, wenn es als Strategieaktivator wirkt und der Schreibprozess durch anspruchsvolle kognitive Prozesse begleitet wird. Als Problemquelle analytischen Schreibens kann sich die Aufgabeninterpretation seitens der Lerner erweisen (Penrose, 1992).

Aus den von Klein (1999) angeführten Studien lassen sich nur mit Einschränkungen didaktische Schlussfolgerungen ableiten, denn die Studie weist Beschränkungen im Hinblick auf Lernervariablen auf. So bleiben z.B. die Variablen *Motivation* und *Interesse*, die sich in der Lernpsychologie als einflussreiche Lernervariablen erwiesen haben, gänzlich unberücksichtigt. Auch wenn diese Variablen schwierig von außen zu beeinflussen sind, sollten sie als Einflussfaktoren bei der Gestaltung einer Lernumgebung zumindest bedacht werden. Studien aus dem Bereich der Lernmotivation und aus der Forschung zu Schreibmotivation zeigen außerdem, dass zumindest das Aufgabeninteresse (hier also das Interesse an der Schreibinstruktion) dahingehend beeinflusst werden kann, dass Lerner Aufgaben als herausfordernd und interessant wahrnehmen. Eine weitere Grenze der Studien besteht darin, dass das Bedingungsgeflecht, das der Wirksamkeit von Schreibinstruktionen zugrunde liegt, nicht ausreichend differenziert untersucht wird (z.B. die Art der Schreibinstruktion und der kognitiven Prozesse).

Als Fazit des Forschungsüberblicks lassen sich abschließend folgende zwei Punkte festhalten:

1. Wenn durch das Schreiben von Texten anspruchsvolle kognitive Strategien

angeregt werden, entwickelt sich die lernförderliche Wirkung des Schreibens.

2. Die Schreibfertigkeit von Personen variiert. Um die Entwicklung zu einer *knowledge transforming*-Strategie zu fördern, ist es nach Bereiter und Scardamalia (1987) notwendig, die Teilprozesse des Schreibens bei ungeübten Textproduzenten zu unterstützen.

Um die Zusammenhänge zwischen Schreiben und Lernen in ihrer Komplexität besser zu erfassen und den Forschungsbedarf herauszuarbeiten, werden im folgenden Kapitel Studien mit Lerntagebüchern vorgestellt. Diese sind aktueller als die bisherigen Studien und nehmen die didaktischen Bedingungen von Schreibinstruktionen differenzierter in den Blick. So wird z.B. untersucht, inwiefern der Zeitpunkt der Schreibinstruktionen innerhalb des Lernprozesses eine Rolle spielt (Thillmann, Künsting, Wirth & Leutner, 2009), oder welche Strategien besonders wirksam sind (Berthold, Nückles & Renkl, 2007).

2.3 Studien mit Lerntagebüchern und Lernprotokollen

Im vorigen Abschnitt wurden Theorien des Lernens durch Schreiben besprochen und empirische Studien referiert, die die Bedeutung kognitiver Prozesse in den Fokus nehmen. Dabei zeigte sich, dass die lernförderliche Funktion des Schreibens in der Aktivierung kognitiver und metakognitiver Strategien liegt. Während die ersten Studien noch überprüften, welche Strategien bei verschiedenen Schreibformen zur Anwendung kommen, gehen aktuellere Studien (ab 2000) in Zusammenhang mit Lerntagebüchern und Lernprotokollen detailliert der Frage nach, durch welche Art der Instruktion sich spezifische Strategien anregen lassen, welche Strategien und welche Kombination von Strategien lernförderlich sind (vgl. Berthold et al., 2007; Nückles, Schwonke, Berthold & Renkl, 2004; Nückles, Hübner & Renkl, 2006, 2009) und welcher Zeitpunkt zur Förderung der Strategien ideal ist (Wirth, 2009). Unter der Methode des Lerntagebuchs wird eine Vielzahl von Ansätzen zusammengefasst, deren Gemeinsamkeit darin besteht, dass Lernende sich schriftlich mit dem eigenen Lernen auseinandersetzen und zwar einerseits in Form einer Dokumentation des Erlernten und andererseits evaluierend, indem das Verstehen des Lernstoffes überprüft wird (vgl. Gläser-Zikuda, 2010; Rambow & Nückles, 2002). Das Lerntagebuch wird in der Regel zur Nachbereitung

von Unterricht eingesetzt, wobei das Schreiben als Lerninstrument dient (Hübner, Nückles & Renkl, 2007). Ein ähnliches Instrument der Nachbereitung stellt ein Lernprotokoll dar. Dieses unterscheidet sich lediglich im Umfang von einem Lerntagebuch und bezieht sich auf jeweils eine Lernsituation.

Lerntagebücher und Lernprotokolle können vorstrukturiert sein (z.B. durch Instruktionen) oder nur aus einer einmaligen Instruktion bestehen. Es zeigte sich jedoch in der Studie von Nückles et al. (2004), dass Lerntagebücher ohne explizite Aufforderungen zu spezifischen kognitiven und metakognitiven Aktivitäten zu einer eher oberflächlichen Auseinandersetzung mit den Lerninhalten und mit dem eigenen Lernen führen. Nückles et al. (2004) stellten bei der Analyse von insgesamt 192 Lerntagebucheinträgen, die Studierende mithilfe einer informellen Anleitung verfasst hatten, fest, dass die Einträge [...] „stark herkömmlichen Zusammenfassungen ähnelten und im Sinne eines knowledge telling Lerninhalte lediglich reproduzierend wiedergaben. Elaborative Strategien wie das Generieren eigener Beispiele und das kritische Prüfen von Argumenten waren vergleichsweise schwach ausgeprägt. Metakognitive Strategien im Sinne einer Überwachung des eigenen Verstehens wurden praktisch gar nicht angewandt“ (Hübner et al., 2007, S. 7-8).

Auch innerhalb von Lerntagebüchern haben sich strukturierende, Tiefenorientierung anregende Instruktionen demnach als hilfreich erwiesen, da sie die Aufmerksamkeit der Lerner auf lernrelevante Aspekte, wie z.B. das Überwachen des Inhaltsverständnisses, lenken und die Anwendung von Strategien anregen können (McCrinkle & Christensen, 1995).

Das Ziel dieses Abschnitts ist es nun, weiteren Aufschluss über das Bedingungsgefüge, das Schreiben als Lerninstrument auszeichnet, zu gewinnen. Dazu werden Studien mit Lerntagebüchern referiert, da diese sich insbesondere mit der Frage auseinandersetzen, wie durch gezielte Schreibinstruktionen solche Lernprozesse unterstützt werden können, die tiefgreifendes und nachhaltiges Lernen nach sich ziehen. Im Fokus pädagogischer und psychologischer Forschung stehen derzeit vor allem Lernformen, bei denen der Lerner Verantwortung für sein Lernen übernimmt und die wesentlichen Entscheidungen im Lernprozess selbst steuert. Diese werden unter dem Stichwort des „selbstregulierten Lernens“ zusammengefasst (vgl. Spörer, 2004). Damit Lerner ihren Lernprozess selbst steuern können, müssen sie Kenntnisse über Lernstrategien und deren Anwendung besitzen, denn

diese sind für die Steuerung des Lernprozesses entscheidend (Hasselhorn & Gold, 2009).

Das Vorhandensein von Strategie-Kenntnissen alleine ist jedoch keine Voraussetzung dafür, dass Lerner diese auch anwenden (vgl. Boekaerts, 1997; Hasselhorn & Gold, 2009; Lompscher, 1995). Dies kann damit zusammenhängen, dass die Anwendung anspruchsvoller Strategien nicht immer notwendig ist, wie etwa, wenn der Lerninhalt nicht besonders schwierig oder bereits in Sinneinheiten eingeteilt ist (wie z.B. in manchen Lehrbüchern). Ein weiterer Grund kann darin bestehen, dass Lerner die Strategien nicht für nützlich halten. Wenn Strategiewissen vorhanden ist und ein Lerner Strategien von sich aus nicht anwendet, spricht man von einem Produktionsdefizit (vgl. Hasselhorn & Gold, 2009). Lerntagebücher stellen einen Versuch dar, dieses Defizit zu überwinden und die einzelnen Teilprozesse des Lernens zu unterstützen.

Betrachtet man die Literatur zu Lerntagebüchern, fällt zum einen auf, dass vergleichsweise wenige empirische Studien zur Arbeit mit Lerntagebüchern mit Erwachsenen vorliegen. Zwar gibt es vereinzelte Berichte, doch die Zahl der Studien ist recht überschaubar. Zum anderen wird innerhalb der Studien mit höchst unterschiedlichen Konzepten gearbeitet, innerhalb derer Schreiben entweder zentraler Bestandteil des Lernprozesses ist oder gar keinen Stellenwert hat. Letzteres ist der Fall bei sogenannten standardisierten Lerntagebüchern, die den Lernprozess, auf unterschiedlichen Prozessmodellen des Selbstregulierten Lernens basierend, begleiten (z.B. Löb, Perels & Schmitz, 2004; Pickl, 2004; Schmitz, 2001). Standardisierte Lerntagebücher bestehen zumeist aus Checklisten, in denen Lernende ihr Lernverhalten beobachten (*self-monitoring*). Zum Bearbeiten dieser Lerntagebücher ist kein Fließtext erforderlich und Schreiben wird hierbei nicht als Lerninstrument verwendet. Im Mittelpunkt steht vielmehr die Beobachtung des eigenen Lernens anhand vorgegebener Kriterien, manchmal durch Abarbeiten von Fragestellungen zur Motivation oder Emotionen beim Lernen, wie bspw. im Experiment von Schmitz (2001). Weitere Unterscheidungsmerkmale der vorliegenden Studien sind der Einsatzbereich (z.B. E-Learning) und der Gegenstand der Förderung (z.B. einzelne Lernstrategien). Gläser-Zikuda, Rohde und Schlomske (2010) unterteilen die aus dem Bereich der Lernförderung vorliegenden Studien in folgende Bereiche:

- Studien zur Förderung selbstregulierten Lernens

- Studien zur Förderung einzelner Lernstrategien
- Studien zur Steigerung des Lernerfolgs
- Spezielle Einsatzbereiche
 - Lehrerbildung
 - E-Learning
 - Lernprozess Diagnose-Instrumente

Für die vorliegende Arbeit sind nur solche Studien relevant, in denen Schreiben als Lerninstrument eingesetzt wird. Aus diesem Grund wird nicht näher auf standardisierte Lerntagebücher eingegangen. Im Fokus stehen insbesondere die Bedingungen, die Schreiben als Lerninstrument wirksam machen.

In einer der ersten empirischen Untersuchungen zu Lerntagebüchern gingen McCrindle und Christensen (1995) der Frage nach, wie sich das Führen eines Lerntagebuchs auf die Lernkonzeption von Studierenden, die Anwendung von Lernstrategien und den Studienerfolg auswirkt. Im Rahmen einer Feldstudie ließen sie innerhalb eines Seminars zwei Gruppen von jeweils 20 Biologie-Studierenden über einen Zeitraum von fünf Wochen Lerntagebuch führen oder Berichte über die Laborsitzungen schreiben.

Die Studierenden der Experimentalgruppe wurden in die Arbeit mit dem Lerntagebuch eingewiesen und erhielten Informationen zur Gestaltung des Inhaltes des Lerntagebuchs („Was habe ich gelernt?“), zu der erwünschten Reflexion („Warum habe ich es gelernt?“) und der Erläuterung der Strategien, mit deren Hilfe die Studierenden den Inhalt erarbeitet hatten. Die Autoren forderten die Studierenden explizit dazu auf, Verbindungen zwischen den erlernten Wissensinhalten und den dazu angewendeten Strategien herzustellen. Die Lerntagebücher wurden vor, während und nach jeder Sitzung geschrieben. Die Dozenten gaben den Studierenden Rückmeldung über die Angemessenheit der Aufzeichnungen und stellten Reflexionsfragen, wenn dies notwendig erschien: „Appropriate work was acknowledged, while questions to prompt reflection were provided where students seemed to be experiencing difficulty (e.g., why did drawing a diagram help here?“ (McCrindle & Christensen, 1995, S. 175). Feedback wurde jedoch nicht als systematische Variable in der Untersuchung berücksichtigt.

Die Teilnehmer der Kontrollgruppe schrieben über denselben Zeitraum Berichte, die vom Umfang her ähnlich ausfielen wie die Lerntagebücher. Auch diese Gruppe wurde instruiert, die Berichte zu schreiben. Im Kontrast zu der Experimentalgruppe wurden die Probanden der Kontrollgruppe jedoch nicht nach Lernstrategien gefragt, sondern sie sollten in den Berichten die Lerninhalte zusammenfassen. Die Studierenden dieser Gruppe erhielten ebenfalls korrigierende Rückmeldungen, wenn es notwendig war.

Nach der Intervention ließen McCrindle und Christensen (1995) die Probanden beider Gruppen eine Lernstrategieaufgabe bearbeiten, bei der es darum ging, sich eine Textpassage anzueignen. Die Studierenden durften sich dazu Notizen machen. Im Anschluss daran wurde ein Test mit offenen Fragen durchgeführt, bei dem die Lernenden die Textinhalte wiedergeben sollten. Um zu ermitteln, wie die Studierenden bei der Bearbeitung der Lernstrategieaufgabe vorgegangen waren, ließen McCrindle und Christensen (1995) sie im Anschluss an die Aufgabe einen Lernstrategiefragebogen ausfüllen, in dem die Probanden auf einer Skala von eins bis fünf beurteilten, welche kognitiven und metakognitiven Strategien beim Erlernen der Textpassage für sie wichtig waren. Ergänzend dazu führten die Autoren ein Interview mit den Probanden durch, in dem diese nach ihrer Lernkonzeption befragt wurden („What does the term *learning* mean to you?“). Im Interview wurden die Lerner außerdem aufgefordert, ihre Vorgehensweise beim Lernen möglichst detailliert zu beschreiben. Auf diese Weise sammelten die Autoren Informationen über das metakognitive Wissen (*metacognitive awareness*) der Lerner.

Ein Multiple-Choice-Test am Ende des Kurses wurde als Kriterium für den Lernerfolg verwendet. Die Untersuchung kam zu folgenden Ergebnissen:

1. Aus den mit dem Interview erhobenen Daten zeigte sich, dass ein signifikant höherer Anteil von Probanden in der Experimentalgruppe anspruchsvolle Lernkonzepte aufwies als Probanden in der Kontrollgruppe. Eine Person mit einem hochentwickelten Lernkonzept versteht unter Lernen „etwas verstehen“, „etwas interpretieren“ im Kontrast zu „auswendig lernen“ oder „Zunahme an Wissen“.
2. Aus dem auf die Lernstrategieaufgabe folgenden Lernstrategiefragebogen und dem ergänzenden Interview ging hervor, dass die Lernenden der Expe-

rimentalgruppe nach eigenen Angaben mehr metakognitive Strategien einsetzen als die Probanden der Kontrollgruppe. Die Untersuchung mittels χ^2 -Test zeigte einen signifikanten Zusammenhang zwischen der Gruppenzugehörigkeit und der Anwendung metakognitiver Strategien. Die Teilnehmer aus der Lerntagebuchgruppe gaben außerdem häufiger als die anderen Probanden an, Elaborationsstrategien eingesetzt zu haben und wendeten nach eigenen Aussagen weniger Wiederholungsstrategien an als die Lerner, die die Berichte verfasst hatten.

3. Die Analyse der mit der Lernstrategieaufgabe produzierten Texte zeigte, dass die Lerntagebuchgruppe komplexere Wissensstrukturen aufgebaut hatte: „The learning journal group, when asked to learn from text, showed a pattern of knowledge structuring which took account of a variety of features embedded in the text as well as a tendency to build an integrated and related understanding of the text. The scientific report group tended to focus appropriately on one or more relevant features of the text but did not build complex, related and meaningful networks of knowledge“ (McCrinkle & Christensen, 1995, S. 183).
4. Beim Multiple-Choice-Test am Ende der Intervention schnitt die Lerntagebuchgruppe signifikant besser ab.

McCrinkle und Christensen (1995) stellen eine Methode zur Förderung eines effektiven Lernprozesses vor und demonstrieren, dass mit Lerntagebüchern anspruchsvolle kognitive Strategien (z.B. Elaborationsstrategien) angeregt werden können. Die Lerner entwickeln durch das Schreiben eines Lerntagebuches ein komplexeres Verständnis von Lernen und die Autoren stellen die Vermutung an, dass sich dieses Verständnis positiv auf den Lernerfolg auswirkt, weil die Lerner anspruchsvolle (z.B. tiefenorientierte) Strategien einsetzen.

In der Untersuchung kommen außerdem zwei Formen des Feedbacks zum Einsatz, die McCrinkle und Christensen (1995) jedoch nicht als systematische Variable berücksichtigen. Dies ist wahrscheinlich der Grund dafür, dass sie nicht konkreter auf die Inhalte und Bedingungen des Feedbacks eingehen, sondern diesbezüglich vage bleiben: „Appropriate work was acknowledged, while questions to prompt reflection were provided where students seemed to be experiencing difficulty (e.g., why did drawing a diagram help here?)“ (McCrinkle & Christensen,

1995, S. 175).

Ausgehend von Studien zu Feedback (vgl. Bangert-Drowns, Kulik, Kulik & Morgan, 1991; Butler & Winne, 1995) ist anzunehmen, dass diese Form der Rückmeldung durchaus einen Einfluss auf den Lernprozess hat. Dieser könnte bspw. auf der motivationalen Ebene festzustellen sein und sich dadurch im Lernerfolg niederschlagen. In weiteren Studien zu Lerntagebüchern oder Lernprotokollen wird Feedback ebenfalls entweder nicht verwendet oder nicht systematisch untersucht. Hier ist ein Forschungsbedarf zu konstatieren, auf den konkreter am Ende des Kapitels eingegangen wird.

Während McCrindle und Christensen (1995) mit einer einmaligen Instruktion zur Förderung von Lernstrategien arbeiten, wird in anderen Studien versucht, Lernstrategien kontrollierter und gezielter anzuregen. Dazu werden sogenannte Prompts, das sind kurze Fragen oder Hinweise, um Lernstrategien anzuregen, verwendet. Verschiedene Studien konnten positive Zusammenhänge zwischen Prompts und Strategieverbrauch zeigen. Diese Studien stehen im Mittelpunkt des nächsten Abschnitts.

2.3.1 Prompts zum gezielten Aktivieren von Lernstrategien beim Schreiben

In neueren Studien mit Lerntagebüchern und Lernprotokollen wird versucht, Lerner durch Prompts innerhalb eines Lernzyklus gezielt dazu anzuregen, verschiedene Kontroll- oder Steuerungshandlungen durchzuführen. Prompts sind kurze Fragen oder Hinweise mit Aufforderungscharakter, mit denen sowohl metakognitive als auch kognitive Strategien evoziert werden und Wissen aktiviert wird. Beim Einsatz solcher Lernhilfen wird vorausgesetzt, dass die Lerner grundsätzlich über das zu aktivierende Wissen verfügen und die zu promptenden Strategien beherrschen, diese jedoch nicht spontan einsetzen (Bannert, 2009).

In Zusammenhang mit Prompts wurde aus unterschiedlichen Perspektiven untersucht, wie deren Ausgestaltung und die Art der Einbindung in den Lernprozess sich auf Aspekte des Arbeitsgedächtnisses auswirkt. So untersuchen bspw. Thillmann et al. (2009) die Rolle des Darbietungszeitpunktes, während sich Glogger, Holzäpfel, Schwonke, Nückles und Renkl (2009) mit Fragen der Spezifität der Prompts befassen. Die Forschergruppe um Nückles, Renkl, Berthold und Hübner geht hingegen in mehreren Untersuchungen der Frage nach, welche Lernstrategien mit Prompts unterstützt werden sollten.

Im folgenden Abschnitt werden die bisherigen Befunde zu Prompts dargestellt mit dem Ziel, diese in die Gestaltung der Lernumgebung zu integrieren und den Forschungsbedarf herauszuarbeiten.

In einem Experiment untersuchten Berthold et al. (2007) im Zusammenhang mit Lernprotokollen verschiedene Kombinationen von Prompts hinsichtlich ihrer Effekte auf den Lernerfolg und den Strategiegebrauch. Dazu wurden 84 Studierende der Psychologie vier Untersuchungsbedingungen zugewiesen. Nach einer Video-Vorlesung schrieben drei Gruppen ein Lernprotokoll mit einer der drei Bedingungen:

1. 6 kognitive Prompts, die Organisations- und Elaborationsstrategien evozieren sollten
2. 6 metakognitive Prompts zur Überwachung des Verstehens
3. 3 kognitive und 3 metakognitive Prompts

Eine vierte Gruppe diente als Kontrollgruppe und schrieb ein Lernprotokoll ohne Prompts. Die Lernprotokolle wurden anschließend auf einer Skala von 1-6 (1= keine Strategien enthalten, 6= sehr gut strukturiertes Protokoll) bezüglich der Anzahl der kognitiven und metakognitiven Strategien eingestuft. Zur Überprüfung des Lernerfolgs wurden direkt im Anschluss an das Experiment ein Verstehenstest und eine Woche später ein Behaltenstest durchgeführt.

Anwendung kognitiver Strategien

Hinsichtlich der Strategieanwendung zeigte die Analyse der Lernprotokolle, dass die Gruppe mit den kognitiven Prompts und die Gruppe mit den gemischten Prompts signifikant mehr kognitive Lernstrategien anwendeten als die beiden anderen Gruppen. Die Lerner, die die metakognitiven Prompts bearbeitet hatten, unterschieden sich in dieser Hinsicht nicht von der Kontrollgruppe.

Anwendung metakognitiver Strategien

Metakognitive Strategien wurden am häufigsten von den Lernern, die mit metakognitiven und kognitiven Prompts (gemischte Untersuchungsbedingung) gearbeitet hatten und denjenigen, die ausschließlich metakognitive Prompts erhalten hatten, angewendet. Die Lernprotokolle dieser Gruppe zeigten signifikant mehr

metakognitive Strategien als die der Kontrollgruppe und die der Lerner mit nur kognitiven Prompts. Es zeigte sich hier auch, dass sogar nur kognitive Prompts zu einer signifikant häufigeren Anwendung metakognitiver Strategien führten als gar keine Prompts.

Lernerfolg

In Bezug auf den Lernerfolg schnitten die Lerner in den beiden Gruppen, die mit kognitiven Prompts gearbeitet hatten, am besten ab (nur kognitive Prompts, metakognitive und kognitive Prompts). Metakognitive Prompts alleine führten zwar zu einer häufigeren Anwendung metakognitiver Strategien im Lernprotokoll, schlugen sich jedoch im Vergleich mit der Kontrollgruppe nicht in einem besseren Lernergebnis nieder.

Da die Ergebnisse darauf hindeuten, dass die kognitiven Prompts am lernförderlichsten sind, überprüften Berthold et al. (2007) mittels Mediatoranalyse, ob die Unterschiede im Lernergebnis über kognitive Strategien vermittelt werden. Dazu aggregierten sie aus den bestehenden Gruppen zwei neue Gruppen (die Kontrollgruppe und die Gruppe mit den metakognitiven Prompts und die Gruppe mit kognitiven Prompts und die Gruppe mit den gemischten Prompts) und kontrastierten diese miteinander. Die Mediatoranalyse bestätigte die Hypothese, der zufolge die kognitiven Strategien für den Lernerfolg verantwortlich waren: „Thus, cognitive learning strategies in the learning protocols mediated the learning outcomes“ (Berthold et al., 2007, S. 572).

Eine Begründung dafür, dass die metakognitiven Prompts und eine Steigerung der metakognitiven Strategie nicht zum Lernerfolg beitrugen, sehen Berthold et al. (2007) in der fehlenden Möglichkeit, Verstehensschwierigkeiten nachzuregulieren. In einem Folgeexperiment gleichen Aufbaus überprüften Hübner, Nückles und Renkl (2006) diese Hypothese. Dazu bildeten sie, zusätzlich zu den bestehenden vier Gruppen, eine weitere Untersuchungsgruppe, die zwei kognitive Prompts und zwei metakognitive Prompts zum Schreiben erhielt und nach dem Verfassen des Lernprotokolls mit zwei zusätzlichen metakognitiven Prompts dazu aufgefordert wurde, als schwierig wahrgenommene Textstellen noch einmal nachzulesen und Verständnislücken zu bearbeiten. Dadurch wurden wesentliche Teilprozesse des Lernens, wie die Verarbeitung des Lernstoffes (kognitive Prompts zum Anregen von Elaboration- und Organisationsstrategien), die metakognitive Überwachung

des Verstehens und die Nachbearbeitung schwieriger Inhalte, unterstützt. Im Vergleich mit den anderen vier Gruppen (1. nur kognitive Prompts, 2. nur metakognitive Prompts, 3. metakognitive und kognitive Prompts, 4. Kontrollgruppe) erzielte die neue Versuchsgruppe das beste Lernergebnis. Aus den beiden Experimenten lässt sich resümieren, dass Lerner am meisten von Prompts profitieren, wenn alle Teilprozesse des Lernens unterstützt werden.

Es stellt sich allerdings die Frage, inwiefern eine derart starke vorgegebene Strukturierung des Lernprozesses bei erwachsenen Lernenden sich für längere Zeiträume eignet. Dieser Frage gingen Nückles et al. (2006) in einer Feldstudie nach, in der eine Gruppe über ein Semester ein Lerntagebuch mit der wirkungsvollsten Kombination von Prompts führte und eine Kontrollgruppe lediglich mit einer einmaligen Anleitung ohne Prompts Lerntagebuch schrieb. Nach der ersten Hälfte des Semesters führten die Forscher einen Wissenstest durch, bei dem die Experimentalgruppe besser abschnitt als die Kontrollgruppe. Eine Inhaltsanalyse zeigte, dass die Lerntagebücher der Experimentalgruppe erwartungskonform auch mehr Strategien aufwiesen. Zusätzlich erhoben Nückles et al. (2006) die Motivation der Untersuchungsteilnehmer und stellten hierbei fest, dass diese wiederum bei den Studierenden mit Prompts höher ausfiel.

Ein bemerkenswertes Ergebnis erbrachte jedoch die zweite Erhebung am Ende des Seminars. Hier schnitt nun die Kontrollgruppe in allen Maßen besser ab als die Experimentalgruppe. Die Forscher schließen aus diesem Befund, dass die Prompts über den Zeitraum eines Semesters eine zu starke Beeinflussung des Lernprozesses von außen darstellten, die die Motivation der Studierenden beeinträchtigte und möglicherweise zu einem Expertise-Umkehr-Effekt (*expertise reversal effect*) (vgl. Kalyuga, 2007) führte. Dieses Konzept beschreibt den Zusammenhang zwischen der Wirksamkeit didaktischer Unterstützungsangebote und dem individuellen Wissensstand von Lernern. In verschiedenen Studien konnte nachgewiesen werden, dass Lerner mit geringem Vorwissen von Unterstützungsangeboten (z.B. strukturierende Erläuterungen zu einem Lerntext) profitieren, wohingegen diese bei Lernenden mit hohem Vorwissen entweder nur einen geringen Einfluss auf deren Lernleistung haben oder diese sogar beeinträchtigen. Kalyuga, Chandler und Sweller (1998) und Yeung, Jin und Sweller (1998) begründen diesen Effekt damit, dass die Verarbeitung zusätzlicher, unterstützend intendierter Informationen das Arbeitsgedächtnis erfahrener Lerner („Experten“) belastet, da sie für diese red-

undant sind. Für Anfänger hingegen stellt dieselbe Information eine Entlastung des Arbeitsgedächtnisses dar, wie die Studie von Glogger et al. (2009) zeigt.

Glogger et al. (2009) verglichen hinsichtlich der Formulierung spezifisch und eher unspezifisch gestaltete Prompts an einer Stichprobe von 14-15-jährigen Schülern. Die spezifischen Prompts enthielten, zusätzlich zu einer Aufgabenstellung (z.B. „Versuche, Verknüpfungen herzustellen zwischen dem, was Du letzte Woche gelernt hast und dem, was Du bereits darüber weißt“), zwei bis drei Bearbeitungsschritte (z.B. „Denke Dir dazu ein Beispiel aus“), die die Aufgabe in Teilschritte zerlegten. In dieser Studie wird die Konkretheit der Prompts vor dem Hintergrund des Expertise-Umkehr-Effektes als Aufgabenschwierigkeit ausgelegt. Der Expertise-Umkehr-Effekt basiert vereinfacht zusammengefasst auf der Tatsache, dass Experten aufgrund ihres Vorwissens eher von weniger strukturierten Aufgaben profitieren, wohingegen Novizen eine stärkere Strukturierung benötigen, um erfolgreich zu lernen (vgl. Kalyuga, 2007). Als Erfolgskriterien definieren die Autoren die Anzahl der eingesetzten Strategien und deren Qualität. Die spezifischen Prompts führten in der Untersuchung zum Einsatz von mehr und vielfältigeren Elaborationsstrategien als die unspezifischen Prompts. Die Schüler bildeten mehr Beispiele für den Lerninhalt und verwendeten zudem mehr Organisationsstrategien zur Gliederung des Lernstoffes. Hinsichtlich der Qualität der Lernstrategien ergaben sich jedoch keine Verbesserungen. Dies wirft die Frage auf, inwiefern nicht nur von einem Produktionsdefizit, sondern auch einem Mediationsdefizit (die Lerner verfügen nicht über die notwendigen Strategien) ausgegangen werden muss. Letzterem könnte man durch gezielte Rückmeldung, die den Lernenden über die Information der Bearbeitungsqualität ihrer Lernjournale hinaus Hilfestellungen bietet, entgegensteuern.

Beim Einsatz von Prompts sollte nach Nückles et al. (2006) daher das Vorwissen der Lerner berücksichtigt werden, damit sich die Intervention nicht motivationsbeeinträchtigend auswirkt.

Ein weiterer Aspekt, der mit dem Arbeitsgedächtnis in Zusammenhang steht, ist nach Thillmann et al. (2009) der Darbietungszeitpunkt der Prompts. Eingebettet in ein Modell selbstregulierten Lernens überprüften Thillmann et al. (2009), inwiefern eine zeitliche Abstimmung der Prompt-Darbietung mit einem idealtypischen Lernzyklus notwendig und lernförderlich ist. In einer experimentellen Studie arbeiteten 95 Schüler der neunten Klasse einer Realschulstufe unter

drei experimentellen Bedingungen mit Prompts, die zum einen die Verarbeitung des Lernmaterials unterstützten und zum anderen das Generieren eigener Inhalte anregten. Dazu bildeten die Forscher drei Gruppen, die die Prompts 1. vor dem Lernzyklus, 2. angepasst an einen idealtypischen Regulationsverlauf und 3. entgegen dem idealtypischen Regulationsverlauf erhielten. Das Regulationsmodell bezieht sich auf das Modell der Lernzielregulation von Wirth und Leutner (2006). Um den Lernzuwachs abzubilden, führten die Forscher einen Multiple-Choice-Test vor und nach der Intervention durch.

Die Studie von Thillmann et al. (2009) führt zu dem Ergebnis, dass eine Einbindung der Prompts in den Lernprozess sich am günstigsten auf den Lernerfolg auswirkt. Dabei spielte es allerdings keine Rolle, ob der optimale Regulationsverlauf berücksichtigt wird oder nicht. Eine Mediatoranalyse zeigte außerdem, dass der Strategienutzung während des Lernens eine Mediatorfunktion zwischen Darbietungszeitpunkt und Lernerfolg zukommt. Hier bestätigt sich demnach wiederholt die Rolle von Lernstrategien beim Lernen durch Schreiben.

In den vorangehenden Abschnitten wurden Studien im Hinblick auf die Gestaltung einer Lernumgebung mit Prompts betrachtet. Dabei kristallisierten sich Merkmale heraus, die bei der Formulierung von Schreibinstruktionen und deren Einbettung in den Lernprozess einzubeziehen sind.

2.3.2 Forschungsbedarf

Die Forschung zu Prompts verfolgt das Ziel, Bedingungen für die gezielte Unterstützung von Lernprozessen zu ergründen. Aus den referierten Studien wird deutlich, dass Prompts für das Lernen durch Schreiben eine vielversprechende Methode darstellen, da verschiedene Strategien und Teilprozesse des Lernens gezielt unterstützt werden.

Typisch für die Studien in diesem Themenfeld ist, dass das Lernen durch Schreiben überwiegend in realen Unterrichtssituationen beobachtet wird. Die Interventionen sind zumeist einmalig gestaltet oder werden über eher kurze Zeiträume durchgeführt. Eine Ausnahme bilden die wenigen Untersuchungen zu Lerntagebüchern, die meist über die Dauer eines Semesters laufen. Um die Effekte von Prompts und Lerntagebüchern zu untersuchen, stützen sich Untersuchungen zum einen auf die Analyse von Lernprotokollen und zum anderen auf Lernerfolgsmessungen, für die in der Regel Multiple-Choice-Tests verwendet werden. Als

Indikatoren für die Wirksamkeit der Prompts werden die Häufigkeit der in den Lernprotokollen vorkommenden Strategien und der Erfolg in einem Wissenstest, der meistens direkt im Anschluss an die Intervention durchgeführt wird, interpretiert. Dieses Untersuchungsdesign ist zwar üblich, es bleibt jedoch fraglich, ob geschlossene Items den Lernzuwachs durch das Schreiben mit Prompts angemessen reflektieren. Auf diese Weise wird zwar das allgemeine Lernen erfasst, doch typischerweise soll durch das Schreiben auch ein besseres Verständnis für den Lernstoff erreicht werden. Idealerweise wird in einem Unterricht, der in einem sozial-konstruktivistischen Ansatz verankert ist, das Wissen durch Schreiben transformiert, wie es Bereiter und Scardamalia (1987) in der *knowledge transforming*-Schreibstrategie ausformulieren.

Charakteristisch für diese Form des Lernens ist die Integration neuer Informationen in das bestehende Wissen, die sich in einem flexiblen Umgang mit den neuen Inhalten niederschlägt. Es ist anzunehmen, dass sich dies in einer entsprechend komplexen Wissensstruktur widerspiegelt, wie z.B. Van Dijk und Kintsch (1983) in Untersuchungen zum Textverständnis belegen. Um Effekte des Schreibens auf das Wissen zu erfassen, scheint es daher lohnenswert, die Wissensstrukturen von Lernenden nach einer Schreibintervention zu analysieren. Hier stoßen die bisherigen Studien zum einen methodisch und zum anderen aufgrund des Untersuchungszeitraumes an ihre Grenzen, da letzterer wahrscheinlich zu kurz angelegt ist, um Wissensveränderungsprozesse initiieren und abbilden zu können. Aus den Ergebnissen der Studien zu Prompts, die über eine Stunde oder zwei Stunden angelegt sind, kann man eher auf Tendenzen schließen, die darauf hindeuten, dass Schreiben auf das Wissen zurückwirkt. Ob allerdings nachhaltige Effekte erzielt werden, ist nicht beantwortet. Auch Bereiter und Scardamalia (1987) gehen davon aus, dass anspruchsvolles Schreiben, das zu einer Wissenstransformation führt, Zeit benötigt.

In ihrer Metastudie zu Effekten von Schreibinterventionen stellen Bangert-Drowns et al. (2004) tatsächlich höhere Effektstärken für längere Untersuchungszeiträume fest und vermuten, dass die Treatment-Länge die Effekte des Schreibens verstärkt: „One might plausibly expect that, if writing is to have a positive influence on learning, the influence will be cumulative over time, as students become more familiar with the assigned writing tasks, as more content is covered, and as, potentially, the writing practice affects learning strategies themselves“

(Bangert-Drowns et al., 2004, S. 51). Allerdings nimmt die Effektstärke bei größeren Stichproben sowie in Kombination mit gewichteten Daten ab: „This pattern was repeated in the weighted analysis but did not achieve the significance criterion. Larger-sampled studies with lower effect sizes were conducted over longer periods; when weighted, these studies pulled down the averages of the smaller-sampled long-term studies“ (Bangert-Drowns et al., 2004, S. 51).

In der vorliegenden Arbeit wird die Schreibintervention mit Prompts daher auf ein Semester angelegt. Aus der Analyse der mit den Prompts produzierten Texte soll Aufschluss darüber gewonnen werden, inwiefern das Schreiben über einen längeren Zeitraum sich auf die Qualität der Wissensstruktur (operationalisiert als Verarbeitungstiefe) auswirkt. Darüber hinaus sollen die Daten dazu genutzt werden, um Beziehungen zwischen der Qualität des Schreibens und Merkmalen der Schreibaufträge zu ergründen.

Fasst man die bisherigen Ergebnisse zusammen, fällt auch auf, dass Lernervariablen bislang wenig Beachtung geschenkt wurde. Ein Aspekt, der im Zusammenhang mit der Bearbeitung von Schreibaufgaben in Lerntagebüchern und Portfolios noch weitgehend offen ist, betrifft die Frage, inwieweit spezifische Voraussetzungen auf Seiten der Lernenden auf die Qualität des Schreibens Einfluss haben, bzw. welchen Einfluss das Lernen durch Schreiben auf die nachfolgenden Aktivitäten der Lernenden hat. Die Feldstudie zum Schreiben mit Prompts im Rahmen von Lerntagebüchern von Nückles et al. (2006) deutet darauf hin, dass die Motivation der Studierenden sich durch das Schreiben verändert. Untersuchungen aus dem Bereich der Lernmotivation decken zudem Zusammenhänge zwischen Lernerfolg und Lernmotivation auf. So wiesen Rheinberg, Vollmeyer und Burns (2001) in einem Experiment, in dem zwei Lerngruppen kontrastiert wurden (fragengeführtes Faktenlernen und selbstgesteuertes Verständnislernen), positive Korrelationen zwischen den Motivationskomponenten *Interesse* und *Herausforderung* und dem Lernerfolg nach. Für die Personen in der Lernbedingung *selbstgesteuertes Verständnislernen* war es demnach ausschlaggebend für den Lernerfolg, wie sehr sie sich für den Lerninhalt interessierten und die Aufgabe als Herausforderung ansahen.

Andere Studien befassen sich mit der Rolle der Motivation für den Lernerfolg und deuten auf positive Zusammenhänge zwischen der Lernzielorientierung, dem Interesse an der Aufgabe und der Anstrengung hin, die wiederum in eine bessere

Lernleistung münden (vgl. Spörer, 2004). Insgesamt lässt sich festhalten, dass die Motivation von Personen beim Lernen eine entscheidende Rolle spielt, da sie ausschlaggebend dafür ist, wie diese sich im Lernprozess engagieren. Lernmotivation ist insofern als Voraussetzung für Lernaktivitäten zu betrachten.

Im empirischen Teil dieser Arbeit werden daher verschiedene Aspekte der Lernmotivation zum einen als Voraussetzung für das Schreiben und zum anderen im Hinblick darauf, wie der Schreibprozess darauf zurückwirkt, betrachtet.

Eine weitere, nicht systematisch untersuchte Variable beim Lernen mit Prompts ist Feedback. Unter dem Begriff des Feedbacks werden ganz allgemein Informationen verstanden, die als Konsequenz auf eine erbrachte Leistung durch verschiedene Instanzen (z.B. einen Lehrer, Peers, ein Buch) zurückgemeldet werden (vgl. Hattie & Timperley, 2007; Winne & Butler, 1994) und sich auf Aspekte der Leistung und des Verstehens beziehen. Diese Information kann der Lerner z.B. verwenden, um seinen Lernprozess zu optimieren oder sie in sein Wissen zu integrieren oder er kann sie verwerfen: „Feedback is information with which a learner can confirm, add to, overwrite, tune, or restructure information in memory, whether that information is domain knowledge, meta-cognitive knowledge, beliefs about self and tasks, or cognitive tactics and strategies“ (Winne & Butler, 1994, S. 5740).

Der Metastudie von Bangert-Drowns et al. (2004) zufolge sind verschiedene Formen des Feedbacks Bestandteil zahlreicher Studien zum Lernen durch Schreiben. Die Autoren beziehen daher Feedback als eigenständige Variable in die Meta-Studie ein, kommen jedoch zu einem überraschenden Ergebnis: „It was surprising to see that the provision of feedback to students on their writing showed no significant relationship with effect size“ (Bangert-Drowns et al., 2004, S. 50). Eine Erklärung für diesen Befund, der den allgemein vorzufindenden Ergebnissen aus der lernpsychologischen Forschung widerspricht (vgl. Bangert-Drowns et al., 1991; Butler & Winne, 1995; Hattie & Timperley, 2007; Kluger & DeNisi, 1996), könnte jedoch nach Angaben der Autoren im Design der untersuchten Studien zu finden sein. Da diese oftmals ungenau bei der Beschreibung des Feedbacks sind (z.B. hinsichtlich der Häufigkeit und der Form) und Bangert-Drowns et al. (2004) sich nicht sicher sind, ob die Bereitstellung von Feedback immer angegeben wurde, könnte das Ergebnis der Studie auf Kodierungsprobleme zurückzuführen sein. Die geringen Effektstärken sind folglich unter diesem Vorbehalt einzuordnen.

Da zahlreiche andere Studien die Wirksamkeit von Feedback in Zusammenhang mit Lernprozessen nachweisen (z.B. Azevedo & Bernard, 1995; Clariana, 1993; Schimmel, 1983), ist es plausibel anzunehmen, dass Feedback auch die Qualität des Schreibprozess positiv beeinflussen kann.

Aus den Studien zu Lernen durch Schreiben und zu Prompts lassen sich vor allem zwei Aspekte ableiten, die unterstützt werden sollten: die Aufgabeninterpretation mit einem Fokus auf anspruchsvollen Strategien und die Motivation der Lernenden.

Dass die Aufgabeninterpretation ein Faktor ist, der die Qualität der Bearbeitung und die Wahl der Strategien beim Schreiben beeinflusst, zeigten Klein (1999) und Penrose (1992). Um angemessene Strategien zur Bewältigung einer Aufgabe zu wählen, müssen Personen zunächst einmal die Intention der Aufgabe verstehen. Interpretieren sie die Aufgabe anders als intendiert, kann dies zu einer nicht optimalen Bearbeitung führen. Um dem entgegenzuwirken, müssen Zielvorgaben durch die Aufgabe und Feedback aufeinander abgestimmt werden. Lernziele müssen den Lernenden aus der Aufgabe deutlich werden und das Feedback sollte sich wiederum darauf beziehen, inwiefern die Anforderungen erreicht wurden und gegebenenfalls Hinweise für weitere Teilziele bei der Bearbeitung geben. Einer Analyse von Hattie und Timperley (2007) zufolge ist Feedback am effektivsten, wenn es aufgabenorientierte Hinweise zur Verbesserung bei der Durchführung der Aufgabe enthält oder die Korrektheit bestätigt.

Zur Abstimmung der Aufgaben und des Feedbacks wird in der vorliegenden Arbeit auf den Ansatz des *constructive alignment* von Biggs (1996) zurückgegriffen. Das Modell des *constructive alignment* basiert auf einem konstruktivistischen Verständnis des Lernens, das durch die Aktivität des Lerners geprägt ist. Das Ziel des didaktischen Modells besteht darin, eine Abstimmung zwischen der Aufgabenstellung, des didaktischen Ansatzes des Lehrenden und den Aktivitäten der lernenden Person vorzunehmen. Dies wird erreicht, indem konkrete (kognitive) Tätigkeiten zur Bearbeitung einer Aufgabe angeregt werden, die durch Verben in die Aufgabenstellung integriert werden (z.B. „Vergleichen Sie das Konzept mit dem Konzept xy“). Das Feedback wiederum bezieht sich darauf, inwiefern es dem Lernenden gelungen ist, diesen Prozess durchzuführen. Um diese Qualität zu beurteilen, bietet sich die von Biggs und Collis (1982) entwickelte SOLO-Taxonomie (*Structure of the observed learning outcome*) an, mit der Lernergebnisse, die aus

offenen Fragestellungen hervorgehen, hinsichtlich ihrer Komplexität eingestuft werden können. Die Taxonomie basiert auf der Annahme, dass die Komplexität von Wissensstrukturen mit fortschreitender Expertise zunimmt, was sich insbesondere in den Vernetzungen abbildet, die ein Lerner beim Beantworten einer Lernaufgabe in einem Themengebiet herstellt, und dem Elaborationsgrad, mit dem er dies tut. Eine detaillierte Beschreibung der SOLO-Taxonomie findet sich in Abschnitt 5.1.

Im Rahmen der vorliegenden Arbeit wird die SOLO-Taxonomie einerseits verwendet, um Zusammenhänge zwischen Aufgabenmerkmalen und Wissensstrukturen der Lerner zu untersuchen, und andererseits, um Feedback bereitzustellen, das den Probanden Auskunft darüber gibt, inwiefern sie die intendierten Anforderungen der Aufgabe erreicht haben. Auf diese Weise wird Fehlinterpretationen der Aufgabe entgegengewirkt und es wird eine Methode erprobt, mit der Lernprozesse beim Schreiben gezielter als bisher gefördert werden können. Zusätzlich bietet die Verwendung der SOLO-Taxonomie den Vorteil, dass die Gestaltung der Aufgaben und die Kriterien für das Feedback aufeinander abgestimmt sind. Anspruchsvolle Strategien können, basierend auf der Hierarchie der SOLO-Taxonomie, sowohl explizit angeregt als auch hinsichtlich ihrer Qualität den Lernenden zurückgemeldet werden.

Die zweite Feedbackkomponente bezieht sich auf motivationale Variablen seitens der Studierenden. So zeigten sich in der Studie von Nückles et al. (2006) Zusammenhänge zwischen der Abnahme der Motivation und der Lernleistung in einer Gruppe von Lernern, die über ein Semester mit Prompts Texte verfassten. Bangert-Drowns et al. (2004) vermuten, dass der Leistungsabfall bei längeren Schreibinterventionen auf Motivationsprobleme bei Lernern zurückzuführen sein kann, für die das Schreiben eine große Herausforderung darstellt. Allerdings ist diese Hypothese nicht empirisch überprüft.

Hinweise in diese Richtung gibt jedoch die Feldstudie von Picard und Imhof (2010), in der sich zeigte, dass die Produktion höherwertiger Texte mit einer positiveren Einstellung zum Lernen korrelierte. Im Rahmen eines einwöchigen Blockseminars erhielten Studierende nach jeder Lerneinheit zwei Prompts zur Nachbearbeitung des Seminarstoffs. Die Texte wurden am Ende der Woche mit der SOLO-Taxonomie (Biggs & Collis, 1980) auf die Qualität der Wissensstruktur hin eingestuft. Um zu untersuchen, ob sich die Erfahrung mit dem Lerntagebuch

auf die Lernkonzepte der Lernenden auswirkt, wurden die über die Lernkonzepte gewonnenen Daten (Fragebogen zur Erfassung subjektiver Lernkonzepte von Nieder (2006)) mit den in den Texten erreichten Niveaustufen (SOLO-Taxonomie) in Beziehung gesetzt. Im Ergebnis zeigte sich, dass „vermehrt negative Gefühle gegenüber dem Lernen wahrgenommen werden, wenn die Lernenden häufig auf Niveaustufe zwei arbeiten“ (Picard & Imhof, 2010, S. 74). Die SOLO-Taxonomie besteht aus insgesamt fünf Stufen, von denen die zweite Stufe Texte mit einem geringen Elaborationsgrad und wenig inhaltlichen Vernetzungen repräsentiert. Die fünfte Stufe repräsentiert das höchste Niveau. In dieser Untersuchung zeigte sich also, dass ein einfacheres Verarbeitungsniveau die Einstellung zum Lernen beeinträchtigte. Konvergent dazu zeigte die Untersuchung auch, dass eine Zunahme an positiven Gefühlen im Lernkonzept mit einer häufigeren Einstufung der Texte auf dem hohen vierten Niveau einherging. Dieses setzt eine intensive Verarbeitung des Lernstoffes voraus, die sich im Text z.B. darin abbildet, dass vielfältige Aspekte eines Themas benannt und miteinander verknüpft werden.

Insgesamt lässt sich aus diesen Befunden ableiten, dass Motivation eine wichtige Komponente in Zusammenhang mit Schreibinstruktionen darstellt, über deren Wirkung bislang jedoch keine eindeutigen Antworten vorliegen. Daher wird in der empirischen Untersuchung auch überprüft, inwiefern verschiedene Formen des Feedbacks motivationsförderliche Wirkungen beim Schreiben mit Prompts entfalten. Der Einsatz von Feedback verfolgt in der vorliegenden Arbeit also zwei Zielrichtungen: Zum einen sollen die Untersuchungsteilnehmer durch die Rückmeldung ihrer Schreibqualität anhand der SOLO-Taxonomie dazu angeregt werden, anspruchsvolle kognitive Strategien einzusetzen, und zum anderen wird erwartet, dass das Feedback dazu beiträgt, die Motivation zum Bearbeiten der Prompts über ein Semester aufrechtzuerhalten.

Vor dem geschilderten Hintergrund besteht das Ziel der empirischen Untersuchung zunächst darin, die Effekte von Schreibinstruktionen (Prompts) mit verschiedenen Anforderungsspezifika auf die Qualität kognitiver Prozesse beim Textproduzieren zu untersuchen. Da über die Qualität der mit Prompts angeregten Texte noch kaum empirische Daten zu finden sind (vgl. Bannert, 2009), ist es außerdem das Ziel der vorliegenden Arbeit, die Einträge in das Lerntagebuch auf die Verarbeitungstiefe hin zu analysieren, die in den Texten erkennbar wird. Dazu wird die SOLO-Taxonomie von Biggs und Collis (1980) verwendet. Die Ergebnisse

können Hinweise auf die Gestaltung solcher Schreibinstruktionen geben, mit denen komplexe Schreibprozesse angeregt werden, die zum Aufbau anspruchsvoller Wissensstrukturen beitragen.

Da Prompts bislang nicht in Zusammenhang mit Feedback, sondern nur als isolierte Intervention untersucht wurden, besteht ein weiteres Ziel darin, zu differenzierten Aussagen darüber zu gelangen, inwiefern und wenn ja, auf welcher Ebene, sich Schreibprozesse durch Rückmeldung optimieren lassen und ob Variablen der Motivation durch die Schreiberfahrung (das schließt sowohl Feedback als auch die Textqualität ein) beeinflusst werden.

Insgesamt werden also Wechselwirkungen zwischen Aufgabenmerkmalen, Feedback, Motivation und der Qualität der Schreibprozesse in den Fokus genommen. Aufbauend auf den bisherigen theoretischen und empirischen Erkenntnissen wird im folgenden Kapitel eine Lernumgebung vorgestellt, mit der anspruchsvolle Schreibprozesse bei Studierenden angeregt werden sollen.

3 Die Konzeption der Lernumgebung

Die im Folgenden konzipierte Lernumgebung mit Prompts basiert auf den in den vorangehenden Kapiteln referierten theoretischen und empirischen Studien zu Lernen durch Schreiben, aus denen Merkmale herausgearbeitet wurden, die Schreiben als Lerninstrument kennzeichnen. Vor dem Hintergrund eines kognitiv-konstruktivistischen Lernparadigmas stehen bei didaktischen Maßnahmen vor allem die Aktivierung des Lerners und die Unterstützung beim Aufbau elaborierter Wissensstrukturen im Vordergrund.

Aus der Textproduktionsforschung wird deutlich, dass Schreiben, wenn es sich im Sinne der *knowledge transforming*-Strategie restrukturierend auf das Wissen auswirken soll, eine anspruchsvolle Aufgabe darstellt, deren Bewältigung in hohem Maße von der kognitiven Verarbeitungskapazität des Textproduzenten abhängt (vgl. Bereiter & Scardamalia, 1987; Schnotz, 2001). Um Lerner nicht zu überfordern und um sie zu solchen Schreibprozessen zu bringen, die den Aufbau von Wissen nach sich ziehen, ist es nach Bereiter und Scardamalia (1987) daher sinnvoll, dem Lerner prozedurale Hilfestellungen zu geben, die „dazu beitragen, das komplexe Problem der Textproduktion in handhabbare Teilprobleme zu untergliedern“ (Schnotz, 2001, S. 155).

In der vorliegenden Arbeit wird dazu eine Kombination aus Prompts und Feedback erprobt. Um diese zu gestalten, wurden aus den bisherigen Erkenntnissen der Textproduktionsforschung und der Feedbackforschung Gestaltungsprinzipien abgeleitet, die Gegenstand des folgenden Abschnitts sind.

Die Lernumgebung umfasst ein durch Prompts vorstrukturiertes Lernjournal, ein Musterbeispiel und verschiedene Formen der Rückmeldung.

3.1 Die Gestaltung der Prompts

In der vorliegenden Arbeit werden Prompts eingesetzt, um den Einsatz kognitiver Lernstrategien beim Schreiben zu fördern. Unter Prompts werden Lernhilfen im weitesten Sinne verstanden, deren Ziel es ist, Wissen zu aktivieren, Lernstra-

tegien anzuregen und auch spezifische Teilprozesse des Schreibens (z.B. Planen und Revidieren) zu initiieren. Dabei wird grundsätzlich angenommen, dass die Lerner bereits über das zu aktivierende Wissen bzw. die zu aktivierenden Strategien verfügen, diese aber nicht einsetzen. Die Bandbreite von Prompts reicht von allgemeinen, kurzgehaltenen Fragestellungen (z.B. „Wie gehst Du dabei vor?“) bis hin zu konkreten Instruktionen (z.B. „Zähle zunächst die einzelnen Elemente zusammen“) (vgl. Bannert, 2009, S. 139). Nach Wirth (2009) hingegen enthalten Prompts nur minimale Informationen und eröffnen dem Lerner durch eine offene Formulierung selbstregulative Freiräume. Die Bandbreite dessen, was unter Prompts verstanden wird, ist also groß. Zudem variieren Prompts hinsichtlich zahlreicher anderer Aspekte, wie z.B. der Strategien, die angeregt werden, des Grades an Konkretheit (vgl. Glogger et al., 2009) und des Darbietungszeitpunktes (vgl. Thillmann et al., 2009). Um die Prompts zu gestalten, wurden auf Basis der bisherigen Forschung zu Prompts drei Faktoren identifiziert, die bei der Entwicklung zu berücksichtigen sind und im Folgenden erläutert werden.

1. Vorwissen und *cognitive load*
2. Strategien, die sich als effektiv erwiesen haben
3. Darbietungszeitpunkt

3.1.1 Vorwissen und *cognitive load*

Ein zentrales Ergebnis des aktuellen Forschungsstandes zu Lernen durch Schreiben ist, dass anspruchsvolle Strategien im Rahmen von Schreibprozessen angeleitet werden müssen. Die bisherigen Befunde zu Prompts deuten jedoch auf ein Spannungsfeld aus einer Unterforderung der Lerner (Expertise-Umkehr-Effekt) bei länger dauernden Interventionen und zunehmender Kompetenz der Textproduzenten einerseits und einem Produktionsdefizit (Lerner wenden die Strategien nicht spontan an) andererseits hin, in dem die inhaltliche Ausgestaltung von Prompts sich bewegt. In diesem scheinen Vorwissen und *cognitive load* zwei der Determinanten darzustellen, die die Wirksamkeit von Prompts beeinflussen.

Von explizit angeleiteten Teilschritten innerhalb der Prompts profitieren insbesondere unerfahrene Textproduzenten und Lerner mit geringem Vorwissen, wie die Studie von Glogger et al. (2009) zeigt. Darin wird die Spezifität verschiedener Prompts variiert. Ein unspezifischer Prompt zur Elaboration besteht z.B. aus

einer Aufgabenformulierung: „Try to build connections between what you have learned last week and what you already know“. Im spezifischen Prompt werden, zusätzlich zu der Aufgabe, Teilziele zur Bearbeitung derselben benannt, da die Autoren davon ausgehen, dass die Probanden nur geringes Vorwissen bezüglich der optimalen Durchführung der Aufgabe aufweisen: „Try to build connections between what you have learned last week and what you already know. For this purpose, write down how you could apply what you have learned this week at home in your spare time. Create some examples. Choose one example and explain the calculations that can be done. Explain it in a way so that a classmate who was absent last week could well understand“ (Glogger et al., 2009, S. 98). In der Studie zeigte sich, dass die Schüler einer neunten Klasse im Fach Mathematik stärker von den spezifischen Instruktionen profitierten. Dies spiegelte sich in der Anzahl und Vielfalt der kognitiven Lernstrategien wieder. Da keine Leistungsmaße erhoben wurden, können jedoch keine Aussagen über Effekte auf den Lernprozess gemacht werden.

Mehr Aufschluss gab in dieser Hinsicht die Feldstudie von Nückles et al. (2006), in der Studierende über ein Semester mit Prompts arbeiteten. Hier zeigte sich nach der Hälfte des Semesters, dass erwachsene Lerner zunächst von der durch die Prompts vorstrukturierten Lernumgebung profitierten. Dieser Befund spricht demnach dafür, mit Prompts spezifische Lernstrategien anzuregen. Allerdings beobachteten die Autoren einen kontinuierlichen Motivations- und Leistungsabfall über das Semester, den sie auf eine zu starke externe Steuerung des Lernprozesses zurückführten. Nückles et al. (2006) nehmen an, dass die Prompts mit der zunehmenden Fertigkeit, die angeregten Strategien eigenständig einzusetzen, von den Studierenden als störend wahrgenommen wurden und eine zusätzliche Belastung für das Arbeitsgedächtnis darstellten (Expertise-Umkehr-Effekt).

Da beide Studien in Bezug auf die Stichprobe, das Fach und die Untersuchungsdauer unterschiedlich ausfallen, lassen sich hieraus keine eindeutigen Schlussfolgerungen zur optimalen Ausgestaltung von Prompts ableiten. Zudem wäre bei Nückles et al. (2006) als Alternativerklärung zum Expertise-Umkehr-Effekt auch denkbar, dass sich fehlendes Feedback bei anspruchsvollen Aufgaben über einen so langen Zeitraum motivationshemmend auswirkt.

Es gilt jedoch als gesicherter Befund, dass Schreiben ohne explizite Anleitung oder Zielvorgabe wirkungslos im Hinblick auf lernförderliche Wirkungen ist (vgl.

Bangert-Drowns et al., 2004; Borsch & Imhof, 2007; Nückles et al., 2004). Werden Schreibaufträge zu „offen“ formuliert, wie es Wirth (2009) zum Zwecke selbstregulativer Freiräume propagiert, besteht die Gefahr, dass Lerner bei der Bearbeitung der Aufgabe eher eine Strategie der Wissensreproduktion als solche Strategien wählen, die eine tiefgreifende Verarbeitung nach sich ziehen. Zu offene Schreibinstruktionen können also wiederum in ein Produktionsdefizit münden.

Da aufgrund dieser Wechselwirkungen keine kausalen Schlussfolgerungen für den Grad der Spezifität der Prompts abgeleitet werden können, wird in der vorliegenden Arbeit eine Kombination aus anspruchsvollen, spezifisch formulierten Schreibinstruktionen und Feedback erprobt. Hinsichtlich der Spezifität der Prompts orientiert sich die Arbeit am Gestaltungsprinzip des *constructive alignment* (Biggs, 1996), das die Lernaktivitäten, die der Textproduzent am Lerninhalt vornehmen soll, benennt und Rückmeldung zur Qualität derselben bereitstellt. Anstelle von Teilschritten werden konkrete Prozesse benannt, z.B. „Bilden Sie Beispiele für das Konzept“ oder „Stellen Sie Bezüge zu anderen Konzepten her“. Die Stärke des Verfahrens liegt in der transparenten Gestaltung der entscheidenden Komponenten des Lernprozesses (Zielsetzung, Bewertung der Erreichung, Feedback) und ihrer Abstimmung aufeinander. Um mehr Aufschluss über die optimale Ausgestaltung der Prompts zu gewinnen, werden Prompts mit verschiedenen Bezügen im Hinblick auf die Qualität der Wissensstruktur analysiert.

3.1.2 Strategien, die sich als effektiv erwiesen haben

In den vorgestellten Studien zu Prompts werden verschiedene Aktivitäten und Strategien durch Prompts unterstützt. Durch kognitive Prompts sollen z.B. Lernaktivitäten zur Informationsverarbeitung stimuliert werden (z.B. Wiederholen, Elaboration, Organisation des Lernmaterials). Metakognitive Prompts zielen hingegen auf das Initiieren von Kontroll- und Planungsprozessen ab, wie z.B. die Überwachung des Lernprozesses im Hinblick auf Verständnislücken und die situationsadäquate Regulation kognitiver Strategien.

Insgesamt lässt sich aus den vorliegenden Befunden ein genereller Lernvorteil von Personen, deren Schreibprozess mit Prompts unterstützt wurde, gegenüber solchen Lernern, die ohne Prompts schrieben, feststellen. Hinsichtlich der Art der Prompts zeigte sich, dass sowohl kognitive als auch metakognitive Prompts sich positiv auf den Wissenszuwachs und den Wissenstransfer auswirken (Bannert,

2003; Berthold et al., 2007; Hübner et al., 2006). Im Vergleich der beiden Arten von Prompts stellten sich in den Studien von Berthold et al. (2007) und Hübner et al. (2006) jedoch besonders kognitive Prompts als lernförderlich heraus. Die Forscher beobachteten in Zusammenhang mit diesen Prompts zudem, dass Lerner nicht nur signifikant mehr kognitive Strategien beim Schreiben einsetzten, sondern auch signifikant mehr metakognitive Strategien verwendeten als Lerner ohne Prompts (vgl. Berthold et al., 2007). Dies schlug sich auch im Lernerfolg nieder.

Für metakognitive Prompts zeigte sich, dass diese separat nur dann wirksam sind, wenn die Lerner die Möglichkeit zur Nachregulation haben (vgl. Bannert, 2003; Hübner et al., 2006). Ist dies jedoch nicht der Fall, trägt der Einsatz metakognitiver Prompts nicht zur Steigerung des Lernerfolgs bei. Eine Mediatoranalyse von Berthold et al. (2007) zeigte für den kombinierten Einsatz metakognitiver und kognitiver Prompts zudem, dass für den Lernerfolg kognitive Strategien ausschlaggebend sind.

Zusammengefasst erscheint es demzufolge sinnvoll, entweder eine Kombination metakognitiver und kognitiver Prompts oder nur kognitive Prompts zu verwenden, da diese auch metakognitive Strategien anzustoßen scheinen.

In der vorliegenden Arbeit wird der Ansatz verfolgt, die Nachbereitung der Lerninhalte aus einem Seminar zu unterstützen. Eine effektive Nachbereitung von Lerninhalten zielt auf die Integration der neuen Inhalte in die bestehende Wissensstruktur ab. Aus diesem Grund richtet sich der Fokus der Arbeit auf Elaborationsstrategien. Nach Wild (2000) sind dies:

- Verknüpfung des neu gelernten Materials mit Alltagsbeispielen und persönlichen Erlebnissen
- die Formulierung des Aufgenommenen in eigenen Worten
- die Bildung von Analogien zu bereits bekannten Zusammenhängen

Die Formulierung der Prompts für die empirische Untersuchung orientiert sich an diesen Richtlinien.

3.1.3 Darbietungszeitpunkt

Zur Förderung bestimmter Teilprozesse durch Prompts erscheint es plausibel, diese zu einem Zeitpunkt zu initiieren, der mit einer idealtypischen Lernepisode übereinstimmt. Verschiedene Modelle konzipieren eine Lernepisode als Abfolge von drei oder vier aufeinanderfolgenden Lernphasen (vgl. Schmitz, 2001; Winne & Hadwin, 1998; Zimmerman, 2000), die durch unterschiedliche kognitive und metakognitive Prozesse und Emotionen geprägt sind. Aufgrund der Rekursivität von Lernprozessen ist es jedoch schwierig, einen optimalen Lernverlauf vorauszusagen und Prompts zeitgenau anzupassen (vgl. Thillmann et al., 2009).

Eine derartige Feinabstimmung ist jedoch für die Wirksamkeit von Prompts auch nicht notwendig, wie die Studie von Thillmann et al. (2009) zeigt. Im Hinblick auf den Lernerfolg machte es innerhalb dieser Feldstudie nur auf deskriptiver Ebene einen Unterschied, ob zuerst Prompts zum Initiieren kognitiver Strategien zur Verarbeitung von Lerninhalten und erst im Anschluss daran Prompts zur Anregung metakognitiver Strategien zur Regulation des Lernprozesses präsentiert wurden. Die Untersuchung erbrachte jedoch, dass es entscheidend ist, die Prompts in den Lernprozess zu integrieren und nicht separat zu stellen. Dies entspricht nach Bannert (2009) den Befunden der Lernstrategieforschung. Demzufolge zeichnen sich erfolgreiche Strategieinstruktionen durch eine Einbettung in den Lernprozess aus.

Das Ziel der vorgestellten Untersuchung ist es, die Nachbereitung eines Seminars durch Prompts zu unterstützen und Studierende dabei anzuleiten, die Lerninhalte mit anspruchsvollen Strategien zur Vertiefung zu wiederholen. Bei der Nachbereitung wird vorausgesetzt, dass Lerner durch das Seminar bereits Wissen aufgebaut haben, das nun weiterverarbeitet werden kann und in die bestehende Wissensstruktur integriert werden soll. Diesem Zweck dienen Elaborationsstrategien, mithilfe derer das Verstehen und Behalten des Lernstoffes gefördert wird. Elaborative Strategien sind z.B. die Verknüpfung des neuen Lerninhaltes mit Alltagsbeispielen oder individuellen Erfahrungen sowie das Herstellen von Bezügen zu bereits Bekanntem (Wild & Schiefele, 1994).

3.1.4 Zusammenfassung und Ableitung der Prompts

Basierend auf den voranstehenden Überlegungen wird ein Lernjournal mit Prompts für eine erste Untersuchung entwickelt. Da Lernstrategien bereicherspezifisch anzuregen sind, werden die Prompts an den Lerninhalt angepasst.

Bezüglich der Spezifität der Prompts hat sich in den Studien gezeigt, dass insbesondere jüngere und unerfahrene Lerner von Teilschritten zur Bearbeitung der Prompts profitieren, dass diese jedoch mit zunehmender Kompetenz zu einem Expertise-Umkehr-Effekt führen können. Da die vorliegende Untersuchung auf einen längeren Zeitraum abzielt, wird auf die explizite Anleitung von Teilschritten verzichtet. Die Prompts werden vielmehr im Sinne von Leitfragen formuliert.

Was den Darbietungszeitpunkt der Lernhilfen angeht, zeigte die Studie von Thillmann et al. (2009), dass Prompts am wirksamsten sind, wenn sie in den Lernzyklus eingebunden werden. Da die Prompts als Unterstützung bei der vertiefenden Nachbereitung der Seminarinhalte dienen sollen, erhalten die Lerner nach einer in sich geschlossenen Lerneinheit (meistens eine Seminarsitzung) jeweils einen oder zwei Prompts (dies variiert in den zwei durchgeführten Untersuchungen). Das Lernjournal wird als integraler Seminarbestandteil konzipiert.

Zur optimalen Vertiefung der Seminarinhalte werden mit den Prompts elaborative Strategien unterstützt, denn diese bewirken, dass neue Inhalte in die bestehenden Wissensstrukturen integriert und somit besser behalten und verstanden werden. Elaborationsstrategien bestehen im Wesentlichen darin, Verknüpfungen zwischen dem Lernstoff und dem eigenen Wissen (z.B. indem der Lerner eigene Analogien entwickelt) bzw. innerhalb des neuen Lerngegenstands (z.B. einzelner Teilkonzepte) zu bilden. Zur Formulierung der Prompts wird das Inventar zur Erfassung von Lernstrategien im Studium (LIST) von Wild und Schiefele (1994) herangezogen. Die Prompts, die in dieser Arbeit verwendet werden, werden in Kapitel 5 ausführlich dargestellt und begründet.

3.2 Konzeptueller und methodischer Hintergrund des Feedbacks

Feedback wird in der Lernpsychologie als grundlegende Komponente und wichtige Einflussgröße effektiver Lernprozesse erachtet, weil sie das zukünftige Verhalten von Personen determiniert (Bangert-Drowns et al., 1991; Butler & Winne, 1995;

Narciss, 2006). So wurden im Zusammenhang mit Feedback nicht nur Effekte auf die kognitive Leistung (z.B. Veränderung des Wissens durch die durch das Feedback angeregten Vergleichsprozesse zwischen Ist- und Soll-Zustand), sondern auch motivationale Auswirkungen beobachtet (vgl. Kluger & DeNisi, 1996). Letztere spielen für zukünftige Lernsituationen eine Rolle, weil sie z.B. die Erfolgserwartung des Lerners beeinflussen und die Art determinieren, wie ein Lerner an Aufgaben herangeht (z.B. Rheinberg et al., 2001).

Unter Feedback werden gewöhnlich Informationen verstanden, die einem Lerner in Folge einer Aufgabe präsentiert werden und diesem über die Qualität seiner Leistung (z.B. im Hinblick auf sein Lernverhalten oder die Qualität seiner Antwort) Aufschluss geben. Diese Informationen sind deshalb von Bedeutung, weil sie es dem Lerner ermöglichen, einen Abgleich zwischen dem eigenen Lernstand und externen oder selbstgewählten Lernzielen herzustellen.

Im Laufe der Zeit haben sich Feedbackformen herausgebildet, die unterschiedliche Funktionen verfolgen. Das Spektrum reicht hier von Formen, die auf die Motivation abzielen (z.B. Reattributionsfeedback) bis hin zu informativem Feedback, in dessen Zentrum Rückmeldungen zur Informationsverarbeitung und zur Korrektheit der Aufgabenbeantwortung stehen (vgl. Kulhavy & Stock, 1989). Darüber hinaus gibt es zahlreiche Klassifikationen von Feedback, bei denen andere Aspekte im Vordergrund stehen: z.B. die Quelle (externes und internes Feedback), der Zeitpunkt und die Inhalte.

Da für das Schreiben mit Prompts bislang keine systematischen Daten vorliegen, die Aufschluss darüber geben, inwiefern Feedback den Schreibprozess beeinflussen kann, werden für die vorliegende Arbeit zwei Feedbackformen gewählt. Das Ziel ist es, Studierende bei der Anwendung von Elaborationsstrategien beim Schreiben zu unterstützen. Durch das Feedback sollen die Lerner über die Qualität ihrer Texte informiert und die Strategien gefestigt werden. Zusätzlich soll das Feedback neben der kognitiven Funktion auch motivationale Funktion haben. Im Hinblick auf die Motivation wird bezweckt, „motivationale Dispositionen der Lernenden positiv zu beeinflussen“ (Huth, 2004, S. 5). Dies gelingt durch eine motivationsförderliche Gestaltung des Feedbacks, die auf empirischen Erkenntnissen beruht.

Die Konzeption des Feedbacks stützt sich auf zwei Modelle: das heuristische Modell zu Informativem tutoriellen Feedback von Narciss (2006), das zur Ablei-

tung von Gestaltungsprinzipien verwendet wird, und die SOLO-Taxonomie von Biggs und Collis (1982), die zur Zielvorgabe für die Studierenden und zur Beurteilung der Qualität der Texte dient.

3.2.1 Funktionen des Feedbacks

Für die systematische Gestaltung von Feedback sind nach Narciss (2006) drei Fragen zentral, die im Folgenden erläutert werden:

1. Was soll mit dem Feedback bewirkt werden (Feedbackfunktionen)?
2. Welche bereichsspezifischen Wissens Elemente sind notwendig, um die Aufgabe erfolgreich lösen zu können (Analyse der Aufgabenanforderungen)?
3. Welche Informationen sind für die Studierenden nützlich, um die angezielten Prozesse in anspruchsvoller Weise umzusetzen (Feedbackinhalte)?

Den Ausgangspunkt bildet die beabsichtigte Funktion: Was soll durch das Feedback auf die Prompts bewirkt werden? Narciss (2006) entwickelt hier drei verschiedene Funktionsebenen. In Anlehnung an Modelle des Selbstregulierten Lernens unterscheidet sie zwischen einer kognitiven, einer metakognitiven und einer motivationalen Ebene.

Feedback übt kognitive Funktionen aus, wenn es bei Problemen interveniert, die sich auf inhaltliche, prozedurale und strategische Wissens Elemente zurückführen lassen. Sind diese Wissens Elemente z.B. falsch verknüpft oder werden falsch angewendet, greift Feedback korrigierend, präzisierend, ergänzend oder restrukturierend (Umstrukturierung falsch verknüpfter Elemente) ein. In der vorliegenden Arbeit werden auf der kognitiven Ebene daher Rückmeldungen zur Qualität der angewandten (und geprompteten) Strategien gegeben, um diese zu festigen und zu fördern.

Auf die metakognitive Funktionsebene zielen Rückmeldungen ab, die das Augenmerk der Lernenden auf die Überwachung des Lernprozesses lenken oder die Informationssuche stimulieren. So können z.B. Hinweise zu Kriterien für die Überwachung des Lernprozesses gegeben werden, die mit den Lernzielen korrespondieren. In dieser Untersuchung wird die metakognitive Feedbackebene insbesondere mit dem Selbstfeedback angesprochen. Dabei werden die Lernenden dazu ange regt, die Qualität ihrer Leistung (z.B. der Antwort oder Aufgabe) anhand vorgegebener Kriterien selbst einzuschätzen.

Die motivationale Funktionsebene des Feedbacks ist als die komplexeste einzustufen, da sie durch ein Wechselspiel aus Aufgabenstruktur und individuellen Lernervoraussetzungen (z.B. Ziele, Selbstwirksamkeit und Attribuierungsmuster) geprägt ist. Sie sollte jedoch deshalb nicht vernachlässigt werden, denn man nimmt an, dass motivationale Erfahrungen beim Lernen die Herangehensweise des Lerners an zukünftige, ähnliche Aufgaben beeinflussen, z.B. hinsichtlich der Einschätzung der Aufgabenschwierigkeit, der Erfolgsaussicht und der Anstrengungsbereitschaft (z.B. Hattie & Timperley, 2007).

Anhaltspunkte zur Gestaltung motivationsförderlichen Feedbacks geben unter anderem die Arbeiten von Deci und Ryan (z.B. Deci & Ryan, 1985; Deci, Koestner & Ryan, 2001; Ryan & Deci, 2000). Diese haben sich intensiv mit Bedingungen der Lernmotivation auseinandergesetzt und betonen die Rolle des Kompetenzerlebens für die Lernmotivation. Nach Deci und Ryan (1985) wirkt sich Feedback dann motivationsförderlich (hier ist von intrinsischer Motivation die Rede) aus, wenn es vom Empfänger als nicht kontrollierend, sondern informativ interpretiert wird. Um die informative Funktion von Feedback zu stützen, sollten aufgabenspezifische Informationen darin enthalten sein, die relevant für den Kompetenzerwerb sind. Dazu zählen Informationen über den erreichten Leistungsstand im Vergleich mit Lernkriterien oder Lernzielen und die Rückmeldung über die individuelle Leistungsentwicklung (z.B. auch im zeitlichen Vergleich). Die Untersuchungen von Narciss (1999, 2001) unterstützen die Annahme, dass sich informatives Feedback positiv auf die Ausdauer und die Zufriedenheit der Lernenden auswirkt (zitiert aus Huth, 2004).

Nach Narciss und Huth (2004) sind informative Feedbackkomponenten

- „Hinweise, die den Abruf von Faktenwissen unterstützen
- Hinweise, die Analogien oder Fallbeispiele anbieten
- Hinweise auf relevante Informationsquellen
- Hinweise auf Korrekturstrategien
- Hinweise auf kognitive und metakognitive Strategien
- Erklärungen zu Art oder/und Ursache von Fehlern
- Erklärungen zu kognitiven oder metakognitiven Strategien

- Demonstration des Lösungsweges an einem Lösungsbeispiel
- Demonstration des korrekten Lösungsweges
- Leitfragen“

(Huth, 2004, S. 31)

Auf der motivationalen Ebene steht bei der vorliegenden Arbeit die Intention im Vordergrund, die Anstrengungsbereitschaft der Studierenden zu erhalten. Angesichts der zeitlichen Dauer der Prompting-Maßnahme und im Hinblick auf die Befunde von Nückles et al. (2006) scheint dies ein zentrales Anliegen. Feedback soll demzufolge motivationsförderlich gestaltet werden. Dies impliziert nach Narciss (2006) aufgabenspezifische Informationen, die den Lernenden Aufschluss über die Qualität ihrer Bearbeitung geben und den persönlichen Kompetenzzuwachs deutlich machen (internale Bezugsnorm) und solche Informationen, die die Aufgabenbearbeitung erleichtern, ohne jedoch bei Problemen die Lösung anzubieten. Stattdessen werden Hinweise gegeben, die eine eigenständige Bearbeitung ermöglichen.

Um die Feedbackinhalte zu bestimmen, schlägt Narciss (2006) eine Anforderungsanalyse vor. Anforderungsanalysen dienen im Vorfeld von Instruktionsangeboten der Bestimmung des Wissens und der (kognitiven) Fertigkeiten, die ein Lerner zur Bearbeitung einer Aufgabe benötigt. Für Lehrende sind sie von Bedeutung, weil auf ihrer Basis Rückmeldung geben können, inkorrekte Lösungen erklären und fehlerspezifische Informationen geben können, die den Lernenden bei der Lösung der Aufgabe helfen.

Um das Vorgehen zu illustrieren, wird im Folgenden eine Anforderungsanalyse, die exemplarisch an einem Prompt vorgenommen wurde, demonstriert. Als Orientierung diene die Vorgabe von Narciss (2006, S. 182-183).

Prompt 3:

Aus Ihrer Schulbiographie kennen Sie vielleicht Situationen, in denen Sie das Gefühl hatten, „nicht das Richtige gefragt worden zu sein“. Testtheoretisch handelt es sich hierbei um die Verletzung der Validität. Beschreiben oder konstruieren Sie ein solches Beispiel. Be-

halten Sie dabei die verschiedenen Formen der Validität im Hinterkopf.

Um diese Frage vollständig beantworten zu können, muss der Proband/die Probandin in der Lage sein,

1. zu wissen, dass „pädagogische“ Fragen bzw. Unterrichtsfragen einen bestimmten Zweck erfüllen sollen, nämlich zu einer bestimmten Erkenntnis zu führen.
2. beurteilen zu können, ob mit einer bestimmten Frage ein bestimmtes Problemfeld angesprochen wird bzw. ob sie zielführend ist.
3. sich an konkrete Situationen der Schulzeit zurück zu erinnern.
4. sich an mindestens eine Situation der Schulzeit zurück zu erinnern, in der nicht „das Richtige“ gefragt wurde, d.h. in der die Frage nicht so gestellt wurde, dass sie auf den Sachverhalt hinführte.
5. zu wissen, was mit Validität im Rahmen der psychologisch-diagnostischen Testtheorie gemeint ist.
6. zu wissen, welche Formen der Validität es gibt (Inhalts-, Kriteriums-, Konstruktvalidität etc.).
7. zu wissen, was die verschiedenen Validitätsformen theoretisch-inhaltlich bedeuten.
8. zu wissen, wie diese Validitätsformen auf die Praxis anzuwenden sind bzw. wie sie sich in der Praxis erkennen lassen
9. verschiedene Validitätsformen auf das ausgesuchte Schulbeispiel retrospektiv anwenden zu können
10. die Validitätsformen richtig auf das ausgesuchte Beispiel anwenden zu können.

3.2.2 Feedbackinhalte und Konstruktion des Feedbacks

Im folgenden Abschnitt werden zwei Formen des Feedbacks konstruiert, die auf den vorgestellten Überlegungen und den daraus abgeleiteten Empfehlungen basieren. Es handelt sich dabei zum einen um Selbstfeedback und zum anderen um Dozentenfeedback. Das Dozentenfeedback setzt sich aus drei Komponenten zusammen, das Selbstfeedback wird daran anschließend gesondert beschrieben.

- Evaluative Komponente
- Informierende Komponente
- Instruierende Komponente

Feedback Variante 1: Dozentenfeedback

In der evaluativen Komponente des Feedbacks wird das individuelle Leistungsniveau, das der Lerner erreicht hat, zurückgemeldet. Dieses wird anhand der SOLO-Taxonomie (eine Beschreibung findet sich in Abschnitt 5.1) bestimmt, mit der die Texte analysiert werden. Da Feedback nur in Verbindung mit Lernzielen einen Informationswert hat (Narciss, 2006), werden in einer Instruktion der Sinn der Prompts und die Gütekriterien für die Einstufung der Texte erläutert.

Die informative Komponente des Feedbacks kommt nur dann zum Tragen, wenn ein Text auf dem ersten oder zweiten Niveau der SOLO-Taxonomie und damit als nicht optimal eingestuft wird. Dem ersten Niveau der SOLO-Taxonomie wird ein Text dann zugeordnet, wenn er keine für die Aufgabe relevanten Komponenten enthält. Dies kann durch Fehlkonzepte oder zahlreiche andere Gründe zustande kommen: „Die zur Lösung der Aufgaben notwendigen inhaltlichen, prozeduralen oder auch strategischen Wissensselemente können fehlen, falsch oder ungenau sein. Des Weiteren können die notwendigen Wissensselemente falsch verknüpft sein oder die Bedingungen für ihre Anwendung falsch oder unklar sein“ (Narciss, 2006, S. 78). In diesem Fall stellt das Feedback Informationen bereit, die ergänzend, korrigierend oder präzisierend gestaltet sind.

Auf dem zweiten Niveau eingestufte Texte enthalten lediglich einzelne, unverbundene Aspekte eines Themas. Sie ähneln eher Aufzählungen einzelner Wissensselemente und spiegeln kein zusammenhängendes Verständnis eines Konzeptes wider. Ähnlich wie bei Texten, die dem ersten Niveau zugeordnet werden,

können hier verschiedene Gründe zum Tragen kommen, die zu einer geringen Verarbeitungstiefe der Lerninhalte führen. Wenn aus dem Text ersichtlich wird, dass die Aufgabe nicht verstanden wurde oder strategische Wissens Elemente fehlen, verweist das Feedback auf das Musterbeispiel, in dem die Bearbeitung eines Elaborationsprompts schrittweise erläutert wird. Das Feedback könnte dann folgendermaßen aussehen:

„Sie haben in Ihrer Antwort auf die Lernfrage überwiegend zusammengefasst und einzelne Aspekte aufgezählt (hier die einzelnen Strategien aufzählen (= evaluative Komponente)). Um den Lernstoff besser zu verstehen, ist es wichtig, dass Sie auch eigene Beispiele formulieren, das neue Wissen mit ihren eigenen Erfahrungen verknüpfen etc. Schauen Sie sich dazu noch einmal das Lösungsbeispiel an. (Da Feedback nur effektiv ist, wenn im Anschluss daran weiter an dem Thema/der Aufgabe gearbeitet werden kann, sollte es auch diese einbeziehen:) Vorschlag: Notieren Sie sich zwei Strategien, die Sie bei der nächsten Lernfrage einsetzen werden.“

Für den Fall, dass ein Mangel an inhaltlichem Wissen oder Fehlkonzepte deutlich werden, greift das Feedback informierend ein, wie es für Texte auf dem ersten Niveau beschrieben wurde. Es werden dann Hinweise zur selbstständigen Nachbearbeitung des Themas gegeben, wie etwa Analogien, Erklärungen, Lösungsbeispiele. Das Dozentenfeedback wird jeweils nach der Bearbeitung eines Prompts bereitgestellt.

Feedback Variante 2: Das Selbstfeedback

Im Rahmen des Selbstfeedbacks werden den Lernenden, gleichzeitig mit der dozentenorientierten Rückmeldung der anderen Lernjournale, Leitfragen zur Bewertung ihrer eigenen Texte zur Verfügung gestellt. Als Bezugsrahmen für das Selbstfeedback werden die Fragen der Softwareanwendung eHELP (Nückles et al., 2004) verwendet. eHELP ist eine an der Universität Freiburg konzipierte elektronische Lernumgebung, in der Lernprotokolle erstellt und verwaltet werden können. Die Software unterstützt Lerner nicht nur durch Strukturvorgaben für ein Lernprotokoll, sondern bietet zusätzlich Fragen an, die bei der Revision der Protokolle helfen. Diese werden in der vorliegenden Arbeit für das Selbstfeedback verwendet.

Mit dem Selbstfeedback wird ein metakognitives Element in den Schreibprozess eingebaut. Die Studierenden werden dazu aufgefordert, sich mit der Qualität

ihrer Arbeit auseinanderzusetzen und deren Qualität einzuschätzen. Anhand dreier Fragen aus eHElp werden die Studierenden nach jeweils einem Prompt dazu aufgefordert, ihren eigenen Text entlang der Fragen einzustufen und im Anschluss an die Selbstdiagnose der eigenen Texte ein Ziel für die Bearbeitung der nächsten Prompts zu formulieren. Ein Beispiel für eine Instruktion zum Selbstfeedback sieht so aus:

„Liebe Studierende,
vielen Dank für die Einreichung Ihres ersten Prompts. Um Ihr Lernen zu optimieren, analysieren Sie bitte anhand der folgenden drei Fragen Ihr erstes Lernprotokoll.

1. Ich habe bei meinen Beispielen kritisch geprüft, ob sie begründet und belegt sind.
2. Die Struktur meines Lernprotokolls gibt die Struktur des Stoffes meiner Meinung nach wieder.
3. Ich habe versucht, mehrere Aspekte des Themas miteinander zu verbinden.

Notieren Sie dann auf der Basis dieser Einschätzung ein Ziel für die Bearbeitung des nächsten Prompts.

Die kompletten Fragen zum Selbstfeedback werden im Anhang (.2 Selbstfeedbackfragen) aufgeführt.

3.3 Das Lösungsbeispiel

Lösungsbeispiele werden klassischerweise im naturwissenschaftlichen Unterricht eingesetzt. Darunter werden ganz allgemein Aufgaben verstanden, die exemplarisch gelöst werden, wobei auf einzelne Schritte des Lösungsweges eingegangen wird. Das Schreiben mit Prompts stellt für die Lerner eine neue Arbeitsform dar. Es ist insofern anspruchsvoll, als dass es ein relativ freies Schreibformat ermöglicht. Wissenschaftliche Inhalte müssen dabei mit eigenem Alltagswissen verknüpft werden, was eine Transformation des Lernstoffes in die Alltagssprache erfordert. Das Ziel des Lösungsbeispiels in der vorliegenden Arbeit ist es,

zu demonstrieren, wie dies gelingen kann. Da die Lerner zu einer vertieften Verarbeitung der Lerninhalte angeregt werden sollen, wird ein Text gewählt, der anspruchsvolle Prozesse, die dem dritten Niveau der SOLO-Taxonomie zugeordnet werden, enthält. Das Lösungsbeispiel wird über die Dauer der Intervention auf der Lernplattform zur Verfügung stehen.

Renkl, Schworm und Hilbert (2004) postulieren, dass Lösungsbeispiele am effektivsten sind, wenn sie vor der Lernepisode eingeführt werden. Dies soll gewährleisten, dass das Potenzial des Lösungsbeispiels genutzt werden kann, indem verhindert wird, dass Lerner sich nur oberflächlich damit auseinandersetzen. Aus diesem Grund wird das Lösungsbeispiel zusammen mit den Prompts in der zweiten Untersuchung eingeführt.

4 Fragestellungen

Gegenstand des folgenden Kapitels sind die Forschungsfragen dieser Arbeit. Im Mittelpunkt des empirischen Teils steht die Erprobung einer Prompting-Maßnahme zur Unterstützung kognitiver Prozesse beim Schreiben. In einer ersten Untersuchung wird zunächst geprüft, inwiefern sich die konzipierten Prompts zur Aktivierung kognitiver Lernstrategien beim Schreiben eignen. Dabei interessiert zum einen, welche kognitiven Prozesse durch die Prompts angeregt werden und zum anderen, inwiefern diese Prozesse in Zusammenhang mit Merkmalen der Prompts stehen. Zu diesem Zweck kommen Prompts zum Einsatz, die sich zum einen auf verschiedene Kontexte beziehen (z.B. auf Lerninhalte aus anderen Seminaren, auf die Berufspraxis) und zum anderen Strategien aus dem Bereich der Elaborationsstrategien ansprechen (z.B. Verknüpfungen zwischen den Inhalten herstellen, Beispiele formulieren).

Ein weiterer Aspekt, der im Zusammenhang mit der Bearbeitung von Prompts in Lerntagebüchern und Portfolios noch weitgehend offen ist, betrifft die Frage, inwieweit spezifische Voraussetzungen auf Seiten der Lernenden auf die Qualität des Schreibens Einfluss haben, bzw. welchen Einfluss das Lernen durch Schreiben auf die nachfolgenden Aktivitäten der Lernenden hat. Prinzipiell ist bei den Voraussetzungen für das Schreiben an Lerner-Variablen wie Vorwissen und Schreibfertigkeit zu denken, aber auch an kognitive und metakognitive Kompetenzen, Selbstregulationskompetenzen und Motivation. Da ein schlüssiges theoretisches Modell zum Lernen durch Schreiben noch nicht vorliegt, aus dem diese Variablen konsistent abzuleiten wären, wird in der zweiten Studie die Motivation in den Blick genommen. Aus verschiedenen Experimenten liegen Hinweise dazu vor, dass Motivationskomponenten einen Einfluss darauf haben, wie sich Lernende im Lernprozess engagieren und welche Lernleistung sie erzielen (vgl. Rheinberg et al., 2001). In der zweiten Untersuchung soll daher geprüft werden, ob Unterschiede in den motivationalen Voraussetzungen auch zu unterschiedlichen Verarbeitungstiefen beim Schreiben mit den Prompts führen.

Überdies kommen Nückles et al. (2006) zu dem Ergebnis, dass sich Prompts

über einen längeren Zeitraum negativ auf die Motivation auswirken. Daher ist auch von Interesse, inwiefern sich dieser Befund replizieren lässt.

Eine weitere Fragestellung der zweiten Untersuchung bezieht sich auf die Gestaltung der Lernumgebung. In dieser Hinsicht interessiert, inwiefern die Lernenden von Feedback, das die Qualität ihres Schreibproduktes informiert und in den Schreibprozess eingebunden ist, profitieren. Um dieser Frage nachzugehen, werden verschiedene Formen des Feedbacks hinsichtlich der Effekte auf die Verarbeitungstiefe und die Lernmotivation erprobt.

4.1 Studie 1: Evaluation der Prompts und des Analyse-Instrumentes

4.1.1 Fragestellungen

Die in der ersten Studie formulierten Fragestellungen beziehen sich auf die Konstruktionsweise der Prompts zur Förderung kognitiver Strategien beim Schreiben und das Analyse-Instrument (SOLO-Taxonomie), mit dem die Komplexität der Texte eingestuft wird. Aus den bisherigen Studien zeigt sich, dass mit Prompts Teilprozesse des Lernens (z.B. Planen und Revidieren) unterstützt werden können und Lernstrategien angeregt werden. Es liegen bislang jedoch keine differenzierten Aussagen zur Qualität der Prozesse vor, die mit den Prompts angeregt werden und aus der Textstruktur erkennbar werden. Zur Aufklärung dieses Sachverhaltes werden die mit den Prompts verfassten Texte mithilfe einer Taxonomie zur Klassifizierung von Lernergebnissen hinsichtlich ihrer Komplexität eingestuft. Dies soll Aufschluss darüber geben, zu welcher Verarbeitungstiefe und zu welcher Wissensstruktur unterschiedliche Prompts führen.

1. Sind die konzipierten Prompts geeignet, um kognitive Strategien zu initiieren?

Wenn durch Prompts unterschiedliche Prozesse angeregt werden können, kann man annehmen, dass eine inhaltliche Differenzierung der Prompts ebenfalls zu unterschiedlich komplexen Wissensstrukturen bzw. Verarbeitungstiefen führt. Um dies zu explorieren, werden sechs Kategorien von Prompts konzipiert. Es wird erwartet, dass die Prompts, die ähnliche Anforderungen aufweisen, zu annähernd kongruenten Verteilungsmustern bei der Einstufung der Texte führen. Damit

könnte gezeigt werden, dass die Ausgestaltung der Prompts entscheidend für die Qualität der Prozesse, die initiiert werden sollen, ist.

2. Lässt sich ein Zusammenhang zwischen der Verarbeitungstiefe in den Texten und spezifischen Merkmalen der Prompts feststellen?

Die dritte Fragestellung nimmt das Analyse-Instrument in den Blick, das erstmals verwendet wird, um Aussagen über die Zusammenhänge zwischen den Prompts und der Qualität der Wissensstruktur zu machen, die mit den Prompts erreicht wird. Betrachtet man die bisherigen Studien, so wird zum einen die Häufigkeit bestimmter Strategien erfasst und zum anderen das Lernergebnis, das in einer Anschlussklausur erreicht wird. Die Prompts werden jedoch nicht in direkten Bezug gesetzt mit der Qualität des Wissens, das mit den eingesetzten Strategien sichtbar wird. In der vorliegenden Arbeit werden mithilfe der SOLO-Taxonomie Merkmale der Prompts und die Qualität des Wissens in Relation zueinander gesetzt, um Empfehlungen für die optimale Gestaltung von Prompts machen zu können. Daher gilt die dritte Fragestellung der Exploration des Instrumentes:

3. Kann die SOLO-Taxonomie aussagekräftige Kategorien für die Einstufung der Texte finden?

4.2 Studie 2: Effekte verschiedener Feedbackformen auf die Verarbeitungstiefe beim Schreiben und auf die Motivation

Die zweite Studie fokussiert auf Wechselwirkungen zwischen Motivation, Verarbeitungstiefe und Feedback. Motivation ist im Kontext von Prompts dann von Interesse, wenn solche Daten Rückschlüsse auf die Gestaltung einer anregenden und leistungsförderlichen Lernumgebung erlauben. Innerhalb der Schreibforschung wurde bislang insbesondere untersucht, inwiefern Interesse und Selbstwirksamkeit sich im Rahmen von Schreibaufträgen auf die Lernleistung und auf die Qualität des Schreibens auswirken (vgl. Hidi, Renninger & Krapp, 2004). Interesse an einem Gegenstand erleichtert demzufolge das Interesse am Schreiben, ist jedoch keine ausreichende Bedingung für eine höhere Schreibmotivation (Boscolo & Cisotto, 1997). Besonders starke Befunde ergaben sich im Zusammenhang mit Selbstwirksamkeit. Das Konzept der Selbstwirksamkeit beschreibt die positive

Erwartungshaltung einer Person, Handlungen aufgrund der eigenen Fertigkeiten oder des eigenen Wissens erfolgreich durchführen zu können. Ein detaillierter Überblick über die Studien findet sich bei Pajares und Valiante (2006). Selbstwirksamkeit stellte sich in unterschiedlichen Studien als bedeutsamer Faktor für die Anstrengungsbereitschaft heraus und ist ein Prädiktor für Schreibkompetenz und die Anwendung von Schreibstrategien (Pajares & Valiante, 2006; Schunk & Swartz, 1993). Zudem zeigten sich Zusammenhänge zwischen Selbstwirksamkeit und der Qualität verschiedener Texte (z.B. Essay, Erzählung).

Aus den bisherigen Studien wurde die Bedeutung motivationaler Aspekte herausgearbeitet. Folgerichtig wird in dieser Arbeit überprüft werden, inwiefern Feedback sich als Bestandteil der Lernumgebung auf die Motivation auswirkt. Motivation wird in diesem Kontext zum einen als Einflussgröße auf die Qualität der Texte und zum anderen als Lerner-Variable, die durch die Qualität der Textproduktionsprozesse beeinflusst wird, untersucht. Es wird also der Frage nachgegangen, ob durch Feedback Einfluss auf die Qualität des Schreibprozesses und auf die Motivation der Studierenden genommen werden kann.

4.2.1 Fragestellungen

Feedback wird in der Lernpsychologie als grundlegende Komponente und wichtige Einflussgröße effektiver Lernprozesse erachtet, weil man davon ausgeht, dass sie das zukünftige Verhalten des Lerners determiniert (vgl. Narciss, 2006) und die Lernleistung positiv beeinflusst. Zahlreiche Studien befassen sich mit den Effekten von Feedback auf akademische Leistung und kommen zu dem Ergebnis, dass Feedback als ein komplexes Zusammenspiel aus Voraussetzungen des Lerners (z.B. Selbstwirksamkeit), der Feedbackart, der Feedbackebene, dem Aufgabenkontext und dem Zeitpunkt des Feedbacks zu verstehen ist, das sowohl leistungsförderliche als auch leistungsmindernde Effekte haben kann (Bangert-Drowns et al., 1991; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Narciss, 2006). Diese Annahme basiert auf einer sehr heterogenen Befundlage der untersuchten Feedbackinterventionen (z.B. Hattie & Timperley, 2007) und lässt den Schluss zu, dass bei der Gestaltung einer Feedbackintervention, soweit möglich, eine genaue Abstimmung dieser Faktoren vorzunehmen ist.

Auch in Zusammenhang mit Schreibinstruktionen scheint Feedback nicht einheitlich positive Effekte zu bewirken. Die Metastudie von Bangert-Drowns et al.

(2004), die die Effekte von Schreibinstruktionen auf die Lernleistung in den Blick nimmt, deutet zwar auf substanzielle Effektstärken hin, allerdings fallen diese nicht signifikant aus. Zudem weisen die Autoren auf eigene methodische Probleme bei der Kodierung des Faktors hin und räumen ein, damit möglicherweise ein unklares Ergebnis erzeugt zu haben. Es lassen sich aus dieser Studie demnach keine gesicherten Schlüsse über die Wirksamkeit von Feedback ableiten. In der vorliegenden Arbeit werden daher auch drei Formen des Feedbacks (Selbstfeedback, Dozentenfeedback und kein Feedback) in Kombination mit der Prompting-Maßnahme untersucht. Das Ziel ist es, zu differenzierten Aussagen darüber zu gelangen, inwiefern und wenn ja, auf welcher Ebene, sich Schreibprozesse durch Rückmeldung optimieren lassen. Aus diesen Überlegungen werden für die Untersuchung im folgenden Abschnitt Fragestellungen entwickelt.

4. Welche Unterschiede bestehen bei drei Gruppen von Studierenden (A, B und C), die in unterschiedlicher Reihenfolge drei Feedbackvarianten erhalten, in Bezug auf die Motivation nach den unterschiedlichen Feedbackformen (Treatments)?

Es wird erwartet, dass die Motivation der Seminarteilnehmer, die Schreibinstruktionen zu bearbeiten, in Abhängigkeit der Feedbackform variiert. In Anlehnung an Bangert-Drowns et al. (2004) wird vermutet, dass sich insbesondere das Dozentenfeedback positiv auf die Motivation der Studierenden auswirkt und zu einer Steigerung des Interesses an der Bearbeitung der Prompts beiträgt. Da die Fragestellung in erster Linie explorativ ist, soll desweiteren geklärt werden, welche Feedbackvariante sich im Zusammenhang mit Schreibinstruktionen als besonders effektiv auszeichnet.

5. Welche Feedbackvariante ist hinsichtlich der Motivation der Studierenden die günstigste?

Feedback wirkt nicht nur auf einer affektiven, sondern im Idealfall auf der kognitiven Ebene und hilft dem Lerner dabei, eine Aufgabe erfolgreicher zu bearbeiten, indem er Hinweise zur Verbesserung seines Lernprozesses erhält. Nach Hattie und Timperley (2007) ist Feedback, das sich auf die Durchführung der Lerntätigkeiten bezieht, besonders effektiv und zwar insbesondere, wenn es mit Zielen verknüpft wird. Es wird erwartet, dass sich die Effekte des Feedbacks auf der kognitiven

Ebene in einer Verbesserung der Textqualität, die mit der SOLO-Taxonomie eingestuft wird, feststellen lassen. Die Studierenden erhalten auf der Grundlage dieser Taxonomie Rückmeldungen zu der Qualität ihrer Texte und erhalten, wenn nötig, zudem Informationen, wie sie ihr Schreiben durch die Anwendung weiterer, aus der Taxonomie abgeleiteter, Arbeitsschritte verbessern können. Da die Feedbackvarianten in jeder Gruppe drei Mal variieren, wird vermutet, dass auch die Textqualität in Abhängigkeit des Feedbacks Veränderungen unterliegt. Aus dieser Annahme ergeben sich folgende Fragestellungen:

6. Führen die unterschiedlichen Feedbackvarianten zu Unterschieden in der Verarbeitungstiefe?
7. Welche Feedbackvariante ist hinsichtlich der Verarbeitungstiefe der Studierenden die günstigste?

Die letzte Fragestellung der Arbeit bezieht sich auf die Wirksamkeit der gesamten Prompting-Maßnahme. Es interessiert, ob sich die Leistung der Studierenden durch die Bearbeitung der Prompts verbessert.

8. Gibt es einen Zusammenhang zwischen der Verarbeitungstiefe der von den Studierenden produzierten Texte und dem Lernzuwachs?

5 Eigene empirische Untersuchungen

Im empirischen Teil wird am Beispiel einer studentischen Lerngruppe die Frage untersucht, inwiefern Prompts die Qualität der Schreibprozesse und damit auch der Texte beeinflussen. Dabei wird auch die Rolle von Feedback und motivationalen Komponenten im Schreibprozess beleuchtet. Wie in Abschnitt 2.2 erläutert, wird davon ausgegangen, dass Prompts zur Aktivierung von Lernstrategien eingesetzt werden können. Der Einsatz von Lernstrategien sollte sich wiederum auf die Qualität der Texte auswirken, die über die Komplexität der in diesen Texten präsentierten kognitiven Strukturen operationalisiert wird, und könnte auch die Lernleistung positiv beeinflussen. Um diese Annahmen zu testen, wurden Prompts entwickelt, mit denen Elaborationsstrategien initiiert werden sollten. Zur Bearbeitung der Forschungsfragen wurden zwei aufeinander aufbauende Untersuchungen durchgeführt, die Gegenstand dieses Kapitels sind.

Das Ziel in der ersten Untersuchung war es, zu testen, inwiefern sich die Prompts zur Förderung kognitiver Prozesse eignen und wie gut die Analyse der Textqualität mit einer Taxonomie zur Beschreibung von Lernergebnissen gelingt. Dazu bearbeiteten Studierende im Rahmen eines Blockseminars über den Zeitraum von einer Woche insgesamt 16 Prompts.

Das Ziel der zweiten Untersuchung war es, zu überprüfen, inwiefern sich Feedback auf die Qualität der Texte und die Motivation von Textproduzenten auswirkt. Um dies zu erproben, schrieben Studierende über einen Zeitraum von drei Monaten mit den aus der ersten Untersuchung als strategieförderlich getesteten Prompts. Zusätzlich wurden verschiedene Formen von Feedback hinsichtlich der Wirksamkeit auf die Motivation und die Qualität der Textproduktion getestet. Um die Qualität der Texte quantitativ abzubilden, kommt ein Analyseinstrument zum Einsatz, mit dem die den Texten inhärenten kognitiven Strukturen ermittelt und hinsichtlich ihrer Komplexität eingestuft werden. Zur Erhebung der Motivation wird der FAM (Rheinberg et al., 2001) verwendet.

Die Untersuchungen sind in thematisch verschiedenen Inhaltsgebieten angesiedelt und erforderten daher eine inhaltliche Anpassung der Prompts an die

Lernthemen. Die erste Untersuchung fand in einer einführenden Veranstaltung für Lehramtstudierende aus dem Bereich der Pädagogischen Psychologie (Einführung in die Sozialpsychologie) statt. Die zweite Untersuchung erfolgte in einer Einführungsveranstaltung in empirische Methoden für Studierende der Erziehungswissenschaft. Insgesamt versucht die vorliegende Arbeit zu klären, welche Charakteristika des Schreibprozesses mithilfe von Prompts gefördert werden können, wie sich die Textqualität mittels verschiedener Formen von Feedback beeinflussen lässt und welche Rolle motivationale Aspekte dabei spielen. Im Folgenden werden zunächst das Verfahren zur Analyse der Texte und im Anschluss daran die beiden aufeinander aufbauenden Untersuchungen beschrieben.

5.1 Beschreibung und Analyse der Lernergebnisse

Im Rahmen von zwei Untersuchungen werden Prompts eingesetzt, um die Komplexität der kognitiven Prozesse beim Verfassen von Lerntagebüchern und Portfolios zu fördern. Damit ist die Erwartung verbunden, dass die Texte, die dabei entstehen, entsprechend anspruchsvoll sind und tiefer greifende, komplexe Lernerfahrungen repräsentieren. Da über die Qualität der mit den Prompts angeregten Texte noch kaum empirische Daten zu finden sind (vgl. Bannert, 2009), ist es das Ziel der nachfolgenden Untersuchungen, die schriftlichen Ausführungen der Untersuchungsteilnehmer auf die kognitiven Strukturen hin zu analysieren, die daraus erkennbar werden. Die Strukturen, die im Text identifiziert werden, werden wiederum als Indikatoren für die kognitiven Prozesse während der Textproduktion interpretiert und die Komplexität der Textstrukturen gibt Aufschluss über den Elaborationsgrad der vorausgehenden kognitiven Prozesse.

Für die Inhaltsanalyse der Texte wird die SOLO-Taxonomie von Biggs und Collis (1980) verwendet, ein Modell zur Beschreibung von Lernergebnissen, das die fachübergreifende Einstufung von Lernergebnissen auf ihre Komplexität hin ermöglicht. Unter dem Begriff der Inhaltsanalyse werden verschiedene Verfahren zur Analyse von schriftlichen Dokumenten zusammengefasst, die dem Zweck dienen, Daten zu vergleichen, zu kontrastieren und zu kategorisieren (Schwandt, 2007). Je nach Fragestellung kommen dabei sowohl interpretative und numerische Verfahren der Datengenerierung als auch Kombinationen beider Vorgehensweisen zum Einsatz.

Die SOLO-Taxonomie basiert auf der Annahme, dass die fortschreitende Entwicklung eines Lerners in einem Lernfach sich in zunehmend komplexeren und abstrakteren Wissensstrukturen abbildet. Dies sollte sich auch in den Texten der Untersuchungsteilnehmer widerspiegeln. Biggs und Collis (1980) postulieren fünf hierarchisch organisierte Niveaus, auf denen sich Lernergebnisse hinsichtlich ihrer Komplexität einstufen lassen. Die Niveaus repräsentieren die Entwicklung themenspezifischen Wissens und beginnen mit jedem Lernzyklus erneut. Je weiter die Entwicklung des Wissens fortschreitet, desto mehr Anforderungen an das Arbeitsgedächtnis und an die Aufmerksamkeitsspanne werden vorausgesetzt. Auf den verschiedenen Niveaus wird dies so reflektiert, dass das Lernergebnis, je höher es in der SOLO-Hierarchie eingestuft wird, umso mehr Aspekte eines Themenfeldes berücksichtigen und umso mehr Vernetzungen zwischen einzelnen Wissens-elementen aufweisen muss. In der SOLO-Taxonomie wird Wissensfortschritt also über die zwei Dimensionen „Anzahl der Vernetzungen“ und „Elaborationsgrad der Vernetzungen“ operationalisiert. Fortgeschrittenere Niveaus implizieren jeweils die vorausgegangenen.

Dem ersten Niveau (*pre-structural*) werden Antworten zugeordnet, die keine relevanten Komponenten oder solche ohne logische Beziehung zur Lernaufgabe enthalten. Texte auf dem zweiten Niveau (*uni-structural*) enthalten lediglich einen Aspekt eines Themas, ohne diesen in Bezug zu anderen zu setzen oder weitere relevante Elemente zu berücksichtigen. Antworten auf dem dritten Niveau (*multi-structural*) zeichnen sich dadurch aus, dass sie mehrere relevante Einzelkomponenten eines Themas auflisten, die allerdings nicht eigenständig strukturiert und auch nicht in einen größeren Zusammenhang integriert werden. Bis zu diesem Punkt ist der Lernprozess durch eine quantitative Zunahme einzelner Wissens-elemente gekennzeichnet, mit denen der Lerner bislang keine anspruchsvollen kognitiven Prozesse durchführen kann, sondern im Bereich des additiven Wissensaufbaus bleibt. Das ändert sich bei Lernergebnissen, die dem vierten Niveau (*relational*) zugeordnet werden. Um auf dem vierten Niveau eingestuft zu werden, muss aus einem Text erkennbar sein, dass vielfältige Aspekte eines Themenkomplexes benannt, miteinander verknüpft und gegeneinander abgewägt werden, sodass der Autor zu kohärenten Schlussfolgerungen gelangt. Auf dem fünften Niveau (*extended abstract*) schließlich zeigt der Lerner, dass er relevante Aspekte eines Themas miteinander in Beziehung setzt und diese eigenständig auf einen neuen Kontext

oder ein neues Phänomen überträgt. Ein Text, der diesem Niveau zugeordnet wird, enthält typischerweise Verallgemeinerungen über die Beispiele im Seminar hinaus, eigene Hypothesen sowie kritische Reflexionen eines Konzeptes, z.B. mit Hinblick auf Anwendungsfragen. Die Textbeispiele in Tabelle 1 illustrieren die verschiedenen Niveaus anhand von Ankerbeispielen.

Tabelle 1: Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)

Niveau	Prompt und Textbeispiel
Niveau 1: pre-structural	<p data-bbox="411 340 481 1482">Prompt: „Inwiefern weist der Lernstoff eine Relevanz für Ihr Alltagsleben auf?“</p> <p data-bbox="481 340 849 1482">Neuer Lernstoff hat eine starke Relevanz für mein Alltagsleben. Erstens wird er von mir über den Tag und zum Teil auch über Nacht verarbeitet und trägt zu Diskussionen in der Familie und mit Freunden bei. Zweitens wird neu Erlerntes auch teilweise mit in den Alltag eingebracht oder eingeflochten. Für mich persönlich ist es wichtig, den neu vermittelten Lernstoff im Alltag direkt anzuwenden oder für mich Situationen herzustellen, in denen ich den Lernstoff anwenden kann. Da ich selbst zwei Kinder habe, kann ich das neu Vermittelte auch oft auf sie reflektieren und in mancher Vorlesung geht mir dann auch mal ein Licht auf. Oft denke ich mir dann, ach wie schade, dass ich das nicht schon früher erfahren habe, es hätte mir Einiges bei der Erziehung meiner Kinder erleichtert oder erklärt (zum Beispiel im Exkurs Einführung in die Psychologie).</p>
Begründung	<p data-bbox="865 340 1082 1482">Der Text wurde auf Niveau 1 eingestuft, weil kein Bezug zum konkreten Inhalt des Lernstoffes genommen wird. Warum der spezifische Lernstoff wichtig ist oder welche konkreten Schlussfolgerungen für das Alltagsleben entstehen, bleibt unerläutert. Es wird nicht einmal minimal erlerntes Wissen wiedergegeben und insgesamt wird die Frage verfehlt. Um auf Niveau 2 eingestuft zu werden, müsste zumindest minimalistisches Verständnis oder eine Wiedergabe einzelner Aspekte des Gelernten stattfinden.</p>
Niveau 2: uni-structural	<p data-bbox="1098 340 1310 1482">Prompt: „Welche Anknüpfungspunkte zwischen dem Gelernten und aus anderen Seminaren/Fächern bereits bekannten Theorien, Methoden oder Befunden sind Ihnen aufgefallen?“ Ich habe bereits in einem anderen Seminar vom „Basking-in-reflected-glory-effect“ gehört. In diesem Zusammenhang ist mir daher bereits bekannt, dass durch die Identifikation von Menschen mit einer starken Gruppe auch die Selbstwertschätzung des Einzelnen steigt.</p>

Tabelle 1: Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)

Begründung

Der Text wurde auf Niveau 2 eingestuft, weil ein minimalistisches Verständnis für eine Komponente des Erlernten erkennbar ist und wiedergegeben wird. Eine Antwort auf dem dritten Niveau würde weitere Aspekte des Lernthemas und mehr Details enthalten.

Niveau 3: multi-structural

Prompt: Welchen Aspekt des Gelernten fanden Sie überraschend / kontraintuitiv und warum?

Fassen Sie diesen kurz und prägnant zusammen.

Überraschend fand ich die Aspekte der Eindrucksbildung, des Reihenfolgeeffekts und der zentralen Rolle der Eigenschaften „warm“ und „kalt“ bei der Wahrnehmung von Personen. Ich fand es interessant, dass Menschen anhand einiger vorgegebenen Eigenschaften einer Person einen abgerundeten allgemeinen Eindruck von dieser bilden. Erstaunlich fand ich, dass die Reihenfolge, in der diese Eigenschaften gegeben werden, einen so großen Einfluss auf die anschließende Eindrucksbildung ausübt. Diesen Reihenfolgeeffekt fand ich sehr interessant, weil ich etwas darüber erfahren habe, wie man unbewusst, ohne dies eigentlich zu beabsichtigen, bei seiner Eindrucksbildung über andere Personen „manipuliert“ wird. Das Ergebnis des Tests, den wir im Seminar durchgeführt haben, hat eindeutig belegt, dass man sich bei seiner Bewertung eines Menschen an den ersten gegebenen Eigenschaften orientiert. Obwohl beide Gruppen dieselben Eigenschaften vorgegeben hatten, nur in anderer Reihenfolge, kamen zwei entgegengesetzte Einschätzungen heraus. Genauso überraschend fand ich die Auswirkungen auf eine Eindrucksbildung der Eigenschaften „warm“ und „kalt“. Ich hätte nicht gedacht, dass diese beiden Eigenschaften so viel mehr Einfluss auf die Eindrucksbildung haben als alle anderen Eigenschaften, und dass sie sogar unser Verhalten beeinflussen (wie in der Untersuchung, in welcher ein Diskussionsleiter den Studenten entweder als unter anderem „kalt“ oder „warm“ vorgestellt wurde, und sich an der anschließenden Diskussion im Fall „warm“ 24 % mehr Studenten beteiligten als im Fall „kalt“).

Table 1: Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)

Begründung

Bei dieser Antwort ist das Verständnis der beiden dargestellten Komponenten erkennbar. Außerdem wird eine sehr detaillierte Beschreibung des Erlernten vorgenommen. Um auf Niveau 4 eingestuft zu werden, müsste aus dieser Antwort ein Verständnis für den Zusammenhang der verschiedenen Bereiche erkennbar werden, Gründe für die gefundenen Ergebnisse genannt oder das Konzept auf bekannte Situationen übertragen werden.

Tabelle 1: Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)

Niveau 4: relational

Prompt: Was wussten Sie schon vorher über das Thema und wo können Sie Bezüge herstellen?

Aus dem Alltagsleben kennt natürlich jede Person die ansteigende Erregung in Gegenwart von anderen Personen, vor allem natürlich bei Präsentationen. Jeder kennt die Situation vor einem Publikum zu stehen und kein Wort mehr aus dem Mund zu bekommen. Dann der typische Blackout bei dem man alles Gelernte wieder vergisst. Auch die negative Auffassung von Konformität ist weitläufig verbreitet. Dies rührt sicherlich auch aus schlechten Erfahrungen in der deutschen Geschichte. Das Asch-Experiment kennen wir bereits aus der Vorlesung „Pädagogische Psychologie“. Hier zeigt sich, dass es aber immer noch einen „Mitläufer“-Trend gibt. Der normative soziale Einfluss wurde bei uns in der Schule, nicht mit der Bezeichnung und nicht so tiefgründig, am Beispiel des Filmes „Die 12 Geschworenen“ thematisiert. In diesem Film geht es darum, dass ein einzelner Geschworener alle anderen 11 Geschworenen von ihrem Urteil (es geht um die Todesstrafe) abbringt. Man merkte, dass sich der Widersetzer in seiner Position deutlich gestärkt fühlt, als die erste Person die Meinung wechselt. Während der Widersetzer davor nur seine Zweifel an der Schuldigkeit des Angeklagten äußert, beginnt er nach dem ersten Meinungswechsel aktiv Argumente zur Unschuld des Angeklagten herauszuarbeiten. Hier lässt sich auch eine Brücke zum Thema Majorität / Minorität schlagen. In diesem Fall ist der Widersetzer nämlich die doppelte Minorität. Er ist erstens in der Minderheit. Außerdem ist er der einzige, der sich aktiv mit der Beweislage auseinandersetzt, d.h. wirkliche Verantwortung für seine Aufgabe als Geschworener zeigt. Er schafft es letztendlich, als Minderheit die Mehrheit zu beeinflussen, da er gute Argumente vorweist, seine Meinung konstant vertritt, einen flexiblen Verhandlungsstil aufweist und der Großteil der Majorität sich seiner Meinung nicht sicher ist.

Table 1: Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)

Begründung

Bei dieser Antwort ist sehr gut zu erkennen, dass ein Verständnis für das Gesamtkonzept des sozialen Einflusses besteht. Die einzelnen Komponenten werden auf ein Filmbeispiel angewandt und Veranschaulichungen der Komponenten werden vorgenommen. Um auf Niveau 5 zu gelangen, müsste in der Antwort allerdings über das im Seminar Gesprochene hinausgegangen werden und weiterführende Fragestellungen konzipiert bzw. eigene Hypothesen aufgestellt werden.

Niveau 5: extended abstract

Prompt: Welche Bezüge weisen die neuen Konzepte zu Ihnen bereits bekannten Theorien und Modellen auf?

Zunächst möchte ich auf die Studie von Rost und Czeschlik eingehen. In dieser stehen sich die beiden Eigenschaften „intelligent“ und „beliebt“ gegenüber. Bei diesem Versuch mussten ungefähr 6500 Kinder einer vierten Klasse verschiedene Intelligenztests machen. Zudem wurden Aussagen der Schüler über ihre Mitschüler aufgrund ihrer Beliebtheit gemacht. Das Ergebnis dieser Studie ist, dass die intelligenten Schüler meistens auch die beliebtesten und nicht abgelehnten Schüler sind. Im Gegensatz dazu sind die unbeliebtesten und abgelehnten Schüler meistens die Schüler mit einer geringeren Intelligenz. Deshalb könnte man die Frage aufwerfen, inwiefern dieses Ergebnis ein Ergebnis der Sich-selbsterfüllenden-Propheteiung ist. Denn könnte es nicht sein, dass die Folgen für die Außenseiter ihr Leben und Lernverhalten beeinflussen? Viele abgelehnte Schüler nehmen nämlich nicht mehr am Austausch (sowohl schulischer als auch sozialer Art) teil, wodurch ihnen Wissen verloren geht. Denn in der Schule helfen sich Schüler untereinander, ob dies bei den Hausaufgaben ist oder ob man auch mal abschreiben darf. Da dieser Kontakt nicht stattfindet und die Mitschüler ja sowieso glauben, dass der abgelehnte Schüler so und so ist, passt er sich dieser Meinung an und somit kommt sein Schulversagen zustande. Auch Fritz Heider stellt die These auf, dass ein Mensch, sobald es eine externe Beeinflussung gibt, nicht mehr so stark auf seine innere, eigene Beeinflussung eingeht.

Table 1: Textbeispiele für die einzelnen Kategorien der SOLO-Taxonomie (wird fortgesetzt)

Begründung

Diese Antwort enthält eine Verknüpfung verschiedener Seminarthemen, wobei selbstständig eine Hypothese über den Grund der gefundenen Ergebnisse aufgestellt wird.

Die SOLO-Taxonomie findet im englischsprachigen Raum breite Anwendung und wurde in mehreren Studien hinsichtlich ihrer Passung im Hinblick auf verschiedene Altersklassen und fachlich unterschiedliche Lerninhalte geprüft (z.B. Chan, Tsui, Chan & Hong, 2002; Chick, 1998; Tang, 1994; Watson, Chick & Collis, 1988). Aus den Studien zeichnet sich ab, dass die Taxonomie sich zum einen für die Einstufung von Lernergebnissen in so unterschiedlichen Fächern wie Mathematik, Biologie und Sprachwissenschaften eignet (vgl. Chan et al., 2002). Hier wurden Beziehungen zu anderen Lernfaktoren wie Schreibstil, Lernstrategien und Vorwissen hergestellt (eine detaillierte Übersicht über die Studien zur SOLO-Taxonomie findet sich bei Chan, Hong und Chan (2001)). Zum anderen wurde die Taxonomie in unterschiedlichen Altersklassen erprobt und erwies sich als geeignet, um altersadäquate Aussagen über Lernleistungen zu machen. Trotz aller Vorzüge deuten die Studien auch auf Probleme im Hinblick auf die Interrater-Reliabilität hin: „The conceptual ambiguity of SOLO’s structure makes categorization unstable along with the problem of low interrater reliability and the accident of mistaking an „extended abstract“ (very good answer) as „prestructural“ (very poor answer)“ (Chan et al., 2002, S. 512). Innerhalb der Studie von Chan et al. (2002) erreichen die Rater Konkordanzen von $r = .60$ und $r = .66$. Boulton-Lewis (1994) hingegen berichtet eine Übereinstimmung von 80%.

Die Unschärfe der SOLO-Hierarchiestufen kann als berechtigte Kritik an der Güte des Instrumentes eingebracht werden, da die möglichst große Übereinstimmung der Beurteiler bei der Anwendung der qualitativen Inhaltsanalyse ein zentrales Gütekriterium darstellt. Sie greift allerdings weiter zurückliegende, grundsätzliche Einwände gegen das Konzept der Taxonomie auf (z.B. Foucault, 1994; Wittgenstein, 2003) und stellt somit kein singuläres Problem der SOLO-Taxonomie dar, sondern ein generelles Problem von Klassifizierungssystemen. Bei der Einteilung von Personen zu künstlichen Gruppen stellt sich immer die Frage, nach welcher Systematik einzelne Kriterien gewichtet werden, und inwiefern die Anordnung der Kriterien zu Gruppen sinnvoll und nicht willkürlich ist. In Bezug auf die SOLO-Taxonomie könnte man konkret fragen, warum sich Wissen entlang der durch die Niveaus repräsentieren Modi entwickeln sollte, und nicht in sich wiederholenden Zyklen innerhalb einer Lernsequenz, in der sich Lerner zwischen dem ersten, zweiten und dritten Niveau hin- und herbewegen, wie Campbell, Watson und Collis (1992) annehmen.

Aus diesen Überlegungen heraus werden die folgenden Schlüsse abgeleitet, vor deren Hintergrund die SOLO-Taxonomie zur Anwendung kommt. Die Taxonomie wird als ein mögliches Interpretationsverfahren verstanden, mit dem die Komplexität von Wissensstrukturen sichtbar gemacht werden kann. Die Validität des Verfahrens kann allerdings nicht als gesichert bewertet werden. Daher ist es entscheidend, die Vorgehensweise bei der Einstufung der Texte explizit zu machen und anhand der Daten zu überprüfen, ob sich im Kontext der vorliegenden Studie logische Entwicklungsstufen im Wissensaufbau abbilden lassen. Zudem werden im Vorfeld der Untersuchung angesichts der berichteten Probleme mit der Interrater-Reliabilität aussagekräftige Ankerbeispiele für die Kategorien entwickelt, die den Ratern bei der Einstufung der Texte helfen sollen.

Innerhalb dieser Arbeit werden mit der SOLO-Taxonomie Aussagen über die Qualität der kognitiven Strukturen, die in den Texten der Studierenden identifiziert werden, gemacht, weil bislang kaum empirische Daten dazu vorliegen. Zusätzlich wird herausgearbeitet, inwieweit die Verarbeitungstiefe, die in den Texten erreicht wird, mit den Merkmalen der eingesetzten Prompts variiert. In den folgenden zwei Untersuchungen wird für die Analyse der Texte daher die SOLO-Taxonomie verwendet.

Beschreibung der Interrater-Reliabilität

Zwei Kodiererinnen wurden zwei Stunden lang in die Kategorisierungsprozedur eingewiesen. Zunächst wurde die SOLO-Taxonomie erklärt. Anschließend kategorisierten die Kodiererinnen 20 Texte parallel, wobei sie eine Übereinstimmung von 73% erreichten. Die Texte, in denen keine Übereinstimmung erlangt wurde, wurden gemeinsam besprochen und die Regeln für die Zuweisung nochmals detaillierter ausformuliert. Um eine höhere Übereinstimmung zu erreichen, erstellten die Kodiererinnen gemeinsam Textbeispiele für die einzelnen Kategorien der Taxonomie (siehe Tabelle 1). Diese wurden für die anschließenden Analysen verwendet. Bei einem weiteren Trainingsdurchlauf von 20 Texten erreichten die Kodiererinnen nun 92% Übereinstimmung. Diese wurde als zufriedenstellend eingestuft.

5.2 Studie 1: Evaluation der Prompts und des Textanalyseinstrumentes

Im Rahmen der ersten Untersuchung wurden 105 Studierende in drei Seminaren angeleitet, mit insgesamt 16 Prompts Lernjournale zu schreiben. Es sollte erprobt werden, welche Prompts sich gut zur Initiierung kognitiver Strategien eignen. Dies wird über die Verarbeitungstiefe der Texte ermittelt, die mit der SOLO-Taxonomie kategorisiert wurden. Die SOLO-Taxonomie stellt ein Textanalyseinstrument dar, mit dessen Hilfe kognitive Strategien in Texten klassifiziert werden können. Ein weiteres Ziel der Untersuchung war es festzustellen, ob sich die SOLO-Taxonomie dazu eignet, aussagekräftige Kategorien für die Einstufung kognitiver Strukturen zu finden.

5.2.1 Methode

Untersuchungsdesign

In einer explorativen, quasiexperimentellen Studie wurden die Schreibprozesse von Studierenden, die durch Prompts unterstützt wurden, untersucht. Die Untersuchung wurde im Sommersemester 2008 am Institut für Pädagogische Psychologie an der Johann Wolfgang Goethe-Universität in Frankfurt durchgeführt. Im Rahmen von drei Wahlpflichtseminaren der Pädagogischen Psychologie aus dem Themengebiet der Sozialpsychologie erhielten Studierende zu jeder thematisch abgeschlossenen Lerneinheit des Seminars zwei Prompts, die direkt im Anschluss an die Lerneinheit bearbeitet werden mussten. Es handelte sich um ein Blockseminar ($n_1 = 60$) über einen Zeitraum von einer Woche und zwei Seminare ($n_2 = 22; n_3 = 23$) mit der Laufzeit über ein Semester. Die Prompts wurden als Voraussetzung für den Scheinerwerb eingeführt. Insgesamt bearbeiteten die Studierenden 16 Prompts, die in den drei Seminaren in derselben Reihenfolge präsentiert wurden. Die Kategorisierung der Schreibprodukte wurde von zwei Raterinnen vorgenommen.

Stichprobe

Für die Befragung wurde aufgrund praktischer Erwägungen eine Ad-hoc-Stichprobe gewählt, da eine einfache Zufallsauswahl nicht praktikabel gewesen

wäre. Bei diesem Verfahren wird die Stichprobe nicht per Zufall gebildet, sondern rekrutiert sich aus einer bereits bestehenden Personengruppe (vgl. Bortz, 2005, S. 89), in diesem Falle also einem Seminar. Für die Verwendung inferenzstatistischer Verfahren werden Ad-hoc-Stichproben als problematisch bewertet, da davon ausgegangen werden muss, dass sie keine repräsentative Auswahl der Zielpopulation darstellen und folglich keine Rückschlüsse von der Stichprobe auf diese zulassen. Angesichts des explorativen Charakters der vorliegenden Studie ist eine Ad-hoc-Stichprobe jedoch als angemessen zu bewerten. Die Stichprobe der vorliegenden Studie besteht aus 105 Studierenden der Lehramtsstudiengänge an der Universität Frankfurt. Alle Probanden befanden sich noch in den Anfangssemestern (1. - 3. Semester) ihres Studiums und gaben an, keinerlei Vorerfahrungen mit Prompts oder dem Führen von Lernjournalen zu haben. Hinsichtlich der Effekte der Prompts auf die Anwendung von Lernstrategien wurde bei den Teilnehmern ein hinlängliches Maß an Lernstrategiewissen und Lernstrategienutzung angenommen.

Instrument

Im ersten Teil der Untersuchung ging es zunächst darum, Aufschluss darüber zu gewinnen, welche Verarbeitungstiefen die Studierenden mit den Prompts in ihren Texten erreichten. Um das Niveau der Texte quantitativ abzubilden, wurden die Texte mithilfe der SOLO-Taxonomie (Biggs & Collis, 1982) von zwei Raterinnen kategorisiert und quantitativ eingestuft. Ein weiteres Ziel der Untersuchung war das Herausarbeiten von Zusammenhängen zwischen inhaltlichen Merkmalen der Prompts und der Qualität der Texte.

Durchführung

Die zentrale Intervention bestand darin, dass 16 unterschiedliche Prompts vorgegeben wurden. Um zu überprüfen, ob es systematische Zusammenhänge zwischen Charakteristika der Prompts und der Verarbeitungstiefe der produzierten Texte gibt, wurden Prompts mit verschiedenen Bezügen entwickelt. Diese wurden sechs Kategorien zugewiesen, die im Folgenden erläutert werden. Prompts in der Kategorie „Alltagsbezug“ fordern den Lerner auf, Bezüge zwischen Lernstoff und persönlichem Alltag herzustellen. Unter der Kategorie „Beruflicher Bezug“ wurden Prompts subsumiert, mit denen berufliche Bezüge hergestellt werden soll-

ten. Die Kategorie „Bewertung“ enthält Prompts, die eine persönliche Bewertung der Lerninhalte vorsehen. In die Kategorie „Offene Fragen“ fallen Prompts, die die Themenwahl freistellen. Die Kategorie „Beispiele“ enthält Prompts, die zum Generieren von Anwendungsbeispielen oder zum Formulieren von Beispielen aus der eigenen Biographie auffordern. Die der Kategorie „Kontextübergreifende Bezüge“ zugeordneten Prompts zielten auf Bezüge zwischen den Lerninhalten aus der aktuellen Veranstaltung und solchen aus anderen Veranstaltungen ab. Für die weiteren Analysen werden die jeweils einer Kategorie zugeordneten Prompts gemeinsam betrachtet. Tabelle 2 gibt einen Überblick über die in den Kategorien gruppierten Prompts, die unterschiedliche kognitive Prozesse anregen sollten.

Tabelle 2: Merkmalsausprägungen der Prompts

Merkmale	Prompts
Alltagsbezug	1: „Inwiefern weist der Lernstoff eine Relevanz für Ihr Alltagsleben auf?“ 3: „An welchen Stellen weist das Gelernte Bezüge zu eigenen Erfahrungen auf?“
Beruflicher Bezug	2: „Welches der vorgestellten Konzepte erscheint Ihnen am wichtigsten für ihren späteren Beruf? Fassen Sie dieses kurz und prägnant zusammen und begründen Sie Ihre Wahl.“ 11: „Welche praktischen Implikationen hat das Gelernte für Ihren späteren Beruf?“
Bewertung	4: „Welchen Aspekt des Gelernten fanden Sie überraschend oder kontraintuitiv und warum? Fassen Sie diesen kurz und prägnant zusammen.“ 14: „Welche Aspekte des Gelernten fanden Sie interessant, welche fanden Sie überzeugend und welche fanden Sie nicht überzeugend?“
Offene Fragen	10: „Welche weiterführenden Fragen wirft das Gelernte auf und wie könnte man diese untersuchen?“ 13: „Was wussten Sie schon vorher über das Thema und wo können Sie Bezüge herstellen?“
Beispiel	7: „Beschreiben Sie möglichst bildlich ein Beispiel ungünstiger nonverbaler Kommunikation aus dem Schulalltag.“ 9: „Denken Sie sich ein positives und ein negatives Beispiel für das Gelernte und zeigen Sie daran die Implikationen für die Beurteilungspraxis auf.“ 16: „Führen Sie Beispiele aus Ihrer eigenen Schulbiographie an, die das Gelernte bestätigen oder ihm widersprechen.“

Tabelle 2: Merkmalsausprägungen der Prompts

Merkmale	Prompts
Kontextübergreifende Bezüge	5: „Welche Beziehungen weist das Gelernte über das Thema Vorurteile zu den Inhalten verwandter Fächer bzw. Lehrveranstaltungen auf?“ 8: „Welche Bezüge sind Ihnen zwischen dem Gelernten und dem Feld der Schauspielerei aufgefallen?“ 12: „Welche Anknüpfungspunkte zwischen dem Gelernten und aus anderen Seminaren/Fächern bereits bekannten Theorien, Methoden oder Befunden sind Ihnen aufgefallen?“ 15: „Welche Bezüge weisen die neuen Konzepte zu Ihnen bereits bekannten Theorien und Modellen auf?“

Die Prompts wurden von 105 Studierenden des Lehramts aus drei Seminaren in der Reihenfolge bearbeitet, in der sie in der Tabelle 2 nummeriert sind. Das erste der drei Seminare lief eine Woche ($n_1 = 60$), das zweite ($n_2 = 22$) und das dritte Seminar ($n_3 = 23$) umfassten den Zeitraum eines Semesters. Die Prompts wurden als Bedingung für den Scheinerwerb eingeführt und in allen drei Seminaren in derselben Reihenfolge (wie in der Nummerierung innerhalb der Tabelle dargestellt) gegeben. Den Studierenden wurde mitgeteilt, dass die Schreibaufträge der Nachbereitung des Seminars und als Vorbereitung auf die Klausur dienen. Nach jeweils einer thematisch in sich geschlossenen Lerneinheit erhielten die Studierenden im einwöchigen Seminar jeweils zwei Prompts, während die Teilnehmer aus dem zweiten und dritten Seminar pro Sitzung einen Prompt erhielten, da sie mit denselben Prompts über einen längeren Zeitraum arbeiteten. Die Prompts wurden auf einem Word-Formular ausgegeben, das die Textmenge auf eine 3/4 DIN A4-Seite beschränkte. Auf der Grundlage der Schreibaufträge verfassten die Studierenden im Anschluss an die Veranstaltung Texte, die in der nächsten Sitzung abgegeben werden mussten.

5.2.2 Ergebnisse

Vorgehensweise bei der qualitativen Analyse der Lerntagebuchttexte

Die Auswertung der Texte erfolgte über eine Zuordnung der Antworten zu den Niveaus der SOLO-Taxonomie. Das Ziel der Analyse bestand darin, das schriftliche

Datenmaterial auf kognitive Prozesse hin zu untersuchen, um zu Erkenntnissen hinsichtlich der Qualität der Schreibprozesse zu gelangen. Zu diesem Zweck wurden die Texte den fünf Ergebnisebenen der SOLO-Taxonomie (Biggs & Collis, 1982) zugeordnet. Es wurden thematische Analyseeinheiten verwendet. Analyseeinheiten bezeichnen die inhaltlichen Segmente, die auf bestimmte Merkmale hin untersucht und festgelegten Kategorien (Bedeutungseinheiten) zugewiesen werden. Da Aussagen über die Komplexität und Verarbeitungstiefe längere und zusammenhängende Textsegmente erfordern, werden die jeweils zu einem Prompt produzierten Texte als eine Kodiereinheit abgegrenzt. Angesichts der Vorstrukturierung der Texte durch die Prompts ist dies sowohl ein praktikables als auch ein methodisch schlüssiges Vorgehen. Es wird damit auch dem Anspruch Rechnung getragen, durch die Inhaltsanalyse ein möglichst differenziertes und realistisches Bild der Daten wiederzugeben (vgl. De Wever, Schellens, Valcke & Van Keer, 2006). Durch diese Art der Segmentierung können nun Aussagen über fragenspezifische Qualitäten (z.B. welche kognitiven Prozesse durch eine bestimmte Lernfrage angeregt werden) und personenspezifische Eigenheiten (z.B. welche Textniveaus ein spezieller Proband im Durchschnitt erzeugt) vorgenommen werden. Ein weiteres Argument für diese Vorgehensweise ist die Struktur der SOLO-Taxonomie, deren Niveaus sich zumeist aus mehreren Prozessen zusammensetzen. Um die Vergleichbarkeit hinsichtlich des Textumfangs zu sichern, wurde jede Lernfrage in ein viertelseitiges Word-Formular (DIN A4) eingebettet und die Studierenden wurden instruiert, dieses Feld auszufüllen. Insgesamt wurden 1058 Aussageeinheiten segmentiert. Die Kategorien und Zuordnungsregeln werden im Kodierleitfaden im Abschnitt .4 des Anhangs dokumentiert. Insgesamt gingen 1058 Lernjourneinträge in die Analyse mit der SOLO-Taxonomie ein. Eine Stichprobe von 20 Lernstagebüchern wurde zur Überprüfung der Reliabilität des Verfahrens von zwei unabhängigen Kodiererinnen doppelt kodiert. Die Interrater-Übereinstimmung lag bei der Überprüfung mit der Platz-zu-Platz-Übereinstimmung bei 92% und kann als ausreichend zuverlässig angesehen werden.

Deskriptive Befunde zur Textqualität

Die erste Fragestellung der Arbeit bezieht sich auf die Qualität der Texte, die über das Kriterium „Verarbeitungstiefe“ definiert wird. Unter der Verarbeitungstiefe wird die Komplexität der kognitiven Prozesse verstanden, die aus den Tex-

ten ermittelt wurde. Mithilfe des Analyse-Instrumentes, der SOLO-Taxonomie, werden die Texte fünf Ergebnisebenen zugeordnet. Die Komplexität der Texte wird folglich als Indikator für den Elaborationsgrad der kognitiven Prozesse während der Textproduktion interpretiert. Die Verarbeitungstiefe wird umso höher eingeschätzt, je deutlicher aus einem Text erkennbar wird, dass der Lerner die Lerninhalte komplex verknüpfen kann. Es wird zunächst betrachtet, welche Verarbeitungstiefen insgesamt in den mit den Prompts produzierten Texten erreicht werden. Zum Überblick wird die prozentuale Verteilung aller Texte (insgesamt 1058) auf den fünf Taxonomie-Niveaus dargestellt (vgl. Abbildung 1). Desweiteren interessiert, inwiefern das Textniveau in Abhängigkeit der Prompt-Kategorien variiert. Um dieser Frage nachzugehen, werden die prozentualen Anteile der gruppierten Prompts auf den Taxonomieniveaus ermittelt und analysiert. Es wird erwartet, dass spezifische Merkmale der Lernfragen (z.B. Bewertung von Lerninhalten vs. Bezüge herstellen) unterschiedliche Verteilungsmuster im Hinblick auf die Taxonomieniveaus erzeugen.

Forschungsfrage 1: Welche Art von kognitiven Prozessen werden mit den ausgewählten Prompts angeregt? Abbildung 1 gibt einen Überblick über die Gesamtverteilung aller Texte aus allen Seminaren auf den SOLO-Niveaus.

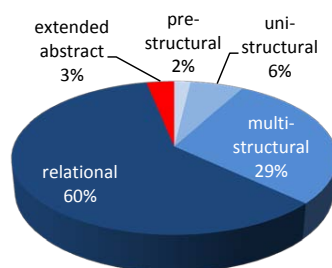


Abbildung 1: Verteilung der Lernjournaltexte über die Taxonomiestufen

Auf den ersten Blick wird deutlich, dass der überwiegende Anteil der Lerntagebücher sich auf den hohen Textniveaus *multi-structural* (Niveau 3) und *relational* (Niveau 4) wiederfindet. Besonders häufig haben die Studierenden Texte produziert, die dem vierten Niveau (*relational*) zugeordnet wurden (60%). Die Voraussetzung für die Zuordnung zu diesem Niveau war, dass ein Text auf ein

weitreichendes Verständnis für eine Theorie oder ein Konzept und deren Zweck schließen ließ. Diese Voraussetzung galt als erfüllt, wenn der Textproduzent demonstrieren konnte, dass er das Konzept auf einen anderen Kontext (z.B. den Berufs- oder Familienalltag) anwenden kann. Dies impliziert wiederum die Fähigkeit, verschiedene Komponenten des Konzeptes oder der Theorie zu integrieren. Zugehörige Prozesse sind Kontrastieren, Vergleichen, Integrieren, Analysieren und in Beziehung setzen. Die Textproduzenten mussten also mehrere anspruchsvolle Aufgaben durchführen, um auf dem *relational*-Niveau eingestuft zu werden. Für die nächsthäufigste Einstufung auf dem dritten Niveau (*multi-structural*) (29%) genügte hingegen ein erkennbares Verständnis für mehrere Komponenten eines Konzeptes. Im Vergleich zu dem vierten Niveau fehlt hier die Integration der einzelnen Elemente, sodass man nicht von einem umfassenden Verständnis eines Konzeptes oder einer Theorie ausgehen kann. Die Ideen werden in einer Art präsentiert, die keine eigenständige Struktur erkennen lässt. Dies tritt in den Texten durch solche Strukturen wie Aufzählen, Klassifizieren, Beschreiben und Kombinieren zutage. Eher selten wurden Texte auf dem ersten (*prestructural*) und fünften Niveau (*extended abstract*) produziert. Dennoch erreichen immerhin drei Prozent (in Häufigkeiten 34 Einträge) aller produzierten Texte das höchste SOLO-Niveau *extended abstract*. Ein Lerner musste hierfür relevante Aspekte eines Themas miteinander in Beziehung setzen und diese auf einen neuen Kontext übertragen. Ein Text wurde dann auf diesem Niveau eingestuft, wenn der Textproduzent darin eigene Hypothesen anstellte oder den Lerninhalt kritisch reflektierte. Das niedrigste Niveau (*pre-structural*) kam nur bei zwei Prozent (in Häufigkeiten 20 Einträge) der Texte vor und impliziert, dass ein Text keine relevanten Komponenten der Lernaufgabe enthält. Summiert man die Prozentwerte der höheren Taxonomieniveaus drei (*multi-structural*), vier (*relational*) und fünf (*extended abstract*), so ergibt sich mit 88 Prozent ein deutlicher Schwerpunkt von Texten auf den anspruchsvollen Niveaus.

Forschungsfrage 2: Inwiefern beeinflussen Merkmale der Prompts die Textqualität? Da von einer Anforderungsspezifität der Prompts ausgegangen wird, sollten sich verschiedene Niveauprofile bezogen auf die fünf Ergebnisebenen der SOLO-Taxonomie auffinden lassen. Es wurde zum einen vermutet, dass die Texte umso besser ausfallen, je spezifischer die Anforderungen formuliert sind und je

konkreter der Bezug ist (Beruflicher Bezug vs. Kontextübergreifende Bezüge). So wurden z.B. für die Prompts der Kategorie „Offene Fragen“ eher niedrige Niveaus erwartet, da sie keine konkreten kognitiven Prozesse instruieren. Die folgenden sechs Abbildungen (Abb. 2, 3, 4, 5, 6, 7) geben einen detaillierteren Überblick über die Ausprägungen der Taxonomieebenen auf den gruppierten Prompts. Es werden die prozentualen Anteile der Texte in Abhängigkeit des Seminars auf den fünf Ergebnisebenen der SOLO-Taxonomie dargestellt.

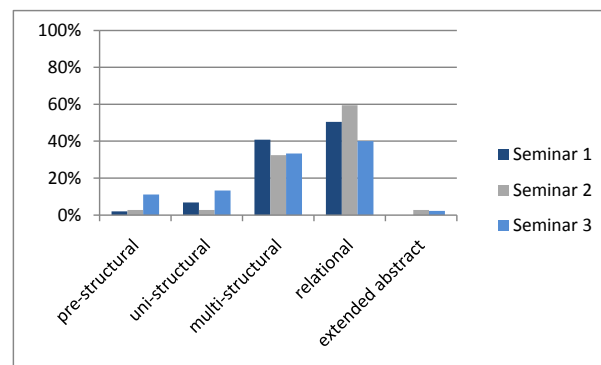


Abbildung 2: Prozentuale Verteilung der Texte in der Kategorie „Alltagsbezug“

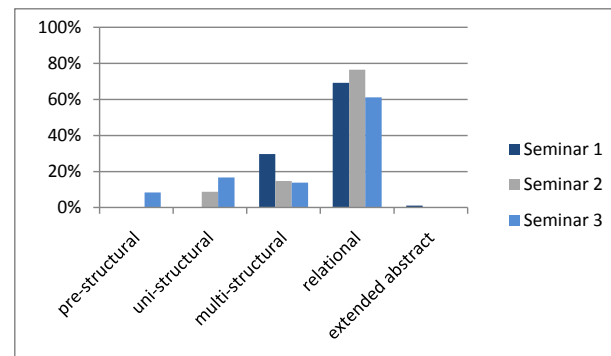


Abbildung 3: Prozentuale Verteilung der Texte in der Kategorie „Beruflicher Bezug“

Aus den Darstellungen der Niveaustufen-Verteilungen über die jeweils einer Kategorie zugeordneten Prompts ist erkennbar, dass, analog zur Gesamtverteilung, die gruppierten Prompts ebenfalls insgesamt durchweg hohe Anteile von Texten auf dem vierten Niveau (*relational*) aufweisen. So dominiert in fünf von

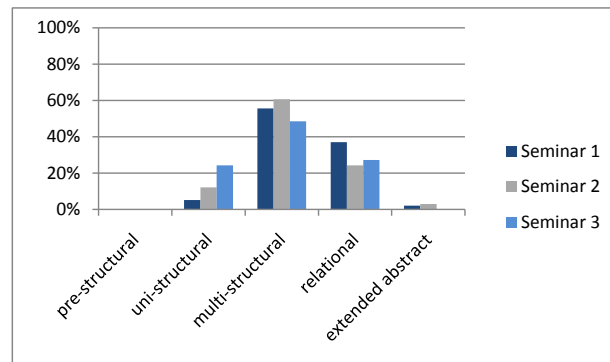


Abbildung 4: Prozentuale Verteilung der Texte in der Kategorie „Bewertung“

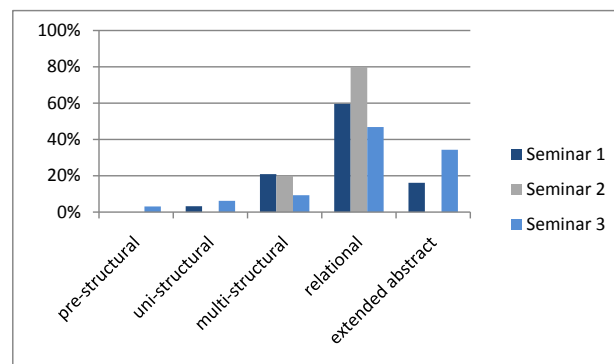


Abbildung 5: Prozentuale Verteilung der Texte in der Kategorie „Offene Fragen“

sechs Kategorien (Alltagsbezug, Beruflicher Bezug, Offene Fragen, Beispiel, Kontextübergreifende Bezüge) über alle drei Seminare hinweg das vierte Taxonomie-niveau. Besonders deutlich wird dies in den Kategorien Beispiel (81% - 97%) und Beruflicher Bezug (61% - 69%). Das höchste Niveau (*extended abstract*) wurde am häufigsten in der Kategorie „Offene Fragen“ erreicht (16% - 34%). Dies trifft sowohl auf das Blockseminar als auch auf eines der beiden auf ein Semester angelegten Seminare zu. Eine Begründung für das häufige Vorkommen des fünften Niveaus in dieser Prompt-Kategorie könnte der relativ späte Zeitpunkt in beiden Seminaren sein, zu dem die Prompts gegeben wurden. Möglicherweise wirkt sich hier die kontinuierliche Nachbearbeitung des Lernstoffs positiv auf die Schreibqualität aus. In den restlichen Kategorien hält sich der Anteil an Texten auf dem höchsten Niveau unter zehn Prozent. Die beiden niedrigen Niveaus *pre-structural* und *uni-structural* wurden anteilmäßig selten produziert (*pre-structural* unter

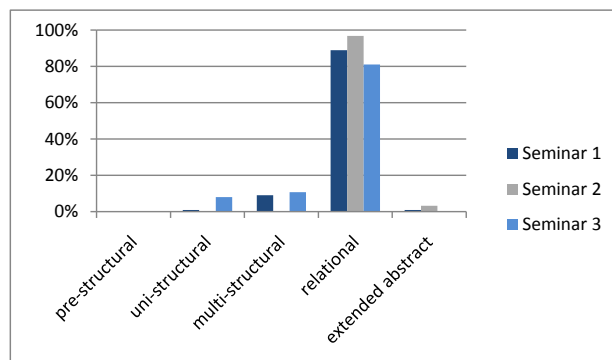


Abbildung 6: Prozentuale Verteilung der Texte in der Kategorie „Beispiele“

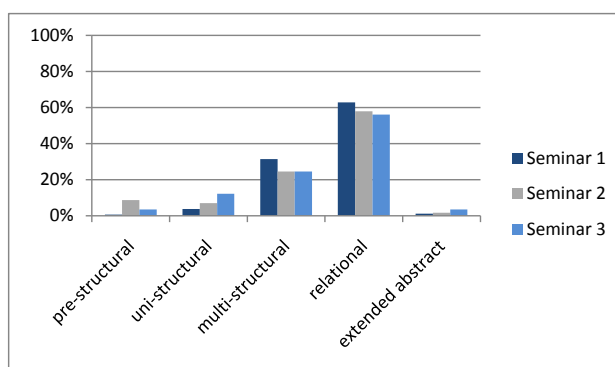


Abbildung 7: Prozentuale Verteilung der Texte in der Kategorie „Kontextübergreifende Bezüge“

zehn Prozent, *uni-structural* unter 25 Prozent).

Bezüglich der formulierten Vorannahmen über die Auswirkungen der Anforderungsspezifität der Prompts auf die Verarbeitungstiefe könnte man aus den voranstehenden Verteilungen vorsichtige Schlüsse ziehen. Im Bezug auf die Spezifität der Prompts scheint sich abzuzeichnen, dass sich ein enger, inhaltlicher Fokus positiv auf die Verarbeitungstiefe beim Schreiben auswirkt. Sichtbar wird dies, wenn man die Kategorien „Alltagsbezug“, „Beruflicher Bezug“ und „Kontextübergreifende Bezüge“ miteinander kontrastiert. In der Kategorie „Beruflicher Bezug“ wurden am häufigsten Texte auf dem vierten Niveau produziert, während die Einstufungen der Texte bei den anderen beiden Kategorien breiter streuten. Die Prompts mit „beruflichem Bezug“ fragten nach Implikationen des Lernstoffes für die berufliche Praxis, womit ein konkreter, inhaltlicher Bezug vorgegeben wur-

de. Die Prompts in der Kategorie „Alltagsbezüge“ hingegen stellten die Art des Bezuges relativ frei, „Kontextübergreifende Bezüge“ forderten auf, Verknüpfungen zu Lerninhalten aus anderen Seminaren herzustellen. Möglicherweise wirkt bei der Kategorie „Berufliche Bezüge“ auch verstärkend, dass Verknüpfungen des Lernstoffes mit einem beruflichen Bezug eine höhere persönliche Relevanz haben. Konform mit den Ergebnissen der Studie von Glogger et al. (2009) scheint es sich auch positiv auszuwirken, wenn spezifische kognitive Prozesse in den Prompts explizit angeleitet werden. Dies wird sichtbar bei den Kategorien „Beispiele“ und „Offene Fragen“. Hier wurden konkrete Beispiele gefordert. In der letzten Kategorie „Offene Fragen“ hingegen sollten die Studierenden weiterführende Fragen generieren bzw. Bezüge zu Bekanntem herstellen. Verantwortlich für den hohen Anteil von Texten auf dem fünften Niveau ist allerdings Prompt 10, der zur Formulierung einer eigenständigen Fragestellung auffordert. In dieser Kategorie wurde der einzig nennenswerte Anteil von Texten auf dem fünften Niveau produziert, wobei die Verteilung gleichzeitig auch relativ breit über die restlichen Niveaus streut. In der Kategorie „Bewertung“ wurde der Hauptanteil der Texte auf dem dritten Niveau eingestuft. Die Bewertung von Lerninhalten scheint eine hohe Anforderung für die Studierenden darzustellen, möglicherweise eine zu hohe Anforderung für Personen mit geringem bereichsspezifischem Vorwissen. Man könnte vorsichtig interpretieren, dass hier die Bandbreite der Kompetenzen der Studierenden am ehesten erkennbar wird.

5.2.3 Zusammenfassung und Konsequenzen

Zusammengefasst können drei zentrale Ergebnisse der Vorstudie festgehalten werden, die für die Förderung und für die Erfassung kognitiver Prozesse beim Schreiben relevant sind.

a) Durch die Unterstützung des Schreibens mit den Prompts werden insgesamt komplexe Verarbeitungsniveaus angeregt. Die Mehrheit der Texte wurde auf dem vierten Niveau eingestuft, das immerhin voraussetzt, dass die Lerner Elemente eines Themenkomplexes miteinander vergleichen, kontrastieren, Begründungen für Sachverhalte finden, das Thema eigenständig analysieren oder das Wissen auf ein anderes Themenfeld anwenden. Um dieses Ergebnis angemessen einzustufen zu können, muss man berücksichtigen, dass die Prompts jeweils unmittelbar nach dem Abschluss einer ersten Lernepisode bearbeitet wurden. Die Studierenden standen

damit am Beginn eines neuen Lernzyklus. Man hätte also erwarten können, dass sich die Lernenden an diesen Lernzyklus noch herantasten, was sich in den Texten in weniger komplexen Strukturen, wie dem Aufzählen oder Klassifizieren einzelner Elemente, niederschlagen sollte. Die vorliegenden Ergebnisse zeigen aber, dass die Studierenden durch die Prompts von Anfang an zu komplexen Prozessen angeregt werden können.

b) Was die Ausgestaltung der Prompts angeht, so eignen sich im Hinblick auf die Verarbeitungstiefe am besten solche Prompts, die einen konkreten, inhaltlichen Bezug implizieren. Desweiteren wirkt sich auch die explizite Benennung spezifischer, kognitiver Prozesse positiv auf die Qualität des Schreibprozesses aus (z.B. Beispiele generieren oder eine eigene Fragestellung entwickeln).

c) Die SOLO-Taxonomie eignet sich, um verschiedene Verstehensniveaus in einem Themengebiet abzubilden. Die Kategorien verdeutlichen dabei die qualitativen Unterschiede kognitiver Strukturen. In der auf der Vorstudie aufbauenden Untersuchung soll geklärt werden, unter welchen Bedingungen der Einsatz von Prompts über einen längeren Zeitraum gelingen kann: Welche Voraussetzungen sind seitens der Studierenden zu beachten? Wie kann die Lernumgebung motivationsförderlich gestaltet werden?

5.3 Studie 2: Effekte verschiedener Feedbackformen auf die Verarbeitungstiefe beim Schreiben und auf die Motivation

Aufbauend auf den Erkenntnissen der ersten Studie werden die als effizient bewerteten Prompts genutzt und mit verschiedenen Formen der Rückmeldung verknüpft, um zu überprüfen, ob die Qualität des Schreibens weiter optimiert werden kann. Die letzte Studie zeigte, dass die Studierenden besonders gute Texte mit Prompts produzierten, die Bezüge zu Lerninhalten aus der aktuellen Lehrveranstaltung erfordern. Erfolgversprechend war auch die Lernstrategie „Beispiele generieren“. Darüber hinaus soll die zweite Studie klären, welche Voraussetzungen seitens der Studierenden zu beachten sind.

5.3.1 Methode

Untersuchungsdesign

Über ein Semester arbeiteten 46 Studierende im Rahmen eines Seminars im Bachelor-Studiengang Erziehungswissenschaft zur Einführung in empirische Methoden mit Prompts. In dieser Studie wurden Effekte verschiedener Formen von Feedback auf die Komplexität der kognitiven Prozesse beim Schreiben mit Prompts untersucht. Ein weiteres Ziel der Untersuchung war es, Zusammenhänge zwischen Aspekten der Motivation der Studierenden, unterschiedlichem Feedback und der Verarbeitungstiefe zu prüfen. Zu diesem Zweck wurden innerhalb eines Seminars drei Gruppen (A, B und C) gebildet, die mit neun identischen Prompts arbeiteten. In Abhängigkeit der ihnen zugewiesenen Gruppe erhielten alle Teilnehmer der drei Gruppen jeweils einmal informatives Feedback durch den Dozenten, einmal eine Anleitung zum Selbstfeedback und einmal kein Feedback (vgl. Tabelle 3). Die aktuelle Motivation wurde mittels Online-Fragebogen zu drei Zeitpunkten erfasst. Dazu wurde der Fragebogen zur Erfassung der aktuellen Motivation (FAM) von Rheinberg et al. (2001) eingesetzt. Zur Analyse der Verarbeitungstiefe kam die SOLO-Taxonomie zum Einsatz, mit der die mit den Prompts verfassten Texte kategorisiert wurden. Das Seminar fand am Institut für Pädagogik der Sekundarstufe statt. Die Prompts wurden als Bedingung für den Scheinerwerb eingeführt.

Tabelle 3: Chronologie der Untersuchung in der zweiten Studie

	Gruppe A	Gruppe B	Gruppe C
1. Erhebung (Z1)	Wissenstest		
	Prompts 1-3	Prompts 1-3	Prompts 1-3
	Selbstfeedback mit Diagnosefragen	Dozentenfeedback	Kein Feedback
2. Erhebung (Z2)	Fragebogen zur Erfassung aktueller Motivation (Rheinberg, Vollmeyer & Burns, 2001)		
	Prompts 4-6	Prompts 4-6	Prompts 4-6
	Kein Feedback	Selbstfeedback mit Diagnosefragen	Dozentenfeedback
3. Erhebung (Z3)	Fragebogen zur Erfassung aktueller Motivation (Rheinberg, Vollmeyer & Burns, 2001)		
	Prompts 7-9	Prompts 7-9	Prompts 7-9
	Dozentenfeedback	Kein Feedback	Selbstfeedback mit Diagnosefragen
4. Erhebung (Z4)	Fragebogen zur Erfassung aktueller Motivation (Rheinberg, Vollmeyer & Burns, 2001) Cognitive load (Paas, 1992) Wissenstest		

Stichprobe

Die Stichprobe setzt sich aus 46 Studierenden des Studiengangs Diplom-Pädagogik zusammen. Aufgrund unvollständiger Datensätze wurden 16 Probanden von den weiteren Analysen ausgeschlossen. Bei den verbleibenden 30 Probanden dominierte der Anteil der Frauen mit 87 %. Das Alter der Probanden variierte zwischen 20 und 30 ($M = 22,8$; $SD = 2,55$). Bei einer Person fehlt die Altersangabe. Die Teilnehmer wurden randomisiert den Gruppen A ($n_A = 11$), B ($n_B = 10$) und C ($n_C = 9$) zugeordnet. Die Gruppen waren hinsichtlich ihres Durchschnittsalters vergleichbar ($M_A = 22,9$; $M_B = 22,4$; $M_C = 23,1$). Gleiches gilt für die Geschlechter-

verhältnisse ($w_A = 82\%$, $w_B = 90\%$, $w_C = 89\%$). Die Daten der Probanden wurden verschlüsselt.

Innerhalb der Unterskalen des FAM kommt es jedoch zu Verschiebungen bei den Stichprobengrößen.

Auswertungsverfahren

Non-parametrische Verfahren

Friedman-Test Zur Analyse der Motivationswerte innerhalb der Gruppen wurde mit dem Friedman-Test auf eine non-parametrische Alternative zur ANOVA mit wiederholten Messungen zurückgegriffen, da die Voraussetzungen für eine Varianzanalyse aufgrund der geringen Stichprobengröße nicht gegeben waren. Der Friedman-Test dient dem Vergleich von mehr als zwei abhängigen Stichproben hinsichtlich ihrer zentralen Tendenz in Bezug auf ein Merkmal (hier Motivation), das zu mehreren Zeitpunkten gemessen wird (vgl. Bortz, Lienert & Boehnke, 2008). Als Voraussetzungen für die Durchführung des Tests sind mehr als zwei abhängige Stichproben und mindestens ordinales Messniveau erforderlich. Für jeden Probanden wird, wie im Messwiederholungsplan dargestellt, eine Rangreihe der Werte erstellt. Dem kleinsten Wert wird der Rangplatz 1 zugeordnet, bei gleichen Werten werden die Ränge addiert und gemittelt. Anschließend werden die Ränge in jeder Stichprobe addiert.

Tabelle 4: Friedman-Messwiederholungsplan

Proband	Messzeitpunkt			
	1	2	3	...
1	x	x	x	...
2	x	x	x	...
3	x	x	x	...
⋮	⋮	⋮		
i	x	x	x	x
n	x_n	x_n	x_n	x_n

Die Nullhypothese (H_0) des Tests besagt, dass zwischen den Stichproben kein Unterschied im Median besteht. Sie wird verworfen, wenn sich wenigstens zwei

Stichproben in ihrem Medianwert unterscheiden. Als Maß für die statistische Signifikanz dient der p-Wert. Fällt dieser $\leq 0,05$ aus, wird die Nullhypothese abgelehnt und stattdessen die Alternativhypothese (H_1) angenommen. Man geht dann davon aus, dass sich der Medianwert von mindestens einer der untersuchten Stichproben signifikant von dem der anderen Stichproben unterscheidet. Die Beschränkung des Friedman-Tests besteht darin, dass nur ein Haupteffekt untersucht wird (ob sich mehr als zwei Messzeitpunkte unterscheiden), und dass große Unterschiede bei der Verteilung trotz vergleichbarer Mediane ebenfalls als signifikante Unterschiede abgebildet werden.

Wilcoxon-Test Zur Analyse der Verarbeitungstiefen über den Zeitraum der Intervention wurde der Wilcoxon-Test oder auch Rangsummentest eingesetzt. Dieser stellt ebenfalls einen verteilungsfreien Test dar (vgl. Bortz et al., 2008) und dient zum Vergleich der zentralen Tendenz zweier abhängiger Stichproben. Zwei Stichproben sind voneinander abhängig, wenn „jedem Wert der einen Stichprobe auf sinnvolle und eindeutige Weise genau ein Wert der anderen Stichprobe zugeordnet werden kann“ (Bühl, 2008, S. 119). Dies ist z.B. dann der Fall, wenn Personen aus einer Stichprobe zu zwei Messzeitpunkten miteinander verglichen werden. Beim Wilcoxon-Test werden anstelle der ursprünglichen Messwerte die Rangwerte der Daten miteinander verglichen. Zu diesem Zweck werden alle Messwerte der zwei oder mehr Messungen durch Rangplätze ersetzt. Bei verbundenen Stichproben wird dazu zunächst pro Fall die Differenz zwischen dem ersten und dem zweiten Messzeitpunkt ermittelt. Sodann werden alle Datenwerte der Größe nach geordnet und es werden ihnen Ränge zugeordnet. Tritt derselbe Datenwert innerhalb einer Reihe oder über die Messreihen hinweg mehrfach auf, so wird das arithmetische Mittel der entsprechenden Rangplätze auf diese Werte verteilt.

Im Anschluss an die Zuweisung der Ränge wird die Rangsumme errechnet. Für die Ermittlung der Rangsumme ist ausschlaggebend, welche Vorzeichen die aus den Datenwerten ermittelte Differenz aufweist. Es werden einerseits die Rangsummen aus den Differenzen mit positivem Vorzeichen und andererseits die mit negativem Vorzeichen errechnet. Nun wird die Nullhypothese, dass keine Mittelwertunterschiede zwischen den Messreihen bestehen, geprüft. Dazu wird die Prüfgröße in Abhängigkeit der Irrtumswahrscheinlichkeit α mit dem kritischen Wert verglichen, der einer Referenztabelle zu entnehmen ist (z.B.

<http://www.socr.ucla.edu/Applets.dir/WilcoxonRankSumTable.html>, zuletzt geprüft am 10.5.2011). Die Prüfgröße entspricht der kleineren Rangsumme. Ist diese kleiner oder gleich dem kritischen Wert, wird die Nullhypothese abgelehnt.

Tabelle 5: Wilcoxon Messwiederholungsplan

Proband	Prompts 2,3,4	Prompts 5,6,7	x2-x1	x1-x2
1	x	x	x	x
2	x	x	x	x
3	x	x	x	x
i	x	x	x	x
n	x_n	x_n	x_n	x_n

Instrument: Der Fragebogen zu aktuellen Motivation (FAM)

Die Integration eines Instrumentes zur Erfassung der Motivation basiert auf dem Verständnis der Prompting-Maßnahme als Lernprozess, dessen Dimensionen durch Lernmotivation beeinflusst werden (z.B. Anstrengung und Ausrichtung). Unter Lernmotivation werden allgemein emotionale und kognitive Prozesse subsummiert, die ausschlaggebend dafür sind, dass ein Lernender absichtlich etwas Neues lernt, um die von ihm antizipierten und mit dem Lernen mehr oder weniger direkt verknüpften Folgen erreichen oder verhindern zu können (vgl. Wegge, 1998). Es wird daher erwartet, dass die Lernmotivation eine entscheidende Rolle bei der Ausgestaltung des Schreibprozesses spielt. Zudem schließt die vorliegende Untersuchung mit der Integration der Lernmotivation an die Studie von Nückles et al. (2006) an, in der deutlich wurde, dass sich Prompts über einen längeren Zeitraum negativ auf verschiedene Komponenten der Motivation auswirken. Nückles et al. (2006) verwendeten die Subskalen Interesse/Freude an der Tätigkeit, Anstrengung/subjektive Bedeutung der Tätigkeit und Kompetenzerfinden aus dem „intrinsic motivation inventory“ (IMI) von Deci und Ryan (2003) und verzeichneten einen Abwärtstrend für alle Subskalen über zwei Messungen innerhalb eines Semesters. Die Abnahme der Motivation ging einher mit einer Abnahme der in den Texten vorgefundenen kognitiven Strategien und einer schlechteren Lernleistung der Experimentalgruppe im Vergleich mit einer Kontrollgruppe ohne Prompts.

In der vorliegenden Studie wird in Erweiterung der bisherigen Studien zu Prompts die Prompting-Maßnahme in Kombination mit verschiedenen Formen des Feedbacks eingesetzt. Es soll darüber Aufschluss gewonnen werden, inwiefern sich dadurch motivationsförderliche Anreize setzen lassen, die sich wiederum positiv auf die Verarbeitungstiefe auswirken sollten. Da die Feedbackform, und damit auch die Lernsituation, variiert, stehen State-Merkmale der Motivation im Mittelpunkt des Interesses. Darunter werden affektive und kognitive Bewertungen der aktuellen Lernsituation verstanden (vgl. Rheinberg et al., 2001). Über die Erfassung der Motivation in einer aktuellen Situation wird nach Winther (2006) Information über die Zielausrichtung gegeben, die durch die situative Anregung der Motive oder Interessen aktiviert wird. Auf der Basis dieser Überlegungen wurde mit dem FAM (Fragebogen zur aktuellen Motivation) von Rheinberg et al. (2001) ein Instrument gewählt, das an die jeweils konkrete Lernsituation adaptiert werden kann und die aktuelle Motivation abbildet. Motivation wird darin als ein Konstrukt konzipiert, das sich aus vier Erlebens- und Handlungskomponenten zusammensetzt, von denen angenommen wird, dass sie in selbstregulierten Lernumgebungen angeregt werden. Die affektive Ebene der aktuellen Situation wird durch die Skalen *Erfolgswahrscheinlichkeit* und *Misserfolgsbefürchtung* abgebildet. Die Skala *Erfolgswahrscheinlichkeit* erfragt die Zuversicht, mit der der Lernende eine Aufgabe angeht. Die Erfolgserwartung eines Lerners entsteht in Folge eines Abgleichs zwischen der Aufgabenschwierigkeit und der subjektiv wahrgenommenen Kompetenz. Eine hohe Erfolgserwartung entsteht demnach, wenn der Lerner die zu bewältigende Aufgabe als relativ leicht einschätzt oder sich selbst in der gegebenen Situation für fähig hält (vgl. Rheinberg et al., 2001). Ein Beispiel-Item ist „Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein“. Die theoretische Basis für das Konstrukt Erfolgswahrscheinlichkeit stellen die Studien zu Erwartung-mal-Wert-Modellen (vgl. Rheinberg, 2006) und Selbstwirksamkeitsüberzeugungen (Bandura, 1982) dar. Mit der Komponente *Misserfolgsbefürchtung* wird erfasst, inwiefern der Lernende befürchtet, bei der aktuellen Aufgabe schlecht abzuschneiden. Zusätzlich erfragt die Skala auch die Folgen des damit verbundenen Drucks. Ein Beispiel-Item lautet „Die konkreten Leistungsanforderungen hier lähmen mich“. Decken diese Skalen die affektive Ebene ab, so bilden *Herausforderung* und *Interesse* die motivationalen Wirkfaktoren auf der kognitiven Ebene ab. Die kognitiven Komponenten sind insofern relevant, als

dass Lernende sich auch um ein gutes Abschneiden bei einer Aufgabe bemühen können, wenn sie keine weitreichenden Folgen erwarten (z.B. durch Benotung). In diesem Fall kann das inhaltliche Interesse an der Aufgabe (Beispiel-Item: „Bei Aufgaben wie dieser brauche ich keine Belohnung, sie machen mir auch so viel Spaß“) ausschlaggebend für die Beschäftigung damit sein oder die kompetenzbezogene Leistungsmotivation (Heckhausen, 1989) zum Tragen kommen. Letztere wird in der Skala *Herausforderung* erfasst. Ein Beispiel-Item heisst „Wenn ich die Aufgabe schaffe, werde ich schon ein wenig stolz auf meine Tüchtigkeit sein“. Der FAM setzt sich aus insgesamt 18 Items zusammen (*Erfolgswahrscheinlichkeit*: 4 Items, *Misserfolgsbefürchtung*: 5 Items, *Herausforderung*: 4 Items, *Interesse*: 5 Items) und das Antwortformat der Items reicht auf sieben Stufen von „trifft nicht zu“ (1) bis „trifft zu“ (7). Der Fragebogen zur aktuellen Motivation wird an den spezifischen Lernkontext des Seminars, in dem die Untersuchung stattfindet, adaptiert (vgl. Anhang). Der Einsatz des Instrumentes erfolgt immer nach einer Feedbackmaßnahme, sodass die Effekte der jeweiligen Intervention erfasst werden. Es wird erwartet, dass sich die Motivation der Studierenden über die Untersuchung hinweg in Abhängigkeit der Feedbackform ändert. Zusätzlich werden auch Wechselwirkungen zwischen Verarbeitungstiefe und Motivation in den Blick genommen.

Die Prompts

Aufbauend auf der Vorstudie wurden für die zweite Untersuchung Prompts anhand der effektiv erwiesenen Kriterien konzipiert. In Tabelle 6 sind die Prompts der zweiten Studie aufgelistet.

Tabelle 6: Die Prompts in der zweiten Studie

Prompts
1. Beschreiben Sie zwei bis drei Beispiele, in denen Sie geprüft worden sind (= Test-situationen) und suchen Sie Situationen, in denen die Durchführungsobjektivität, Auswertungsobjektivität oder die Interpretationsobjektivität verletzt worden ist.
2. Stellen Sie sich folgende Situation vor: Sie entwickeln einen Sprachtest für Vorschulkinder. Wenn dieser Test schlecht ausfällt, werden die Kinder einer Fördermaßnahme zugewiesen. Wie stellen Sie die Reliabilität des Tests fest und welches Problem entsteht, wenn die Reliabilität nicht gegeben ist?

Tabelle 6: Die Prompts in der zweiten Studie

Prompts

3. Aus Ihrer Schulbiographie kennen Sie vielleicht Situationen, in denen Sie das Gefühl hatten, „nicht das Richtige gefragt worden zu sein“. Testtheoretisch handelt es sich hierbei um die Verletzung der Validität. Beschreiben oder konstruieren Sie ein solches Beispiel. Behalten Sie dabei die verschiedenen Formen der Validität im Hinterkopf.

4. Nehmen Sie das Konstrukt der Kreativität und überlegen Sie: Gegen welches andere relevante Konstrukt würden Sie es abgrenzen wollen?

5. Ein Kind hat im Vergleich zu seinen Klassenkameraden schlechte Leistungen im Fach Deutsch. Welche Konstrukte würden Sie untersuchen und an welchen Merkmalen würden Sie diese festmachen wollen?

6. Wählen Sie ein für Ihren angestrebten Beruf relevantes Thema und formulieren Sie eine Forschungsfrage sowie eine operationale und eine statistische Hypothese zu dieser Forschungsfrage.

7. In einer Studie wurden Kinder mit Aufmerksamkeitsstörungen untersucht. Man hat in einer Kontrollgruppe und in einer Experimentalgruppe getestet, ob die visuelle Eingrenzung der Lernumgebung durch Pappwände dazu führt, dass die Kinder besser arbeiten als in einer ablenkungsreichen Umgebung. Gemessen wird die Zeit in Sekunden, die die Kinder an der Aufgabe (Abschreibeaufgabe) bleiben und die Zahl der Schreibfehler, die die Kinder machen.

Die Unterschiedshypothesen lauten: Kinder in einer abgeschirmten Umgebung bleiben länger an den Aufgaben. Kinder in einer abgeschirmten Umgebung machen weniger Fehler beim Abschreiben. Die erste Hypothese wird bei einem α - Fehler-Niveau von $p < .05$ signifikant (β - Fehler = .90). Die zweite Hypothese wird bei einem α - Fehler-Niveau von $p < .10$ nicht signifikant (β - Fehler = .40) Reflektieren Sie an diesem Beispiel, welche Konsequenzen und Auswirkungen solche Fehler (alpha und beta) haben können.

8. Entwickeln Sie eine praktische Fragestellung, die mit dem Verfahren der einfaktoriellen Varianzanalyse bearbeitet werden kann. Machen Sie bitte konkrete Angaben zu den experimentellen Bedingungen und der Stichprobengröße. Bitte schreiben Sie etwa 250 Wörter.

9. Sie haben im Seminar verschiedene statistische Verfahren kennengelernt. Wie bewerten Sie den fachlichen Nutzen dieses Wissens für Ihr weiteres Studium?

Durchführung

Die Untersuchung fand in einem Seminar mit dem Thema „Einführung in das empirische Arbeiten“ an der Johann Wolfgang Goethe-Universität Frankfurt im Fachbereich Erziehungswissenschaften statt. In der ersten Sitzung wurden die Prompts als Unterstützungsinstrument des Lernens eingeführt. Die Studierenden wurden über den Sinn und das Ziel der Prompting-Maßnahme informiert und es wurde anhand eines Musterbeispiels (vgl. Anhang) verdeutlicht, wie die Prompts bearbeitet werden sollten. Vor der Bearbeitung des ersten Prompts führten die Seminarteilnehmer einen Wissenstest durch und wurden anschließend drei Gruppen (A, B und C) zugewiesen. Die Zuordnung erfolgte ohne Wissen der Teilnehmer und diente ausschließlich der Untersuchung. Sodann wurde über die Dauer von neun Sitzungen nach jeder Sitzung ein Prompt auf der Lernplattform Moodle bereitgestellt, den die Seminarteilnehmer innerhalb von drei Tagen bearbeiten sollten. Auf die eingereichten Prompts erhielt nun jede der drei Gruppen über die Dauer von drei Prompts eine Variante des Feedbacks. Letztlich bekam also jeder Seminarteilnehmer drei Mal jede Feedbackform. Der genaue Ablauf wird in Tabelle 7 dargestellt.

Tabelle 7: Feedbackreihenfolge in den Gruppen

Gruppe A	Selbstfeedback	kein Feedback	Dozentenfeedback
Gruppe B	Dozentenfeedback	Selbstfeedback	kein Feedback
Gruppe C	kein Feedback	Dozentenfeedback	Selbstfeedback

Um die Effekte der Intervention zu erfassen, erfolgte jeweils nach einem Feedbackblock und drei Prompts eine Erhebung der aktuellen Motivation mit dem FAM (Rheinberg et al., 2001). Die Verarbeitungstiefe der Texte wurde über eine qualitative Analyse der Texte erfasst. Im Anschluss an die Intervention erhielten die Studierenden einen Fragebogen zur Erfassung der *cognitive load* (Paas, 1992), in dem die Studierenden beurteilen sollten, wie anstrengend sie die einzelnen Prompts wahrgenommen haben. Es sollte außerdem eingeschätzt werden, wie nützlich die Prompts waren und wie viel Zeit im Schnitt für die Bearbeitung nötig war. Die gesamte Untersuchung umfasste den Zeitraum von Oktober 2009 bis März 2010. Die Analyse der Texte nahmen zwei Raterinnen vor.

5.3.2 Ergebnisse

Für die Datenaufbereitung wurden die Programme Excel (2007) und SPSS (18) verwendet. Mit der SOLO-Taxonomie wurde nach Rücksprache (via E-Mail) mit Biggs (siehe .8) auf metrischem Skalenniveau verfahren. Biggs bestätigt in der E-Mail, dass die SOLO-Taxonomie durchaus metrisch behandelt werden kann. Da die Ausfälle von Studierenden innerhalb des Seminars so groß waren, wurden zur statistischen Auswertung der Daten non-parametrische Verfahren herangezogen. Zu Beginn des Seminars nahmen 51 Teilnehmer teil, zur Mitte des Semesters nur noch 34 und 19 bearbeiteten am Ende des Seminars noch den abschließenden Test. Um Zusammenhänge zwischen Schreibqualität und Vorwissen zu untersuchen, wurde vor Beginn der Untersuchung zunächst das Vorwissen der Teilnehmer überprüft. Der Test enthielt 15 Multiple-Choice-Fragen zu Ausschnitten des im Seminar zu vermittelnden Lernstoffes. Bei der statistischen Auswertung der Daten wurden zuerst die Reliabilitäten des Wissenstests und des FAM geprüft. Anschließend wurden die Verteilungen der Texte auf den Ergebnisebenen der SOLO-Taxonomie betrachtet. Die weiteren Analysen orientieren sich an der Reihenfolge der Fragestellungen in Kapitel 4.2.1.

Reliabilität des FAM

Insgesamt weisen die Skalen des FAM zufriedenstellende Konsistenzen auf ($N=30$). Die Reliabilität (Cronbachs α) der Subskalen liegt zwischen .70 (*Misserfolgsbefürchtung*) und .84 (*Interesse*). Für die Skala *Erfolgswahrscheinlichkeit* ergibt sich ein Wert von $\alpha = .71$, die *Herausforderung* liegt bei $\alpha = .76$. Insgesamt erwies sich damit die Zuverlässigkeit des FAM als erwartungskonform mit den Studien von Rheinberg et al. (2001). In Tabelle 8 sind die Reliabilitäten aus den Studien von Rheinberg et al. (2001) abgebildet.

Wissenstest vor und nach dem Seminar

Vor der Prompt-Intervention absolvierten die Studierenden einen Vorwissenstest mit dem Ziel, das konzeptuelle und inhaltsspezifische Wissen der Teilnehmer zu ermitteln. Die Rolle des Vorwissens beim Lernen wird in zahlreichen Studien betont, wenngleich eine inkonsistente und aufgrund methodischer Heterogenität schwer zu interpretierende empirische Befundlage vorliegt (vgl. Stark, Flender &

Tabelle 8: Reliabilitäten (Cronbachs α) des FAM

Skala	Rheinberg et al. (2001)	vorliegende Studie (n= 38)
Misserfolgsbefürchtung	.71 - .85	.70
Erfolgswahrscheinlichkeit	.68 - .88	.71
Herausforderung	.66 - .81	.76
Interesse	.71 - .90	.84

Mandl, 2001). Im Vorwissenstest in der vorliegenden Studie wurden Ausschnitte des im Seminar zu vermittelnden Lernstoffes abgebildet, dessen inhaltliches Verständnis für die Bearbeitung der Prompts notwendig war. Ziel ist es zu überprüfen, inwiefern Zusammenhänge zwischen dem inhaltsspezifischen Vorwissen und der Verarbeitungstiefe in den Texten bestehen. Der Test setzte sich aus 15 Multiple-Choice-Items zusammen (siehe Anhang .6). Für jedes Item wurden fünf Distraktoren formuliert und es waren jeweils unterschiedlich viele korrekte Antworten möglich. Pro Aufgabe konnte ein Punkt erreicht werden. Dieser wurde dann vergeben, wenn ein Seminarteilnehmer sowohl die richtigen Antwortvorgaben angekreuzt hatte als auch die falschen Antwortvorgaben frei gelassen hatte. In Tabelle 9 werden die Verteilungskennwerte für die Items dargestellt sowie die maximal und minimal erreichte Punktzahl. Da in der empirischen Literatur Befunde dazu gefunden wurden, dass Vorwissen die Qualität des nachfolgenden Lernprozesses beeinflusst, wird auch hier der Zusammenhang zwischen dem inhaltsspezifischen Vorwissen und der Verarbeitungstiefe in den Texten überprüft. Dazu wurden Korrelationen für die Mittelwerte der Verarbeitungstiefen und des Vorwissenstests gerechnet. Es konnte in dieser Hinsicht allerdings kein statistisch bedeutsamer Zusammenhang festgestellt werden ($r = .022$, n.s.). Man kann also festhalten, dass das Vorwissen keinen nachweisbaren Einfluss auf die Komplexität der kognitiven Prozesse im nachfolgenden Schreibprozess hat. Tabelle 9 lässt erkennen, dass die Item-Schwierigkeiten im Vorwissenstest als relativ hoch einzustufen sind.

Tabelle 9: Itemschwierigkeiten P und Standardabweichungen der Items aus dem Vorwissenstest (n=19) und dem Nachtest

Item	Vortest P (SD)	Nachtest P (SD)	Trennschärfe VW	Trennschärfe NW
Item 1	42 (.51)	35 (.49)	.088	.168
Item 2	26 (.45)	6 (.24)	-.029	-.075
Item 3	11 (.32)	6 (.24)	.257	.622
Item 4	11 (.32)	29 (.47)	-.335	.365
Item 5	21 (.42)	29 (.47)	-.157	.072
Item 6	21 (.42)	41 (.51)	-.157	.278
Item 7	0 (.00) ¹	6 (.24)	0	-.075
Item 8	0 (.00) ¹	0 (.00) ¹	0	0
Item 9	11 (.32)	0 (.00) ¹	-.335	0
Item 10	16 (.38)	35 (.49)	-.180	-.233
Item 11	5 (.23)	0 (.00)	-.053	0
Item 12	26 (.45)	41 (.51)	-.399	.278
Item 13	5 (.23)	0 (.00) ¹	-.427	0
Item 14	16 (.38)	24 (.44)	.369	.077
Item 15	21 (.42)	0 (.00) ¹	.507	0

¹ Anmerkung: keiner der Probanden konnte das Item lösen

Zwei Items wurden von keinem Probanden gelöst. Der Schwierigkeitsindex (Itemschwierigkeit) wird über das Verhältnis der richtigen zur Anzahl der falschen Lösungen berechnet und gibt Auskunft darüber, wie viel Prozent der Personen das Item korrekt lösen konnten (vgl. Moosbrugger & Kelava, 2007). Er errechnet sich aus der Formel:

$$P_i = \frac{\bar{x}_i}{\max(x_i)} \cdot 100$$

Dabei ist \bar{x}_i der Itemmittelwert und $\max(x_i)$ der maximal erreichbare Wert in einem Item i . Ein Schwierigkeitsindex von 42 drückt demnach aus, dass 42% der Personen die Aufgabe korrekt gelöst haben. Beim vorliegenden Test erreicht nur ein Item (1) eine mittlere Schwierigkeit. Zwei Items (7 und 8) wurden gar nicht gelöst. Die durchschnittliche Itemschwierigkeit liegt bei $P = 18$, womit ein

Bodeneffekt konstatiert wird. Tests mit hoher Itemschwierigkeit erlauben keine Differenzierung innerhalb der Gruppe. Um gut zu differenzieren, sollten sich Itemschwierigkeiten im mittleren Bereich häufen ($P = 50$) (Lienert & Raatz, 1998). Das Vorwissen in der Gruppe ist somit als niedrig einzustufen. Im Wissenstest nach dem Seminar wurde das selbe Verfahren mit allen 15 Items erneut eingesetzt. Im Nachtest ist die durchschnittliche Itemschwierigkeit mit $P = 25$ nur unwesentlich höher. Auffällig ist, dass im Nachtest fünf Items von keiner Person richtig beantwortet wurden. Es lassen sich aber aufgrund der Testschwierigkeit keine differenzierten Aussagen über gute bzw. schwache Teilnehmer machen. Der Wissenszuwachs in diesem Seminar konnte demzufolge durch den Test nicht abgebildet werden.

Zur weiteren Untersuchung der Effekte des Feedbacks muss vorausgesetzt werden, dass die Gruppen, die verschiedene Feedbackvarianten erhalten sollten, den Lernprozess mit vergleichbarem Vorwissensniveau beginnen. Daher wurden mit Hilfe non-parametrischer Varianzanalysen (Kruskal-Wallis) Vergleiche des Vorwissensniveaus durchgeführt. Es ergab sich auf dem 5%-Niveau jedoch kein signifikanter Unterschied für die Gruppen A, B und C. Die Gruppen gehen demnach mit gleichen Voraussetzungen hinsichtlich ihres Vorwissens in die Intervention.

Befunde zur Verarbeitungstiefe der Texte über ein Semester

Das Auswertungsverfahren mit der SOLO-Taxonomie wurde bereits in Kapitel 5.2.2 beschrieben. Die Auswertung der Texte in der zweiten Studie wurde erneut von zwei trainierten und supervidierten Raterinnen vorgenommen, die nach einem intensiven Training 91% konkordante Übereinstimmungen erreichten (Platz-zu-Platz-Übereinstimmung).

Um an die Vorstudie anzuschließen, wurde zunächst die Verteilung der Kategorien auf den SOLO-Niveaus betrachtet. In der zweiten Studie wurden insgesamt 182 Texte analysiert und mit Hilfe der SOLO-Taxonomie eingestuft. Der Hauptanteil von Texten (60%) wurde auf dem *multi-structural*-Niveau eingestuft. Im Vergleich zur Vorstudie fällt auf, dass sich kein so klarer Schwerpunkt bezüglich der Verarbeitungstiefe feststellen lässt, sondern dass die Verteilung der Texte eher einer Normalverteilung ähnelt. Insgesamt wurden in der zweiten Studie mehr Texte auf dem ersten (*pre-structural*, 15%) und dem fünften (*extended abstract*, 8%) Niveau erzeugt als in der ersten Studie. Eine Begründung für die Unterschiede in

beiden Verteilungen könnte sein, dass das Thema der zweiten Untersuchung sich abstrakter gestaltete und daher insgesamt weniger intuitiv auf den Alltag bezogen werden konnte. Zudem war ein großer Ausfall der Teilnehmerzahlen zu verzeichnen. Das könnte zu systematischen Selektionseffekten führen. Abbildung 8 stellt beide Verteilungen zum Überblick gegenüber und verdeutlicht die Unterschiede zwischen beiden Untersuchungen.

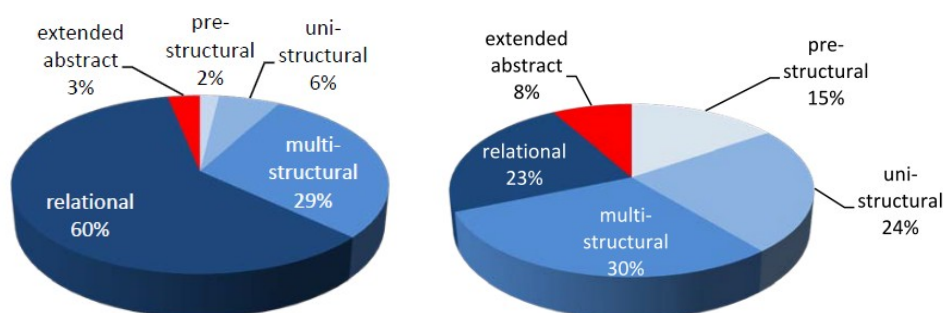


Abbildung 8: Verteilungen der Texte auf der SOLO-Taxonomie in der ersten Studie (links) und in der zweiten Studie (rechts)

Deskriptive Befunde zur aktuellen Motivation in Abhängigkeit von der Feedbackform

Fragestellung: Welche Unterschiede bestehen bei den Gruppen A, B und C in Bezug auf die Motivation nach den unterschiedlichen Feedbackformen?

Zunächst wurde auf deskriptiver Ebene betrachtet, wie sich die Motivation in den Gruppen A (n= 11), B (n= 10) und C (n= 7) über das Semester hinweg entwickelt. Da aus Gruppe C nur zwei Probanden über das gesamte Semester an der Prompting-Maßnahme teilnahmen, ist der dritte Messzeitpunkt nicht aussagekräftig. Die Mittelwerte zeigen, wie die Teilnehmer die Lernsituation zu den drei Messzeitpunkten bewerten (vgl. Tabelle 10). Auf den ersten Blick lässt sich feststellen, dass die Komponente *Erfolgswahrscheinlichkeit* über das Semester hinweg bei allen Gruppen zeitlich relativ konstant bleibt. Gleiches gilt für die *Misserfolgsbefürchtung* in den Gruppen A und B. Insgesamt beurteilen die Seminarteilnehmer die Lernsituation zu Beginn des Seminars (und nach der Bearbeitung der ersten drei Prompts) stark leistungsthematisch. Über alle drei Gruppen weist die Komponente *Herausforderung* hohe, deutlich oberhalb der Skalenmitte ange-

siedelte, Mittelwerte auf. Dies lässt darauf schließen, dass die Studierenden die Aufgabe anfangs als anspruchsvoll einstufen. Dennoch sind die Teilnehmer über das gesamte Semester hinweg zuversichtlich, die Aufgabe bewältigen zu können, wie der durchgängig hohe Mittelwert der Komponente *Erfolgswahrscheinlichkeit* zeigt. Die *Misserfolgsbefürchtung* liegt bei den Gruppen A und B zu Beginn der Intervention knapp unter der Skalenmitte ($M_A = 2.89$, $M_B = 3.40$), bei Gruppe C etwas höher ($M_C = 4.0$). Das *Interesse* der Seminarteilnehmer am Aufgabeninhalt ist bei allen Gruppen relativ hoch ausgeprägt ($M_A = 4.25$, $M_B = 4.40$, $M_C = 3.77$). Aus der Betrachtung der Motivationswerte zum ersten Messzeitpunkt lässt sich schließen, dass gute motivationale Voraussetzungen für die Lernsituation nach den ersten drei Prompts vorliegen. Nach dem ersten Wechsel des Feedbacks ist dann ein Abwärtstrend für die Skalen *Herausforderung* und *Interesse* zu verzeichnen. Hier zeichnen sich auch Unterschiede zwischen den Gruppen ab. Um zu überprüfen, ob sich die Motivationskomponenten des FAM während der drei Messzeitpunkte signifikant verändern, wurde mit dem Friedman-Test für abhängige Stichproben eine Rangvarianzanalyse vorgenommen.

Tabelle 10: Mittelwerte und Standardabweichung (in Klammern) der Gruppen A, B und C zu drei Messzeitpunkten mit dem FAM

Skala	Gruppe A (Z1-Z3)		Gruppe B (Z1-Z3)		Gruppe C (Z1-Z3)				
Misserfolgs- befürchtung	2.89 (1.29)	2.53 (.89)	2.88 (1.06)	3.40 (1.04)	3.28 (1.32)	2.89 (.87)	4.00 (.80)	2.89 (1.15)	4.30 (2.12)
Erfolgs- wahrscheinlichkeit	5.79 (.84)	6.06 (.74)	5.78 (.80)	5.55 (.96)	5.08 (1.14)	4.86 (1.04)	4.71 (.57)	5.18 (.94)	4.88 (1.95)
Herausforderung	4.68 (.99)	3.97 (1.39)	3.03 (1.02)	4.65 (1.70)	3.85 (1.56)	3.11 (1.64)	4.82 (1.0)	3.89 (1.16)	3.50 (1.41)
Interesse	4.25 (1.02)	3.60 (1.40)	3.43 (1.33)	4.40 (1.66)	3.24 (1.28)	2.54 (1.00)	3.77 (1.24)	3.57 (1.37)	3.80 (.28)

Der Friedman-Test zeigt folgende signifikante Änderungen der Motivationswerte: In Gruppe A ändert sich die *Misserfolgsbefürchtung* über die drei Erhebungszeitpunkte signifikant ($\chi^2 = 6.0$, $df = 2$, $n = 7$, $p = .05$). Die *Herausforderung* verändert sich ebenfalls signifikant ($\chi^2 = 9.85$, $df = 2$, $n = 7$, $p = .007$). Ein Blick auf die Mittelwerte zeigt, dass die Studierenden die Aufgabe zum Semesterende hin als weniger herausfordernd einschätzen. In Gruppe B verändert sich die *Herausforderung* in ähnlicher Weise ($\chi^2 = 11.63$, $df = 2$, $n = 7$, $p = .003$) und sinkt ab. Hier ergibt sich auch für die Motivationskomponente *Interesse* eine signifikante Abnahme ($\chi^2 = 10.18$, $df = 2$, $n = 7$, $p = .006$).

Zusammengefasst zeigen die Befunde zum einen, dass über alle drei Gruppen gute motivationale Voraussetzungen zum ersten Erhebungszeitpunkt vorlagen. Die hohen Mittelwerte des Faktors *Herausforderung* belegen, dass alle Studierenden die Aufgabe, Prompts zu bearbeiten, als anspruchsvoll einschätzen. Die Kombination mit ebenfalls hohen Ausprägungen der Komponenten *Interesse* und *Erfolgswahrscheinlichkeit* und einer moderaten *Misserfolgserwartung* zeigt, dass die Lernsituation als motivational anregend wahrgenommen wurde. Über den weiteren Verlauf des Semesters hinweg zeigen sich einige deskriptive und auch signifikante Effekte des Feedbacks. Aufgrund der niedrigen Zahl an Datensätzen lassen sich aus diesen Befunden jedoch nur vorsichtige Schlüsse bezüglich der optimalen Gestaltung des Feedbacks im Hinblick auf die Lernmotivation ziehen. Aus dem Motivationsverlauf der Gruppe B lässt sich mit Einschränkung interpretieren, dass das Ausbleiben des Dozentenfeedbacks negative Auswirkungen auf das Aufgabeninteresse und auf die Anstrengungsbereitschaft (*Herausforderung*) der Lerner hat. Dies könnte so interpretiert werden, dass es im Falle von Dozentenfeedback günstiger ist, es konsequent durchzuführen und nicht mit anderen Feedbackformen zu mischen bzw. es auszusetzen. Der Verlauf von Gruppe A hingegen könnte darauf hindeuten, dass Selbstfeedback über einen gewissen Zeitraum genügen könnte, um die Anstrengungsbereitschaft beim Bearbeiten der Prompts und das Interesse für diese Aufgabe konstant zu halten.

Der FAM macht auf der Basis der Bewertung eines Lerners der aktuellen Lernsituation Prognosen zum Lernverhalten und zur Lernleistung. Rheinberg et al. (2001) weisen nach, dass für die erfolgreiche Bewältigung einer Aufgabe bei unterschiedlichen Lernbedingungen (selbstgesteuertes versus fragengeführtes Faktenlernen, Aufgabendauer) entweder affektive (*Misserfolgsbefürchtung*, *Erfolgswahr-*

scheinlichkeit) oder kognitive Motivations-Faktoren (*Interesse, Herausforderung*) eine Rolle spielen. Mit den Prompts wurde eine Lernsituation geschaffen, die als selbstgesteuert eingeordnet werden kann, da die Seminarteilnehmer das im Seminar erarbeitete Themengebiet selbstständig vertiefen mussten. Sowohl diese Aktivität als auch das anschließende Schreiben mit den Prompts erforderte eine eigenständige Organisation der Lernaktivitäten. Darüber hinaus war die Dauer der Lernsituation mit drei Monaten relativ hoch angesetzt. Nach Rheinberg et al. (2001) sollten in einer vergleichbaren Lernumgebung besonders die Faktoren *Interesse* und *Herausforderung* zum Tragen kommen, denn wenn sich ein Lerner auch für den Inhalt der Aufgabe interessiert, so ist zu erwarten, dass auch bei längerfristigen und komplexen Aufgaben das Aufgabenengagement aufrecht erhalten wird. Eine ähnliche Wirkung sollte die Herausforderung ausüben: „Solange die Aufgabe noch irgendwie lösbar, also nicht als extrem schwierig erscheint, sollten Personen mit hohen Herausforderungswerten bei auftretenden Schwierigkeiten sich besonders engagieren und relativ gute Leistungen erzielen“ (Rheinberg et al., 2001, S. 11). Vor diesem theoretischen Hintergrund kann das signifikante Absinken des Motivationsfaktors *Herausforderung* in den Gruppen A und B darauf hindeuten, dass die Anstrengungsbereitschaft über einen so langen Zeitraum unabhängig von der Reihenfolge des Feedbacks abnimmt. Möglicherweise hängt dies mit dem Neuheitswert der Aufgabe zusammen, der sich ebenfalls abnutzt. Andererseits kann das Interesse an der Aufgabe durch Selbstfeedback offenbar über ein Semester hinweg stabil gehalten werden, wohingegen ein Ausbleiben des Dozentenfeedbacks sich negativ darauf auswirkt.

Anknüpfend an die Befunde von Nückles et al. (2006) ist zu erwarten, dass sich das zeitlich stabile Interesse positiv in der Verarbeitungstiefe bzw. idealerweise im Lernerfolg widerspiegelt.

In den Abbildungen 9, 10, 11 und 12 wird die Entwicklung der Motivation auf den Subskalen des FAM in den Gruppen A, B und C über das Semester dargestellt. Es werden dabei die mittleren Ränge für jeden der drei Messzeitpunkte abgebildet.

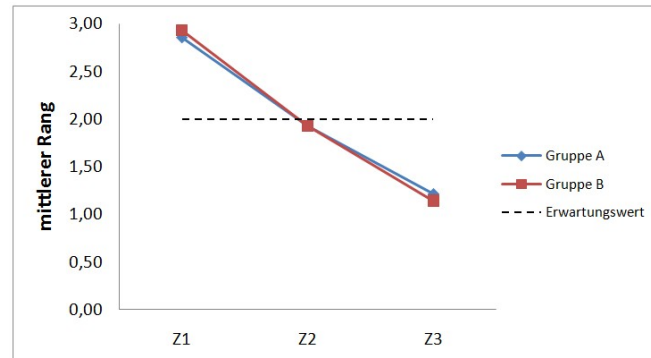


Abbildung 9: Verlauf der Motivationskomponente *Herausforderung* über die drei Messzeitpunkte

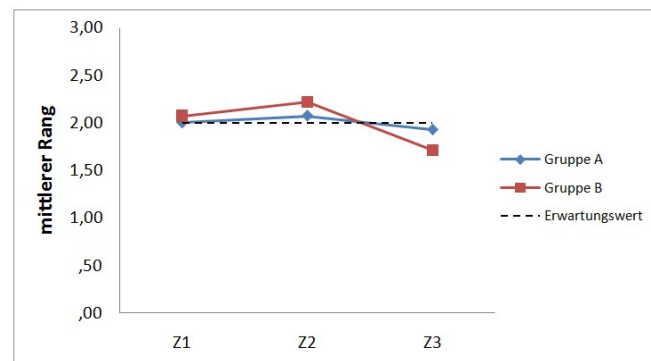


Abbildung 10: Verlauf der Motivationskomponente *Erfolgserwartung* über die drei Messzeitpunkte

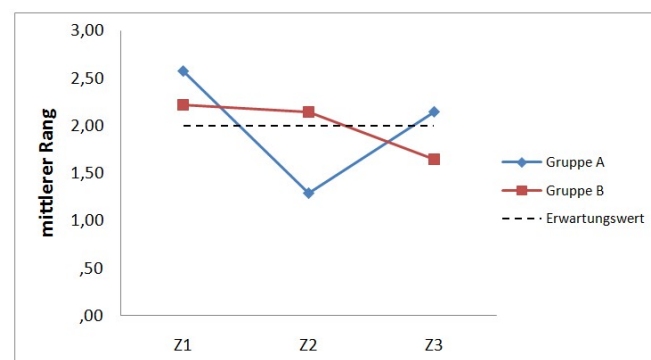


Abbildung 11: Verlauf der Motivationskomponente *Misserfolgsbefürchtung* über die drei Messzeitpunkte

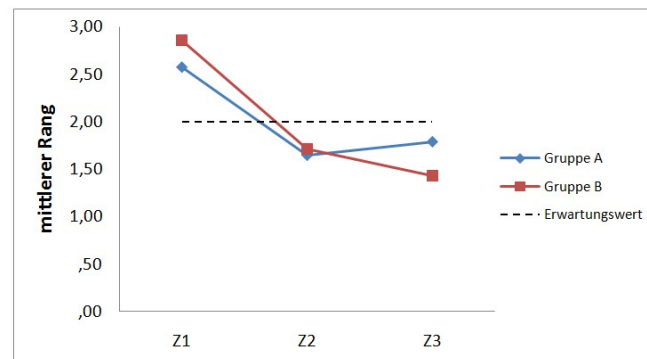


Abbildung 12: Verlauf der Motivationskomponente *Interesse* über die drei Messzeitpunkte

Feedbackform und Verarbeitungstiefe

Zunächst wurde betrachtet, wie sich innerhalb der Gruppen A, B und C die Verarbeitungstiefen über das Semester verändern. Der neunte Prompt wurde aus den Analysen ausgeschlossen, da dieser von zu wenig Seminarteilnehmern bearbeitet wurde. In Tabelle 11 sind die Mittelwerte pro Gruppe dargestellt.

Auf deskriptiver Ebene ist erkennbar, dass sich die Textqualität bei allen Gruppen zum Semesterende hin mäßig und mit Schwankungen verbessert. Die Gruppen A und B produzieren schon zu Beginn der Intervention gute Texte (M_A Prompt 1= 2.64, $SD = 1.43$; M_B Prompt 1= 2.90, $SD = .99$) und verbessern sich, lässt man Prompt 8 bei Gruppe A außer Acht, um fast einen Skaleneinheit. Das bedeutet, dass die Teilnehmer aus diesen Gruppen am Ende der Intervention Texte produzieren, die eine Niveaustufe höher eingestuft wurden als zu Beginn der Intervention. Bei Gruppe C ist ein kontinuierlicher Aufwärtstrend nach dem Dozentenfeedback (Prompts 4-6) zu beobachten und die Gruppe verbessert sich um fast zwei Niveaustufen.

Um Effekte des Feedbacks zu prüfen, wurden im Anschluss daran die Verarbeitungstiefen innerhalb der Gruppen A, B und C auf signifikante Änderungen untersucht. Da den Seminarteilnehmern nicht mitgeteilt wurde, wie oft und zu welchem Zeitpunkt sie welche Feedbackform empfangen würden, wurde davon ausgegangen, dass nach einmaligem Feedback noch keine Effekte sichtbar werden. Die Erfahrung über drei Prompts sollte jedoch genügen, dass sich die Wirksamkeit einer Feedbackform über diesen Zeitraum entwickelt. Des Weiteren wird davon ausgegangen, dass die Erfahrung des Feedbacks darüber hinaus noch einen

Tabelle 11: Mittelwerte und Standardabweichungen (in Klammern) der Einstufungen der Texte mit der SOLO-Taxonomie (P= Prompt)

	P 1	P 2	P 3	P 4	P 5	P 6	P 7	P 8
Gruppe A	2.64 (1.43)	3.11 (1.17)	2.5 (.97)	2.78 (1.20)	3.22 (.97)	3.11 (1.17)	3.57 (.79)	2.4 (1.14)
Gruppe B	2.9 (.99)	2.88 (1.64)	2.5 (1.41)	2.8 (1.48)	3.2 (1.03)	2.11 (1.05)	3.38 (1.06)	3.71 (1.11)
Gruppe C	2.0 (.89)	2.13 (.84)	1.83 (.75)	2.0 (1.27)	2.88 (1.25)	2.33 (1.51)	3.2 (.84)	3.75 (1.5)

Prompt fortwirkt. Die Effekte des Feedbacks sollten sich innerhalb der Gruppen anhand signifikanter Unterschiede zwischen den Ausprägungsgraden der Verarbeitungstiefe zum ersten Erhebungszeitpunkt (operationalisiert als der summierte Mittelwert aus Prompts 2, 3 und 4) und zum zweiten Erhebungszeitpunkt (operationalisiert als der summierte Mittelwert aus Prompts 5, 6 und 7) zeigen. In Tabelle 12 wird dieses Verfahren exemplarisch dargestellt.

Tabelle 12: Vorgehen bei der Untersuchung der Feedbackeffekte auf die Textqualität

	Gruppe A	Gruppe B	Gruppe C
Prompt 1		...	
<i>Selbstfeedback</i>		...	
Prompt 2		...	
<i>Selbstfeedback</i>		...	
Prompt 3		...	
<i>Selbstfeedback</i>		...	
Prompt 4			
<i>Dozentenfeedback</i>			
...			

Die Effekte für die Feedbackform Selbstfeedback wurden an diesen Prompts überprüft

Mit dem Wilcoxon-Test wurde ermittelt, inwieweit sich die mittleren Tendenzen in der Verteilung der Verarbeitungstiefen zu den zwei Messzeitpunkten unterscheiden.

Der Wilcoxon-Test zeigt bei einer zweiseitigen Prüfung, dass sich die Verarbeitungstiefe in den Gruppen A und C tendenziell signifikant verändert. Die Teil-

nehmer aus Gruppe A erreichen in der zweiten Hälfte des Semesters (Prompts 5,6 und 7) höhere Verarbeitungstiefen als in der ersten Hälfte ($Z = -1.79; p = .074$, Prompts 2,3 und 4). Gleiches gilt für die Teilnehmer der Gruppe C ($Z = -1.79, p = .073$). Gruppe B hingegen weist ungefähr gleich viele positive wie negative Ränge auf.

Wie auch bei der vorigen Fragestellung lassen sich nur vorsichtige Schlüsse formulieren. So sind auf der Basis der vorliegenden Befunde keine eindeutigen und generalisierbaren Aussagen über die Vorzüge einer Feedbackform gegenüber einer anderen möglich. Es wird sichtbar, dass sich alle Seminarteilnehmer im Laufe des Semesters in ihren Verarbeitungstiefen verbessern. Dies wird besonders bei Gruppe C deutlich. Die Teilnehmer dieser Gruppe begannen auf einem eher niedrigen Niveau, verbesserten sich dann aber zum Semesterende hin um fast zwei Niveaustufen. Diese Entwicklung könnte ein Hinweis auf einen Expertise-Umkehreffekt sein, bei dem davon ausgegangen wird, dass insbesondere ungeübte Lerner von den strukturierten Prompts profitieren, wohingegen geübtere Schreiber bzw. Lerner eher von offenen Arbeitsaufträgen profitieren.

Zusätzlich wurde geprüft, inwiefern Anstrengung beim Abschneiden in den Prompts eine Rolle spielt. Dazu wurden bivariate Korrelationen nach Pearson zwischen dem durchschnittlichen *cognitive-load*-Wert pro Prompt (die Studierenden schätzten für jeden Prompt ein, wie schwierig sie ihn fanden) und der Verarbeitungstiefe berechnet. Es zeigten sich jedoch keine signifikanten Zusammenhänge zwischen der Verarbeitungstiefe und der subjektiv empfundenen Anstrengung pro Prompt.

Verarbeitungstiefe und Motivation

In einem weiteren Schritt wurde geprüft, ob zwischen der Motivation und der Verarbeitungstiefe ein Zusammenhang besteht. Dazu wurden Korrelationen nach Pearson zwischen den mittleren Verarbeitungstiefen aus den Prompts 1, 2 und 3 (Prompt123), den Prompts 4, 5 und 6 (Prompt456) und den Prompts 7 und 8 (Prompt78) mit den Mittelwerten der FAM-Subskalen zu Z1, Z2 und Z3 gerechnet. Im Ergebnis zeigte sich, dass sich weder die vorausgehende Verarbeitungstiefe auf die nachfolgende Motivation auswirkt, noch dass sich die aktuelle Motivation auf die nachfolgende Verarbeitungstiefe auswirkt. In Tabelle 13 sind exemplarisch die Berechnungen dargestellt.

Tabelle 13: Korrelationen zwischen den Subskalen des FAM zu Z1 und der Verarbeitungstiefe

		Prompts 123	Prompts 456
Erfolg Z1	Pearson	-.201	-.151
	Sig. (2-seitig)	.315	.452
	N	27	27
Interesse Z1	Pearson	.012	.057
	Sig. (2-seitig)	.951	.777
	N	27	27
Misserfolg Z1	Pearson	.283	.187
	Sig. (2-seitig)	.153	.351
	N	27	27
Herausforderung Z1	Pearson	.094	.058
	Sig. (2-seitig)	.642	.774
	N	27	27

Die Korrelationen sind äußerst gering und erreichen in keinem Fall die Signifikanzgrenze. Für alle anderen Messzeitpunkte ergibt sich ein ähnliches Befundmuster.

Verarbeitungstiefe und Lernzuwachs

Was die Beziehung zwischen Verarbeitungstiefe und Lernleistung angeht, so war vermutet worden, dass das Niveau der kognitiven Strukturen in den Texten positive Bezüge zur Lernleistung aufweist. Um diese Annahme zu überprüfen, wurden der Lernzuwachs und die durchschnittliche Verarbeitungstiefe der Seminarteilnehmer miteinander in Beziehung gesetzt. Es wurden Korrelationen zwischen der Differenz aus dem Wissenstest vor und nach dem Seminar mit den Verarbeitungstiefen bei den einzelnen Prompts berechnet. Es zeigt sich jedoch auch hier kein substanzieller Zusammenhang. Einschränkend muss gesagt werden, dass aufgrund der vorweg ermittelten Testschwierigkeit davon ausgegangen wird, dass der Wissenszuwachs im Seminar durch den Test nicht gut abgebildet werden konnte und daher auch weiterführende Analysen nicht sinnvoll sind.

Verarbeitungstiefe und Seminarerfolg

Da mit dem Wissenstest aufgrund der Testschwierigkeit kein Lernzuwachs abgebildet werden konnte, wurde als externes Leistungsmaß die Klausur am Ende des Seminars herangezogen. Um zu prüfen, inwiefern die Verarbeitungstiefe in den Texten zu einem besseren Seminarergebnis führt, wurde die Korrelation zwischen dem Mittelwert in der Verarbeitungstiefe über alle Prompts hinweg und der Klausurnote berechnet. Für diese Berechnung wurde der Spearman- Rangkorrelationskoeffizient herangezogen, weil die Zahl der Probanden mit $n = 14$ eher gering war und den Klausurnoten Ordinalskalenniveau unterstellt wird. Dabei zeigte sich ein signifikanter Zusammenhang von $r_{ho} = -.57$ ($p < .05$). Je höher also das durchschnittlich erreichte Komplexitätsniveau in den Schreibaufgaben eingestuft wurde, desto besser fiel die Klausurleistung in Noten von 1-6 aus. Vergleicht man die erreichte Verarbeitungstiefe in den Lerntagebüchern bei den Teilnehmern, die die Klausur mitgeschrieben haben, und den Seminarabbrechern, so zeigen sich die folgenden Effekte: Die Klausurteilnehmer erreichten eine Verarbeitungstiefe von durchschnittlich $M = 3.10$, $SD = .49$, während die Abbrecher eine durchschnittliche Verarbeitungstiefe von $M = 2.35$, $SD = .75$ erzielten. Dieser Unterschied ist signifikant ($t(df = 28) = -3.11$, $p < .01$).

Damit unterscheiden sich die erfolgreichen Studierenden im Seminar in der Verarbeitungstiefe von den weniger erfolgreichen und innerhalb der Gruppe der erfolgreichen Studierenden scheint die erreichte Verarbeitungstiefe ein Prädiktor für Seminarerfolg zu sein.

6 Diskussion

Aufbauend auf Forschungsarbeiten zu kognitiven und affektiven Prozessen beim Schreiben wurden im Rahmen dieser Arbeit Schreibaufgaben zu Lerninhalten (Prompts) mit verschiedenen Anforderungsmerkmalen daraufhin getestet, wie sie kognitive Prozesse beim Schreiben anstoßen, die sich in entsprechend anspruchsvollen Schreibergebnissen niederschlagen. Dazu bearbeiteten Lehramtstudierende in zwei aufeinander aufbauenden Untersuchungen mehrere Schreibaufträge in Form von kurzen, spezifischen Prompts, die sich jeweils auf einen Aspekt einer Seminareinheit bezogen.

Studie 1: An der ersten Studie nahmen 105 Studierende des Lehramts aus drei Seminaren teil. Eines der drei Seminare erstreckte sich über einen Zeitraum von einer Woche, die zwei weiteren in die Untersuchung eingehenden Seminare liefen über ein Semester. Thematisch handelte es sich um einführende Seminare aus dem Bereich der Pädagogischen Psychologie mit den Themenbereichen „Einführung in die Sozialpsychologie“ und „Einführung in die Entwicklungspsychologie“. Die Teilnehmer bearbeiteten insgesamt 16 Schreibaufträge. Das Ziel war, zu testen, inwiefern sich die Prompts zur Förderung kognitiver Prozesse eignen und wie gut die Analyse der Textqualität mit einer Taxonomie zur Beschreibung von Lernergebnissen (SOLO-Taxonomie) gelingt. Die Studierenden erhielten nach jeder thematischen Lerneinheit zwei Prompts, auf deren Grundlage sie im Anschluss an die Veranstaltung einen Text verfassten. Aus den Ergebnissen konnte herausgearbeitet werden, welche Form der Prompts lernstrategieförderlich sind und zu besonders anspruchsvollen Texten anregen bzw. welche Art von Prompts wahrscheinlich weniger ergiebig sind.

Studie 2: Die zweite Untersuchung umfasste einen Zeitraum von drei Monaten und war in ein Seminar für Studierende des Studienganges Diplom-Pädagogik eingebettet. Thematisch behandelte das Seminar die Einführung in empirische Methoden für Pädagogen. Mittels der in der ersten Untersuchung als strategieförderlich getesteten Prompts verfassten 46 Studierende Texte zu den Lerninhalten aus dem Seminar. Es wurden insgesamt neun Prompts verwendet, von denen

die Studierenden nach jeder Veranstaltung einen bearbeiteten. Die Bearbeitung erfolgte online mittels der Lernplattform Moodle. Die Texte wurden direkt im Anschluss an die Bearbeitung mit der SOLO-Taxonomie eingestuft. Diese Einstufung wurde genutzt, um den Studierenden Rückmeldung zu geben. Die Studierenden erhielten eine von drei Formen des Feedbacks (Anleitung zum Selbstfeedback, Dozentenfeedback, kein Feedback). Um zu untersuchen, inwiefern sich verschiedene Formen von Feedback auf die Qualität der Texte auswirken, wurden innerhalb des Seminars drei Gruppen (A, B und C) gebildet. Alle Teilnehmer erhielten in derselben Reihenfolge die gleichen Prompts. In den verschiedenen Gruppen wurde jedoch die Reihenfolge des Feedbacks variiert, wie in Tabelle 14 dargestellt. Darüber hinaus interessierten Zusammenhänge zwischen Aspekten der Motivation der Studierenden, unterschiedlichem Feedback und der Verarbeitungstiefe.

Tabelle 14: Feedbackreihenfolge in den Gruppen

Gruppe A	Selbstfeedback	kein Feedback	Dozentenfeedback
Gruppe B	Dozentenfeedback	Selbstfeedback	kein Feedback
Gruppe C	kein Feedback	Dozentenfeedback	Selbstfeedback

In beiden Studien wurden die mit den Prompts produzierten Texte anhand einer Taxonomie zur Beschreibung von Lernergebnissen auf ihre Komplexität hin überprüft. Dies gab Aufschluss über die Struktur des Wissens und erlaubte Rückschlüsse auf die Verarbeitungstiefe. Auf der Grundlage der eingestuften Texte wurde analysiert, ob sich Feedback positiv auf die Qualität der Schreibprodukte auswirkt. Dieses wiederum wird als Verbesserung des Schreibprozesses interpretiert. Zusätzlich wurde geprüft, ob die Motivation der Probanden einen Einfluss auf die Bearbeitungsqualität der Aufgaben hat.

Zur Gewinnung von Daten setzt auch die vorliegende Untersuchung die Methode der Inhaltsanalyse mit der SOLO-Taxonomie zur Abbildung der kognitiven Struktur der Textprodukte ein. Es wird außerdem ein standardisierter Fragebogen zur Erhebung der aktuellen Motivation verwendet.

Ein Ziel der Arbeit war es, Texte von Studierenden nach Qualitätsmerkmalen zu unterscheiden. Dazu wurde die SOLO-Taxonomie eingesetzt, da diese differenzierte Aussagen über kognitive Strukturen ermöglicht. Wie in Abschnitt 2.3.2 dargestellt wurde, fokussieren die Beschreibungskategorien aus den bisherigen Studien

innerhalb der Textproduktionsforschung zum einen auf die Menge erinnerter Informationen beim Schreiben mit Texten und zum anderen auf die Textkohärenz als Maß des Verstehens. Diese Maße sind zwar interessant, geben jedoch wenig Aufschluss über Zusammenhänge zwischen der Qualität des Schreibprozesses und der Wissensstruktur des Textproduzenten. Wissenseffekte von Schreibprozessen mit Prompts lassen sich hingegen gut abbilden, wenn man analysiert, inwiefern und in welcher Qualität Textproduzenten einzelne Aspekte eines Themenfeldes in ihren Texten vernetzen (z.B. indem die Schreiber Analogien bilden oder elaborierte Vergleiche anstellen). Diese Vernetzungen können als Qualität des Wissens interpretiert werden (Biggs & Collis, 1982).

Da die bisherigen Beschreibungskategorien hier zu undifferenziert sind, sollte in der ersten Untersuchung gezeigt werden, dass diese Strukturen mittels Textanalyse zuverlässig abgebildet und kategorisiert werden können. In der zweiten Untersuchung sollten dann darauf aufbauend der Einfluss motivationaler Voraussetzungen und verschiedener Rückmeldeformen mit in den Blick genommen werden.

Im Folgenden werden die Ergebnisse der Untersuchungen zusammengefasst und diskutiert. Zunächst wird die SOLO-Taxonomie als Diagnose-Instrument zur Einstufung der Texte erörtert und beurteilt. Daran schließt sich eine Bewertung der Ergebnisse zu den Strukturanalysen subjektiver Lernkonzepte an. Die Diskussion schließt mit einer Zusammenfassung und einem Ausblick, der weitere Einsatzmöglichkeiten von Prompts zur Förderung tiefenorientierten Lernens in Seminaren aufzeigt.

6.1 Zur Kategorisierung der Texte mit der SOLO-Taxonomie

Um Aussagen über die Effekte der Prompts auf die Verarbeitungstiefe beim Produzieren von Texten zu machen, wurden die Texte in beiden Studien, die im Rahmen dieser Arbeit durchgeführt wurden, mithilfe der SOLO-Taxonomie auf ihre kognitive Struktur hin analysiert und eingestuft. Im Folgenden werden zunächst die Eigenschaften dieser Methode zur Datengewinnung dargestellt und diese anschließend im Hinblick auf ihre Tauglichkeit bewertet.

Eine gängige Methode zur Beschreibung der Zusammenhänge zwischen kognitiven Prozessen und der Qualität von Schreibprodukten in der Textproduktions-

forschung ist die Analyse von Denkprotokollen. Textproduzenten berichten dazu während des Schreibens oder in kurzen Schreibpausen, was sie in einer spezifischen Situation denken und wie sie vorzugehen planen. Diese Aussagen werden aufgezeichnet und anschließend ausgewertet.

Wie bereits angedeutet wurde, gibt es jedoch zahlreiche Kritikpunkte an dieser Methode. Diese beziehen sich zusammengefasst auf die unzureichende Validität von Selbstaussagen, Verzerrungen des Lernprozesses durch die Unterbrechung und kognitive Überlastung der Lernenden und zuletzt auf die Interpretation der verbalen Daten, deren Fehleranfälligkeit betont wurde (z.B. Garner, 1988).

Die Analyse der Texte anhand der SOLO-Taxonomie stellt hier eine Alternative dar. Zum einen ermöglicht die Textanalyse die Durchführung komplexer kognitiver Prozesse ohne Unterbrechung. Die Lernprozesse sind also authentisch und nicht von außen beeinflusst. Zum anderen stellt sich hinsichtlich der Interpretation der Daten die Frage nach der Relevanz einzelner Informationen nicht, denn die SOLO-Taxonomie stellt Beschreibungskategorien für die Einstufung zusammenhängender Texte zur Verfügung. Ein Selektionsverfahren bei der Analyse der Texte nach wichtigen und nicht wichtigen Teilen erübrigt sich somit. Zuletzt beschränkt sich die Erfassung des Lernverhaltens im Unterschied zu Denkprotokollen auf spezifische, angewandte Strategien der Textverarbeitung und nicht auf Selbstberichte.

Ein weiteres, für diese Arbeit relevantes, methodisches Problem betrifft die in den bisherigen Studien vorgenommene Art der Kategorisierung der Texte. So wurden als Lernmaße bislang häufig Textkohärenz und die Menge erinnerter Informationen (beim Lernen mit Texten) verwendet oder Multiple-Choice-Tests zur Überprüfung des Wissenszuwachses durchgeführt. Über die Qualität des Wissens sind mit diesen Maßen jedoch nur wenig differenzierte Aussagen möglich.

Mit der SOLO-Taxonomie wurde ein Analyse-Instrument gewählt, das fachübergreifende Kategorien für die qualitative Klassifikation von Wissensstrukturen bietet. Die Grundannahme des Modells besteht darin, dass Lerner mit fortschreitender Entwicklung in einem Lernfach zunehmend komplexere und abstraktere Wissensstrukturen aufbauen, die sich in den schriftlichen Ausführungen der Lerner (z.B. in schulischen Texten) widerspiegeln. Zur Abbildung der Komplexität der Wissensstrukturen postulieren Biggs und Collis (1982) fünf hierarchisch organisierte kognitive Niveaus, auf denen sich Lernergebnisse einstufen lassen. Diese

repräsentieren die Entwicklung themenspezifischen Wissens und beginnen mit jedem Lernprozess erneut. Die Entwicklung des Wissens geht in dem Modell einher mit zunehmenden Anforderungen an das Arbeitsgedächtnis und die Aufmerksamkeitsspanne: Je höher das Niveau in der SOLO-Hierarchie angesiedelt ist, umso mehr Aspekte eines Themenfeldes werden in die Antwort mit einbezogen und umso mehr Bezüge werden zwischen den Wissensselementen hergestellt. Das Wissen wird zunehmend konsistenter und dies drückt sich aus durch einen verstärkten Gebrauch von Verbindungen zwischen Einzelementen. Die SOLO-Taxonomie ist multimodal angelegt und fortgeschrittenere Niveaus implizieren jeweils die vorausgegangenen.

In der ersten Untersuchung sollte zunächst geklärt werden, ob die SOLO-Taxonomie sich zur Einstufung der Wissensqualität eignet. Die mit den Prompts produzierten Texte wurden mit der Taxonomie den fünf Niveaus zugewiesen. Dabei wurde gemäß Biggs und Collis (1982) davon ausgegangen, dass, je komplexer die im Text vorgefundenen Verbindungen zwischen einzelnen Text-Elementen sind, dem Text auch umso komplexere, kognitive Prozesse vorausgegangen sind. Bei der Kategorisierung der Texte stand die Frage im Vordergrund, ob die Zuordnung der Texte reliabel erfolgte. Die in der ersten Untersuchung berechnete prozentuale Beurteilerübereinstimmung von 92 % spricht für eine sehr gute Konkordanz. Chan et al. (2002) berichten geringere Übereinstimmungen von $r = 0.60$ und $r = 0.66$. Ein Grund dafür, dass die Übereinstimmung in der vorliegenden Arbeit sehr gut ausgefallen ist, könnte in den eigens für die Untersuchung entwickelten Text-Beispielen liegen. Diese wurden für das Training der Raterinnen verwendet und standen diesen während des Kategorisierens der Texte zur Verfügung. Ohne Kodierbeispiele ist die SOLO-Taxonomie möglicherweise zu abstrakt. Die Zuordnung zu den Kategorien kann damit als reliabel und objektiv bewertet werden.

Das Ausmaß der Verteilung der Texte auf die verschiedenen Niveaus kann zudem als Indikator dafür betrachtet werden, dass sich mit der SOLO-Taxonomie qualitative Unterschiede der Texte abbilden lassen. Insgesamt ist zwar ein Schwerpunkt auf dem dritten bis vierten Niveau festzustellen, jedoch zeigen die Ergebnisse der ersten Untersuchung, dass fast bei allen Prompts alle Niveaus vorkommen.

Ein kritischer Punkt bei der Einstufung von Lernergebnissen mit der Taxonomie ist der Umgang mit Texten, die zwar ein hohes Maß an Abstraktion und

inhaltlicher Auseinandersetzung widerspiegeln, inhaltlich jedoch das vorgegebene Thema verfehlen. Hierbei handelt es sich möglicherweise um ein Problem der Aufgabeninterpretation, bei dem die Frage geklärt werden muss, wie die entsprechenden Texte zu bewerten sind. In Folge dieses Problems wurden die Texte in der zweiten Untersuchung in zwei Durchläufen bewertet, nämlich in Bezug auf inhaltliche Passung und auf das Abstraktionsniveau.

Insgesamt kann die Analyse der Texte mit der SOLO-Taxonomie als zufriedenstellend bewertet werden. Es lassen sich qualitative Unterschiede in den Texten abbilden, die zuverlässig erfasst werden. Sowohl in der ersten als auch in der zweiten Untersuchung gelingt mit der SOLO-Taxonomie eine stabile, inhaltsunabhängige Kategorisierung der Lernergebnisse.

6.2 Können Prompts die Verarbeitungstiefe beim Schreiben beeinflussen?

Ausgehend von den Arbeiten zu Lerntagebüchern und Prompts wurde untersucht, inwiefern mit Prompts die Anwendung kognitiver Strategien, insbesondere Elaborationsstrategien, beim Schreiben unterstützt werden kann. Dazu wurden zwei Studien durchgeführt, die Prompts mit unterschiedlicher inhaltlicher Ausgestaltung über zwei verschiedene Zeiträume (1. eine Woche, 2. drei Monate) untersuchen.

In der ersten Studie wurde erprobt, ob sich unterschiedliche Merkmale von Prompts auf die Bearbeitungsqualität beim Schreiben auswirken. Zu diesem Zweck wurden 16 Prompts entwickelt, mit denen Elaborationsstrategien initiiert werden sollten. Diese unterschieden sich zum einen hinsichtlich des inhaltlichen Bezugs (z.B. kontextübergreifende Bezüge, beruflicher Bezug) und zum anderen in der Explizitheit, mit der sie zu einer spezifischen kognitiven Aktivität aufforderten (z.B. offen: „Inwiefern weist der Lernstoff eine Relevanz für Ihr Alltagsleben auf?“ oder spezifisch: „Führen Sie Beispiele aus Ihrer eigenen Schulbiographie an, die das Gelernte bestätigen oder ihm widersprechen“).

Hierbei zeigte sich, dass vor allem solche Prompts hohe Schreibniveaus erzeugen, die spezifische kognitive Prozesse explizit anregen. So schrieben die Studierenden besonders elaborierte Texte, wenn sie explizit dazu aufgefordert wurden, eigene Beispiele zu entwickeln und eine eigene Fragestellung zu einem Lernthema

zu formulieren. Diese Ergebnisse decken sich mit den Befunden von Glogger et al. (2009).

Im Weiteren wurde hinsichtlich der inhaltlichen Bezüge deutlich, dass die Qualität des Schreibens besser ausfiel, wenn die Prompts sich auf Wissensinhalte aus dem aktuellen Seminar bezogen und einen beruflichen Bezug aufwiesen. Sollten hingegen Verknüpfungen zu Wissensinhalten aus anderen Seminaren oder zum Alltag der Lernenden hergestellt werden (*kontextübergreifende Bezüge*), führte dies zu einer größeren Varianz in der Textqualität.

Zusammengefasst zeigte die erste Untersuchung, dass durch das Schreiben mit Prompts Elaborationsstrategien angeregt werden können. Die ersten Befunde deuten darauf hin, dass die Schreibqualität durch Merkmale wie die Spezifität der Formulierung und das erforderliche Wissen (Bezüge) beeinflusst wird. Dies könnte Gegenstand einer weiterführenden Untersuchung sein, wurde jedoch in der vorliegenden Arbeit nicht weiter verfolgt. Die Befunde aus der ersten Studie wurden als Ausgangspunkt für eine weitere Untersuchung verwendet, die motivationale Prozesse und Feedback als Vermittlungsprozesse in den Blick nahm.

6.3 Feedback und Motivation

In der zweiten Studie wurde untersucht, inwieweit sich Feedback zur Optimierung der Lernwirksamkeit des Schreibens mit Prompts eignet, wobei zwei Aspekte in den Blick genommen wurden. So wurde zum einen vermutet, dass der Einsatz von Feedback sich positiv auf die Motivation der Studierenden auswirkt und das Ausbleiben desselben den gegenteiligen Effekt haben würde. Zum anderen wurde angenommen, dass Feedback auch einen positiven Effekt auf die Qualität der Texte haben würde. Die Probanden erhielten daher nach jedem eingereichten Prompt eine von drei Feedbackvarianten (Selbstfeedback, Feedback vom Dozenten und kein Feedback). Dies wurde über den Zeitraum der Untersuchung so variiert, dass jeder Seminarteilnehmer alle drei Varianten bekam. Im Zusammenhang mit Feedback und Motivation interessierten zwei Forschungsfragen, deren Bearbeitung in diesem Abschnitt zusammenfassend dargestellt wird:

1. Welche Unterschiede bestehen bei den Gruppen A, B und C in Bezug auf die Motivation nach den unterschiedlichen Feedbackformen (Treatments)?

2. Welche Feedbackvariante ist hinsichtlich der Motivation der Studierenden die günstigste?

Da die Zahl der Probanden starken Schwankungen unterlag (bis zum Ende der Untersuchung brachen circa 75% der Studierenden das Seminar ab), können zum einen nur vorsichtige Schlüsse aus den Ergebnissen gezogen werden. Zum anderen wurde Gruppe C aufgrund der geringen Gruppengröße ($n=2$) aus den weiteren Analysen ausgeschlossen. Für die Analyse der Daten wurden infolgedessen non-parametrische Verfahren verwendet. Zusätzlich wurden die Motivationsverläufe auf deskriptiver Ebene betrachtet. Vor dem theoretischen Hintergrund des FAM lassen sich aus den Ergebnissen instruktionale Implikationen ableiten.

Auf deskriptiver Ebene zeigte sich zunächst, dass zum ersten Messzeitpunkt insgesamt gute motivationale Voraussetzungen von Seiten der Studierenden vorlagen. Dies manifestierte sich auf den vier Subskalen des FAM unter anderem in einem hohen Mittelwert des Faktors *Herausforderung*, der darauf hindeutet, dass alle Studierenden die Aufgabe als anspruchsvoll einschätzten. Hinzu kommen hohe Ausprägungen der Komponenten *Interesse* und *Erfolgswahrscheinlichkeit* und eine moderate *Misserfolgserwartung*. Insgesamt zeigen diese Werte, dass die Lernsituation als motivational anregend wahrgenommen wurde.

Bei Gruppe A ließ sich sodann beobachten, dass durch das Selbstfeedback die Anstrengungsbereitschaft beim Bearbeiten der Prompts und das *Interesse* für diese Aufgabe konstant blieben. Signifikante Veränderungen ergaben sich für diese Gruppe bei der *Misserfolgsbefürchtung*. Beim Betrachten der Mittelwerte findet man allerdings keine konstante Zu- oder Abnahme der *Misserfolgsbefürchtung*, sondern lediglich eine Schwankung. Ebenfalls signifikant ist die kontinuierliche Abnahme der *Herausforderung*. Die Studierenden schätzen die Prompts folglich im Laufe der Studie als immer weniger schwierig ein und ihre Anstrengungsbereitschaft sinkt.

Bei Gruppe B zeigte sich nach Ausbleiben des Dozentenfeedbacks eine signifikante Abnahme des *Interesses* und der Anstrengungsbereitschaft (*Herausforderung*). Auf deskriptiver Ebene nehmen die beiden Faktoren *Misserfolgsbefürchtung* und *Erfolgswahrscheinlichkeit* ebenfalls ab. Hier stellt sich die Frage, ob es im Falle von Dozentenfeedback günstiger ist, es konsequent durchzuführen und nicht mit anderen Feedbackformen zu mischen oder ob es günstiger ist, es ganz auszusetzen.

Basierend auf den Befunden von Rheinberg et al. (2001) sind in einer Lernumgebung, die von den Lernern selbstreguliertes Lernen verlangt, besonders die motivationalen Faktoren *Interesse* und *Herausforderung* von Bedeutung. Das Aufgabeninteresse ist insofern wichtig, als dass man davon ausgehen kann, dass ein an der Aufgabe interessierter Lerner sich stärker bei der Bearbeitung der Aufgabe engagiert als ein nicht interessierter Lerner. Nimmt er die Aufgabe zusätzlich als Herausforderung wahr, ist dies ein guter Prädiktor dafür, dass auch moderate Hindernisse (z.B. fehlendes Wissen, um eine Aufgabe weiterzubearbeiten) bewältigt werden.

Vor diesem theoretischen Hintergrund könnte das signifikante Absinken des Motivationsfaktors *Herausforderung* in den Gruppen A und B und die ebenfalls signifikante Abnahme des *Interesses* in Gruppe B als motivational ungünstiger Befund bewertet werden. Es wäre aber auch denkbar, dass die Seminarteilnehmer, unabhängig von der Reihenfolge des Feedbacks, die anfänglich als schwierig eingeschätzte Aufgabe nach einer gewissen Bearbeitungszeit realistischer und einfacher einschätzten. Das würde erklären, dass nach Einschätzung der Probanden keine besondere Anstrengung (*Herausforderung*) mehr zur Bewältigung der Aufgabe erforderlich war. Ein Hinweis hierfür könnte sein, dass die Mittelwerte nicht bis auf den Minimalwert abfallen. Darüber hinaus könnte auch der Neuheitswert der Aufgabe zum Absinken des *Interesses* und der Anstrengungsbereitschaft beigetragen haben, da sich dieser nach drei Monaten abgenutzt hat.

Die Frage, welche Feedbackvariante im Hinblick auf die Motivation die günstigste ist, kann anhand der vorliegenden Ergebnisse nur unter Vorbehalt beantwortet werden, da die Gruppengrößen verhältnismäßig klein sind. Angesichts des starken Abbruchs der Teilnehmerzahlen ist zudem ein Selbst-Selektionseffekt nicht auszuschließen. Dennoch deuten die Daten auf deskriptiver Ebene daraufhin, dass sich das Dozentenfeedback zwar nicht positiver auf die Motivation auswirkt als das Selbstfeedback. Es lässt sich jedoch eine signifikante Abnahme des *Interesses* nach Ausbleiben des Dozentenfeedbacks beobachten (Gruppe B), wohingegen bei Gruppe A, die zuerst Selbstfeedback erhielt, das *Interesse* an der Aufgabe stabil bleibt. Dies könnte als Hinweis dafür gedeutet werden, dass Selbstfeedback ausreicht, um das Aufgabeninteresse über einen Zeitraum von drei Monaten stabil zu halten. Weiteren Aufschluss darüber könnte eine Prompt-Maßnahme derselben Dauer geben, in deren Rahmen ausschließlich mit Selbstfeedback gearbeitet wird.

Zusammenfassend lässt sich somit für die Forschungsfrage 5 (Welche Feedbackvariante ist hinsichtlich der Motivation der Studierenden die günstigste?) vorsichtig konstatieren, dass der Abbruch von externem Feedback sich nicht günstig auf das Aufgabeninteresse auswirkt. Ansonsten bleiben jedoch bei allen Gruppen die für eine selbstregulierte Lernumgebung wichtigen Faktoren (vgl. Rheinberg et al., 2001) Aufgabeninteresse und Herausforderung relativ konstant. Wenn in Betracht gezogen wird, dass es in der Untersuchung keine richtige Nicht-Feedbackkondition gab und alle Teilnehmer Feedback erhalten haben, könnte dies als möglicher Beleg für die motivationsförderliche Bedeutung von Feedback gedeutet werden.

Generalisierbare Aussagen zum Zusammenhang zwischen Feedback und Motivation können auf der Basis dieser Ergebnisse jedoch nicht gemacht werden.

6.4 Feedback und Verarbeitungstiefe

Es wurde weiter vermutet, dass Feedback sich auf die Schreibleistung, die in Form der Verarbeitungstiefe operationalisiert wurde, auswirkt. Um diese Fragestellung zu untersuchen, wurde die Entwicklung der Schreibtiefe innerhalb jeder Gruppe betrachtet. Im Mittelpunkt standen dabei die Fragestellungen:

1. Führen die unterschiedlichen Feedbackvarianten zu Unterschieden in der Verarbeitungstiefe?
2. Welche Feedbackvariante ist hinsichtlich der Verarbeitungstiefe der Studierenden die günstigste?

Auf deskriptiver Ebene zeigte sich, dass sich die Textqualität bei allen Gruppen mäßig und mit Schwankungen verbessert. Gruppe A und B verfassen schon zu Beginn der Untersuchung gute Texte und werden am Ende der Studie fast eine Niveaustufe höher eingestuft als zu Beginn. Bei Gruppe C ist ein kontinuierlicher Aufwärtstrend nach dem Dozentenfeedback (Prompts 4-6) zu beobachten und die Gruppe verbessert sich um fast zwei Niveaustufen.

Zur Untersuchung spezifischer Effekte des Feedbacks wurden jeweils um einen Prompt versetzt drei mit Prompts produzierte Texte nach einer Feedbackform analysiert. In Abbildung 13 ist das Vorgehen dabei exemplarisch dargestellt. Es wurde auf diese Weise mit den Daten verfahren, weil die Seminarteilnehmer nicht wussten, wie oft und zu welchem Zeitpunkt sie welche Feedbackform erhalten

würden. Man kann daher davon ausgehen, dass nach einmaligem Feedback noch keine Effekte sichtbar werden. Die Erfahrung derselben Feedbackform über drei Prompts sollte jedoch genügen, um in irgendeiner Form zu wirken. Es wurde außerdem angenommen, dass die Wirkung noch über einen weiteren Prompt fortwirkt, da die Seminarteilnehmer sich erst auf eine neue Feedbackform einstellen müssen.

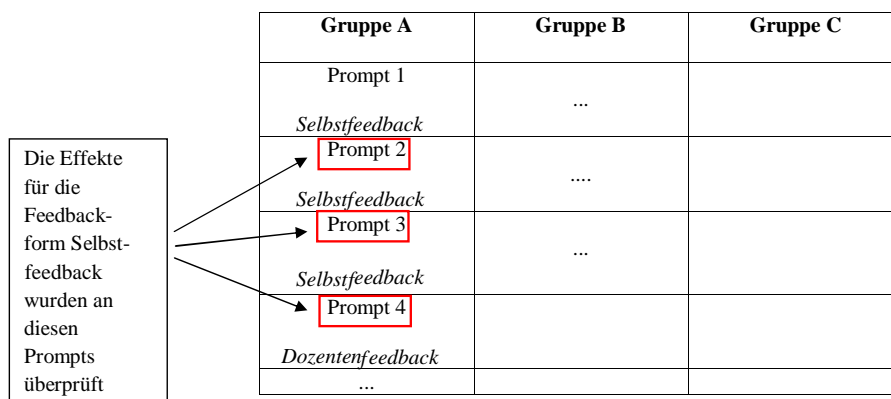


Abbildung 13: Vorgehensweise bei der Untersuchung der Feedbackeffekte

Innerhalb der Gruppen wurden nun die Ausprägungsgrade der Verarbeitungstiefe zum ersten Erhebungszeitpunkt und zum zweiten Erhebungszeitpunkt auf signifikante Veränderungen geprüft. Der Wilcoxon-Test zeigt bei einer zweiseitigen Prüfung eine tendenziell signifikante Veränderung der Verarbeitungstiefe in den Gruppen A und C. Sowohl Gruppe A als auch Gruppe C produzieren in der zweiten Hälfte der Untersuchung in den Prompts 5,6 und 7 höhere Verarbeitungstiefen als in der ersten Hälfte in den Prompts 2,3 und 4. Es ist jedoch schwierig, aus diesen Befunden Rückschlüsse auf die Wirksamkeit des Feedbacks zu ziehen. So kann es sein, dass aufgrund des berichteten Ausfalls von Seminarteilnehmern nur die Teilnehmer, denen die Bearbeitung der Prompts gut gelang, bis zum Schluss durchhielten. Des Weiteren sind auf der Basis der vorliegenden Ergebnisse keine eindeutigen und generalisierbaren Aussagen über die Vorzüge einer Feedbackform gegenüber einer anderen möglich.

6.5 Implikationen für die instruktionale Praxis

Aus den Ergebnissen der Untersuchungen lassen sich für die instruktionale Praxis relevante Schlussfolgerungen ableiten. In der vorliegenden Arbeit konnte gezeigt werden, wie mit Schreibinstruktionen in Form von Prompts komplexe kognitive Prozesse gleich zu Beginn eines Lernzyklus erfolgreich angeregt werden können. Der Schwerpunkt lag auf der Förderung von Elaborationsstrategien, also Strategien mit dem Ziel, neue Informationen mit dem Vorwissen zu verknüpfen (Wild, 2000). Mit Hilfe der Prompts konnten die Lernenden dazu angeregt werden, beim Schreiben eigene Beispiele für Lerninhalte zu generieren und Theorien selbstständig auf andere Kontexte anzuwenden. Zahlreiche Trainingsprogramme befassen sich damit, wie Lernprozesse durch die Förderung von Lernstrategien verbessert werden können. Dabei wird versucht, Strategien kontextungebunden zu vermitteln. Es zeigt sich jedoch in der Regel, dass solche Programme nicht den vermehrten Einsatz der erlernten Strategien nach sich ziehen, weil Lerner die Lernstrategien nicht spontan einsetzen. Ähnliche Befunde finden sich in Lerntagebuch-Studien (Borsch & Imhof, 2007; Nückles et al., 2004) und Studien aus der Textproduktionsforschung, in denen Lerner ohne explizite Benennung der gewünschten Bearbeitungsstrategien Texte verfassen (Bangert-Drowns et al., 2004). Hier zeigten die Analysen der Texte, dass Lerner vornehmlich oberflächliche Strategien zur Bearbeitung der Lerninhalte einsetzten, deren Nützlichkeit im Hinblick auf den Lernerfolg fraglich ist.

Prompts stellen eine Möglichkeit dar, diese Lücke zwischen Kenntnis und Nutzung zu schließen, da Lerner in einem konkreten Kontext dazu aufgefordert werden, Lerninhalte mithilfe spezifischer Strategien zu bearbeiten. Wenn die Zielgruppe, wie hier, aus erwachsenen Lernern besteht und man davon ausgehen kann, dass diese im Prinzip über diese Lernstrategien verfügen, dann erscheint die didaktische Maßnahme des Prompting erfolgversprechend.

Zentral bei der Ausgestaltung der Prompts sind vor dem Hintergrund der Ergebnisse zweierlei Komponenten. So ist es für das Evozieren anspruchsvoller Prozesse entscheidend, dass diese explizit in der Schreibinstruktion angesprochen werden. Diese Idee weist große Überschneidungen mit dem Konzept des *constructive alignment* von Biggs (1996), Biggs und Tang (2007) auf. Das Konzept beschreibt die Ausrichtung des Lernprozesses am angestrebten Ergebnis. Im Kern geht es

beim *constructive alignment* darum, dass die Lehrperson Lernziele klar formuliert und die Lernaufgaben auf die Ziele abstimmt. Dies wird dadurch erreicht, dass die Lerner in den Lernaufgaben mittels entsprechender Verben dazu aufgefordert werden, die erwünschten Prozesse durchzuführen (z.B. vergleichen, kontrastieren etc.).

Die zweite Komponente besteht in der Art der Lerninhalte, die angesprochen werden. Auf der Basis der dargestellten Befunde ist es empfehlenswert, Inhalte aus der aktuellen Lehrveranstaltung zu verwenden. Zudem sind Verknüpfungen, wie sie als Bestandteil von Elaborationsstrategien aufgefasst werden, innerhalb des Lernstoffs vielversprechender als Verknüpfungen zu anderen Lehrveranstaltungen. Letztere erscheinen zwar wünschenswert, dennoch kann aufgrund des heterogenen Vorwissens der Studierenden nur schwer abgeschätzt werden, inwiefern externe Verknüpfungen sinnvoll und möglich sind.

Hinsichtlich der Motivation legen die Ergebnisse nahe, dass Dozentenfeedback nicht ausreicht, um das Interesse der Lerner an den Prompts über einen so langen Zeitraum aufrecht zu erhalten. Eine Möglichkeit, um die Arbeit mit Prompts motivationsförderlicher zu gestalten, könnte darin bestehen, den Lernenden Leitfragen für das Selbstfeedback bereitzustellen, denn in der Untersuchungsbedingung Selbstfeedback blieb das *Interesse* der Lerner an den Prompts stabil. Alternativ wäre auch denkbar, dass die externe Regulation des Lernprozesses mit zunehmender Kompetenz der Lerner, die geprompteten Prozesse eigenständig durchzuführen, sukzessive zurückgenommen wird (Nückles, Hübner & Renkl, 2008), um einen Expertise-Umkehr-Effekt zu vermeiden.

Insgesamt lässt sich vor dem Hintergrund der Ergebnisse festhalten, dass die in der Arbeit erprobten Prinzipien hinsichtlich der Gestaltung der Prompts und des Feedbacks inhaltsunabhängig auch auf andere Seminare übertragen werden können. Die SOLO-Taxonomie (Biggs & Collis, 1980) lässt sich dabei gut nutzen, um 1. eine Passung zwischen Lernzielen, 2. Aufgaben, die die Lerner bei den Prozessen unterstützen, um die Ziele zu erreichen und 3. aufgabenspezifischem Feedback vorzunehmen. Die Analysen der Lernjournale zeigen, dass es sich für Lehrende lohnt, anspruchsvolle Strategien anzuregen, da diese zum Aufbau komplexer Wissensstrukturen bei den Lernern beitragen. Mithilfe der SOLO-Taxonomie können individuelle Rückmeldungen gestaltet werden, die den Lerner über die Qualität der Prozesse informieren und Empfehlungen zur Verbesserung zukünftiger Lern-

prozesse geben (z.B., mehr zu vernetzen).

6.6 Grenzen der Untersuchung und Perspektiven für zukünftige Forschung

In der vorliegenden Arbeit wurde eine auf theoretischen und empirischen Erkenntnissen basierende Lernumgebung mit Prompts konzipiert und erprobt. Das Ziel der Arbeit war es, Erkenntnisse zu Zusammenhängen zwischen der Qualität der Schreibprozesse und den daraus resultierenden Wissensstrukturen zu gewinnen.

In der ersten Untersuchung wurden Prompts mit unterschiedlichen Bezügen konzipiert und ihre Auswirkungen auf die Wissensstruktur untersucht. Dabei wurde zum einen der Frage nachgegangen, inwiefern Zusammenhänge zwischen den inhaltlichen Bezügen und der Verarbeitungstiefe in den Texten bestehen. Zum anderen interessierte, inwiefern die qualitativen Unterschiede in den Wissensstrukturen mit einer Taxonomie zur Beschreibung der Komplexität von Lernergebnissen (SOLO-Taxonomie) abgebildet werden können. Es bestätigte sich die Annahme, dass insbesondere spezifisch formulierte Prompts mit Bezug zur aktuellen Lernveranstaltung gute Verarbeitungstiefen erzeugen. Mit der SOLO-Taxonomie gelang es zudem, unterschiedliche Verstehensniveaus themenunabhängig abzubilden.

Auf den Befunden der ersten Studie wurde in der zweiten Untersuchung zusätzlich zu den Prompts Feedback eingesetzt. Hier stand im Fokus, inwiefern Feedback und motivationale Voraussetzungen sich auf die Schreibqualität auswirken. Basierend auf den Befunden der Feedbackforschung wurde vermutet, dass Feedback kognitive und motivationale Wirkungen hat, die wiederum die Qualität der Schreibprozesse beeinflussen. Die Ergebnisse zeigten unterschiedliche Effekte für die zwei Feedbackformen auf die Motivation. So nahm das *Interesse* der Lernergruppe, die zuerst Dozentenfeedback erhalten hatte, nach dem Ausbleiben des Dozentenfeedbacks signifikant ab, hielt sich jedoch bei der Gruppe, die zuerst mit Selbstfeedback arbeitete, stabil. Es deutet sich also an, dass Feedback den Lernprozess auf der motivationalen Ebene beeinflusst und es ist anzunehmen, dass dadurch auch die kognitive Leistung der Lerner beeinflusst wird. Diese Zusammenhänge lassen sich allerdings nicht in dieser Deutlichkeit aus den Ergebnissen ableiten, da die Stichprobe sehr instabil und zu klein ausfiel.

Um der Frage nachzugehen, wie durch Feedback die Motivation der Teilnehmer aufrecht erhalten werden kann und welche kognitiven Wirkungen Feedback nach sich zieht, bietet die Arbeit Ansatzpunkte für weitere Forschung. So wäre es für zukünftige Untersuchungen wichtig, die Effekte des Feedbacks und der Motivation auf die Verarbeitungstiefe getrennt zu untersuchen, um so zu differenzierteren Aussagen über die Rolle von Lernervoraussetzungen und Feedback beim Schreiben und den Wechselwirkungen zwischen beiden zu kommen.

Methodisch ist es wahrscheinlich sinnvoll, eine Feedbackform über einen längeren Zeitraum zu erproben. Um einen Reihenfolge-Effekt zu vermeiden, wurde in der zweiten Untersuchung jeweils über die Dauer von drei Prompts eine Rückmelde-Form beibehalten, damit jeder Teilnehmer ein Mal jede Feedbackform erhält. Auf diese Weise sollten differenzielle Wirkungen des Feedbacks erfasst werden. Diese Vorgehensweise eignet sich bei größeren Stichproben, erwies sich jedoch bei der geringen Teilnehmerzahl als Problem, da, unabhängig von der Reihenfolge des Feedbacks, immer eine Nachwirkung zu erwarten ist. Dies führt in einer kleinen Stichprobe vermutlich zu einem verzerrten Ergebnis, könnte jedoch mit einer größeren Stichprobe gelöst werden. Es ist außerdem prinzipiell denkbar, die Dauer jeder Feedbackform zu verlängern, da drei Feedbackinterventionen möglicherweise ein zu kurzer Zeitraum sind, um Rückschlüsse auf die Wirkung einer einzelnen Form zuzulassen. Die Einführung einer Kontrollgruppe, die gar kein Feedback erhält, könnte zusätzlichen Aufschluss über die tatsächliche Stärke der Effekte geben.

Um zu validen Ergebnissen bezüglich der Wirkungen von Feedback zu kommen, ist außerdem denkbar, die Wirkung einzelner Rückmeldungen auf individueller Ebene zu kontrollieren, da die Feedbackforschung auf komplexe Wirkungen von Feedback hindeutet (Winne & Butler, 1994). Die Zusammenhänge zwischen Feedbackrezeption und den kognitiven bzw. motivationalen Effekten wurden in der vorliegenden Arbeit nicht untersucht.

Um die Wirkung der Prompting-Maßnahme auf den Wissenszuwachs zu überprüfen, wurde ein Wissenstest konzipiert, der vor und nach der Intervention durchgeführt wurde. Es stellte sich jedoch heraus, dass die Testschwierigkeit zu hoch angesetzt war, sodass die Abschluss-Klausur herangezogen wurde, um Informationen über die Leistungseffekte der Prompting-Maßnahme zu gewinnen. Diese bilden jedoch nicht den Wissenszuwachs der Lerner ab, sondern zeigen Zu-

sammenhänge zwischen der Verarbeitungstiefe und dem Lernerfolg. Zur Klärung der Frage nach den Effekten der Prompts wäre jedoch auch von Interesse, inwiefern diese tatsächlich den Wissenszuwachs begünstigen. Dazu sollte im Vorfeld ein Test entwickelt werden, der die Lerner nicht überfordert und sich für einen solchen Vergleich eignet.

Ein letzter Punkt betrifft die Frage, inwiefern sich Prompts zur Förderung leistungsheterogener Lerner eignen. Insbesondere bei der Überprüfung der Lerneffekte der Prompting-Maßnahme anhand der Klausur und der Verarbeitungstiefen wurde deutlich, dass vor allem Lerner, die gute Texte produzierten, von den Prompts profitierten, wohingegen Lerner mit geringeren Verarbeitungstiefen nicht an der Klausur teilnahmen und noch gegen Ende des Seminars abbrachen. Dies wirft die Frage auf, ob durch Prompts nur solche Lerner unterstützt werden, die von vornherein ein gutes Leistungsniveau aufweisen. Die Bearbeitung dieser Frage ist ein wichtiger Ansatzpunkt für zukünftige Forschung, der in dieser Form bislang nicht untersucht wurde.

Am Beginn der Arbeit stand die Frage, inwiefern Schreiben sich als Lerninstrument eignet und welche Merkmale Schreibaufträge aufweisen müssen, um den Aufbau komplexer Wissensstrukturen bei erwachsenen Lernern zu fördern.

Es konnte mit den Ergebnissen gezeigt werden, dass Motivation und Lernervoraussetzungen eine wichtige Rolle beim Produzieren von Texten zukommt, da diese den Lernprozess positiv beeinflussen. So schnitten z.B. die Lerner, deren Texte anspruchsvolle Verarbeitungstiefen widerspiegeln, in der Klausur ebenfalls gut ab. Des Weiteren konnte gezeigt werden, dass die epistemische Funktion des Schreibens, nämlich die Weiterverarbeitung des eigenen Wissens, durch Schreibaufträge gelingen kann. Mit den Prompts produzierten die Seminarteilnehmer zu einem relativ frühen Zeitpunkt im Lernprozess anspruchsvolle Texte. Eine Frage für zukünftige Forschung wäre, inwiefern Texte, die komplexe Wissensstrukturen aufweisen, die Schreibstufe repräsentieren, die Bereiter (1980) als die höchste Form der Schreibkompetenz einstuft. Diese ist dann erreicht, wenn sich Schreiben und reflexives Denken verbinden und der Textproduzent erkennt, dass Schreiben nicht nur der Ausdruck des Gedachten ist, sondern ein „Instrument des Präzisierens, Erweiterns, ja des Entwickelns von Wissen“ (Eigler, Jechle, Merziger & Winter, 1987, S. 383), ein Denkwerkzeug an sich. Dies kann in der vorliegenden Arbeit nicht endgültig beantwortet werden, weil z.B. nicht immer

klar ist, inwiefern z.B. bei Prompts, die zum Generieren eigener Beispiele aufforderten, die Texte eigenständig oder mit Rückgriff auf externe Quellen verfasst wurden.

Ein weiterer Ansatzpunkt weiterführender Studien ist die strengere Kontrolle der Lernvoraussetzungen beim Schreiben. Die vorliegenden Befunde deuten auf variierende Motivationsverläufe über die Dauer der Intervention hin, es können jedoch auf der Basis der kleinen Stichprobe keine Aussagen über die differenziellen Effekte der Motivation gemacht werden. Es besteht somit Bedarf an Untersuchungen, die die Motivation ebenfalls zeitnah erheben und Zusammenhänge zwischen der Qualität der Schreibprozesse und der Motivation in den Blick nehmen. Mithilfe solcher Befunde könnte man auch der Frage nachgehen, wie eine Lernumgebung mit Prompts so gestaltet werden kann, dass auch schwächere Textproduzenten davon profitieren.

Insgesamt unterstützt die vorliegende Arbeit die Annahme, dass durch Schreiben anspruchsvolle Prozesse gefördert werden können, die zum Aufbau komplexer Wissensstrukturen beitragen. Es wird ein methodischer Zugang aufgezeigt, wie die Komplexität von Wissensstrukturen abgebildet und spezifische kognitive Prozesse durch Prompts angeregt werden können.

Anhang

.1 Prompts der zweiten Studie

1. Überlegen Sie Beispiele, in denen Sie geprüft worden sind (= Testsituationen) und suchen Sie Situationen, in denen die Durchführungsobjektivität, Auswertungsobjektivität oder die Interpretationsobjektivität verletzt worden ist.
2. Stellen Sie sich folgende Situation vor: Sie entwickeln einen Sprachtest für Vorschulkinder. Wenn dieser Test schlecht ausfällt, werden die Kinder einer Fördermaßnahme zugewiesen. Wie stellen Sie die Reliabilität des Tests fest und welches Problem entsteht, wenn die Reliabilität nicht gegeben ist?
3. Aus Ihrer Schulbiographie kennen Sie vielleicht Situationen, in denen Sie das Gefühl hatten, „nicht das Richtige gefragt worden zu sein“. Testtheoretisch handelt es sich hierbei um die Verletzung der Validität. Beschreiben oder konstruieren Sie ein solches Beispiel. Behalten Sie dabei die verschiedenen Formen der Validität im Hinterkopf.
4. Nehmen Sie das Konstrukt der Kreativität und überlegen Sie: Gegen welches andere relevante Konstrukt würden Sie es abgrenzen wollen?
5. Ein Kind hat im Vergleich zu seinen Klassenkameraden schlechte Leistungen im Fach Deutsch. Welche Konstrukte würden Sie untersuchen und an welchen Merkmalen würden Sie diese festmachen wollen?
6. Wählen Sie ein für Ihren angestrebten Beruf relevantes Thema und formulieren Sie eine Forschungsfrage, eine operationale und eine statistische Hypothese zu dieser Forschungsfrage. Bitte schreiben Sie etwa 250 Wörter.
7. In einer Studie wurden Kinder mit Aufmerksamkeitsstörungen untersucht. Man hat in einer Kontrollgruppe und in einer Experimentalgruppe getestet, ob die visuelle Eingrenzung der Lernumgebung durch Pappwände dazu führt, dass die Kinder besser arbeiten als in einer ablenkungsreichen Umgebung. Gemessen wird die Zeit in Sekunden, die die Kinder an der Aufgabe (Abschreibeaufgabe) bleiben und die Zahl der Schreibfehler, die die Kinder machen.

Die Unterschiedshypothesen lauten:

Kinder in einer abgeschirmten Umgebung bleiben länger an den Aufgaben.

Kinder in einer abgeschirmten Umgebung machen weniger Fehler beim Abschreiben.

Die erste Hypothese wird bei einem alpha-Fehler-Niveau von $p < .05$ signifikant (β -Fehler = .90). Die zweite Hypothese wird bei einem alpha-Fehler-Niveau von $p < .10$ nicht signifikant. (β -Fehler = .40)

Reflektieren Sie an diesem Beispiel, welche Konsequenzen und Auswirkungen solche Fehler (alpha und beta) haben können. Bitte schreiben Sie etwa 250 Wörter.

.2 Selbstfeedbackfragen

Selbstfeedback 1

Liebe Studierende, vielen Dank für die Einreichung Ihres ersten Prompts. Um Ihr Lernen zu optimieren, analysieren Sie bitte anhand der folgenden drei Fragen Ihr erstes Lernprotokoll.

1. Ich habe bei meinen Beispielen kritisch geprüft, ob sie begründet und belegt sind.
 2. Die Struktur meines Lernprotokolls gibt die Struktur des Stoffes meiner Meinung nach wieder.
 3. Ich habe versucht, mehrere Aspekte des Themas miteinander zu verbinden.
- Notieren Sie dann auf der Basis dieser Einschätzung ein Ziel für die Bearbeitung des nächsten Prompts.

Selbstfeedback 2

Liebe Studierende, vielen Dank für die Einreichung Ihres zweiten Prompts. Um Ihr Lernen zu optimieren, analysieren Sie bitte anhand der folgenden Fragen Ihr zweites Lernprotokoll.

1. Ich habe das Konzept der Reliabilität auf ein konkretes Problem angewendet.
 2. Ich habe versucht, den Seminarinhalt zum Thema Reliabilität nicht nur wiederzugeben, sondern auch anzuwenden.
- Notieren Sie dann auf der Basis dieser Einschätzung ein Ziel für die Bearbeitung des nächsten Prompts.

Selbstfeedback 3

Liebe Studierende, vielen Dank für die Einreichung Ihres dritten Prompts. Bitte analysieren Sie anhand der folgenden Fragen Ihr drittes Lernprotokoll.

1. Ich habe zur Erläuterung des Konzeptes der Variabilität ein gut verständliches Beispiel formuliert, mit dem ich das Konzept auch einer Person mit wenig Vorwissen zu dem Themenbereich erklären könnte.
2. Das Beispiel illustriert die Bedeutung und das Gewicht der Variabilität für die Testgüte.

Notieren Sie dann auf der Basis dieser Einschätzung ein Ziel für die Bearbeitung des nächsten Prompts.

Selbstfeedback 4

Liebe Studierende, vielen Dank für die Einreichung Ihres vierten Prompts. Um Ihr Lernen zu optimieren, analysieren Sie bitte anhand der folgenden zwei Fragen Ihr letztes Lernprotokoll.

1. Ich habe versucht, den Seminarinhalt zum Thema Konstrukte nicht nur wiederzugeben, sondern auch auf ein konkretes Problem anzuwenden.
2. Ich habe bei meinem Beispiel kritisch geprüft, ob es begründet und gut belegt ist.

Notieren Sie dann auf der Basis dieser Einschätzung ein Ziel für die Bearbeitung des nächsten Prompts.

Selbstfeedback 5

Liebe Studierende, vielen Dank für die Einreichung Ihres fünften Prompts. Um Ihr Lernen zu optimieren, analysieren Sie bitte anhand der folgenden zwei Fragen Ihr letztes Lernprotokoll.

1. Ich habe versucht, das Problem in einen breiten Kontext einzubetten und mehrere Aspekte des vorgegebenen Problems zu beleuchten.
2. Ich habe dabei versucht, die Tragweite der einzelnen Aspekte herauszuarbeiten.

Notieren Sie dann auf der Basis dieser Einschätzung ein Ziel für die Bearbeitung des nächsten Prompts.

Selbstfeedback 6

Liebe Seminar Teilnehmerinnen und Seminar Teilnehmer, vielen Dank für die Einreichung Ihres sechsten Prompts. Um Ihr Lernen zu optimieren, analysieren Sie bitte anhand der folgenden drei Fragen Ihr letztes Lernprotokoll.

1. Ich habe versucht, die Hypothesen in einen beruflichen Kontext einzubetten.
2. Ich habe versucht, die Kontraste zwischen den verschiedenen Arten von Hypothesen herauszuarbeiten.

3. Ich habe bei meiner Darstellung darauf geachtet, die Hypothesen voneinander abzugrenzen.

Notieren Sie dann auf der Basis dieser Einschätzung ein Ziel für die Bearbeitung des nächsten Prompts.

Selbstfeedback 7

Liebe Seminar Teilnehmerinnen und Seminar Teilnehmer, vielen Dank für die Ein-

reichung Ihres siebten Prompts. Um Ihr Lernen zu optimieren, analysieren Sie bitte anhand der folgenden drei Fragen Ihr letztes Lernprotokoll.

1. Ich habe bei der Bearbeitung des Prompts versucht, einen Bezug zur Praxis herzustellen und darzustellen, welche Konsequenzen eine Fehlannahme für die Förderung der Kinder haben kann.
2. Ich habe versucht, die Kontraste zwischen den verschiedenen Arten von Fehlern herauszuarbeiten.

Notieren Sie dann auf der Basis dieser Einschätzung ein Ziel für die Bearbeitung des nächsten Prompts.

Selbstfeedback 8

Liebe Seminarteilnehmerinnen und Seminarteilnehmer, vielen Dank für die Einreichung Ihres achten Prompts. Um Ihr Lernen zu optimieren, analysieren Sie bitte anhand der folgenden zwei Fragen Ihr letztes Lernprotokoll.

1. Ich habe zur Erläuterung der Varianzanalyse ein gut verständliches Beispiel formuliert, mit dem ich das Konzept auch einer Person mit wenig Vorwissen zu dem Themenbereich erklären könnte.
2. Ich habe versucht, die Besonderheit der Methode der Varianzanalyse in meinem Beispiel herauszuarbeiten.

Notieren Sie dann auf der Basis dieser Einschätzung ein Ziel für die Bearbeitung des nächsten Prompts.

.3 Fragen aus eHElp

1. Ich habe im Lernprotokoll Beziehungen zu den Inhalten anderer Fächer hergestellt.
2. Beim Schreiben des Lernprotokolls habe ich meine Lerntechnik dem Schwierigkeitsgrad des Lernstoffes angepasst (z.B. eigene Beispiele für ein theoretisch komplexes Thema generiert).
3. Wenn ich etwas nicht gleich verstanden habe, habe ich diese Lücken festgehalten und versucht, sie zu beheben (z.B. Seminarleiter/ Studienkollegen gefragt, Lehrbuchtext gelesen,...).
4. Für neue Konzepte habe ich praktische Anwendungen gesucht.
5. In meinem Lernprotokoll habe ich kritisch geprüft, ob die einzelnen Themen ausreichend begründet und belegt sind.
6. Fehlende Informationen habe ich mir aus verschiedenen Quellen zusammengesucht (z.B. Mitschriften, Bücher, Fachzeitschriften).
7. Neue Begriffe oder Theorien habe ich auf mir bereits bekannte Begriffe bzw. Theorien bezogen.
8. Bevor ich mit dem Lernprotokoll begonnen habe, habe ich mir überlegt, wie ich am effektivsten vorgehen kann.
9. Ich habe den Stoff so geordnet, dass ich ihn mir gut einprägen kann.
10. Im Lernprotokoll habe ich versucht, aus dem bearbeiteten Stoff eigene Ideen zu entwickeln.
11. Ich habe mir aus Mitschrift, Skript oder Literatur kurze Zusammenfassungen mit den Hauptideen zusammengestellt.
12. Um sicherzugehen, dass ich alles verstanden habe, habe ich mir Fragen zum Stoff gestellt.
13. In Texten und Mitschriften habe ich wichtige Stellen markiert.

14. Zu bestimmten Lerninhalten habe ich in meinem Lernprotokoll konkrete Beispiele festgehalten.
15. Die wichtigsten Unterlagen für das Lernprotokoll hatte ich an meinem Arbeitsplatz griffbereit.
16. Die Struktur meines Lernprotokolls gibt die Struktur des Stoffes meiner Meinung nach am besten wieder.
17. Neu Gelerntes habe ich auf meine eigenen Erfahrungen bezogen.
18. Ich habe die Vor- und Nachteile verschiedener theoretischer Konzeptionen verglichen.
19. Wichtige neue Fachausdrücke und Definitionen habe ich markiert.
20. Das Lernprotokoll zeigt die Relevanz des Lernstoffes für mein Alltagsleben auf.

.4 Kodierleitfaden

Tabelle 15: Taxonomie-Niveaus und Indikatoren

SOLO Levels	Kodierungsanweisung	Indikatoren
Prestructural (1)	Die Antwort enthält keine relevanten Komponenten oder solche ohne logische Beziehung zur Lernaufgabe.	- Paraphrasieren der Lernfrage
Unistructural (2)	Die Antwort enthält einen Aspekt eines Themenkomplexes. Weitere mögliche, relevante Aspekte werden nicht berücksichtigt.	- Aufführen eines einzelnen Aspektes - Identifizieren einer Komponente eines Themenkomplexes
Multistructural (3)	Die Antwort bezieht sich auf mehrere relevante Einzelkomponenten eines Themas. Diese werden allerdings selektiv ausgewählt und es besteht kein wirkliches Verständnis für deren Zusammenhänge oder die Relevanz in Bezug auf das Gesamtkonzept.	- Aufzählen und Klassifizieren mehrerer, relevanter Aspekte - Keine Verknüpfung der Aspekte untereinander - Stellt keine Beziehungen der Aspekte zum Gesamtkonzept her
Relational (4)	Die Antwort berücksichtigt nahezu alle relevanten Aspekte. Diese werden richtig kombiniert und es wird eine kohärente Schlussfolgerung gezogen.	- Anwenden von Konzepten auf Alltags- oder Arbeitssituationen - Generieren eigener Beispiele - Vergleichen, kontrastieren, Begründungszusammenhänge herstellen

Tabelle 15: Taxonomie-Niveaus und Indikatoren

SOLO Levels	Kodierungsanweisung	Indikatoren
Extended Abstract (5)	Relevante Aspekte werden miteinander in Beziehung gesetzt und können auf einen neuen Kontext oder ein neues Phänomen übertragen werden.	<ul style="list-style-type: none">- Anwenden des Konzeptes auf einen ganz neuen Kontext- Hypothesen formulieren- Verallgemeinern eines Konzeptes- Das Konzept kritisch reflektieren

.5 Musterbeispiel

Der folgende Text stammt aus einem Lernjournal in einem Seminar zur Sozialpsychologie. Der Prompt fordert dazu auf, Beziehungen zwischen dem neuen Lernstoff und dem eigenen Vorwissen herzustellen. Der kursive Text stellt eine Analyse der einzelnen Textabschnitte dar und beschreibt die darin enthaltenen Lernaktivitäten.

LERNJOURNALAUSZUG

Seminar: Einführung in die Sozialpsychologie

Thema: Interaktionen in der Schule und deren Einfluss auf die Leistung

Lernfrage: Welche Bezüge weisen die neuen Konzepte zu Ihnen bereits bekannten Theorien und Modellen auf?

Antwort: Zunächst möchte ich auf die Studie von Rost und Czeschlik eingehen. In dieser stehen sich die beiden Eigenschaften intelligent und beliebt gegenüber. Bei diesem Versuch mussten ungefähr 6500 Kinder einer vierten Klasse verschiedene Intelligenztests machen. Zudem wurden Aussagen der Schüler über ihre Mitschüler aufgrund ihrer Beliebtheit gemacht. Das Ergebnis dieser Studie ist, dass die intelligenten Schüler meistens auch die beliebten und nicht abgelehnten Schüler sind. Im Gegensatz dazu sind die unbeliebten und abgelehnten Schüler meistens die Schüler mit einer geringeren Intelligenz.

Hier wird zunächst der Kontext beschrieben, auf den sich die Antwort bezieht. Der Autor oder die Autorin hat dazu eine Studie aus dem Bereich der Psychologie ausgewählt und das zentrale Ergebnis dargestellt.

Deshalb könnte man die Frage aufwerfen, inwiefern dieses Ergebnis ein Ergebnis der Sich-selbst-erfüllenden-Prophezeiung ist. Könnte es nicht sein, dass die Folgen für die Außenseiter ihr Leben und Lernverhalten beeinflussen? Viele abgelehnte Schüler nehmen nämlich nicht mehr am Austausch (sowohl schulischer als auch sozialer Art) teil, wodurch ihnen Wissen verloren geht. Denn in der Schule helfen

sich Schüler untereinander, ob dies bei den Hausaufgaben ist oder ob man auch mal abschreiben darf. Da dieser Kontakt nicht stattfindet und die Mitschüler ja sowieso glauben, dass der abgelehnte Schüler so und so ist, passt er sich dieser Meinung an und somit kommt sein Schulversagen zustande. Auch Fritz Heider stellt die These auf, dass ein Mensch, sobald es eine externe Beeinflussung gibt, nicht mehr so stark auf seine innere, eigene Beeinflussung eingeht.

Es wird auf der Basis der Studie selbstständig eine weiterführende Fragestellung entwickelt und eine Hypothese formuliert. Der Lernstoff wird in einen neuen Kontext übertragen. Zudem werden eigene Erfahrungen aus dem Schulalltag herangezogen.

All diese Lernaktivitäten (Hypothese formulieren, Verknüpfung zu eigenen Erfahrungen und dem eigenen Vorwissen herstellen, ein Konzept in einen neuen Kontext übertragen) zählen zu den Elaborationsstrategien.

.6 Vorwissenstest

Liebe Studierende,

zur Vorbereitung auf das Seminar möchten wir Sie bitten, einen kurzen Test zu machen. Bitte versuchen Sie, alle Fragen zu beantworten, so gut es Ihnen möglich ist.

Im Folgenden sind 15 Multiple-Choice-Fragen mit jeweils 5 Antwortalternativen zu beantworten. Achtung: Bei manchen Fragen trifft nur eine, bei anderen treffen mehrere Aussagen zu.

Welche Aussagen über Merkmale treffen zu?

- a) Latente Merkmale sind nicht direkt messbar
- b) Psychische Eigenschaften sind meistens manifeste Merkmale
- c) Die Operationalisierung definiert das Vorgehen zur Merkmalerfassung
- d) Latente Merkmale sind stets normalverteilt
- e) Alle Aussagen sind zutreffend

Welche der Aussagen sind auf Intervallskalenniveau nicht zulässig?

- a) Ein Messwert ist größer als ein anderer
- b) Die Differenz eines Messwertpaares ist kleiner als die eines anderen Messwertpaares
- c) Ein Messwert ist halb so groß wie ein anderer
- d) Zwei Messwerte sind ungleich
- e) Der Unterschied zweier Messwerte ist doppelt so groß wie der zwischen zwei anderen

Bitte beurteilen Sie, welche der folgenden Aussagen zum Median richtig sind.

- a) Der Median ist derjenige Messwert mit der größten Häufigkeit
- b) Der Median sollte erst ab Intervallskalenniveau angegeben werden
- c) Der Median ist nicht immer eindeutig bestimmbar
- d) Die positiven und negativen Abweichungen vom Median heben sich auf
- e) Die Summe der quadrierten Schätzfehler ist kleiner als bei konstanter Schätzung eines beliebigen anderen Wertes

Welche der folgenden Statistiken sind auf Ordinalskalenniveau problemlos verwendbar?

- a) Median
- b) Mittlerer Quartilabstand
- c) Standardabweichung
- d) Spearman-Rang-Korrelation
- e) Mean

Welche der folgenden statistischen Kennwerte werden nicht zu den Dispersionsmaßen gezählt?

- a) Streubereich
- b) Interquartilbereich
- c) Korrelation
- d) Mittlerer Quartilabstand
- e) Boxplot

Welche Aussagen können durch die Angabe von Prozenträngen gemacht werden?

- a) Aus jedem einzelnen Messwert kann man ablesen, wie viel Prozent der Werte gleich oder niedriger waren.
- b) Nur aus einem Teil der Messwerte kann man ablesen, wie viel Prozent der Werte gleich waren.
- c) Aus jedem einzelnen Messwert kann man ablesen, wie viel Prozent der Werte gleich oder größer waren.
- d) Nur aus einem Teil der Messwerte kann man ablesen, wie viel Prozent gleich oder größer waren.
- e) Keine der bisher ausgeführten Antwortalternativen trifft zu

Welche Aussagen zur Hypothesentestung sind zutreffend?

- a) Die Teststärke lässt sich a priori oder a posteriori bestimmen
- b) Um eine Teststärke bestimmen zu können, muss eine Effektstärke verfügbar sein
- c) Je größer die Streuung der AV, desto geringer ist die Teststärke bei sonst

gleichen Bedingungen

- d) Bei ungerichteten Hypothesen ist die Teststärke unter sonst gleichen Bedingungen größer
- e) Die Teststärke ist von der Richtung der zu prüfenden Hypothesen unabhängig

Im Rahmen der Regressionsrechnung gelte $y_i = a + bxi + ei$ mit den Beobachtungen $i = 1, \dots, n$. Welche Aussagen sind korrekt?

- a) Das R^2 ist der Anteil der durch die unabhängigen Variablen (UV) erklärten Varianz der abhängigen Variablen (AV) an ihrer Gesamtvarianz.
- b) Der Determinationskoeffizient R^2 und der Korrelationskoeffizient stehen in einem direkten Zusammenhang: Beide können Werte zwischen -1 und 1 annehmen.
- c) Die Koeffizienten der Regressionsgerade erhält man durch die Maximierung der Summe der quadrierten Abweichungen (SQA).
- d) Bei gegebener Streuung der beiden Variablen UV und AV verläuft die Regressionsgerade um so flacher, je schwächer der lineare Zusammenhang zwischen den Variablen ist.
- e) Keine der bisher ausgeführten Antwortalternativen trifft zu.

Welche der folgenden Aussagen stimmen?

- a) Für einen Mittelwertvergleich zwischen zwei Gruppen muss die abhängige Variable intervallskaliert sein.
- b) Für experimentelle Untersuchungen ist die Zufallsauswahl der Probanden charakteristisch.
- c) Externe Validität ist die Voraussetzung für interne Validität.
- d) In Experimenten wird höchstens eine unabhängige Variable untersucht.
- e) Experimentelle Laboruntersuchungen haben eine geringere externe, dafür aber eine hohe interne Validität.

Welche der folgenden Behauptungen sind zutreffend? Mangelnde Reliabilität hat ihren Grund in der Empfindlichkeit eines Messinstruments für solche Faktoren,

- a) deren Einwirkungen auf die Größe der Messwerte von systematischer Natur sind

-
- b) deren Einwirkung auf die Größe der Messwerte von Messzeitpunkt zu Messzeitpunkt gleich groß ist
 - c) die in keiner Weise auf die Größe der Messwerte einwirken
 - d) deren Einwirkung auf die Größe der Messwerte unsystematisch wechselt
 - e) die mit der Messgröße korrelieren

Ein Proband A hat in einem Intelligenztest mit einer Inter-Itemkonsistenz von $r_{tt} = .64$ einen Z-Wert von 110 erzielt. Innerhalb welcher Grenzen liegt mit einer Wahrscheinlichkeit von $p < .05$ der wahre z-Wert des Probanden?

- a) Etwa zwischen 98 und 122
- b) Etwa zwischen 104 und 116
- c) Etwa zwischen 100 und 120
- d) Etwa zwischen 103.5 und 116
- e) Etwa zwischen 94 und 126

Suchen Sie den Begriff, der in dem folgenden Satz fehlt: Bei der ——— muss man davon ausgehen können, dass das Testverhalten eine repräsentative Stichprobe aus den fraglichen Verhaltensweisen bzw. dem relevanten Merkmalspool ist.

- a) Konstruktvalidität
- b) Inhaltsvalidität
- c) Differentielle Validität
- d) Logische Validität
- e) Faktorielle Validität

Der Zusatznutzen, den ein Test zur Prognoseverbesserung beiträgt, läuft unter der Bezeichnung

- a) Synthetische Validität
- b) Diskriminate Validität
- c) Faktorielle Validität
- d) Inkrementelle Validität
- e) Differentielle Validität

Unter den vereinfachten Annahmen der klassischen Testtheorie bezeichnet ein Reliabilitätskoeffizient den Anteil

- a) der wahren Varianz an der Fehlervarianz
- b) der Fehlervarianz an der wahren Varianz
- c) der Gesamtvarianz an der wahren Varianz
- d) der Fehlervarianz an der Gesamtvarianz
- e) der wahren Varianz an der Gesamtvarianz

Ein Pädagoge gibt einen neuen Power-Test in einer 6. Klasse vor. Ein Jahr später führt er den Test in der Parallelklasse durch. Aufgrund der erhobenen Daten kann nun welche der folgenden Reliabilitätsarten bestimmt werden?

- a) Stabilität
- b) Äquivalenz
- c) Inter-Itemkonsistenz
- d) Jede der vorgenannten Reliabilitätsarten
- e) Keine der vorgenannten Reliabilitätsarten

.8 Austausch mit John Biggs bezüglich des Skalenniveaus der SOLO-Taxonomie

Dear Mr Biggs,

I hope you are doing well! Let me introduce myself shortly: I am Konstanze Jenderek, psychology-student at the Goethe-University at Frankfurt/Germany. I have encountered your SOLO-Taxonomy, which is very interesting in my opinion. I would like to please you to answer a few questions about it:

1. May I use it as a metrical scale? I found an article using the SOLO-Taxonomy which correlated the findings, so I subsume it as being a metrical scale. The article: Chan, C. C., Tsui, M. S & Chan, M. Y. C. (2002). Applying the Structure of the Observed Learning Outcomes (SOLO) Taxonomy on Student's Learning Outcomes: an empirical study. *Assessment & Evaluation in Higher Education*, 27, 511-527.
2. If it is a metrical scale, can you tell me out of which reasons?
3. In how far does it depend on Marton's categorisation? I would be deeply grateful, if you would be so kind to help me answering these questions!

I wish you all the best so far!

Sincerely yours,

Konstanze Jenderek.

Dear Konstanze,

1. May I use it as a metrical scale?

If you mean can you count uni-structural as 1, multi-structural as 2, relational as 3 and Extended abstract as 4 (say 5) then yes, I have done that myself. It's not a true equal interval scale of course it suffices to use parametric statistics.

2. If it is a metrical scale, can you tell me out of which reasons (I am sure my tutor will ask me about it).

I think I've answered that above. Treating it as a scale is an arbitrary convenience but it works. To be really strict you should use nonparametric statistics.

3. In how far does it depend on Marton's categorisation?

Not at all. Marton categorises responses as they come. SOLO is generic. I once called phenomonography „an infinitely variable SOLO“ to a phemonenographer. She was not amused.

SOLO was born from Piagetian theory of development. But it is true that I saw some of Marton's results about learning and saw the similarity and concluded that we were dealing with learning not development. However the structure and theory of SOLO is quite different.

Hope that helps.

John

Literatur

- Applebee, A. (1984). Writing and reasoning. *Review of Educational Research*, 54, 577–596.
- Azevedo, R. & Bernard, R. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research*, 13, 111–127.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37, 122–147.
- Bangert-Drowns, R., Hurley, M. & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, 74, 29–58.
- Bangert-Drowns, R., Kulik, C., Kulik, J. & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61, 213–238.
- Bannert, M. (2003). Effekte metakognitiver Lernhilfen auf den Wissenserwerb in vernetzten Lernumgebungen. *Zeitschrift für Pädagogische Psychologie*, 17, 13–25.
- Bannert, M. (2009). Promoting self regulated-learning through prompts: A discussion. *Zeitschrift für Pädagogische Psychologie*, 23, 139–145.
- Beaugrande, R. (1982). Psychology and composition: Past, present, and future. In M. Nystrand (Ed.), *What writers know* (pp. 211–267). New York: Academic Press.
- Benton, S., Kiewra, K., Whitfill, J. & Dennison, R. (1993). Encoding and external-storage effects on writing processes. *Journal of Educational Psychology*, 85, 267–280.

- Bereiter, C. (1980). Development in writing. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 73–96). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bereiter, C. & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Berthold, K., Nückles, M. & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17*, 564–577.
- Bühl, A. (2008). *SPSS 16: Einführung in die moderne Datenanalyse*. München: Pearson Studium.
- Biggs, J. (1988). Assessing student approaches to learning. *Australian Psychologist, 23*, 197–206.
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*, 347–364.
- Biggs, J. & Collis, K. (1980). The solo taxonomy. *Education News, 17*, 19–23.
- Biggs, J. & Collis, K. (1982). *Evaluating the quality of learning: The solo taxonomy (structure of the observed learning outcome)*. New York: Academic Press.
- Biggs, J. & Tang, C. (2007). *Teaching for quality learning at university*. Buckingham: The Society for Research into Higher Education and Open University Press.
- Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction, 7*, 161–186.
- Borsch, F. & Imhof, M. (2007). Evaluation der Methode des Lerntagebuchs beim Einsatz in Seminaren der Pädagogischen Psychologie für Lehramtsstudierende. In M. Krämer, S. Preiser & K. Brusdeylins (Hrsg.), *Psychologiedidaktik und Evaluation VI* (S. 217–225). Göttingen: Vandenhoeck & Ruprecht.

- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- Bortz, J., Lienert, G. & Boehnke, K. (2008). *Verteilungsfreie Methoden in der Biostatistik*. Heidelberg: Springer.
- Boscolo, P. & Cisotto, L. (1997). *Making writing interesting in elementary school*. (Paper presented at the seventh biannual meeting of the European Association for Research on Learning and Instruction, Athens, Greece)
- Boulton-Lewis, G. (1994). Tertiary students' knowledge of their own learning and a SOLO taxonomy. *Higher Education*, 28, 387–402.
- Britton, J. (1980). Shaping at the point of utterance. In A. Freedman & I. Pringle (Eds.), *Reinventing the rhetorical tradition* (pp. 61–65). Ottawa: Canadian Council of Teachers of English.
- Butler, D. & Winne, P. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65, 245–281.
- Campbell, K., Watson, J. & Collis, K. (1992). Volume measurement and intellectual development. *Journal of Structural Learning*, 11, 279–298.
- Chan, C., Hong, J. & Chan, M. (2001). *Applying the Structure of the Observed Learning Outcomes (SOLO) taxonomy on students learning outcomes: a comparative review*. (Unpublished manuscript, Hong Kong, Polytechnic University.)
- Chan, C., Tsui, M., Chan, M. & Hong, J. (2002). Applying the structure of the observed learning outcomes (SOLO) taxonomy on student's learning outcomes: an empirical study. *Assessment & Evaluation in Higher Education*, 27, 511–527.
- Chick, H. (1998). Cognition in the formal modes: research mathematics and the SOLO Taxonomy. *Mathematics Education Research Journal*, 10, 4–26.
- Clariana, R. (1993). A review of multiple-try feedback in traditional and computer-based instruction. *Journal of Computer-Based Instruction*, 20, 67–74.

- Deci, E., Koestner, R. & Ryan, R. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research*, 71, 1–27.
- Deci, E. & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Deci, E. & Ryan, R. (2003). *Intrinsic motivation inventory (IMI)*. Verfügbar unter: <http://www.psych.rochester.edu/SDT/measures/intrins.html>. (16.03.2009)
- De Wever, B., Schellens, T., Valcke, M. & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education*, 46, 6–28.
- Edelmann, W. (2000). *Lernpsychologie*. Weinheim: BeltzPVU.
- Eigler, G. (1990). *Wissen und Textproduzieren*. Tübingen: Gunter Narr.
- Eigler, G. (1996). Methoden der Textproduktionsforschung. In H. Günther & O. Ludwig (Hrsg.), *Schrift und Schriftlichkeit: Ein interdisziplinäres Handbuch internationaler Forschung* (S. 992–1004). Berlin: de Gruyter.
- Eigler, G., Jechle, T., Merziger, G. & Winter, A. (1987). Über Beziehungen von Wissen und Textproduzieren. *Unterrichtswissenschaft*, 15, 382–395.
- Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication*, 28, 122–128.
- Entwistle, N. (1995). Frameworks for understanding as experienced in essay writing and in preparing for examination. *Educational Psychologist*, 30, 47–54.
- Feilke, H. (1996). Die Entwicklung der Schreibfähigkeiten. In H. Günther & O. Ludwig (Hrsg.), *Schrift und Schriftlichkeit: Ein interdisziplinäres Handbuch internationaler Forschung* (S. 1178–1191). Berlin: de Gruyter.
- Fellows, N. (1994). A window into thinking: Using student writing to understand conceptual change in science learning. *Journal of Research in Science Teaching*, 31, 985–1002.

- Flower, L. & Hayes, J. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32, 365–387.
- Foucault, M. (1994). *Die Ordnung der Dinge. Eine Archäologie der Humanwissenschaften*. Frankfurt/M.: Suhrkamp.
- Garner, R. (1988). Verbal-report data on cognitive and metacognitive strategies. In C. Weinstein, E. Goetz & P. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (p. 63-76). San Diego, CA: Academic Press.
- Glogger, I., Holzäpfel, L., Schwonke, R., Nückles, M. & Renkl, A. (2009). Activation of learning strategies in writing learning journals. The specificity of prompts matters. *Zeitschrift für Pädagogische Psychologie*, 23, 95-104.
- Gläser-Zikuda, M. (2010). *Lerntagebuch und Portfolio aus empirischer Sicht*. Landau: Verlag Empirische Pädagogik.
- Gläser-Zikuda, M., Rohde, J. & Schlomske, N. (2010). Empirische Studien zum Lerntagebuch- und Portfolio-Ansatz im Bildungskontext - ein Überblick. In R. S. Jäger & U. Lissmann (Hrsg.), *Lerntagebuch und Portfolio aus empirischer Sicht* (S. 3–34). Landau: Verlag Empirische Pädagogik.
- Gould, J. & Boies, S. (1978). Writing, dictating, and speaking letters. *Science*, 201, 1145–1147.
- Hasselhorn, M. & Gold, A. (2009). *Pädagogische Psychologie: Erfolgreiches Lehren und Lernen*. Stuttgart: Kohlhammer.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81–112.
- Hayes, D. (1987). The potential for directing study in combined reading and writing activity. *Journal of Reading Behavior*, 19, 333–352.
- Hayes, J. & Flower, L. (1980). Identifying the organization of writing processes. In E. R. Gregg, L. W. & Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale: Erlbaum.

- Hayes, J., Flower, L., Schriver, K., Stratman, J. & Carey, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.), *Reading, writing and language learning (advances in applied psycholinguistics)* (Vol. 2, pp. 176–240). Cambridge: Cambridge University Press.
- Hübner, S., Nückles, M. & Renkl, A. (2006). Prompting cognitive and metacognitive processing in writing-to-learn enhances learning outcomes. In R. Sun, N. Miyake & C. Schunn (Eds.), *Proceedings of the 28th annual conference of the cognitive science society* (pp. 357–362). Mahwah: Erlbaum.
- Hübner, S., Nückles, M. & Renkl, A. (2007). Lerntagebücher als Medium des selbstgesteuerten Lernens. Wie viel instruktionale Unterstützung ist sinnvoll. *Empirische Pädagogik*, 21, 119–137.
- Heckhausen, H. (1989). *Motivation und Handeln*. Berlin: Springer.
- Hidi, S., Renninger, K. & Krapp, A. (2004). Interest, a motivational variable that combines affective and cognitive functioning. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion and cognition: integrative perspectives on intellectual development and functioning* (pp. 89–115). Mahwah, NJ: Lawrence Erlbaum Associates.
- Huth, K. (2004). *Entwicklung und Evaluation von fehlerspezifischem informativem tutoriellem Feedback (ITF) für die schriftliche Subtraktion*. Dissertation, Technische Universität Dresden.
- Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, 19, 509–539.
- Kalyuga, S., Chandler, P. & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40, 1–17.
- Klein, P. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review*, 11, 203–270.
- Kluger, A. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254–284.

- Kruse, O. (1997). Wissenschaftliche Textproduktion und Schreibdidaktik. Schreibprobleme sind nicht einfach Probleme der Studierenden; sie sind auch die Probleme der Wissenschaft selbst. In E. M. Jakobs & D. Knorr (Hrsg.), *Schreiben in den Wissenschaften* (S. 141–158). Frankfurt: Lang.
- Kruse, O., Jakobs, E. & Ruhmann, G. (1999). *Schlüsselkompetenz Schreiben*. Kriftel: Luchterhand.
- Kulhavy, R. & Stock, W. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, 1, 279–308.
- Langer, J. & Applebee, A. (1987). *How writing shapes thinking: A study of teaching and learning*. Urbana, IL: National Council of Teachers of English.
- Löb, M., Perels, F. & Schmitz, B. (2004). *Eine Prozessstudie auf intraindividuell- und aggregierter Ebene zum Einfluss eines standardisierten Lern-tagebuches auf die Selbstregulation bei Schülern der 8. Jahrgangsstufe*. Verfügbar unter: <http://psydok.sulb.uni-saarland.de/volltexte/2004/312/>. (3.04.2010)
- Lienert, G. & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Beltz-PVU.
- Lompscher, J. (1995). Erfassung von Lernstrategien mittels Fragebogen. In *LFF-Berichte* (Vol. 10, pp. 80–136). Potsdam: Universität Potsdam.
- Mandl, H. & Friedrich, H. (2006). *Handbuch Lernstrategien*. Göttingen: Hogrefe.
- Marton, F. & Säljö, R. (1984). Approaches to learning. In F. Marton, D. Hounsell & N. Entwistle (Eds.), *The experience of learning* (pp. 36–55). Edinburgh: Scottish Academy Press.
- Mayer, R. E. (1999). *The promise of educational psychology*. Upper Saddle River, NJ: Prentice-Hall.
- McCrindle, A. & Christensen, C. (1995). The impact of learning journals on metacognitive and cognitive processes and learning performance. *Learning and Instruction*, 5, 167–185.

- McNamara, D., Kintsch, E., Songer, N. & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1–43.
- Molitor-Lübbert, S. (1996). Schreiben als mentaler und psychischer Prozess. In H. Günther & O. Ludwig (Hrsg.), *Schrift und Schriftlichkeit: ein interdisziplinäres Handbuch internationaler Forschung* (S. 1005–1027). Berlin: de Gruyter.
- Moosbrugger, H. & Kelava, A. (2007). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer.
- Narciss, S. (1999). Motivational effects of the informativeness of feedback. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 19, 139–158.
- Narciss, S. (2001). *Informative feedback as a bridge from instruction to learning in computer-based trainings*. (Vortrag auf der 9. Konferenz der European Association for Research in Learning and Instruction, Fribourg, Switzerland.)
- Narciss, S. (2006). *Informatives tutorielles Feedback*. Münster: Waxmann.
- Narciss, S. & Huth, K. (2004). How to design informative tutoring feedback for multi-media learning. In H. M. Niegemann, D. Leutner & R. Brünken (Eds.), *Instructional design for multimedia learning* (pp. 181–195). Münster: Waxmann.
- Nückles, M. & Fries, S. (2004). Lernprotokolle schreiben und in Lernpartnerschaften kommentieren: Eine Maßnahme zur Förderung selbstgesteuerten Lernens in der Hochschullehre. *Newsletter der Fachgruppe Pädagogische Psychologie*, 1, 2–8.
- Nückles, M., Hübner, S. & Renkl, A. (2006). The Pitfalls of Overprompting in Writing-to-Learn. In R. Sun, N. Miyake & C. Schunn (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (p. 2575). Mahwah: Erlbaum.

- Nückles, M., Hübner, S. & Renkl, A. (2008). Short-term versus long-term effects of cognitive and metacognitive prompts in writing-to-learn. In G. Kanselaar, V. Jonker, P. A. Kirschner & F. J. Prins (Eds.), *Proceedings of the 8th International Conference of the Learning Sciences* (pp. 124–131). Utrecht, NL: ICLS.
- Nückles, M., Hübner, S. & Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning and Instruction, 19*, 259–271.
- Nückles, M., Schwonke, R., Berthold, K. & Renkl, A. (2004). The use of public learning diaries in blended learning. *Learning, Media and Technology, 29*, 49–66.
- Newell, G. (1984). Learning from writing in two content areas: A case study/protocol analysis. *Research in the Teaching of English, 18*, 265–287.
- Nieder, T. (2006). *Subjektive Lernkonzepte: Analysen zu Struktur, Variabilität und Relevanz*. Hamburg: Dr. Kovac.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology, 84*, 429–434.
- Pajares, F. & Valiante, G. (2006). Self-efficacy beliefs and motivation in writing development. In C. A. MacArthur, S. Graham & J. Fitzgerald (Eds.), *Handbook of Writing Research* (pp. 158–170). New York: The Guilford Press.
- Penrose, A. (1992). To write or not to write. *Written Communication, 9*, 465–500.
- Perrig, W. & Kintsch, W. (1985). Propositional and situational representations of text. *Journal of Memory and Language, 24*, 503–518.
- Picard, C. & Imhof, M. (2010). Prompts zur Anleitung tiefenorientierten Schreibens. In M. Gläser-Zikuda (Hrsg.), *Lerntagebuch und Portfolio aus empirischer Sicht* (S. 59–80). Landau: Verlag Empirische Pädagogik.
- Pickl, C. (2004). *Selbstregulation und Transfer: Entwicklung und Evaluation eines Trainingsprogramms zum selbstregulierten Lernen und die Analyse von Transferdeterminanten in Trainingskontexten*. Weinheim: BeltzPVU.

- Quinn, K. (1995). Teaching reading and writing as modes of learning in college: a glance at the past; a view to the future. *Reading Research and Instruction*, 34, 295–314.
- Rambow, R. & Nückles, M. (2002). Der Einsatz des Lerntagebuchs in der Hochschullehre. *Das Hochschulwesen*, 3, 113–120.
- Renkl, A., Schworm, S. & Hilbert, T. (2004). Lernen aus Lösungsbeispielen: Eine effektive, aber kaum genutzte Möglichkeit, Unterricht zu gestalten. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (S. 77–92). Münster: Waxmann.
- Rheinberg, (2006). *Motivation*. Stuttgart: Kohlhammer.
- Rheinberg, F., Vollmeyer, R. & Burns, B. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern und Leistungssituationen. *Diagnostica*, 47, 57–66.
- Ryan, R. & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68–78.
- Schimmel, B. (1983). *A meta-analysis of feedback to learners in computerized and programmed instruction*. (Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada. (ERIC Document Reproduction Service No. ED233708)
- Schmalhofer, F., Boschert, S. & Kühn, O. (1990). Der Aufbau allgemeinen Situationswissens aus Text und Beispielen. *Zeitschrift für Pädagogische Psychologie*, 4, 177–186.
- Schmalhofer, F. & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representations. *Journal of Memory and Language*, 25, 279–294.
- Schmitz, B. (2001). Self-monitoring zur Unterstützung des Transfers einer Schulung in Selbstregulation für Studierende. *Zeitschrift für Pädagogische Psychologie*, 15, 181–197.

- Schnotz, W. (2001). Textproduktions- und Textverstehensforschung. In M. Haspelmath, E. König, W. Oesterreicher & W. Raible (Hrsg.), *Handbücher zur Sprach- und Kommunikationswissenschaft: Sprachtypologie und sprachliche Universalien* (S. 154–162). Berlin: de Gruyter.
- Schunk, D. H. & Swartz, C. W. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology*, 18, 337–354.
- Schwandt, T. (2007). *The Sage dictionary of qualitative inquiry*. Thousand Oaks, CA: Sage.
- Säljö, R. (1979). Learning about learning. *Higher Education*, 8, 443–451.
- Spivey, N. (1990). Transforming texts. *Written Communication*, 7, 256–287.
- Spörer, N. (2004). *Strategie und Lernerfolg: Validierung eines Interviews zum selbstgesteuerten Lernen*. Dissertation, Universität Potsdam.
- Stark, R., Flender, J. & Mandl, H. (2001). *Lösungsbeispiel pur oder angereichert? Bedingungen und Effekte erfolgreichen Lernens mit einem komplexen Lösungsbeispiel im Bereich empirischer Forschungsmethoden und Statistik* (Forschungsbericht). Universität München.
- Tang, C. (1994). Assessment and student learning: Effects of modes of assessment on students' preparation strategies. In G. Gibbs (Ed.), *Improving student learning: Theory and practice* (pp. 151–170). Oxford, UK: Oxford Brookes University.
- Thillmann, H., Künsting, J., Wirth, J. & Leutner, D. (2009). Is it merely a question of "what" to prompt or also "when" to prompt? The role of point of presentation time of prompts in self-regulated learning. *Zeitschrift für Pädagogische Psychologie*, 23, 105–115.
- Tulving, E. & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology*, 77, 593–601.
- Tulving, E. & Thomson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352–373.

- Tynjälä, P. (1998). Writing as a tool for constructive learning: Students' learning experiences during an experiment. *Higher Education*, 36, 209–230.
- Tynjälä, P., Mason, L. & Lonka, K. (2001). *Writing as a learning tool: Integrating theory and practice*. Berlin: Springer.
- Van Dijk, T. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Wallace, D. & Hayes, J. (1991). Redefining revision for freshmen. *Research in the Teaching of English*, 25, 54–66.
- Watson, J., Chick, H. & Collis, K. (1988). Applying the SOLO Taxonomy to errors on area problems. In J. Pegg (Ed.), *Mathematical interfaces* (pp. 260–281). Armidale: Australian Association of Mathematics Teachers.
- Wegge, J. (1998). *Lernmotivation, Informationsverarbeitung, Leistung*. Münster: Waxmann.
- Wild, K. (2000). *Lernstrategien im Studium*. Münster: Waxmann.
- Wild, K. & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 15, 185–200.
- Wiley, J. & Voss, J. (1996). The effects of playing historian on learning in history. *Applied Cognitive Psychology*, 10, 63–72.
- Wiley, J. & Voss, J. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, 91, 301–311.
- Winne, P. & Butler, D. (1994). Student cognition in learning from teaching. *International Encyclopedia of Education*, 2, 5738–5775.
- Winne, P. & Hadwin, A. (1998). Studying as self-regulated learning. In D. Hacker, J. Dunlosky & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Hillsdale, NJ: Erlbaum.

- Winther, E. (2006). *Motivation in Lernprozessen: Konzepte in der Unterrichtspraxis von Wirtschaftsgymnasien*. Dissertation, Universität Göttingen.
- Wirth, J. (2009). Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie*, 23, 91–94.
- Wirth, J. & Leutner, D. (2006). Selbstregulation beim Lernen in interaktiven Lernumgebungen. In H. Mandl & H. F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 172–184). Göttingen: Hogrefe.
- Wittgenstein, L. (2003). *Philosophische Untersuchungen*. Frankfurt/M.: Suhrkamp.
- Yeung, A., Jin, P. & Sweller, J. (1998). Cognitive load and learner expertise: Split-attention and redundancy effects in reading with explanatory notes. *Contemporary Educational Psychology*, 23, 1–21.
- Zimmerman, B. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, 81, 329–339.
- Zimmerman, B. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego, CA: Academic Press.