
**Entwicklung und Evaluation von fehlerspezifischem
informativem tutoriellem Feedback (ITF)
für die schriftliche Subtraktion**

Katja Huth

von der Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

zur Erlangung des akademischen Grades
Doctor rerum naturalium
(Dr. rer. nat.)

genehmigte Dissertation

Vorsitzender: Prof. Dr. habil. B. Brocke
Gutachter: Prof. Dr. habil. H. Körndle
Prof. Dr. habil. T. Goschke
Prof. Dr. habil. E. Stern

Tag der Verteidigung: 7. Dezember 2004

Ohne dem zahlreichen informativen Feedback, welches ich im Verlauf der Fertigstellung dieser Arbeit erhalten habe, läge diese Arbeit nicht in ihrer jetzigen Form vor. Mein besonderer Dank gilt daher Herrn Prof. Körndle für die wissenschaftliche Betreuung dieser Dissertation. Seine stetige Förderung und sein fachlicher Rat haben das Entstehen dieser Arbeit überhaupt erst möglich gemacht. Ausgesprochen dankbar bin ich Frau Dr. Narciss für ihre konstruktiven Hinweise und kritischen Anregungen sowie ihre moralische Unterstützung. Antje Proske und Grit Reimann möchte ich für die anregenden Diskussionen und für ihren emotionalen Beistand vor allem in der Endphase dieser Arbeit danken. Mein besonderer Dank gilt Kathrin Schleicher, welche garantierte, dass sich meine inhaltlichen Vorstellungen auch programmiertechnisch umsetzen ließen.

1	Problemstellung und Ziele	1
2	Empirische Befunde und theoretische Konzeptionen zu Feedback in Lehr-Lernsituationen	4
2.1	Begriffsbestimmung - Feedback in Lehr-Lernsituationen	4
2.2	Empirische Befunde zur Lernwirksamkeit von Feedback in Lehr-Lernsituationen	6
2.3	Heuristisches Modell zur Analyse und Untersuchung von Feedbackeffekten (Narciss, 2004)	10
2.3.1	Inhaltliche und Formale Feedbackqualität	12
2.3.2	Lernereigenschaften	20
2.4	Gestaltungsprinzipien für ITF (Narciss, 2004; Narciss & Huth, 2004)	25
2.4.1	Determinanten von ITF	26
2.4.2	Auswahl des Feedbackinhalts auf der Grundlage von Anforderungs- und Fehleranalysen	29
2.4.3	Prinzipien für die formale Gestaltung von ITF	32
2.5	Evaluationsprinzipien für ITF (Narciss, 2004)	37
2.6	Zusammenfassung	41
3	Entwicklung von fehlerspezifischem ITF für die schriftliche Subtraktion	42
3.1	Theorien zum Erwerb korrekten und fehlerhaften Wissens in der schriftlichen Subtraktion	42
3.1.1	Die induction hypothesis	43
3.1.2	Die repair theory	44
3.2	Aufgabenanforderungen und systematische Fehler beim schriftlichen Subtrahieren	47
3.2.1	Abziehen mit Borgen vs. Ergänzen mit Erweitern	48
3.2.2	Aufgabenanforderungen und systematische Fehler nach Gerster (1982)	51
3.2.3	Aufgabenanforderungen und systematische Fehler nach Kühnhold und Padberg (1986)	56
3.2.4	Aufgabenanforderungen und systematische Fehler – Eine theoretische und empirische Reanalyse	58
3.3	Entwicklung von fehlerspezifischem ITF für das schriftliche Subtrahieren	72
3.3.1	Selektion der ITF-Inhalte	72
3.3.2	Selektion von Form und Modus der ITF-Präsentation	74
3.3.3	Resultierender fehlerspezifischer ITF Algorithmus	75
3.4	Zusammenfassung	77
4	Evaluation des fehlerspezifischen ITFs in seinem Einfluss auf Lern- und Motivationsparameter – Studie 1	78
4.1	Fragestellungen und Forschungshypothesen	78
4.2	Methode	81
4.2.1	Versuchspersonen	81
4.2.2	Untersuchungsdesign	82
4.2.3	Programmbeschreibung der Übungssoftware „Subtratino“	85
4.2.4	Messinstrumente	91
4.2.5	Untersuchungsablauf	99
4.2.6	Analysevorbereitung und Methoden der Datenanalyse	101

4.3	Ergebnisdarstellung	107
4.3.1	Treatmentleistung	107
4.3.2	Nachtestleistung.....	108
4.3.3	Korrekturwahrscheinlichkeitsindex	109
4.3.4	Einschätzung des wahrgenommenen Feedbacknutzen	111
4.3.5	Post-aktionales Kompetenzerleben.....	112
4.4	Diskussion	113
5	Evaluation des fehlerspezifischen ITFs im Vergleich zu unterschiedlich komplexen Feedbackformen – Studie 2	117
5.1	Fragestellungen und Forschungshypothesen	118
5.2	Methode.....	121
5.2.1	Versuchspersonen	122
5.2.2	Untersuchungsdesign	122
5.2.3	Programmbeschreibung der Übungssoftware „Subtratio“	126
5.2.4	Messinstrumente	127
5.2.5	Untersuchungsablauf.....	133
5.2.6	Analysevorbereitung und Methoden der Datenanalyse.....	134
5.3	Ergebnisdarstellung	139
5.3.1	Nachtestleistung.....	139
5.3.2	Korrekturwahrscheinlichkeitsindex	145
5.3.3	Treatmentleistung	150
5.3.4	Post-aktionales Kompetenzerleben.....	154
5.4	Diskussion	160
5.4.1	Lernwirksamkeit der untersuchten Feedbackformen	161
5.4.2	Motivationale Wirksamkeit der untersuchten Feedbackformen.....	169
6	Zusammenfassung und abschließende Diskussion	174
6.1	Grundlagen- und anwendungsbezogene Implikationen.....	180
6.2	Forschungsd desiderata.....	182
7	Verzeichnisse	186
7.1	Literaturverzeichnis	186
7.2	Abbildungsverzeichnis	197
7.3	Tabellenverzeichnis	199
7.4	Abkürzungsverzeichnis	201
8	Anhänge	202
8.1	Anhang A: Fragebögen.....	202
8.2	Anhang B: Feedbackinhalte für ITF, KR-KCR, KR-KM und KR-KH Feedback	217

1 PROBLEMSTELLUNG UND ZIELE

Obwohl das Angebot von Feedback in Lehr-Lernsituationen eine wichtige instruktio-nale Maßnahme darstellt, um Lernende in ihrem Lernprozess zu unterstützen (Gagné, 1985; Merrill, 1983), wird dessen Potenzial in multimedialen Lernumgebungen nur selten optimal genutzt. Feedback, welches in aller Regel im Anschluss an die Bearbeitung von Lernaufgaben bereitgestellt wird und den Lerner¹ über die Qualität seiner Aufgabenbearbeitung informiert, beschränkt sich häufig auf die Bereitstellung einer richtig/falsch Rückmeldung mit sich anschließenden, zusätzlichen Lösungsversuchen bzw. der Präsentation der richtigen Lösung. Darüber hinausgehende Informationen, wie z.B. das Angebot korrekturrelevanter Informatio-nen verbunden mit einem erneuten Lösungsversuch, werden selten bereitgestellt.

Als prototypisch für diese Art der Feedbackgestaltung erweist sich das in mathemati-scher Grundschulsoftware implementierte Feedback (z.B. Alfons-Lernsoftware, Adi-Lern-
software). Auf eine fehlerhafte Lösung wird dem Lerner entweder unmittelbar die korrekte Lösung präsentiert oder diese wird ihm nach einem zweiten fehlerhaften Lösungsversuch dar-
geboten. Der Lerner erhält weder Informationen zu dem Ort und der Art des Fehlers noch einen Hinweis auf die korrekte Lösungsstrategie, noch die Gelegenheit, dieses Wissen in ei-nem erneuten Lösungsversuch anzuwenden, um selbständig zur korrekten Lösung zu gelan-gen.

Vor diesem Hintergrund bestand das zentrale Ziel der vorliegenden Arbeit in der Ent-wicklung und Evaluation von fehlerspezifischem informativem tutoriellem Feedback für die schriftliche Subtraktion. Dieses Themengebiet wurde ausgewählt, da dieser Gegenstandsbe-reich dank seines algorithmischen Charakters und den Problemen, die das Lösen schriftlicher Subtraktionsaufgaben vielen Schülern bereitet, auf eine lange Forschungstradition im Bereich computergestützter Wissensdiagnostik zurückblicken kann. Verantwortlich für dieses For-schungsinteresse war die Beobachtung, dass sich die fehlerhaften Antwortmuster der Schüler durch eine gewisse Systematik auszeichneten, die darauf schließen ließen, dass die auftreten-den Fehler nicht zufällig zustande kamen im Sinne von „slips“ (Norman, 1981) oder „unsy-
stematischen Fehlern“, sondern auf ein regelhaftes, systematisches Vorgehen zurückgeführt

¹ Aus Gründen der Lesbarkeit wird in der vorliegenden Arbeit auf die Nennung der männlichen und weiblichen Schreibweise verzichtet. Die männliche Schreibweise schließt immer auch die weibliche ein.

werden konnten (Brown & VanLehn, 1980; Cox, 1975; Langley & Ohlsson, 1984; VanLehn, 1986, 1990; Young & O'Shea, 1981). So liegen eine Reihe von Erkenntnissen zu systematischen Fehlern in der schriftlichen Subtraktion vor, die eine Diagnostik nicht beherrschter Subprozeduren des Rechenalgorithmus ermöglichen (Gerster, 1982; Kühnhold & Padberg, 1986; VanLehn, 1990) und damit eine wesentliche Voraussetzung für die Entwicklung von fehlerspezifischem informativem tutoriellem Feedback darstellen.

Fehlerspezifisches informatives tutorielles Feedback, welches eine spezielle Form informativen tutoriellen Feedbacks (ITF, Narciss, 2004) repräsentiert, zeichnet sich dadurch aus, dass bei auftretenden Fehlern fehlerspezifische Informationen zur Fehlerkorrektur angeboten werden, ohne dem Lernenden die korrekte Lösung unmittelbar zu präsentieren. Die Entwicklung dieses fehlerspezifischen ITFs für die schriftliche Subtraktion erfolgte auf der Grundlage des heuristischen Feedbackmodells zur Analyse und Untersuchung von Feedbackeffekten von Narciss (2004) und den aus diesem Feedbackmodell abgeleiteten Prinzipien zur Gestaltung und Evaluation von ITF (Narciss, 2004).

Neben der Entwicklung von fehlerspezifischem ITF bestand ein weiteres zentrales Ziel der vorliegenden Arbeit in der Überprüfung der Effektivität des fehlerspezifischen ITFs hinsichtlich Lern- und Motivationsparameter und damit auch in einer empirischen Validierung der von Narciss (2004) entwickelten Gestaltungs- und Evaluationsprinzipien. Dabei wurden vor dem Hintergrund der Ergebnisse experimenteller Studien zu informativem tutoriellem Feedback (Nagata, 1993; Nagata & Swisher, 1995; Collins, Carnine & Gersten, 1987) und basierend auf theoretischen Überlegungen positive Wirkungen des fehlerspezifischen ITFs sowohl in kognitiver als auch in motivationaler Hinsicht erwartet. So sollte zum einen der Erwerb von Wissen über eine aktive Auseinandersetzung mit den bereitgestellten Feedbackinformationen unterstützt werden. Zum anderen sollten bei dem Lerner positive Lernerfahrungen, die aus einer erfolgreichen Anwendung der korrektiven Informationen resultieren, zu einem positiven Kompetenzerleben und zu einer positiven Selbstbewertung führen, die wiederum das Engagement in zukünftige Lernaktivitäten beeinflussen sollten.

Diese Annahmen wurden in zwei Studien mit einem entwickelten computergestützten Übungsprogramm zum Themengebiet der schriftlichen Subtraktion überprüft.

Die vorliegende Arbeit gliedert sich in fünf Teile. Im ersten Teil (Kapitel 2) wird kurz auf die empirische Befundlage zur Lernwirksamkeit von Feedback eingegangen. Vor diesem Hintergrund wird auf Grundlage des heuristischen Feedbackmodells zur Analyse und Untersuchung von Feedbackeffekten nach Narciss (2004) die Bedeutung individueller Faktoren als auch die Bedeutung der inhaltlichen und formalen Feedbackqualität für die Feedbackeffektivität herausgearbeitet und diskutiert. Mit dem Wissen, dass die Feedbackwirksamkeit von multiplen Faktoren determiniert ist, werden in einem nächsten Schritt Gestaltungsprinzipien von informativem tutoriellem Feedback vorgestellt, die Narciss (2004; Narciss & Huth, 2004) aus ihrem heuristischen Feedbackmodell abgeleitet hat. Abschließend werden Evaluationsprinzipien (Narciss, 2004) vorgestellt, in denen beschrieben ist, wie die Wirksamkeit des entwickelten ITF erfasst und evaluiert werden kann.

Im zweiten Teil (Kapitel 3) wird dezidiert beschrieben, wie die von Narciss (2004; Narciss & Huth, 2004) aus ihrem heuristischen Feedbackmodell abgeleiteten Gestaltungsprinzipien genutzt wurden, um fehlerspezifisches ITF für die schriftliche Subtraktion zu entwickeln.

Im dritten und vierten Teil (Kapitel 4 und 5) dieser Arbeit wird von zwei Studien berichtet, welche den Einfluss des fehlerspezifischen ITFs auf Lern- und Motivationsparameter untersucht haben.

Im Anschluss an eine separate Diskussion der Ergebnisse beider Studien werden diese im letzten Teil der vorliegenden Arbeit (Kapitel 6) in einer abschließenden Zusammenfassung und Diskussion integrierend dargestellt und diskutiert sowie grundlagen- und anwendungsbezogene Implikationen und offene Forschungsfragen abgeleitet.

2 EMPIRISCHE BEFUNDE UND THEORETISCHE KONZEPTIONEN ZU FEEDBACK IN LEHR-LERNSITUATIONEN

Ziel dieses Kapitels ist es, empirische Befunde und theoretische Konzeptionen zu Feedback vorzustellen. Um eine für diese Arbeit relevante Definition von Feedback zu finden, wird in einem ersten Schritt der Versuch einer Begriffsbestimmung und -eingrenzung unternommen (Kapitel 2.1). Anschließend werden Ergebnisse verschiedener Metaanalysen präsentiert, welche hinsichtlich der interessierenden Frage zur Lernwirksamkeit von Feedback inkonsistente, teilweise sogar widersprüchliche Befunde lieferten. Mit dem Ziel, diese inkonsistenten Ergebnisse zu erklären, werden nachfolgend Ursachen benannt, die verdeutlichen, dass die Wirksamkeit von Feedback von multiplen Faktoren determiniert wird (Kapitel 2.2). Vor diesem Hintergrund wird in einem nächsten Schritt das heuristische Feedbackmodell zur Analyse und Untersuchung von Feedbackeffekten von Narciss (2004) vorgestellt, welches mit dem Ziel entwickelt wurde, sowohl die Bedeutung individueller Faktoren als auch die Bedeutung der inhaltlichen und formalen Feedbackqualität für die Feedbackeffektivität herauszuarbeiten. Anhand dieses Modells werden ausgewählte Faktoren, welche die Feedbackeffektivität beeinflussen, und die eine Rolle bei der Entwicklung des fehlerspezifischen ITFs für die schriftliche Subtraktion gespielt haben, vorgestellt und diskutiert (Kapitel 2.3). Anknüpfend an diese Darstellung werden Gestaltungsprinzipien für informatives tutorielles Feedback vorgestellt, die Narciss (2004; Narciss & Huth, 2004) aus ihrem heuristischen Feedbackmodell abgeleitet hat (Kapitel 2.4). Abschließend werden Evaluationsprinzipien (Narciss, 2004) vorgestellt, welche beschreiben, wie die multiplen Wirkungen des entwickelten ITF erfasst und evaluiert werden können (Kapitel 2.5).

2.1 Begriffsbestimmung - Feedback in Lehr-Lernsituationen

Mory (1996) definierte Feedback als jede Art von Rückmeldung, die dem Lernenden im Anschluss an eine Aufgabenbearbeitung durch die Lernumgebung bereitgestellt wird und den Lernenden über die Qualität seiner Antwort informiert. Diese Information dient dabei der Bestätigung oder Veränderung bereits bestehenden Wissens, gleichwohl die zentrale Funktion von Feedback eher in seiner korrigierenden als in seiner bestätigenden Funktion gesehen wird. Diese Definition, welche ein in der Feedbackforschung allgemeines lernpsychologi-

sches Verständnis von Feedback widerspiegelt, stellt aber nur eine Facette von Feedback in Lehr-Lernsituationen dar.

Eine weitere Facette von Feedback wurde im Rahmen der Forschung zum selbstregulierten Lernen herausgearbeitet und in Abgrenzung zu dem klassischen Feedbackverständnis (externes Feedback) als internes Feedback bezeichnet. Beiden Facetten von Feedback gemeinsam ist die Annahme, dass Feedback Informationen beinhaltet, die einen Vergleich des tatsächlichen Resultates (Ist-Zustand) mit einem gewünschten Resultat (Soll-Zustand) ermöglichen. Nur die Quellen, aus denen Informationen zu diesem Ist-Soll-Vergleich gewonnen werden können, differieren. Im Unterschied zum klassischen Feedbackverständnis, wo diese Informationen durch die Lernumgebung bereitgestellt werden, wird das interne Feedback auf der Basis fortwährender Monitoring-Prozesse durch den Lerner selbst generiert (Butler & Winne, 1995). So kann auf der Basis planender Aktivitäten das Erreichen gesetzter Ziele überprüft oder die Qualität eines selbst verfassten Essays anhand selbst aufgestellter Kriterien bewertet werden.

Eine dritte Facette von Feedback in einem Lehr-Lernkontext beschreibt Feedback, welches primär nicht den Lernprozess, sondern den Motivationsprozess regulieren soll. Das heißt, es soll nicht bestehendes Wissen durch interne oder externe Vergleichsprozesse bestätigt oder verändert werden, sondern Feedback soll vordergründig motivieren. So wurden in der pädagogisch-psychologisch orientierten Feedbackforschung verschiedene Formen motivationalen Feedbacks untersucht (Dresel, 2002; Hoska, 1993; Schunk & Rice, 1993). Motivationale Feedbackformen zielen dabei primär auf eine motivationsförderliche Gestaltung von Feedback, um motivationale Dispositionen der Lernenden positiv zu beeinflussen. Sie haben nicht primär das Ziel, die Aufgabenbearbeitung zu unterstützen.

Im Rahmen der vorliegenden Arbeit spielt das klassische Feedbackverständnis eine zentrale Rolle. Feedback, wie es in dieser Arbeit verwendet wird, bezeichnet das Bereitstellen von Informationen, die dem Lernenden während oder nach der fehlerhaften Bearbeitung von Lernaufgaben durch eine externe Informationsquelle angeboten werden, mit dem Ziel, eine korrekte Lösung dieser Aufgaben in der aktuellen oder auch in künftigen Lernsituationen zu ermöglichen. In Abgrenzung zu den beiden anderen Facetten von Feedback, dem internen und dem motivierenden Feedback, wird dieses Feedback auch als informatives Feedback bezeichnet. Wird in der vorliegenden Arbeit nur von Feedback gesprochen, ist damit immer informatives Feedback gemeint.

Eine spezielle Form von informativem Feedback repräsentiert informatives tutorielles Feedback (ITF, Narciss, 2004). Der Fokus des ITFs liegt neben der Bereitstellung korrekturrelevanter Informationen (informative Komponente) vor allem auf der tutoriellen Komponente, welche beinhaltet, dass die korrekturrelevanten Informationen dargeboten werden, ohne dem Lernenden die korrekte Lösung gleichzeitig zu präsentieren. Der Lernende sieht sich vielmehr dazu aufgefordert, die bereitgestellten strategischen Informationen in einem erneuten Lösungsversuch unmittelbar anzuwenden. Im Sinne klassischer Lerner-Tutor-Szenarien wird der Lerner darin unterstützt, die korrekte Lösung *selbständig* unter Nutzung der bereitgestellten Informationen zu finden.

2.2 Empirische Befunde zur Lernwirksamkeit von Feedback in Lehr-Lernsituationen

Die Effektivität von Feedback gegenüber keinem Feedback konnte in zahlreichen Metaanalysen zu Feedbackeffekten bestätigt werden (z.B. Azevedo & Bernard, 1995; Clariana, 1993; Schimmel, 1983). Danach erwies sich sogar Feedback mit dem geringsten Informationswert (richtig/falsch Rückmeldung) immer noch als lernwirksamer als die alleinige Bearbeitung von Lernaufgaben, bei welcher der Lerner nicht über die Qualität seiner Aufgabenbearbeitung informiert wurde. Dieses Ergebnis, welches nicht überraschend war, gilt in der Feedbackforschung mittlerweile als ein gesicherter Befund (Smith & Ragan, 1993). Hinsichtlich der Frage, welche Art von Feedback denn nun am lernförderlichsten sei, erweist sich die Befundlage hingegen als nicht so eindeutig. So konnte in den Metaanalysen zu Feedback (z.B. Azevedo & Bernard, 1995; Bangert-Drowns, Kulik, Kulik & Morgan, 1991; Clariana, 1993; Kulhavy & Stock, 1989; Schimmel, 1983) auf die Frage der Lernförderlichkeit verschiedener Feedbackarten keine zufriedenstellende Antwort gegeben werden. In den in diesen Metaanalysen enthaltenen Studien wurden typische Arten von Feedback, die hinsichtlich ihres Komplexitäts- und Elaborationsgrades variieren, in ihrer Lernwirksamkeit miteinander bzw. mit einer No-Feedback Bedingung verglichen. Solche typischen Feedbackarten waren

1. Knowledge of Response/Result Feedback (KR Feedback): Der Lerner wird darüber informiert, ob seine Aufgabenlösung richtig oder falsch ist. Er erhält keine Information zur korrekten Lösung.
2. Knowledge of the Correct Response Feedback (KCR Feedback): Nach der Aufgabenbearbeitung wird dem Lerner die richtige Lösung mitgeteilt.

3. Answer Until Correct Feedback (AUC Feedback), Try-again Feedback oder Multiple-Try Feedback (MTF): Der Lerner erhält solange Knowledge of Response mit einem erneuten Lösungsversuch, bis er die richtige Lösung gefunden hat.
4. Elaborated Feedback (EF): Diese Feedbackart beinhaltet alle zusätzlichen Informationen, die über die bloße Mitteilung des korrekten Ergebnisses hinausgehen.
5. Delayed Feedback (DF): Das Feedback wird nicht sofort, sondern verzögert präsentiert.

Trotz bestimmter Kriterien, die für alle Metaanalysen definiert waren und die erfüllt sein mussten, um in die Metaanalyse aufgenommen zu werden, zeigten sich in allen vier Metaanalysen (Azevedo & Bernard, 1995; Bangert-Drowns, Kulik, Kulik & Morgan, 1991; Clariana, 1993; Schimmel, 1983) sehr große Schwankungen der Effektstärken (siehe Tabelle 1). Ursachen für diese großen Streuungen der Effektstärken können in zwei zentralen Aspekten gesehen werden. Eine Ursache für diese inkonsistenten, teilweise sogar widersprüchlichen Ergebnisse liegt in der hohen methodischen Variabilität der Studien begründet. So differierten die in die Metaanalysen aufgenommenen Studien hinsichtlich einer Anzahl weiterer Faktoren, die nur vereinzelt als Ausschlusskriterien bei der Auswahl der Studien berücksichtigt worden waren.

Tabelle 1:

Effektstärken verschiedener Feedbackarten im Vergleich

Metaanalysen	Effektstärken	Feedbackarten im Vergleich				
Azevedo & Bernard (1995)	.03 bis 2.12	No-Feedback	< Computer-gestütztes Feedback			
Bangert-Drowns et al. (1991)	-.58 bis .38 ^a -.83 bis 1.42 ^b .00 bis .81 ^c .05 bis 1.24 ^d	KR	< No-Feedback	< KCR	< MTF (AUC)	= EF
Clariana (1993)	.07 bis 1.22 ^c -.08 bis .57 ^e .33 bis .34 ^f -.94 bis .45 ^g	No-Feedback	< KR	= KCR	= DF	= MTF (AUC)
Schimmel (1983)	-1.25 bis 1.14	No-Feedback	< KCR	< KR		

^aKR zu No-Feedback, ^bKCR zu No-Feedback, ^cMTF zu No-Feedback, ^dEF zu No-Feedback, ^eMTF zu KR, ^fMTF zu DF, ^gMTF zu KCR

Solche weiteren Faktoren sind beispielsweise:

- (a) Feedbackzeitpunkt: Wurde das Feedback unmittelbar nach jeder Antwort, nach jeder Testphase oder verzögert, d.h. 1 Stunde bis 7 Tage nach Ende der Lehreinheit, angeboten?
- (b) Aufgabentyp: Hatten die Lernenden Lückentextaufgaben, Multiple-Choice-Aufgaben, Kurzantwortaufgaben oder gemischte Aufgabentypen zu bearbeiten?
- (c) Fehlerrate während der Instruktion: Machten die Lernenden überhaupt Fehler, so dass das Feedback wirksam werden konnte?
- (d) Kontrolle der *presearch availability* (Kulhavy, 1977): Wurde die *presearch availability* kontrolliert? Dabei beschreibt die *presearch availability* das Phänomen der Antwortzugänglichkeit vor der eigentlichen Aufgabenbearbeitung. So wurde in vielen Untersuchungen (Lublin, 1965; Rosenstock, Moore & Smith, 1965; zitiert nach Anderson, Kulhavy & Andre, 1971) die korrekte Antwort unmittelbar unter dem Instruktionstext präsentiert, so dass die Lernenden nur die korrekte Antwort kopieren und in das entsprechende Antwortfeld übernehmen mussten, ohne den Instruktionstext gelesen zu haben.
- (e) Bekanntheit der Posttestitems: Wurden im Posttest andere Items präsentiert als während der Treatmentphase?
- (f) Präsentation der Posttestitems: Wurde der Posttest unmittelbar nach der Lehreinheit präsentiert oder erhielten die Lernenden den Posttest verzögert dargeboten?
- (g) Treatmentdauer: Wie lange dauerte das Treatment?
- (h) Instruktionkontext: Wurde das Feedback im Unterricht oder computergestützt präsentiert? Wenn das Feedback computergestützt präsentiert wurde, um welche Form des computergestützten Lernens handelte es sich, um ein drill-and-practice Programm oder um ein (intelligentes) tutorielles System?
- (i) Versuchspersonen: Welche Lerner nahmen an der Untersuchung teil, Grundschüler, Sekundarschüler oder Studenten?
- (j) Art der Lehrinhalte: Welche Inhalte aus welchen Wissensgebieten wurden gelehrt, Mathematik, Naturwissenschaften, Geisteswissenschaften, Sprachwissenschaften?

Diese Aufzählung von Faktoren, in denen die untersuchten Feedbackstudien variierten, verdeutlicht, dass eine Vergleichbarkeit der Studien selten gegeben war, von daher das be-

kannte „Äpfel-und-Birnen“ Problem auch für diese Metaanalysen gilt und die hohe Streuung in den Effektstärken nicht verwunderlich ist.

Neben der methodischen Vielfalt kann ein weitere Ursache der inkonsistenten Befunde in der Konstruktion des Feedbacks gesehen werden. So erklärten bereits Kulhavy und Stock (1989) die uneinheitlichen Ergebnisse ihrer Metaanalyse über eine häufig nicht theoriegeleitete sondern eher intuitive Konstruktion des Feedbacks. Kulhavy und Stock (1989) nahmen an, dass in nur wenigen Studien die Elaborationskomponente gezielt manipuliert worden sei. Die meisten Studien schienen vielmehr den Eindruck zu erwecken, als seien die Informationen nach Augenscheinvalidität ausgewählt worden. Die inkonsistenten Befunde der Feedbackliteratur können damit auch als eine Folge dieser unsystematischen Herangehensweise an die Konstruktion von Feedback betrachtet werden, da aus dieser eine geringe Passung zwischen den zu bearbeitenden Lernaufgaben und dem bereitgestellten Feedback resultierte, welche wiederum eine geringe Wirksamkeit zur Folge gehabt haben dürfte.

Die Aufzählung von Faktoren, welche in den Feedbackstudien variierten, sollte verdeutlicht haben, dass die Lernwirksamkeit von Feedback von multiplen Faktoren moderiert wird. Besteht das Ziel darin, theoriegeleitet Feedback zu entwickeln, erweist es sich als unerlässlich, Faktoren zu identifizieren, welche die Effektivität des Feedbacks beeinflussen und deren Wirkungsweise differenziert zu beschreiben, um diese dann bei der Feedbackentwicklung gezielt manipulieren zu können. Obwohl verschiedene Ansätze vorliegen, welche die Rolle vor allem individueller Faktoren in ihrem Einfluss auf die Feedbackwirksamkeit beschreiben („mindful processing“, Bangert-Drowns et al., 1991; „internal feedback“, Butler & Winne, 1995; „feedback intervention theory“ von Kluger & DeNisi, 1996; „response confidence model“, Kulhavy & Stock, 1989), fehlte in der Feedbackforschung bisher ein Ansatz, welcher versuchte, die identifizierten Faktoren in einem gemeinsamen Modell zu integrieren. Mit der Erkenntnis, dass das in den Feedbackstudien implementierte Feedback häufig nicht theoriegeleitet konstruiert wurde, sollte dieses Feedbackmodell aber auch Faktoren beinhalten, welche die inhaltliche und formale Qualität des Feedbacks determinieren. Mit dem heuristischen Feedbackmodell zur Analyse und Untersuchung von Feedbackeffekten von Narciss (2004) liegt jetzt ein Ansatz vor, welcher bisherige Erkenntnisse der Feedbackforschung integriert, die Identifikation von Faktoren, welche die Feedbackeffektivität beeinflussen, erlaubt und deren Wirkungsweise dezidiert beschreibt.

2.3 Heuristisches Modell zur Analyse und Untersuchung von Feedbackeffekten (Narciss, 2004)

Narciss (2004) liefert mit ihrem heuristischen Modell zur Analyse und Untersuchung von Feedbackeffekten einen Ansatz, auf dessen Grundlage sich Faktoren identifizieren und klassifizieren lassen, welche die Effektivität von Feedback beeinflussen und somit einen Erklärungsbeitrag hinsichtlich der inkonsistenten Feedbackbefunde leisten. Dieses heuristische Modell zu den Bedingungen und Wirkungen informativen Feedbacks (siehe Abbildung 1) basiert zum einen auf zentralen Erkenntnissen von Lehr-Lern-Modellen (Boekarts, 1996) und den bisherigen Erkenntnissen der Feedbackforschung, zum anderen aber auch auf systemtheoretischen Annahmen.

Zentral für dieses Modell ist die Unterscheidung zwischen einem internen und einem externen Regelkreis. Der externe Regelkreis repräsentiert das jeweilige Lehrmedium (z.B. Lehrperson, computergestützte Lernumgebung). Der interne Regelkreis wird durch den Lernenden repräsentiert. Wirkungen informativen Feedbacks hängen von dem Zusammenwirken beider Regelkreise ab.

In dem externen Regelkreis werden über die Lehrziele und Lehrinhalte die an den Lerner zu stellenden Anforderungen definiert. Diese Anforderungen repräsentieren gleichzeitig auch den externen Soll-Wert, das heißt, was soll der Lernende am Ende der Lehrinheit können. Ausgangspunkt für den internen und externen Regelungsprozess stellen die für die Bewältigung der Anforderungen relevanten Regelgrößen dar, welche kognitive, meta-kognitive und/oder motivationale Leistungskriterien beinhalten können. Der aktuelle Wert der Regelgrößen wird sowohl über einen internen Messfühler (der lernenden Person) und über einen externen Messfühler (das Lehrmedium) erfasst. Der über das Lehrmedium (z.B. Lehrer) erfasste Ist-Wert wird nun in der externen Regeleinrichtung mit dem externen Soll-Wert verglichen. Aus der Diskrepanz des Ist-Soll-Wert Vergleichs wird nach mehr oder weniger explizit definierten Regeln externes Feedback generiert, welches an den internen Regelkreis (den Lernenden) weitergeleitet wird.

Neben dem externen Feedback wird auch der interne Ist-Wert (= internes Feedback), welcher über den internen Messfühler (den Lernenden) erfasst wurde, an die interne Regeleinrichtung weitergegeben.

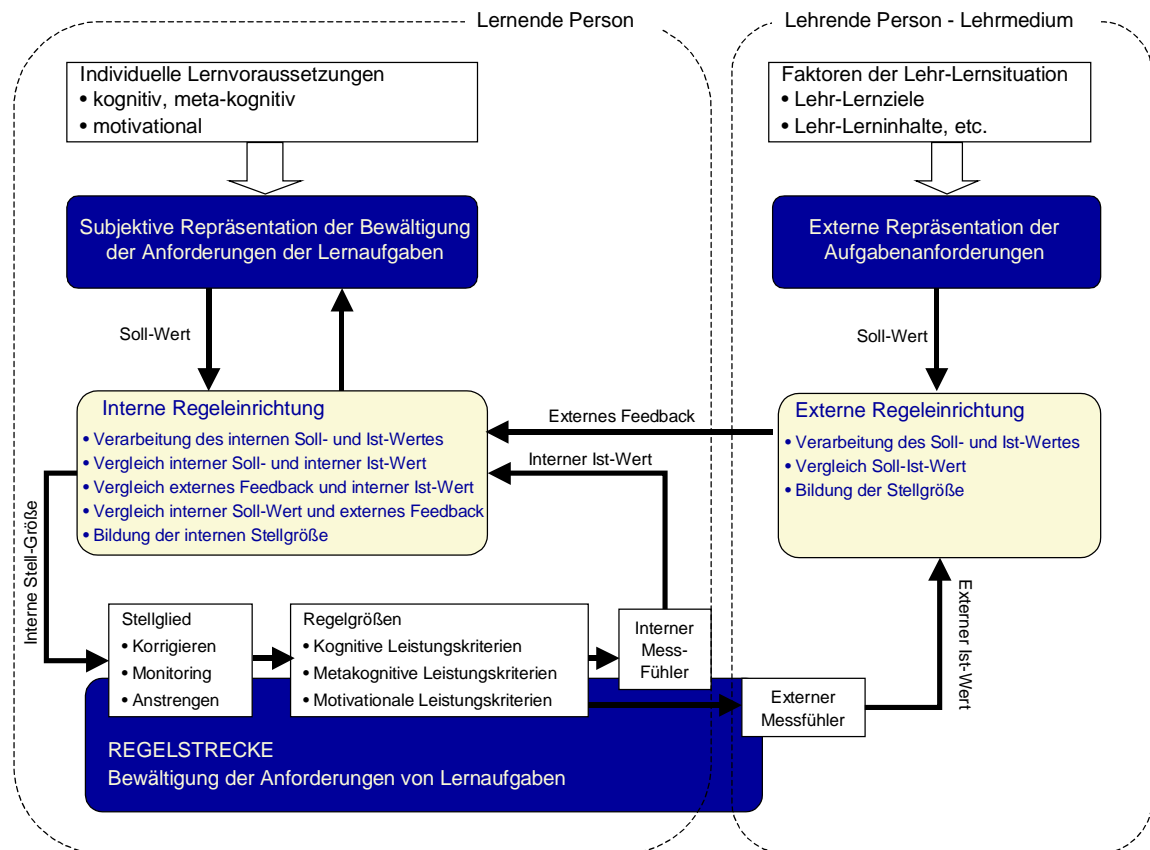


Abbildung 1: Heuristisches Modell zur Analyse und Untersuchung von Feedbackeffekten in Lehr-Lernsituationen nach Narciss (2004)

In der Regeleinrichtung wird sowohl das interne als auch das externe Feedback verarbeitet und eine neue interne Stellgröße gebildet. Diese interne Stellgröße wird an den Eingang der Regelstrecke, an das Stellglied, weiter geleitet und dient dort der Auswahl und Aktivierung korrigierender Maßnahmen. Die sowohl in der internen Regeleinrichtung als auch im Stellglied ablaufenden Prozesse hängen dabei primär von der subjektiven Repräsentation der Bewältigung der Anforderungen der Lernaufgaben ab, welche wiederum durch die individuellen (motivationalen, kognitiven, meta-kognitiven) Lernvoraussetzungen der lernenden Person determiniert sind. Das Diskrepanzmaß zwischen der internen und externen Aufgabenrepräsentation bestimmt wiederum, wie häufig die Regelkreise durchlaufen werden müssen, bevor eine relativ stabile Anpassung der internen an die externe Aufgabenrepräsentation erfolgt ist.

Anhand dieses Feedbackmodells können verschiedene zentrale Faktoren diskutiert werden, von denen vermutet wird, dass diese die Feedbackeffektivität signifikant beeinflussen. Erweist sich beispielsweise das Herstellen einer Übereinstimmung zwischen dem externen und dem internen Soll-Wert als schwierig, kann dies unter anderem daran liegen, dass am externen Regelkreis Defizite bei einzelnen Komponenten vorhanden sind, welche letztendlich zu Lasten der Qualität des dem Lernenden bereitgestellten Feedbacks gehen. Dies soll im Kapitel 2.3.1 detailliert diskutiert werden. Eine andere Erklärung für Schwierigkeiten beim Herstellen einer Kongruenz zwischen internem und externem Soll-Wert kann hingegen auf Probleme im internen Regelkreis zurückgeführt werden, das heißt auf Faktoren, die im Lerner zu lokalisieren sind. Dies soll im Kapitel 2.3.2 ausführlicher beschrieben werden.

2.3.1 Inhaltliche und Formale Feedbackqualität

Inkonsistente Feedbackbefunde können im Sinne des heuristischen Modells von Narciss (2004) darüber erklärt werden, dass das dem Lerner bereitgestellte Feedback gravierende Mängel aufweist, so dass dessen Lernwirksamkeit nicht nachgewiesen werden konnte. Feedbackmängel wiederum können ihre Ursache in einem zu groben externen Messfühler, einer zu groben Festlegung des externen Soll-Wertes sowie einer mangelnden Qualität der externen Regeleinrichtung haben.

So erfordert die Festsetzung eines externen Messfühlers, welcher den Ist-Wert des Lernenden erfasst bzw. diagnostiziert, dass dieser Messfühler ein detailliertes Modell des Wissenszustandes des Lernenden zu bilden in der Lage ist. Dies bedeutet, dass der an die externe Regeleinrichtung gelieferte Ist-Wert Informationen zu fehlenden bzw. unvollständigen Wissens-elementen des Lernenden enthalten sollte.

Weiterhin sollte eine Festlegung des externen Soll-Wertes detaillierte Informationen über die für eine erfolgreiche Bewältigung der Aufgabenanforderungen erforderlichen Wissens-elemente beinhalten. Auf der Basis eines Ist-Soll-Wert Vergleiches können dann Diskrepanzen identifiziert werden, die wiederum Anhaltspunkte liefern, welche Informationen im externen Feedback bereitgestellt werden sollten. Voraussetzung sowohl für die Messung des Ist-Wertes als auch für die Festlegung des Soll-Wertes stellt eine kognitive Anforderungs-analyse oder Lehrzielanalyse dar, welche die für eine erfolgreiche Aufgabenbearbeitung notwendigen Wissens-elemente identifiziert.

Wurden auf der Basis solcher kognitiven Anforderungsanalysen (oder Lehrzielanalysen) die für eine erfolgreiche Bewältigung der Lernaufgaben erforderlichen Wissensselemente identifiziert, muss in einem nächsten Schritt die aus dem Ist-Soll-Wert Vergleich resultierende Diskrepanz in eine Stellgröße, dem externen Feedback, transformiert werden, welche für die Bewältigung der gegebenen Anforderungen einen hohen Informationswert besitzen sollte. Die Identifikation der für eine erfolgreiche Bewältigung der Lernaufgaben erforderlichen Wissensselemente und deren Transformation in das externe Feedback beinhaltet auch eine Entscheidung darüber, wie diese Inhalte dem Lernenden präsentiert werden, damit eine aktive Auseinandersetzung mit diesen über das Feedback bereitgestellten Informationen angeregt wird. Das heißt, neben der inhaltlichen Feedbackqualität spielt auch die formale Feedbackqualität eine zentrale Rolle hinsichtlich der intendierten Feedbackwirksamkeit.

Nachfolgend werden in einem ersten Schritt einzelne Positiv- und Negativbeispiele für Feedbackstudien vorgestellt, welche die zentrale Rolle, die Anforderungsanalysen für die Gewährleistung der inhaltlichen Feedbackqualität spielen, demonstrieren sollen (Kapitel 2.3.1.1). In einem nächsten Schritt wird die Feedbackforschung dahingehend analysiert, welche Rolle die formale Gestaltung des Feedbacks für die Verarbeitung der über das externe Feedback bereitgestellten Informationen hat (Kapitel 2.3.1.2).

2.3.1.1 Inhaltliche Feedbackqualität – Die Bedeutung von Anforderungs- und Lehrzielanalysen

Die Analyse einer Anzahl von Feedbackstudien, die keine Überlegenheit informativer Feedbackformen gegenüber weniger informativen Feedbackformen nachweisen konnte, lässt vermuten, dass in diesen Studien Defizite bei der Konstruktion des Feedbacks vorliegen. So wurde in vielen Studien bei der Entwicklung informativer Feedbackformen das Lehrzielniveau und damit die über das Lehrzielniveau definierten Anforderungen nicht berücksichtigt.

In der Mehrzahl dieser Feedbackstudien wurde die Effektivität unterschiedlich informativer Feedbackformen auf der Basis von Instruktionsmaterial untersucht, welches vom Lerner das Wiedererkennen oder die Reproduktion von deklarativem Wissen erforderte (Mory, 1996; Schimmel, 1988). Aber gerade für Aufgaben, welche das Wiedererkennen oder die Reproduktion deklarativen Wissens verlangen, scheint mehr Information, die über das Feedback bereitgestellt wird, nicht zwingend in einer besseren Leistung zu resultieren. Dass Feedbackinhalte nicht unabhängig von dem jeweiligen verfolgten Lehrziel entwickelt werden sollten,

zeigen vor allem die Ergebnisse der Studien von Kulhavy, White, Topp, Chan und Adams (1985) und Dempsey und Litchfield (1993), die an dieser Stelle exemplarisch vorgestellt werden.

Feedbackstudie von Kulhavy et al. (1985)

Kulhavy et al. (1985) variierten in ihrer Studie die Komplexität des Feedbacks und untersuchten den Einfluss unterschiedlich komplexer Feedbackformen auf die Lernleistung. Sie vermuteten, dass mit Zunahme des Informationsgehaltes des Feedbacks die Wahrscheinlichkeit wächst, dass die Lerner erkennen, warum sie die Aufgabe im ersten Versuch falsch gelöst haben. Kulhavy et al. (1985) nutzten in ihrer Studie als Instruktionsmaterial Texte über die Struktur und Funktionen der U.S. Navy. Diese Texte bestanden aus 16 Abschnitten. Im Anschluss an jeden Abschnitt wurde dem Lerner eine Multiple-Choice-Aufgabe präsentiert. Nach Beantwortung dieser Aufgabe erhielt der Lerner unter Anwendung einer additiven Konstruktions-Strategie eine der folgenden 4 Feedbackformen:

1. Aufgabenstamm mit KCR
2. Aufgabenstamm mit KCR + den falschen Antwortalternativen
3. Aufgabenstamm mit KCR + den falschen Antwortalternativen + 4 Sätzen, die erklären, warum die falschen Antworten inkorrekt waren
4. Aufgabenstamm mit KCR + den falschen Antwortalternativen + 4 Sätzen, die erklären, warum die falschen Antworten inkorrekt waren + der Präsentation der relevanten Textpassage, in der die korrekte Antwort gegeben wurde.

Der Posttest enthielt dieselben 16 Multiple-Choice-Aufgaben der Instruktionsphase. Die Ergebnisse dieser Studie zeigten eine inverse Beziehung zwischen Feedbackkomplexität und Fehlerkorrektur. Lerner der Feedbackgruppe 3 lösten signifikant weniger Aufgaben der in der Instruktionsphase inkorrekt beantworteten Aufgaben im Posttest korrekt als die Lerner der Feedbackgruppe 1. Kulhavy et al. (1985) führten für dieses unerwartete Ergebnis verschiedene Erklärungen an. Zum einen vermuteten sie, dass die Hervorhebung der fehlerhaften Antworten dazu geführt habe, dass diese besser erinnert wurden. Zum anderen erklärten sie die kontraintuitiven Befunde mit dem Ansatz der Verarbeitungstiefe (Mistler-Lachmann, 1974, zitiert nach Kulhavy et al., 1985). Im Sinne dieses Ansatzes hätten die Lerner die ausführlicheren Feedbackinformationen nur oberflächlich verarbeitet, da sich diese Informationen für

die Identifikation der korrekten Antwort als redundant erwiesen hätten, da die richtige Antwort ja zeitgleich mit den zusätzlichen Informationen präsentiert wurde.

Eine dem Ansatz der Verarbeitungstiefe angelehnte Erklärung der nicht erwartungskonformen Ergebnisse erschließt sich aus der Betrachtung des in dieser Untersuchung implementierten Lehrzielniveaus. Die Versuchspersonen in dieser Untersuchung hatten deklaratives Wissen über die Strukturen und Funktionen der U.S. Navy zu erwerben. Dieses Wissen wurde dann über einen Posttest abgefragt, welcher wiederum identische Items aus der Instruktionsphase enthielt. Die Versuchspersonen mussten sich die korrekte Lösung nur einprägen. Darüber hinaus gehende Hinweise, warum bestimmte Antworten falsch sind, oder ein Verweis auf die relevante Textstelle, beinhalteten für eine korrekte Aufgabenbearbeitung redundante Informationen. Für eine richtige Aufgabenlösung stellte allein die Präsentation der korrekten Antwort eine notwendige und vor allem auch hinreichende Bedingung dar.

Dass diese Studie keinen Einzelfall darstellt, belegen weitere Untersuchungen von Phye (1979), Phye, Gugliemella und Sola (1976) sowie Siegel und Misselt (1984), in denen der Einfluss unterschiedlich komplexer Feedbackformen auf den Erwerb deklarativen Wissens (u.a. Vokabellernen) untersucht wurde.

Feedbackstudien, in denen der Einfluss unterschiedlich komplexer Feedbackformen auf den Erwerb von Konzepten und Regeln untersucht wurde, liegen in geringerem Umfang vor. Aber auch hier zeigen sich inkonsistente Befunde, die sich über eine qualitativ unterschiedliche Feedbackkonstruktion erklären lassen. So konnten Dempsey und Litchfield (1993) in ihrer Studie keine signifikanten Unterschiede in einem Behaltenstest zwischen ihren vier Feedbackgruppen nachweisen, im Gegensatz zu der Studie von Collins, Carnine und Gersten (1987).

Feedbackstudie von Dempsey und Litchfield (1993)

In dieser Untersuchung hatten die Lernenden Konzepte und Regeln zum Thema Substanzmissbrauch über die Bearbeitung von Multiple-Choice Aufgaben zu erwerben. Den Lernenden wurden in Abhängigkeit der Feedbackbedingung folgende Informationen auf eine fehlerhafte Antwort bereitgestellt: (1) KCR Feedback, (2) KCR Feedback und Buchstabe der korrekten Lösung eingeben, (3) KCR Feedback plus Erklärung, warum gewählte Antwort falsch war, (4) KCR Feedback und zweiter Antwortversuch.

Nach einer genaueren Betrachtung der Feedbackbedingungen scheint das in dieser Studie gefundene nichtsignifikante Ergebnis nicht verwunderlich. So erfordern sowohl Feedbackbedingung 2 als auch 4 vom Lerner nach der Bereitstellung der richtigen Antwort lediglich eine nochmalige Eingabe des betreffenden Buchstabens der richtigen Lösung. Worin bei diesen Feedbackbedingungen der Mehrwert an Information liegt, scheint nicht ganz nachvollziehbar. Auch in dieser Studie wurden letztendlich bei allen vier Feedbackbedingungen nur assoziative Verknüpfungen zwischen dem Itemstamm und der korrekten Antwort erworben. Da in dieser Studie das Lehrziel als Erwerb konzeptuellen und regelbasierten Wissens beschrieben werden kann, wäre ein Feedback effektiver gewesen, welches begründet, warum die Antwort falsch bzw. korrekt ist, d.h. auf für eine korrekte Aufgabenbearbeitung relevante Merkmale und Regeln hinweist.

Wird Feedback auf der Basis einer Anforderungsanalyse für komplexere Lehrziele wie dem Erwerb von Konzepten und Regeln entwickelt, kann sehr wohl eine Überlegenheit informativer Feedbackformen gegenüber weniger informativen Feedbackformen nachgewiesen werden, wie Studien von Nagata (1993), Nagata und Swisher (1995), Collins, Carnine und Gersten (1987) und viele Studien vor allem aus dem Forschungsbereich zu intelligenten tutoriellen Systemen belegen (Albacete & VanLehn, 2000; Heffernan, 2001).

Nachfolgend wird die Studie von Collins, Carnine und Gersten (1987) exemplarisch vorgestellt, um an dieser zu demonstrieren, dass für komplexere Lehrziele das Angebot elaborierter Feedbackinformationen der alleinigen Mitteilung der korrekten Lösung überlegen sein kann.

Feedbackstudie von Collins, Carnine und Gersten (1987)

In ihrer Studie untersuchten Collins, Carnine und Gersten (1987) die Effektivität zweier Feedbackformen in einer computergestützten Lernumgebung zum Ziehen korrekter Schlüsse aus Syllogismen. Die eine Feedbackbedingung präsentierte nach einem fehlerhaften Lösungsversuch die richtige Lösung (KCR), die andere Feedbackbedingung stellte elaboriertes Feedback in Form von konkreten Hinweisen und Regeln zusammen mit der korrekten Lösung bereit.

In einem Nachtest und einem Follow-up-Test, welcher zwei Wochen nach dem Treatment durchgeführt wurde, wie auch in einem Transfer-Test, in welchem Syllogismen in Textaufgaben angeboten wurden, erwies sich die elaborierte Feedbackgruppe der KCR Feed-

backgruppe signifikant überlegen. Diese Überlegenheit elaborierten Feedbacks gegenüber der KCR Feedbackbedingung kann darauf zurückgeführt werden, dass für die erfolgreiche Bearbeitung neuer Syllogismen im Nachtest und Transfertest die Bereitstellung des korrekten Ergebnisses nicht hinreichend war. Der Lerner musste das Prinzip bzw. die Regel verstanden haben und musste diese anwenden können, um nachfolgende Aufgaben korrekt lösen zu können. Im Unterschied zu der beschriebenen Studie von Kulhavy et al. (1985) war in dieser Studie nicht deklaratives Wissen, sondern Regelwissen erforderlich, so dass über die korrekte Lösung hinausgehende Informationen, die dieses Regelwissen vermittelten, sich als erfolgreich erwiesen.

Eine Schlussfolgerung, die aus dieser Darstellung verschiedener Studien gezogen werden kann, ist, dass die Wirksamkeit des bereitgestellten Feedbacks wesentlich von der inhaltlichen Qualität des Feedbacks determiniert wird, welche in vielen Feedbackstudien nicht gegeben war. So wurde in nur wenigen Studien die Informationskomponente gezielt manipuliert. Viele Studien erwecken vielmehr den Eindruck, als seien die Informationen intuitiv ausgewählt worden, so dass viele inkonsistente Befunde der Feedbackliteratur eine Folge dieser unsystematischen Herangehensweise an die Konstruktion von Feedback zu sein scheinen.

2.3.1.2 Formale Feedbackqualität – Die Bedeutung einer aktiven Auseinandersetzung mit den präsentierten Feedbackinformationen

Auch wenn informatives Feedback auf der Basis kognitiver Anforderungs- und Lehrzielanalysen entwickelt wurde und somit das Potenzial einer hohen Wirksamkeit bei der Korrektur von Fehlern besitzt, stellt dies nur eine notwendige aber keine hinreichende Bedingung für eine nachweisbare Lernwirksamkeit informativen Feedbacks dar. Letztendlich hängt die Wirksamkeit des Feedbacks davon ab, wie der Lerner die im Feedback bereitgestellten Informationen verarbeitet, das heißt, welche Prozesse im internen Regelkreis ablaufen. Diese im internen Regelkreis ablaufenden Verarbeitungsprozesse können aber auch durch die formale Gestaltung des externen Feedbacks beeinflusst werden, wie beispielsweise Erkenntnisse aus der Forschung zur presearch availability (Bangert-Drowns et al., 1991) überzeugend belegen.

So erwies sich in der Metaanalyse von Bangert-Drowns et al. (1991) das Phänomen der presearch availability (Kulhavy, 1977), welches die Zugänglichkeit der korrekten Antwort vor der eigentlichen Aufgabenbearbeitung beschreibt, als einen die Feedbackwirksamkeit beeinflussenden zentralen Faktor. So zeigten Lerner, welche auf die korrekte Lösung vor der ei-

gentlichen Aufgabenbearbeitung zugriffen, eine schlechtere Posttestleistung, niedrigere Fehlerraten während der Instruktionsphase und geringere Bearbeitungszeiten der Aufgaben als die Lerner der No-Feedback Bedingung (Bangert-Drowns et al., 1991). Allein über die Kontrolle der presearch availability konnte die Streuung der Effektstärken stark reduziert und die durchschnittliche Effektstärke erhöht werden. Die geringeren Effektstärken bei gegebener presearch availability können darüber erklärt werden, dass die Lernenden die korrekte Antwort nur zu kopieren und in das entsprechende Antwortfeld zu übernehmen hatten, ohne den Instruktionstext gelesen zu haben. Eine aktive Auseinandersetzung mit den in dem Instruktionstext präsentierten Informationen war daher nicht erforderlich. In Anbetracht dieser Erkenntnisse postulierten Bangert-Drowns et al. (1991), dass Feedback dann lernförderlich wirken sollte, wenn durch den Lerner eine aktive, aufmerksame und bewusste Verarbeitung der präsentierten Feedbackinformationen erfolgt („mindfulness“, Bangert-Drowns et al., 1991). Ein Ziel jeder Instruktion sollte daher darin bestehen, einen solchen Prozess der aktiven Auseinandersetzung mit den präsentierten Feedbackinformationen anzuregen.

In vielen Feedbackstudien wurde Feedback jedoch so konstruiert und implementiert, dass eine vertiefte Verarbeitung der Feedbackinformationen von dem Lernenden nicht gefordert war, auch wenn die presearch availability kontrolliert wurde. So wurden in den Studien von Kulhavy et al. (1985), Dempsey und Litchfield (1993), Lee (1985, zitiert nach Mory, 1996) und Phye (1979) elaborierte Feedbackinformationen kombiniert mit KCR Feedback präsentiert. Da sich für eine erfolgreiche Posttestbearbeitung das Wissen der korrekten Lösung als hinreichend erwies, ist davon auszugehen, dass der Lerner keine Notwendigkeit sah, die zusätzlich bereitgestellten Informationen zu nutzen, d.h. vertiefend zu verarbeiten. Ein signifikanter Unterschied zugunsten der Feedbackbedingung „elaborierte Feedbackinformationen plus KCR“ konnte in diesen Studien nicht gefunden werden.

Ein weiterer Beleg dafür, dass Lerner bereitgestellte Informationen, die gemeinsam mit der korrekten Lösung angeboten werden, nicht in dem intendierten Maße vertiefend bearbeiten, kann aus der Forschung zu ausgearbeiteten Lösungsbeispielen (worked-out examples) gewonnen werden. Von einem ausgearbeiteten Lösungsbeispiel kann nach Renkl (2000) dann gesprochen werden, wenn dieses aus einer Problemstellung, Lösungsschritten und schließlich der Lösung selbst besteht. Voraussetzung für eine effektive Verarbeitung ausgearbeiteter Lösungsbeispiele stellt nach Chi, Bassok, Lewis, Reimann und Glaser (1989) die Qualität der Selbsterklärungen dar, die der Lerner während der Beispielverarbeitung generiert. Erfolgrei-

che Lerner zeichnen sich hierbei dadurch aus, dass sie die Lösungsbeispiele tiefer und verstehensförderlicher bearbeiten, d.h. sie generieren mehr und qualitativ hochwertigere Selbsterklärungen und investieren auch mehr Zeit in die Beispielverarbeitung im Gegensatz zu weniger erfolgreichen Lernern (Stark, 1999). Diese produzieren deutlich weniger und suboptimale Selbsterklärungen, d.h. die Beispielverarbeitung erfolgt insgesamt oberflächlicher. Dass die Mehrzahl der Lerner zu einer solch oberflächlichen Verarbeitung bereitgestellter Informationen tendiert, welche mit einem geringen Lernerfolg einhergeht, konnte in verschiedenen Studien zu ausgearbeiteten Lösungsbeispielen gezeigt werden (Renkl, 2000). Aufgrund dieser Befunde wurden nachfolgend verschiedene Studien mit dem Ziel durchgeführt, die Lerner bei ihrer Beispielverarbeitung durch gezielte instruktionale Maßnahmen zu unterstützen. Auch diese Studien zeigten inkonsistente Ergebnisse. So fanden Alevén und Koedinger (2000) in ihrer Studie, dass die alleinige Aufforderung an den Lerner, Selbsterklärungen zu generieren, ohne diese auf ihre Korrektheit zu analysieren, nicht zu den erwarteten Lernzuwächsen führte. So kamen die Lernenden häufig der Aufforderung nicht nach bzw. wenn sie der Aufforderung nachkamen, erwies sich die Anzahl generierter korrekter Selbsterklärungen als gering. Ähnliche Ergebnisse fanden auch Conati und VanLehn (1999) in ihrer Studie.

Die Ergebnisse einer weiteren Studie von Alevén und Koedinger (2001) zeigen aber auch, dass sogar wenn die korrekte Lösung *nicht* gleichzeitig mit den elaborierten Informationen angeboten wird, dies nicht garantiert, dass die bereitgestellten elaborierten Informationen vertieft verarbeitet werden, solange von den Lernenden keine unmittelbare Anwendung dieser Informationen gefordert ist. In dieser Studie wurden den Lernenden acht Hilfelevel mit zunehmendem Detailliertheitsgrad der präsentierten Informationen angeboten. Diese Hilfen konnten von den Lernenden jederzeit sukzessiv aufgerufen werden, befanden sich demzufolge unter Lernerkontrolle und repräsentierten sogenannte on-demand Hilfen. Alevén und Koedinger konnten beobachten, dass die Lernenden, wenn sie Hilfe anforderten, in der Mehrzahl (über 80%) alle Hilfen nacheinander aufriefen. Das heißt, sie fragten solange nach Hilfe, bis sie das letzte Hilfelevel erreicht hatten, welches ihnen die für die korrekte Lösung erforderliche mathematische Gleichung präsentierte, so dass sie diese nur noch aufzulösen hatten. Die Lernenden nutzten demnach nicht die Möglichkeit, die über die Hinweislevel bereitgestellten Informationen vertiefend zu verarbeiten, um über deren Anwendung zu einer eigenständigen Aufgabenlösung zu gelangen. Stattdessen versuchten die Lernenden, so effizient wie möglich

zum letzten Hinweislevel vorzudringen, da dort ein Hinweis auf die korrekte Lösung gegeben wurde.

Basierend auf diesen Ergebnissen kann angenommen werden, dass viele Lerner die über die Instruktionsumgebung bereitgestellten Informationen, sei dies über elaboriertes Feedback plus KCR, ausgearbeitete Lösungsbeispiele oder on-demand Hilfen, nicht in dem intendierten Sinne einer vertieften Auseinandersetzung nutzen. Vor diesem Hintergrund stellt sich natürlich die Frage, wie Lerner zu einer solch intendierten vertieften Auseinandersetzung angeregt werden können. Verschiedene Möglichkeiten werden im Kapitel 2.4.3 „Ableitung von Gestaltungsprinzipien für ITF“ vorgestellt.

2.3.2 Lernereigenschaften

Auch wenn informatives Feedback so gestaltet ist, dass eine aktive Auseinandersetzung mit den präsentierten Informationen angeregt wird, wird letztendlich die Feedbackwirksamkeit darüber determiniert, wie der Lernende mit diesen Informationen interagiert. Mit der Suche nach Ursachen für die inkonsistente Befundlage der Feedbackforschung wird von daher auch zunehmend der Einfluss unterschiedlicher kognitiver, motivationaler und metakognitiver Lernercharakteristiken auf die Verarbeitung der präsentierten Feedbackinhalte diskutiert (Bangert-Drowns et al., 1991; Butler & Winne, 1995; Kluger & DeNisi, 1996; Kulhavy & Stock, 1989). Unter Nutzung der Terminologie des heuristischen Feedbackmodells von Narciss (2004) bedeutet dies, dass die im internen Regelkreis ablaufenden Prozesse die Wirksamkeit des extern bereitgestellten Feedbacks determinieren. Im folgenden werden kognitive und metakognitive Lernereigenschaften, welche die Feedbackeffektivität moderieren, gezielt besprochen. Hierbei wird in einem ersten Schritt auf die Rolle des aufgabenspezifischen Vorwissens für die Feedbackwirksamkeit rekurriert (Kapitel 2.3.2.1), bevor in einem nächsten Schritt auf die Rolle (meta-)kognitiver Lernvoraussetzungen eingegangen wird (Kapitel 2.3.2.2).

2.3.2.1 Aufgabenspezifisches Vorwissen

Dass das aufgabenspezifische Vorwissen die Wirksamkeit von Feedback moderiert, erscheint evident vor dem Hintergrund, dass Lerner, die über ein hohes aufgabenspezifisches Vorwissen verfügen, Feedback seltener in Anspruch nehmen sollten, da sie weniger Fehler machen. Konnte in Feedbackstudien, welche das aufgabenspezifische Vorwissen im Vorfeld

nicht erhoben haben, keine Lernwirksamkeit von Feedback nachgewiesen werden, kann dies darauf zurückgeführt werden, dass die bereitgestellten Feedbackinformationen bei vorwissensreichen Lernern nicht wirksam werden konnten. Tatsächlich vorhandene Feedbackeffekte hätten auf diese Weise überdeckt werden können. Trotz dieser naheliegenden Erkenntnis der Bedeutsamkeit des aufgabenspezifischen Vorwissens für die Feedbackeffektivität wurde die Erhebung des aufgabenspezifischen Vorwissens in vielen Feedbackuntersuchungen vernachlässigt. Inkonsistente Befunde zur Wirksamkeit von Feedback sind von daher nicht verwunderlich, vor allem auch vor dem Hintergrund, dass das aufgabenspezifische Vorwissen den bedeutendsten Prädiktor nachfolgender Leistungen repräsentiert (Zielinski, 1998). Eine Erklärung dafür, dass das Vorwissen in vielen Feedbackstudien nicht erhoben wurde, sieht Jacobs (2002) darin, dass Vortests eine Art Aktivierung von domänenspezifischem Wissen darstellen und daher den potenziellen Lernzuwachs durch das Feedback vermindern. So berichten Bangert-Drowns et al. (1991) in ihrer Metaanalyse, dass die Lernwirksamkeit des Feedbacks bei Versuchsplänen ohne Vortests deutlicher ausfiel als bei solchen mit Vortests. Allerdings steht dieser Befund im Widerspruch zu der bereits genannten Hypothese, dass Studien ohne Vortests geringere Effektstärken erwarten lassen, da die Lerner möglicherweise bereits über ein hohes Vorwissen verfügen und von daher wenig Fehler während der Instruktionsphase zeigen, Feedback demnach nicht wirksam werden kann.

Einer weiteren interessierenden Fragestellung in diesem Kontext widmete sich vor allem die Arbeitsgruppe um Clariana, in dem sie untersuchte, welche Rolle das aufgabenspezifische Vorwissen bei der Auseinandersetzung mit spezifischen Feedbackinhalten spielt und welche Feedbackformen für welches Vorwissensniveau empfehlenswert sind (Clariana, 1993, 2003; Clariana & Lee, 2001; Clariana & Smith, 1989). Zentrale Frage war nicht mehr der prinzipielle Einfluss des aufgabenspezifischen Vorwissens auf die Feedbackeffektivität. In das Forschungsinteresse war vielmehr die Rolle möglicher Wechselwirkungen zwischen dem Vorwissensniveau und der präsentierten Feedbackform gerückt.

In seiner Metaanalyse analysierte Clariana (1993) mögliche Interaktionseffekte zwischen dem Vorwissensniveau und der präsentierten Feedbackform. Er fand, dass primär vorwissensreiche Lerner im Vergleich zu vorwissensarmen Lernern beim Wissenserwerb von Multiple-Try Feedback (MTF = AUC) profitierten, wohingegen KCR Feedback sich als lernförderlicher gegenüber MTF bei Lernern mit geringem Vorwissen erwies. Eine Erklärung für diese Aptitude-Treatment-Interaction bietet sicherlich nicht der in den beiden Feedbackfor-

men enthaltene Informationsgehalt, da auch MTF letztendlich wie das KCR Feedback die korrekte Antwort präsentiert. Clariana (2003) offeriert hingegen zwei alternative Erklärungsansätze, einen affektiven und einen kognitiven.

Der affektive Erklärungsansatz betont die Möglichkeit, dass Lerner, die viele Fehler machen, durch die fortwährende Aufforderung „try again“ frustriert werden könnten (Dick & Latta, 1970, Clariana, 1999, beide zitiert nach Clariana, 2003). So konnten auch Morrison, Ross, Gopalakrishnan und Casey (1995) zeigen, dass mit steigender Anzahl erfolgloser Lösungsversuche optionale Hilfen seltener aufgerufen wurden. Die Autoren vermuten in diesem Verhalten eine aus den wiederholten fehlerhaften Lösungsversuchen resultierende Apathie oder Entmutigung, welche auch als erlernte Hilflosigkeit (Seligman, 1975) interpretiert werden kann. Lerner mit hohem Vorwissen hingegen würden MTF häufiger präferieren, da ihnen so die Möglichkeit gegeben werde, die richtige Lösung selbständig zu finden (Clariana & Lee, 2001). So begründeten Lerner die Bevorzugung von MTF gegenüber KCR Feedback mit Kommentaren wie folgt „I learn more when I get it on my own.“ und „Trying a second time makes me think more about the question and what it is really asking.“ (Clariana, 2003).

Der kognitive Erklärungsansatz betont die unterschiedliche Verarbeitung der im MTF und im KCR Feedback bereitgestellten, im wesentlichen identischen Informationen. Wird beim KCR Feedback die korrekte Antwort extern durch die Lernumgebung bereitgestellt, besteht beim MTF die Option, sich die korrekte Lösung durch eine eigeninitiierte Auseinandersetzung mit der präsentierten Aufgabenstellung zu erschließen. So erlaubt MTF bei Multiple-Choice Aufgaben die glaubwürdigsten Distraktoren zu eliminieren und den Fokus auf die verbleibenden Antworten zu richten. Unter der Prämisse, dass der Lerner motiviert ist, die richtige Lösung eigenständig und nicht mittels einer Versuch-und-Irrtum-Strategie zu finden, ist zu vermuten, dass MTF auf diese Weise zu einer vertieften Auseinandersetzung mit den verbleibenden Antwortalternativen anregt. MTF sollte daher vor allem Lerner mit hohem Vorwissen begünstigen, da diese über eine elaboriertere Wissensbasis verfügen, die es ihnen ermöglicht, sich nach einer fehlerhaften Antwort die korrekte Lösung durch Bewertung der verbleibenden Antwortmöglichkeiten zu erschließen.

Eine Reanalyse vorliegender Dissertationen (Clariana, 1990a; Nielson, 1990; Noonan, 1984, zitiert nach Clariana, 1993) konnte zeigen, dass eine Überlegenheit von KCR Feedback gegenüber MTF bei vorwissensarmen Lernern nicht mehr zu beobachten war, wenn diesen Lernern MTF kombiniert mit elaboriertem Feedback präsentiert wurde. Dieses Ergebnis

unterstützt die Annahme eines umgekehrt proportionalen Zusammenhangs zwischen der Höhe des Vorwissensniveaus und dem erforderlichen Ausmaß an instruktionaler Unterstützung. Lerner mit einer gering elaborierten Wissensbasis können auf diese Weise durch die Bereitstellung externer instruktionaler Unterstützung in Form zusätzlicher Hinweise bei einer eigenständigen Korrektur auftretender Fehler unterstützt werden.

2.3.2.2 Kognitive und meta-kognitive Strategien

Da die Verarbeitung von Feedbackinformationen auch den Einsatz (meta-) kognitiver Strategien erfordert, ist zu vermuten, dass auch Strategiedefizite eine adäquate Verarbeitung der präsentierten Informationen behindern. Leider liegen zu dem vermuteten Einfluss (meta-) kognitiver Strategien auf die Feedbackverarbeitung zum gegenwärtigen Zeitpunkt nur wenige Studien vor.

In einer dieser wenigen Studien konnten Marzano und Arthur (1977) zeigen, dass Lerner die von der Lehrperson an ihre schriftlichen Arbeiten annotierten Kommentare nicht lasen, bzw. wenn sie diese lasen, nicht versuchten, die Empfehlungen umzusetzen und ihre Fehler zu korrigieren. In einer ähnlichen Studie untersuchte Cohen (1987), wie Lernende mit dem durch die jeweilige Lehrperson bereitgestellten Feedback interagierten. Er fand, dass ein Fünftel der Lerner die Lehrerkommentare nur partiell las und den Korrekturen der Lehrperson nur vereinzelt ihre Aufmerksamkeit widmete. In dieser Studie schienen zudem die Lerner nur über ein begrenztes Repertoire an Strategien zu verfügen, um das Lehrerfeedback zu verarbeiten. So wurden die Annotationen des Lehrers zwar von den Lernern registriert, aber ohne dass der Text unter Nutzung der Lehrerkommentare einem Revisionsprozess unterzogen wurde.

Welche Rolle (meta-)kognitive Strategien bei der Verarbeitung präsentierter Informationen spielen, verdeutlicht wiederum die Forschung zu ausgearbeiteten Lösungsbeispielen, welche gezeigt hat, dass die Mehrzahl der Lerner nicht über angemessene (meta-)kognitive Strategien verfügt (z.B. Generierung von Selbsterklärungen, Monitoringstrategien) (Renkl, 2000). (Meta-)kognitive Strategiedefizite haben dann häufig eine oberflächliche Verarbeitung der bereitgestellten Informationen zur Folge.

In den letzten Jahren versuchte eine Reihe von Forschern eine Antwort auf die Frage zu erhalten, ob die Kontrolle über die Bereitstellung von Feedback beim Lerner verbleiben sollte oder ob das Lehrsystem entscheiden sollte, wann welcher Feedbackinhalt dem Lerner präsentiert werden soll. Eine Beantwortung dieser Frage sollte auch davon abhängen, ob Lerner über die entsprechenden (meta-)kognitiven Fähigkeiten verfügen, um angemessene Entscheidungen bezüglich des Bedarfs an Feedback treffen zu können.

Ein Plädoyer für reine Lernerkontrolle ist im Rahmen der Forschung zu intelligenten tutoriellen Systemen verschiedentlich wieder zu finden (Alevén & Koedinger, 2000). So verkünden Alevén und Koedinger (2000, S.292) ein allgemeines Einvernehmen des Forschungsfeldes zu intelligenten tutoriellen Systemen hinsichtlich des Postulats, dass „the system should let students control and organize their own learning processes; the system should intervene as little as possible [Burton and Brown, 1982] and help should be given on request only.“ Sie begründen diese Forderung damit, dass der Lerner am besten wisse, wann er Hilfe benötige, und dass das System wiederum nicht in der Lage sei, ein angemessenes Lernermodell zu erstellen, welches adäquates Feedback präsentieren könne. Daher wird in einigen intelligenten tutoriellen Systemen Feedback nur auf Lerneranforderung bereitgestellt, z.B. in Belvedere (Paolucci, Suthers & Weiner, 1996, zitiert nach Alevén & Koedinger, 2000) und in Sherlock (Katz, Lesgold, Hughes, Peters, Eggan, Gordin & Greenberg, 1998, zitiert nach Alevén & Koedinger, 2000). In einer eigenen Studie untersuchten Alevén und Koedinger (2000) das Hilfesuchverhalten der Lerner während deren Arbeit mit dem PACT Geometrie Tutor. Die erhaltenen Daten interpretierten Alevén und Koedinger (2000) dahingehend, dass ihre Untersuchungspopulation, die sich aus 15 Jahre alten Schülern zusammensetzte, nicht über angemessene (meta-)kognitive Fertigkeiten verfügte, um ihren Lernprozess eigenständig zu kontrollieren. So nutzten die Lerner das Online Glossar selten, warteten lange, trotz mehrerer Fehler, bevor sie Hilfe anforderten und richteten ihre Aufmerksamkeit vor allem auf spezifische Hinweise, unter Ignorierung allgemeinerer Hinweislevel. Korrespondierend zu Alevén und Koedinger (2000) fanden auch Suthers, Connelly, Lesgold, Paolucci, Toth, Toth und Weiner (2002), dass die Lerner on-demand Feedback nur in sehr seltenen Fällen anforderten. In der Studie von Morrison et al. (1995) riefen die Lerner nach einem Fehler nur in 15% der Fälle den Lehrtext noch einmal auf und verweilten im Durchschnitt weniger als 4 Sekunden auf der aufgerufenen Seite. Im Widerspruch zu diesen Befunden konnte Heift (2001) in ihrer Studie zeigen, dass in 80% aller fehlerhaft bearbeiteten Aufgaben die Lerner fehlerspezifische Hinweise vom System anforderten. Nur in 6% der Fälle einer fehlerhaften

Aufgabenbearbeitung ließen sich die Lerner die korrekte Lösung ohne einen nochmaligen Lösungsversuch anzeigen.

Weitere Belege, die vor allem bei jüngeren Lernern gegen eine Lernerkontrolle sprechen, stammen aus dem Forschungsbereich der „reading comprehension“. So konnten Baker (1984, zitiert nach Butler & Winne, 1995) und Wagoner (1983, zitiert nach Butler & Winne, 1995) zeigen, dass vor allem junge, wenig erfahrene Lerner Probleme beim Entdecken von Fehlern in gelesenen Texten offenbarten, was wiederum auf ineffektive Monitoringstrategien bei diesem Lernertyp zurückzuführen ist. Steinberg (1989) konnte in seiner Studie zeigen, dass Lerner, denen Feedback über ihren momentanen Wissensstand präsentiert wurde, seltener die Lehreinheit abbrachen, im Gegensatz zu Lernern, die diese Information nicht erhielten.

Diese exemplarischen Ergebnisse zeigen, dass Lerner, denen die Möglichkeit geboten wird, ihren Lernprozess selbst zu regulieren, leider häufig nicht über die dazu erforderlichen meta-kognitiven Fertigkeiten verfügen.

2.4 Gestaltungsprinzipien für ITF (Narciss, 2004; Narciss & Huth, 2004)

Ein zentrales Ergebnis des bisher dargelegten theoretischen und empirischen Erkenntnisstandes der Feedbackforschung ist, dass in zahlreichen Studien elaborierte Feedbackinhalte und -formen eher intuitiv als theoretisch begründet entwickelt wurden. So wurden in vielen Untersuchungen die im Feedback präsentierten Informationen nicht auf Basis einer kognitiven Anforderungsanalyse abgeleitet. Zudem wurden die elaborierten Informationen häufig so präsentiert, dass sich der Lerner nicht zu einer aktiven Auseinandersetzung mit den präsentierten Informationen aufgefordert sah. Weiterhin ist davon auszugehen, dass in vielen Feedbackstudien motivationale, kognitive und meta-kognitive Lernvoraussetzungen bei der Feedbackentwicklung nicht in angemessener Weise berücksichtigt wurden.

Diese Erkenntnisse dienen in einem nächsten Schritt der Ableitung konkreter Gestaltungsprinzipien für die Entwicklung informativer tutorieller Feedbackformen (Narciss, 2004; Narciss & Huth, 2004). Diese Gestaltungsprinzipien unterscheiden sich von den in der Feedbackliteratur vorliegenden Richtlinien zur Feedbackkonstruktion (Schimmel, 1988; Smith & Ragan, 1993) dahingehend, dass zum einen detailliert beschrieben wird, wie der Feedbackdesigner auf der Grundlage von Anforderungsanalysen zu relevanten Feedbackinhalten gelangt. Zum anderen werden aber auch Empfehlungen ausgesprochen, wie die als relevant identifizierten Feedbackinhalte präsentiert werden sollten. Diese Gestaltungsprinzipien werden nach-

folgend beschrieben. Dabei wird in einem ersten Schritt auf Determinanten eingegangen, welche die Wirksamkeit von ITF beeinflussen und von daher bei der Konstruktion von ITF zu berücksichtigen sind.

2.4.1 Determinanten von ITF

Auf Basis des heuristischen Modells zur Analyse und Untersuchung von Feedbackeffekten identifizierte Narciss (2004) folgende Faktoren (siehe Abbildung 2), die bei der Entwicklung von ITF-Formen zu berücksichtigen sind:

- (a) Funktion, Inhalt und Form des Feedbacks
- (b) Merkmale der Lehr-Lernsituation, insbesondere Lehrziele, Lehrinhalte und Lernaufgaben und die damit verknüpften Anforderungen
- (c) Individuelle Lernermerkmale wie z.B. Vorwissen oder meta-kognitives Wissen sowie motivationale Faktoren

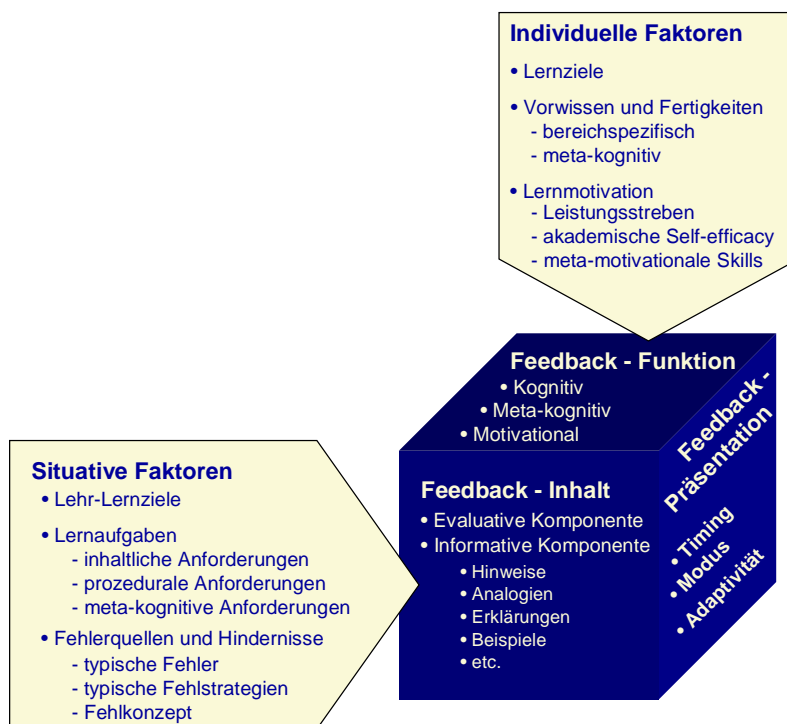


Abbildung 2: Determinanten des Informationswertes von Feedback nach Narciss (2004)

(a) Funktion, Inhalt und Form von ITF

Bereits Smith und Ragan (1993) haben auf die Schwierigkeit verwiesen, die große Anzahl vorliegender Feedbackstudien miteinander zu vergleichen. Eine wesentliche Ursache hierfür sahen sie in der Vielzahl manipulierbarer Feedbackmerkmale (z.B. Inhalt, Zeitpunkt, Funktion).

Nach Narciss (2004) wird eine Klassifikation der Feedbackformen dadurch erschwert, dass es sich bei Feedback um keine eindimensionale, sondern um eine multidimensionale Instruktionsmaßnahme handelt. Die Vielfalt an Gestaltungsmöglichkeiten kann darauf zurückgeführt werden, dass zu einer präzisen Beschreibung von Feedback mindestens drei Facetten berücksichtigt werden müssen:

1. eine funktionale Facette, die eng mit den Lehrzielen assoziiert ist. Wurde z.B. als Lehrziel definiert, dass die Lernenden nicht nur Fertigkeiten in der schriftlichen Subtraktion erwerben sollen, sondern Erfolge beim Bearbeiten solcher Aufgaben auch internal variabel, d.h. auf ihre investierte Anstrengung, attribuieren, muss Feedback nicht nur eine kognitive (im Sinne einer korrigierenden) sondern auch eine motivationale Funktion haben.
2. eine formale Facette, die beschreibt, wie das Feedback präsentiert wird. Feedback kann beispielsweise unmittelbar oder verzögert (Timing), visuell und/oder auditiv (Modalität), abgestuft oder einmalig, system- oder lernerkontrolliert, piktoral oder sprachlich (Code) präsentiert werden.
3. eine semantische oder inhaltliche Facette, welche die bereitgestellten Informationen beinhaltet. Narciss unterscheidet hierbei zwischen einer evaluativen und einer informativen Komponente, wobei die evaluative Komponente Informationen über den erreichten Leistungsstand oder über die Korrektheit der Lösung liefert (z.B. richtig/falsch, Prozentsatz korrekt gelöster Aufgaben). Die informative Komponente beinhaltet alle darüber hinausgehenden Informationen, die sich entsprechend den intendierten Funktionen auf kognitive Aspekte, meta-kognitive Strategien oder auch motivationale Aspekte beziehen können. Diese Klassifikation von Feedback in eine evaluative und eine informative Komponente haben bereits Kulhavy und Stock (1989) in einer ähnlichen Form in ihrem Feedbackmodell vorgeschlagen. Sie unterschieden zwischen einer Verifikations- und einer Elaborationskomponente, wobei die Verifikationskomponente die dichotome Information

(richtig/falsch) über die Korrektheit der Aufgabenbearbeitung enthält und die Elaborationskomponente alle über das bloße Statement der fehlerhaften Aufgabenbearbeitung hinausgehenden Informationen beinhaltet.

(b) Situative Faktoren

Merkmale des instruktionalen Kontextes wie die Lehrziele, die Lernaufgaben und die bei der Bearbeitung der Lernaufgaben auftretenden Fehler beeinflussen, welche Inhalte, welche Präsentationsform und welche Funktionen bei der Feedbackgestaltung gewählt werden sollten. Mit der Frage, welche Feedbackinhalte für welches Lehrziel bereitgestellt werden sollten, haben sich nur wenige Forscher (Schimmel, 1988; Smith & Ragan, 1993) auseinandergesetzt. Diese haben für verschiedene Lehrziele konkrete Empfehlungen für die Gestaltung von Feedback ausgesprochen. So leiteten Smith und Ragan (1993) in Anlehnung an Gagnés Lehrzielhierarchie (1985) für folgende Lehrziele konkrete Feedbackrichtlinien ab: deklaratives Wissen, intellektuelle Fertigkeiten, kognitive Strategien, Einstellungen und psychomotorische Fertigkeiten. Dabei untergliederten sie die Kategorie intellektuelle Fertigkeiten in drei weitere Subkategorien: Konzepte, Regeln und Problemlösen. Allerdings bleiben diese Empfehlungen dahingehend einem relativ groben Niveau verhaftet, dass sie zwar beschreiben, welche Feedbackinhalte bereitgestellt werden sollten, nicht aber, wie der Feedbackdesigner zu diesen Feedbackinhalten (bspw. über eine kognitive Anforderungs- und Fehleranalyse) gelangt.

Eine weitere wichtige Rolle bei der Feedbackgestaltung stellen aber auch die Lernaufgaben dar. Beispielsweise kann zwischen gut und schlecht strukturierten Lernaufgaben unterschieden werden. Feedbackinhalte für gut strukturierte Lernaufgaben (z.B. Programmierungsaufgaben, Aufgaben zu Gleichungssystemen), die auf Basis einer Lernermodellierung Aussagen zu typischen Fehlern ermöglichen, können von daher sehr konkrete, fehlerspezifische Korrekturhinweise bereitstellen. Bei schlecht strukturierten Lernaufgaben (z.B. Argumentationsaufgaben), die häufig keine eindeutige korrekte Lösung besitzen, sollte eher auf Prompts im Sinne ungerichteter Fragen zurückgegriffen werden, welche zum Einsatz (meta-)kognitiver (z.B. Monitoring-) Strategien anregen. Ziel der Feedbackentwicklung sollte sein, eine Passung zwischen den Lehrzielen/Lernaufgaben und dem Feedback herzustellen.

(c) Individuelle Faktoren

Die besondere Bedeutung individueller Lernvoraussetzungen für die Verarbeitung präsentierter Feedbackinformationen durch den Lerner wurde bereits im Kapitel 2.3.2 herausgearbeitet. Aus den dargestellten Erkenntnissen können konkrete Empfehlungen für die Feedbackgestaltung abgeleitet werden. So sollte mit zunehmendem aufgabenspezifischem Vorwissen die instruktionale Unterstützung abgebaut werden, d.h. beispielsweise, dass für einen Lerner mit einem hohen Kenntnisniveau die Angabe des Fehlerortes für eine Fehlerkorrektur hinreicht, wohingegen ein Lerner mit einem geringen Kenntnisniveau elaboriertere Informationen (Hinweis auf Lösungsstrategie) für seine Fehlerkorrektur benötigt. Auch die Ausprägung meta-kognitiver Fertigkeiten (z.B. Monitoringfähigkeit) der Lerner sollte einen Einfluss auf die Entscheidung des Feedbackdesigners haben, das Feedback lerner- oder systemkontrolliert bereit zu stellen. Je schlechter die (meta-) kognitiven Fähigkeiten des Lerners, desto weniger Lernerkontrolle sollte ermöglicht werden und desto stärkere (meta-) kognitive Unterstützung sollte das Feedback bereitstellen. Ziel der Feedbackentwicklung sollte sein, eine Passung zwischen Lernvoraussetzungen und Feedback herzustellen.

2.4.2 Auswahl des Feedbackinhalts auf der Grundlage von Anforderungs- und Fehleranalysen

Bei der Darstellung des heuristischen Feedbackmodells von Narciss (2004) wurde bereits auf die Bedeutung der inhaltlichen Feedbackqualität für die Feedbackwirksamkeit dezidiert eingegangen. Eine zentrale Frage, welche sich dem Feedbackdesigner stellen sollte, lautet daher, wie er zu seinen konkreten Feedbackinhalten unter Berücksichtigung des intendierten Lehrzielniveaus gelangt. In der Feedbackliteratur ist bezüglich dieser Frage keine Antwort zu finden. Obwohl Smith und Ragan (1993) Empfehlungen aussprechen, welche Feedbackinhalte für welches Lehrziel angeboten werden sollten, bleibt die Frage, wie der Feedbackdesigner zu diesen Inhalten gelangt, unbeantwortet. Eine Möglichkeit, relevante Feedbackinformationen theoriegeleitet zu identifizieren, bieten Anforderungs- und Fehleranalysen. Nach Jonassen, Tessmer und Hannum (1999) dienen Anforderungsanalysen unter anderem dem Ziel, die Lehrziele und die für das Erreichen dieser Lehrziele erforderlichen inhaltlichen und kognitiven Voraussetzungen zu definieren. Mit der Kenntnis der für eine erfolgreiche Aufgabenbearbeitung erforderlichen inhaltlichen und kognitiven Voraussetzungen können in einem nächsten Schritt Probleme identifiziert werden, die auftreten, wenn

bestimmte inhaltliche und kognitive Voraussetzungen nicht gegeben sind. Mit diesem Wissen können dann konkrete Feedbackinhalte abgeleitet werden. Vor diesem Hintergrund formulierte Narciss (2004) für die Auswahl relevanter Feedbackinformationen folgende Arbeitsschritte (siehe auch Abbildung 3):

1. In einem ersten Schritt sollten auf der Basis von Lehrplänen oder Curricula allgemeine Lehrziele definiert werden (z.B. Erwerb und Anwenden des schriftlichen Subtraktionsalgorithmus im Zahlenraum bis 1000). Solche Lehrziele vermitteln eine Groborientierung und legen die Merkmalsklasse jener Aufgaben fest, die nach Zielerreichung beherrscht werden sollen. Dieses übergeordnete Lehrziel sollte in einem nächsten Schritt über die Festlegung untergeordneter Ziele (Subziele) spezifiziert werden. Für den Erwerb des Subtraktionsalgorithmus könnten die untergeordneten Ziele folgendermaßen lauten: (a) Erwerb konditionalen Wissens darüber, unter welchen Bedingungen die Regel anzuwenden ist, (b) Erwerb der einzelnen Prozedurschritte, (c) Erwerb der korrekten Reihenfolge der auszuführenden Prozedurschritte, (d) Erwerb von Monitoringfertigkeiten zur Überprüfung der korrekten Prozedurausführung, (e) Erklären des Algorithmus, (f) Bewerten des Algorithmus. Eine solche Spezifikation der übergeordneten in untergeordnete Lehrziele kann auf der Basis sogenannter Lehrzieltaxonomien vorgenommen werden, wie sie von Bloom (1972), Gagné (1985) und Merrill (1983) formuliert wurden.
2. Da Feedback in der Regel im Anschluss an eine Aufgabenbearbeitung angeboten wird, sollten in einem nächsten Schritt auf Basis der über- und untergeordneten Lehrziele typische Lernaufgaben definiert werden. Solche typischen Lernaufgaben für die schriftliche Subtraktion könnten auf der Basis von Lehrbuchanalysen gewonnen werden.
3. In einem nächsten Schritt sollten unter Nutzung verschiedener Aufgabenanalysemethoden Wissens Elemente identifiziert werden, die für eine korrekte Bearbeitung der typischen Lernaufgaben notwendig sind. Hierfür liegen eine Reihe verschiedener Methoden vor, z.B. learning analysis (learning hierarchie analysis, information processing analysis, learning contingency analysis), cognitive task analysis (GOMS, PARI, DNA), content analysis (conceptual graph analysis, master design chart, matrix analysis, repertory grid technique, fault tree analysis) (nach Jonassen, Tessmer & Hannum, 1999). Für den Bereich der schriftlichen Subtraktion nutzten beispielsweise Gagné, Briggs und Wager (1992) die Methode der Lernhierarchieanalyse (learning hierarchie analysis), um auf deren Grundlage verschiedene kognitive Aufgabenanforderungen zu identifizieren.

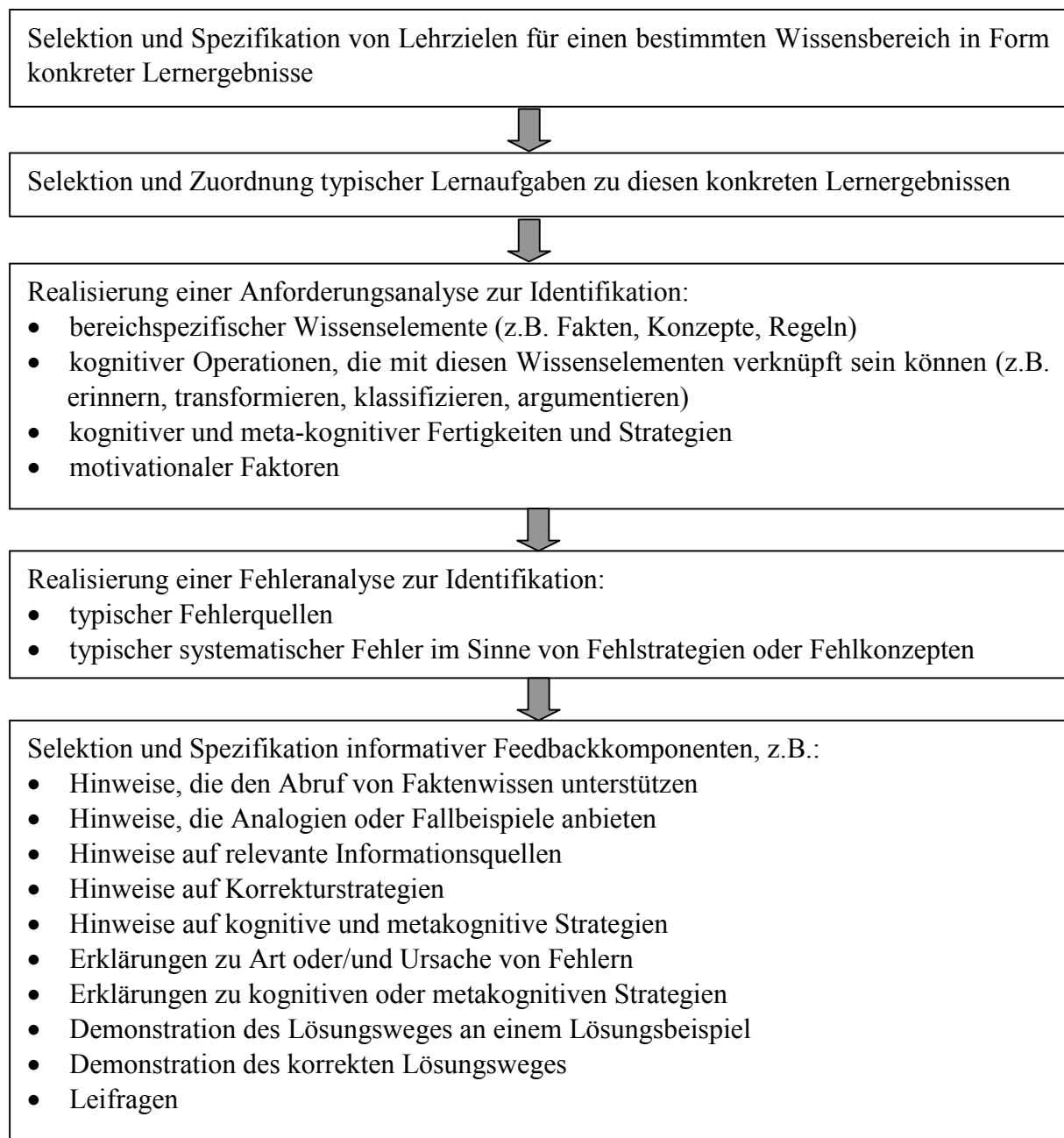


Abbildung 3: Auswahl und Spezifikation von ITF-Komponenten nach Narciss (2004; Narciss & Huth, 2004)

4. Wurden durch eine Aufgabenanalyse die für eine korrekte Aufgabenlösung erforderlichen Wissensselemente identifiziert, sollte in einem nächsten Schritt analysiert werden, welche Fehler wie häufig bei bestimmten Aufgabenmerkmalen/Anforderungen auftreten und worauf diese Fehler zurückgeführt werden können (z.B. Fehlkonzepte oder Fehlstrategien oder unsystematische Fehler). Dies geschieht in der Regel anhand von Fehleranalysen, die typische Fehler diagnostizieren. Solche typischen Fehler weisen im Sinne beobachtbarer

typischer Lösungsmuster (Performanz) darauf hin, welches Wissenselement vom Lerner noch nicht bzw. nur fehlerhaft beherrscht wird (Kompetenz).

5. Die Ergebnisse der Anforderungs- und Fehleranalyse sollten in einem nächsten Schritt genutzt werden, um für die Fehlerkorrektur relevante Informationen zu identifizieren. Solche relevanten Informationen sollten je nach Art der Lernaufgaben und der über diese Lernaufgaben definierten Anforderungen stark variieren. Relevante Informationen können beispielsweise sein: allgemeine oder spezifische Korrekturhinweise, Hinweise auf allgemeine meta-kognitive Strategien, allgemeine oder spezifische Lösungsstrategien, ausgearbeitete Lösungsbeispiele, Musterlösungen, Simulationen und gerichtete Fragen im Sinne des sokratischen Dialogs.

2.4.3 Prinzipien für die formale Gestaltung von ITF

Bereits im Kapitel 2.3.1 wurde darauf hingewiesen, dass die Auswahl relevanter Feedbackinformationen auf Basis kognitiver Anforderungs- und Lehrzielanalysen eine notwendige, aber keine hinreichende Bedingung darstellt, um die Wirksamkeit von informativem tutoriellem Feedback zu gewährleisten. Neben den bereits in Kapitel 2.3.2 beschriebenen kognitiven und meta-kognitiven Voraussetzungen des Lernenden beeinflusst auch die Form der Feedbackpräsentation, wie die bereitgestellten Feedbackinformationen verarbeitet werden. So konnte in Kapitel 2.3.1 gezeigt werden, dass die durch die Instruktionsumgebung bereitgestellten elaborierten Informationen vom Lerner häufig nicht vertiefend verarbeitet werden, so dass eine Lernwirksamkeit informativer Feedbackformen gegenüber weniger informativen Feedbackformen nicht nachgewiesen werden konnte. Um eine vertiefte Auseinandersetzung mit den präsentierten Feedbackinformationen anzuregen, bieten sich verschiedene Möglichkeiten, die an dieser Stelle ausführlich vorgestellt werden, bevor aufbauend auf diesen theoretisch und empirisch begründeten Überlegungen konkrete Prinzipien für die formale Gestaltung von ITF abgeleitet werden.

Eine Möglichkeit, eine aktive Auseinandersetzung zu initiieren, besteht darin, den Lerner im Anschluss an eine Feedbackpräsentation dazu aufzufordern, die in diesem Feedback enthaltenen Informationen direkt anzuwenden, sei dies an der gleichen oder an einer ähnlichen Aufgabe. Eine höhere Lernwirksamkeit einer wiederholten Aufgabenbearbeitung kann einerseits über eine zusätzliche Übungsphase erklärt werden. Eine alternative Erklärung wäre aber auch, dass das Wissen um eine nochmalige Aufgabenbearbeitung den Lerner zu einer

aktiveren Auseinandersetzung mit den präsentierten Feedbackinformationen anspricht, da eine erneute Aufgabenbearbeitung vom Lerner als eine Testphase, die eine Anwendung des erworbenen Wissens erfordert, interpretiert wird.

Diese nochmalige Aufgabenbearbeitung kann beispielsweise realisiert werden über das Angebot elaborierter Feedbackinformationen im Rahmen einer wiederholten Aufgabenbearbeitung (MTF), gekoppelt mit einem KCR Feedback, welches erst im Anschluss an den letzten fehlerhaften Lösungsversuch präsentiert wird. Leider liegen zu dieser Form der Feedbackpräsentation im Bereich der klassischen Feedbackforschung nur wenige Untersuchungen vor (Nielson, 1990; Noonan, 1984, zitiert nach Mory, 1996). In der Studie von Noonan (1990, zitiert nach Mory, 1996) konnten keine signifikanten Unterschiede zwischen der Feedbackbedingung „KR Feedback mit Erklärung und einem zweiten Lösungsversuch“ sowie der Feedbackbedingung „KCR Feedback“ und der Feedbackbedingung „KCR Feedback und ein zweiter Versuch“ gefunden werden.

Eindrucksvolle Belege für die Wirksamkeit solch elaborierten MTF kann hingegen aus der Forschung zu menschlichen Tutoren (Merrill, Reiser, Merrill & Landes, 1995) und zu intelligenten tutoriellen Systemen (ITS) (Albacete & VanLehn, 2000; Gertner, Conati & VanLehn, 1998) gewonnen werden. Menschliche Tutoren zeichnen sich dadurch aus, dass „tutors have the opportunity to pursue a given topic or problem until the students have mastered it.“ (Chi, Siler, Jeong, Yamauchi & Hausmann, 2001, S. 472). Ein solch typischer tutorieller Dialog kann nach Graesser, Person und Magliano (1995) über folgende fünf Schritte beschrieben werden: (1) der Tutor stellt eine Frage, (2) der Lerner gibt eine Antwort, (3) der Tutor gibt KR Feedback, (4) der Tutor unterstützt den Lerner bei seiner eigenständigen Lösungsfindung im wechselseitigen Dialog (scaffolding), (5) der Tutor bewertet des Lerners Verständnis. Das zentrale Merkmal eines solchen tutoriellen Dialogs kann darin gesehen werden, dass dem Lerner die richtige Antwort nicht mitgeteilt wird, sondern dass der Lerner in seinem Aufgabenlösungsprozess durch den Tutor darin unterstützt wird, die richtige Lösung eigenständig zu finden. Zeigt der Lerner im Verlauf dieses Dialogs Probleme beim Bearbeiten einer Aufgabe, wird ihm von Seiten des Tutors entsprechend den Bedürfnissen des Lernenden adaptiv sequentielle Hilfe angeboten. Dies bedeutet, dass der Tutor progressiv immer detailliertere und explizitere Hilfe gibt, bis der Lerner in der Lage ist, das Problem selbständig zu lösen (Wood & Wood, 1999).

Eine weitere Möglichkeit, welche auf dem Konzept einer wiederholten Aufgabenbearbeitung basiert, bietet der Ansatz des zielerreichenden Lernens nach Bloom („mastery learning“, 1968). Dieser Ansatz postuliert, dass ein Fortschreiten im Lehrstoff bzw. das Bearbeiten der nächsten Lehreinheit erst ermöglicht werden sollte, wenn der Lerner ein vorher festgelegtes Lernkriterium erreicht hat. Im Gegensatz zum tutoriellen Feedback, welches unter Bereitstellung elaborierter Feedbackinformationen einen erneuten Lösungsversuch der fehlerhaft gelösten Aufgabe ohne vorherige Bekanntgabe der richtigen Lösung (ohne KCR Feedback) ermöglicht, beinhaltet der mastery learning Ansatz eine erneute Bearbeitung der gleichen fehlerhaft gelösten Aufgabe bzw. von Aufgaben des gleichen Aufgabentyps im Anschluss an die Mitteilung der korrekten Lösung (KCR Feedback), bis ein durch die Instruktionsumgebung im Vorfeld definiertes Lehrziel erreicht wird.

Empirische Evidenz für die Lernförderlichkeit zielerreichenden Lernens (mastery learning) konnte in verschiedenen Feedbackstudien geliefert werden (Collins, Carnine & Gersten, 1987; Kline, Schumaker & Deshler, 1991). In diesen Studien, in denen eine erfolgreiche Bewältigung der Lehreinheit über das Erreichen eines Lernkriteriums definiert war, konnte eine statistisch signifikante Überlegenheit elaborierter Feedbackformen gegenüber einfacheren Feedbackformen nachgewiesen werden. Möglicherweise realisierten die Lerner, dass das Ende der Lehreinheit nur zu erreichen war, wenn die Aufgaben eigenständig korrekt bearbeitet wurden. Dies wiederum motivierte sie, sich intensiver mit den elaborierten Hinweisen auseinander zu setzen (Jacobs, 2002). In vielen Studien zum zielerreichenden Lernen konnte zudem gezeigt werden, dass sich die Implementierung eines Lernkriteriums auch dahingehend motivierend auf den Lerner auswirkte, dass sie ein höheres Motivationslevel und eine positivere Fähigkeitswahrnehmung zeigten (Mory, 1996).

Auf der Grundlage dieser theoretisch und empirisch fundierten Überlegungen sowie vor dem Hintergrund bereits dargestellter weiterer Feedbackkenntnisse (z.B. presearch availability) entwickelten Narciss und Huth (2004) folgende fünf Gestaltungsprinzipien für informatives tutorielles Feedback:

1. Feedback, vor allem KCR Feedback, sollte erst im Anschluss an eine Aufgabenbearbeitung präsentiert werden. So konnten Anderson, Kulhavy und Andre (1971) in ihrer Studie zeigen, dass Lerner die korrekte Antwort einfach kopierten, wurde diese gemeinsam mit der Aufgabenstellung angeboten (presearch availability). Eine vertiefte Auseinanderset-

zung mit der präsentierten Feedbackinformation dürfte nicht stattgefunden haben, was wiederum in einem geringeren Lernzuwachs resultierte.

2. Die identifizierten elaborierten Feedbackinformationen sollten nicht gemeinsam mit der korrekten Antwort (KCR) dargeboten werden. Auch hier besteht die Gefahr, dass infolge des gleichzeitigen Angebotes der richtigen Antwort eine intensive Auseinandersetzung mit den zusätzlich präsentierten elaborierten Feedbackinformationen nicht stattfindet (Kulhavy et al., 1985).
3. Die identifizierten Feedbackinformationen sollten schrittweise, mit zunehmendem Informationswert, dargeboten werden, gekoppelt mit einem erneuten Lösungsversuch, der eine Anwendung der präsentierten Feedbackinformationen ermöglicht. Dieses adaptive Vorgehen sollte zum einen zu einer vertieften Auseinandersetzung mit den präsentierten Informationen führen, zum zweiten eine kognitive Überlastung des Lernenden vermeiden und zum dritten nur so viel Informationen bereitstellen, wie der Lernende für eine korrekte Aufgabenlösung benötigt. Einen empirischen Beleg für ein solch adaptives Vorgehen liefert die Studie von Day und Cordon (1993), die zeigen konnten, dass Feedback, welches sich an das Kenntnisniveau des Lerners anpasst, effizienter ist als ausführliches Feedback, welches unabhängig vom Kenntnisstand des Lerners bei jedem Fehler präsentiert wird.
4. Mastery learning sollte über die Definition eines Lernkriteriums (mastery level) gewährleistet sein, um zum einen eine aktive Auseinandersetzung mit den präsentierten Feedbackinformationen anzuregen. So sollte das Wissen, dass das Ende der Lehreinheit nur zu erreichen ist, wenn die Aufgaben eigenständig korrekt bearbeitet werden, die Lernenden dahingehend motivieren, sich intensiver mit den elaborierten Hinweisen auseinander zu setzen (Jacobs, 2002). Zum anderen sollte die Definition eines Lernkriteriums eine kognitive Überlastung und mögliche Interferenzeffekte der präsentierten Feedbackinformationen vermeiden, die sich negativ auf die Feedbackwirksamkeit auswirken könnten. Das heißt, erst wenn davon ausgegangen werden kann, dass eine bestimmte Teilfertigkeit stabil erworben wurde, sollte ein neuer Aufgabentyp präsentiert werden.
5. Werden die identifizierten Feedbackinformationen computergestützt präsentiert, sollte eine bewusste Entscheidung im Hinblick auf die zu wählende Modalität und Codalität getroffen werden, in welcher die Informationen präsentiert werden. In aktuelleren Arbeiten wird für eine bimodale Präsentation verbaler und piktoraler Informationen plädiert

(multimedia aids, Mayer & Moreno, 2002). So konnte in zahlreichen Studien, in denen verbale und piktorale Informationen rein visuell bzw. visuell und auditiv dargeboten wurden, eine Überlegenheit der bimodalen Kombination auditiv-verbaler und visuell-piktoraler Informationen gegenüber der rein monomodalen, visuellen Darbietung nachgewiesen werden (Mayer & Anderson, 1991, 1992; Mayer & Moreno, 1998). Eine aktuelle Theorie, welche die Überlegenheit multimodaler Informationsdarbietung zu erklären versucht, repräsentiert die Multimedia Theorie von Mayer und Moreno (2002). Entsprechend dieser Theorie und der cognitive-load Theorie (Sweller, Van Merriënboer & Paas, 1998) reduziert die kombinierte Präsentation auditiv-verbaler und visuell-piktoraler Informationen die kognitive Belastung des Lerners aufgrund der Nutzung des sowohl verbalen Informationsverarbeitungssystems (artikulatorische Schleife, Baddeley, 1986) als auch des visuellen Informationsverarbeitungssystems (visuell-räumlicher Notizblock, Baddeley, 1986) und verbessert darüber die Lernleistung (Modalitätseffekt). Werden verbale und piktorale Informationen simultan visuell präsentiert, müssen beide Informationsarten im gleichen System verarbeitet werden, was eine Aufteilung der verfügbaren Aufmerksamkeitsressourcen erfordert. So konnte in verschiedenen Arbeiten (Mayer & Anderson, 1991, 1992; Mayer & Moreno, 1998; Mousavi, Low & Sweller, 1995) bei Lernenden ein stärkerer Lernfortschritt beobachtet werden, wenn ihnen die verbalen Informationen auditiv gemeinsam mit den piktoralen Informationen (z.B. Animation) präsentiert wurden, als wenn sie die verbalen Informationen gemeinsam mit den piktoralen Informationen am Bildschirm präsentiert bekamen („split-attention effekt“). Zum anderen postulieren Mayer und Moreno (2002), dass über das Angebot auditiv-verbaler und visuell-piktoraler Informationen der Lernende zu einer aktiven und vertieften Auseinandersetzung angeregt wird, da für ein vertieftes Verständnis der präsentierten Informationen referentielle Verbindungen zwischen korrespondierenden verbalen und visuellen Repräsentationen aufgebaut werden müssen. Werden auditiv-verbale und visuell-piktoraler Informationen präsentiert, so sollten diese eher simultan als sukzessiv angeboten werden (contiguity aids, Mayer & Moreno, 2002). Werden nur verbale Informationen dem Lerner präsentiert, sollten diese nach Mayer und Moreno (2002) eher auditiv als am Bildschirm angeboten werden (modality aids), obwohl diese Art der rein auditiven Präsentation von Informationen auch klare Nachteile aufweist, wie dessen Flüchtigkeit und die Nichtbeeinflussbarkeit des Informationsdarbietungstempos.

2.5 Evaluationsprinzipien für ITF (Narciss, 2004)

Wurde ITF unter Verwendung der vorgestellten Gestaltungsprinzipien entwickelt, sollte in einem nächsten Schritt eine Überprüfung der Wirksamkeit des entwickelten ITFs auf der Grundlage theoretisch und empirisch begründeter Annahmen zu Funktionen und Wirkungsweisen von Feedback erfolgen. Narciss (2004) fasste die in der Feedbackliteratur diskutierten Erklärungsansätze zur Wirkungsweise von Feedback in drei Hauptfunktionen zusammen – der kognitiven, der meta-kognitiven und der motivationalen Funktion. Jede dieser drei Funktionen differenzierte sie in weitere Unterfunktionen auf, welche in Tabelle 2 zusammengefasst dargestellt sind. Diese Unterfunktionen wurden primär für informatives tutorielles Feedback formuliert, welches aufgrund seiner tutoriellen Komponente im Vergleich zu klassischen Feedbackformen zusätzliche Funktionen beinhaltet. So sollte ITF, welches korrekturrelevante Informationen bereitstellt ohne unmittelbar die richtige Lösung zu präsentieren, dahingehend motivierend wirken, dass Erfolge auf die eigene Anstrengung zurückgeführt, also intern attribuiert werden und somit zu einem erhöhten Kompetenzerleben führen.

Die in Tabelle 2 dargestellten Funktionen von ITF illustrieren, dass es sich bei informativem tutoriellem Feedback nicht um eine unifunktionale sondern um eine multifunktionale Instruktionsmaßnahme handelt, welche nicht nur Einfluss auf kognitive, sondern auch auf meta-kognitive und motivationale Faktoren des Lernprozesses hat. Diese Multifunktionalität gilt es besonders hervorzuheben, da in der bisherigen Feedbackforschung primär der Einfluss von informativem Feedback auf Leistungs- und Lernparameter untersucht wurde. So beinhalteten die in Kapitel 2.2 vorgestellten Metaanalysen nur Studien, welche die Lernwirksamkeit von Feedback untersucht haben. Feedbackstudien, welche die motivationale Wirksamkeit von informativem Feedback zum Untersuchungsgegenstand erklärt haben, liegen hingegen in einer sehr geringen Anzahl vor. Erste empirische Belege für die Annahme einer motivationsförderlichen Wirkung von informativem Feedback konnte Narciss in ihren Arbeiten (1999; 2001) liefern. Sie fand, dass Lernende, welche informatives Feedback angeboten bekamen, sowohl eine höhere Ausdauer bei der Aufgabenbearbeitung zeigten, als auch zufriedener mit ihrer Leistung waren.

Tabelle 2:

Kognitive, meta-kognitive und motivationale Funktionen von ITF nach Narciss (2004)

Kognitive Funktionen
<ul style="list-style-type: none"> • informierende Funktion, wenn Anzahl, Ort, Art der Fehler und/oder Fehlerursachen unbekannt sind • ergänzende Funktion, wenn der Fehler aufgrund <i>fehlender</i> Wissensselemente zustande kam und das Feedback Informationen zu den fehlenden Elementen anbietet • korrigierende Funktion, wenn der Fehler aufgrund <i>fehlerhafter</i> Wissensselemente zustande kam und das Feedback Informationen anbietet, die zur Korrektur genutzt werden können • diskriminierende Funktion, wenn der Fehler aufgrund <i>ungenauer</i> Wissensselemente zustande kam und das Feedback Informationen anbietet, die zur Präzisierung genutzt werden können • restrukturierende Funktion, wenn der Fehler aufgrund <i>fehlerhafter</i> Verknüpfungen der Wissensselemente zustande kam und das Feedback Informationen anbietet, die zur Umstrukturierung genutzt werden können
Meta-kognitive Funktionen
<ul style="list-style-type: none"> • informierende Funktion, wenn meta-kognitive Strategien oder Bedingungen für den Einsatz meta-kognitiver Strategien unbekannt sind und das Feedback Informationen zu meta-kognitiven Strategien oder Bedingungen anbietet • ergänzende Funktion, wenn durch das Feedback z.B. zielrelevante Kriterien für das monitoring angeboten werden oder Bedingungen für den Einsatz meta-kognitiver Strategien expliziert werden • korrigierende Funktion, wenn Fehler beim Einsatz meta-kognitiver Strategien aufgetaucht sind und das Feedback Informationen anbietet, die zur Korrektur genutzt werden können • lenkende Funktion, wenn Lernende z.B. durch Feedback-Leitfragen dazu animiert werden, selbst zielrelevante Kriterien für das monitoring oder die Evaluation zu generieren, oder die Angemessenheit ihrer Lösungsstrategien zu reflektieren
Motivationale Funktionen
<ul style="list-style-type: none"> • Feedback aktiviert das mit dem subjektiven Kompetenzerleben verknüpfte Anreizfeld, indem es Ergebnisse der Aufgabenbearbeitung sichtbar macht • Feedback reduziert die (subjektive) Aufgabenschwierigkeit, wenn es Informationen anbietet, die zur Bewältigung der Aufgabenanforderungen genutzt werden können • Feedback erhöht die subjektive Erfolgswahrscheinlichkeit, wenn es tutorielle Komponenten, aber nicht die Lösung anbietet • Feedback ermöglicht es, Erfolge der eigenen Anstrengung zuzuschreiben, also intern zu attribuieren, wenn es tutorielle Komponenten aber nicht die Lösung anbietet • Informatives tutorielles Feedback erhöht damit die Chance, positiven Kompetenzzuwachs zu erleben

Welche Wirkungen letztendlich wann und über welchen Zeitraum beim Lernen mit informativem Feedback auftreten, hängt neben der inhaltlichen und formalen Qualität von Feedback auch von den individuellen Lernvoraussetzungen und situativen Bedingungen ab (siehe Kapitel 2.4). Neben kognitiven Lernereigenschaften wie das aufgabenspezifische Vorwissen beeinflussen auch motivationale Lernereigenschaften (z.B. self efficacy, Hancock, Thurman & Hubbard, 1995) die Feedbackverarbeitung und darüber dessen Wirksamkeit. Aus diesem Grund sollten bei der Durchführung von Feedbackstudien sowohl kognitive als auch

motivationale Lernereigenschaften erfasst werden. Eine zusammenfassende Darstellung zu den Bedingungen und Wirkungen von Feedback in Lehr-Lernsituationen findet sich in Abbildung 4. Diese Darstellung verdeutlicht noch einmal, dass zum einen individuelle und situative Faktoren bei der Entwicklung von ITF berücksichtigt werden sollten, zum anderen diese Faktoren aber auch die Verarbeitung des präsentierten Feedbacks und darüber dessen Wirksamkeit beeinflussen.

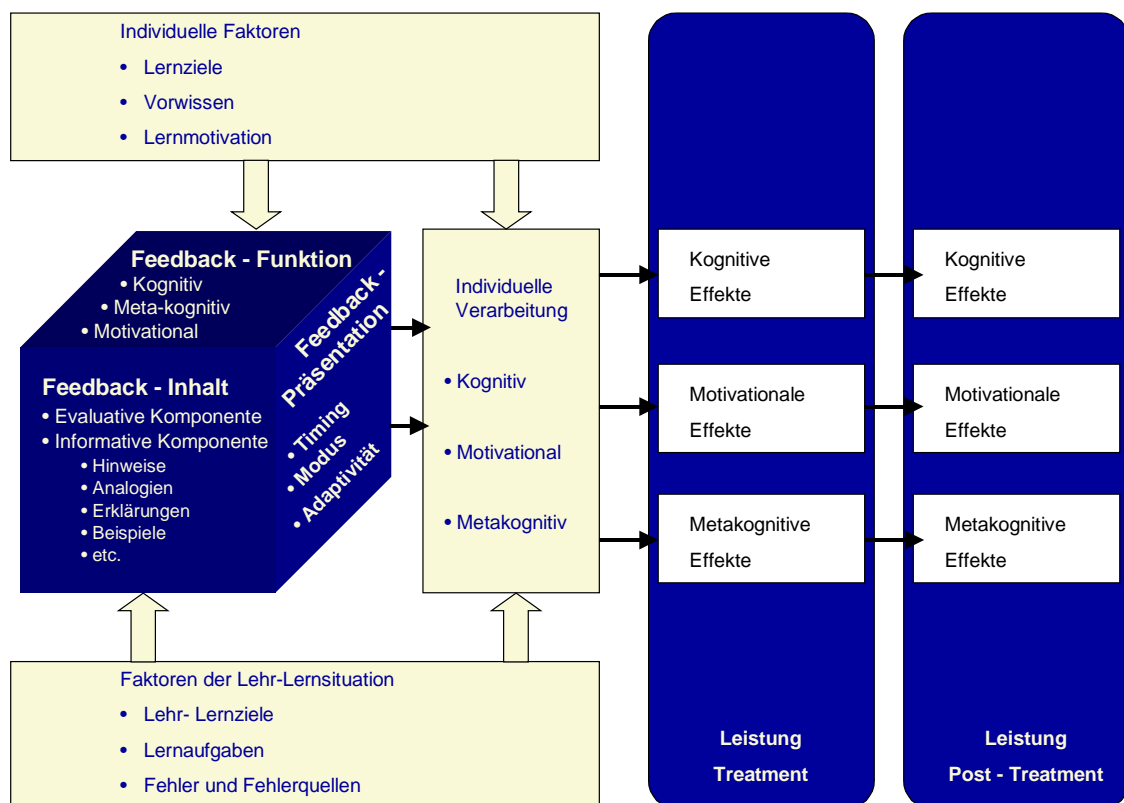


Abbildung 4: Bedingungen und Wirkungen von Feedback in Lehr-Lernsituationen (nach Narciss, 2004)

Unter der Annahme, dass ITF multifunktional wirkt, definierte Narciss (2004) in einem nächsten Schritt Indikatoren, mit denen die Wirksamkeit von ITF auf kognitive, metakognitive und motivationale Faktoren erfasst werden kann. Dabei unterscheidet sie zum einen zwischen beobachtbaren Lern- und Motivationsparametern und zum anderen zwischen berichtbaren Indikatoren von Lern- und Motivationsaspekten wie z.B. Einschätzungen bezüglich der Selbstbewertung der Leistung. Die von Narciss (2004) abgeleiteten Indikatoren für die Untersuchung von kognitiven, meta-kognitiven und motivationalen Feedbackeffekten sind in Tabelle 3 dargestellt.

Tabelle 3:

Ableitung von Indikatoren für die Untersuchung von kognitiven, meta-kognitiven und motivationalen Feedbackeffekten nach Narciss (2004)

Feedback-Funktion	Erwartete Wirkungen	Indikatoren für die erwarteten Wirkungen	
		Berichtbar, d.h. über Fragen erfassbar	Beobachtbar
Kognitiv			
<ul style="list-style-type: none"> • Informieren • Ergänzen • Korrigieren • Diskriminieren • Restrukturieren 	<ul style="list-style-type: none"> • Fehler werden erkannt • Fehlendes Wissen wird ergänzt • Falsches Wissen wird korrigiert • Falsch verknüpfte Elemente werden neu zugeordnet 	<ul style="list-style-type: none"> • Antwortsicherheit • Einschätzung bzgl. möglicher Fehler • Einschätzung bzgl. der Nützlichkeit der angebotenen Feedback-informationen 	<ul style="list-style-type: none"> • Anzahl korrigierter Fehler • Anzahl korrekt/falsch gelöster Aufgaben • Anzahl von Aufgaben, bei denen das Lernkriterium erreicht wurde • Anzahl von Aufgaben, die benötigt wurden, um das Lernkriterium zu erreichen • Abruf von Informationen • Spezifische Lösungsschritte • Spezifische Strategien
Meta-kognitiv			
<ul style="list-style-type: none"> • Informieren • Ergänzen • Korrigieren • Lenken 	<ul style="list-style-type: none"> • Falsche Strategien werden erkannt • Fehlende Strategien werden ergänzt • Falsche Strategien werden korrigiert • Strategien werden mehr beachtet 		
Motivational			
<ul style="list-style-type: none"> • Anreiz erhöhen • Schwierigkeit reduzieren • Erfolgswahrscheinlichkeit erhöhen • Anstrengung und Erfolg verknüpfen • Chancen auf Erleben von Kompetenzzuwachs erhöhen 	<ul style="list-style-type: none"> • Attraktivität von Aufgaben nimmt zu • Bereitschaft Aufgaben zu bearbeiten, nimmt zu • Erfolgszuversicht nimmt zu • Anstrengungsbereitschaft nimmt zu • Ausdauerbereitschaft nimmt zu • Selbstbewertung wird positiver 	<ul style="list-style-type: none"> • Anreizwerte von Aufgaben • Bereitschaft, künftig Aufgaben zu bearbeiten • Subjektive Aufgabenschwierigkeit • Erwartungen (z.B. Self-Efficacy) • Kausalattributionen • Subjektive Leistungseinschätzung • Zufriedenheit mit der Leistung 	<ul style="list-style-type: none"> • Anzahl bearbeiteter Aufgaben • Investierte Zeit für Aufgaben • Investierte Zeit für Aufgaben- bzw. Feedbackbearbeitung • Anzahl von Aufgaben, bei denen die Lernenden aufgegeben haben

2.6 Zusammenfassung

Empirische Erkenntnisse zu Feedback in Lehr-Lernsituationen sind durch eine hohe Inkonsistenz gekennzeichnet. Gründe hierfür sind zum einen in der hohen methodischen Variabilität der Feedbackstudien zu suchen. Zum anderen können Ursachen für die Uneinheitlichkeit der Feedbackbefunde aber auch darin gesehen werden, dass in vielen Studien Feedback nicht theoriegeleitet entwickelt wurde. Eine Analyse der Feedbackliteratur zeigt aber auch, dass bisher kein Feedbackmodell existierte, welches versucht hat, die verschiedenen Erkenntnisse der Feedbackforschung integrierend darzustellen, um daraus Empfehlungen für die zukünftige Untersuchung von Feedbackeffekten abzuleiten. Ein weiteres Problem, welches sich vor allem potenziellen Feedbackdesignern bisher stellte, bestand darin, dass keine Gestaltungsvorschläge vorlagen, die detaillierte Handlungsanweisungen beinhalteten, wie Feedback konstruiert werden sollte. Mit dem heuristischen Feedbackmodell zur Analyse und Untersuchung von Feedbackeffekten von Narciss (2004) liegt nun ein Feedbackmodell vor, welches die Bedeutung individueller Faktoren als auch die Bedeutung der inhaltlichen und formalen Feedbackqualität für die Feedbackwirksamkeit herausgearbeitet hat. An diesem Feedbackmodell lässt sich zum einen sehr gut die Komplexität der die Feedbackeffektivität beeinflussenden Faktoren aufzeigen, zum anderen ist es mit diesem Feedbackmodell möglich, verschiedene Feedbackbefunde zu erklären. Dieses Feedbackmodell und die an diesem Modell aufgezeigten Feedbackerkennnisse dienen zudem als Grundlage zur Ableitung konkreter Gestaltungs- und Evaluationsprinzipien (Narciss, 2004; Narciss & Huth, 2004), welche vor allem zukünftigen Feedbackdesignern Unterstützung und Anleitung bei der Entwicklung und Evaluation von qualitativ hochwertigen Feedback bieten sollen.

Mit der Formulierung konkreter Gestaltungsprinzipien war es nun möglich, fehlerspezifisches ITF für den Gegenstandsbereich der schriftlichen Subtraktion zu entwickeln. Die Entwicklung des fehlerspezifischen ITFs für die schriftliche Subtraktion auf der Grundlage der formulierten Gestaltungsprinzipien ist Thema des nächsten Kapitels.

3 ENTWICKLUNG VON FEHLERSPEZIFISCHEM ITF FÜR DIE SCHRIFTLICHE SUBTRAKTION

Die Entscheidung, fehlerspezifisches ITF für die schriftliche Subtraktion zu entwickeln, basierte auf der Überlegung, dass dieses Wissensgebiet zum einen theoretisch sehr gut untersucht ist, im Hinblick auf seine praktische Bedeutung aber auch brachliegendes Wissen repräsentiert, welches bisher keinen Einzug in die Gestaltung von Feedback für computergestützte Lern- und Übungsprogramme gefunden hat. Feedback in mathematischer Grundschulsoftware gestaltet sich in aller Regel so, dass nach einem fehlerhaften Ergebnis der Schüler die Rückmeldung erhält, dass seine Lösung falsch ist und ein neuer Lösungsversuch bereitgestellt wird. Nach einer wiederholten fehlerhaften Aufgabenbearbeitung wird dem Schüler die richtige Lösung präsentiert. Unterstützung von Seiten des Programms erfolgt im allgemeinen über die Bereitstellung eines ausgearbeiteten Lösungsbeispiels, welches sich der Lerner über einen Programmbutton aufrufen kann. Der Schüler erfährt auf diese Weise weder etwas über die von ihm nicht beherrschte Teilprozedur, noch erhält er Informationen zu dem Ort und der Art seines Fehlers.

Ziel dieses Kapitels ist es, auf der Grundlage der vorgestellten Gestaltungsprinzipien die Entwicklung des fehlerspezifischen ITFs für die schriftliche Subtraktion detailliert zu beschreiben. In Kapitel 3.1 werden Theorien zum Erwerb korrekten und fehlerhaften Wissens in der schriftlichen Subtraktion vorgestellt. Aufbauend auf diesen Theorien werden im Kapitel 3.2 typische Anforderungen und systematische Fehler bei der schriftlichen Subtraktion dargestellt. Dieses Wissen wird in Kapitel 3.3 genutzt, um fehlerspezifisches ITF für die schriftliche Subtraktion basierend auf den Designprinzipien für ITF zu entwickeln.

3.1 Theorien zum Erwerb korrekten und fehlerhaften Wissens in der schriftlichen Subtraktion

Die schriftliche Subtraktion kann dank ihres algorithmischen Charakters und den Problemen, die das Lösen schriftlicher Subtraktionsaufgaben vielen Schülern bereitet, auf eine lange Forschungstradition im Bereich computergestützter Wissensdiagnostik zurückblicken. So wurden vor allem in den 70/80er Jahren verschiedene computergestützte Systeme für die Rekonstruktion und Diagnose fehlerhafter Aufgabenlösungen in der schriftlichen Subtraktion

entwickelt (BUGGY, DEBUGGY, IDEBUGGY, Brown & Burton, 1978; Burton, 1982; ACM, Langley & Ohlsson, 1984). Grundlage hierfür war die Beobachtung, dass sich die manifesten fehlerhaften Antwortmuster der Schüler durch eine gewisse Systematik auszeichneten, die darauf schließen ließen, dass die auftretenden Fehler nicht zufällig zustande kamen im Sinne von „slips“ (Norman, 1981) oder „unsystematischen Fehlern“, sondern auf ein regelhaftes, systematisches Vorgehen zurückgeführt werden konnten (Brown & VanLehn, 1980; Cox, 1975; Langley & Ohlsson, 1984; VanLehn, 1986, 1990; Young & O’Shea, 1981). Fehler, die aus der mehr oder weniger konsistenten Anwendung einer fehlerhaften Prozedur („bug“) resultieren, werden als „systematische Fehler“ (VanLehn, 1990) oder als „rational errors“ (Ben-Zeev, 1995) bezeichnet. Diese Erkenntnis um die Systematik arithmetischer Fehler stellte zugleich die Voraussetzung dar für eine computerbasierte Wissensdiagnostik.

Brown und Burton (1978) schlugen mit ihrem sogenannten prozeduralen Netzwerk eine Form der Wissensrepräsentation vor, welche neben der Abbildung der richtigen Prozedur auch die Abbildung fehlerhafter oder unvollständiger Routinen erlaubte. Eine Diagnose war nun nicht mehr nur über die quantitative Feststellung, wie viele Aufgaben ein Schüler richtig gelöst hatte, möglich. Darüber hinaus war es möglich geworden, ein detailliertes Bild vom Wissen des Schülers zu erhalten, d.h. Aussagen darüber, welche Prozeduren der Schüler in der richtigen Reihenfolge beherrschte und welche noch nicht.

Diese Erkenntnis um die Systematik von Aufgabenlösungen in der schriftlichen Subtraktion stellte die Grundlage für die Entwicklung fehlerspezifischer ITF Formen für diesen Gegenstandsbereich dar. Da aus dem Wissen über die Entstehung systematischer Fehler Konsequenzen für die inhaltliche Gestaltung informativer tutorieller Feedbackformen resultieren, wird nachfolgend die von VanLehn (1986) entwickelte Theorie des induktiven Lernens, die den Aufbau von Rechenalgorithmen anhand ausgewählter und sequenzierter Beispielaufgaben und ihrer Lösungen beschreibt, vorgestellt. In einem zweiten Schritt wird die repair theory von Brown und VanLehn (1980) beschrieben, welche das Auftreten fehlerhaften algorithmischen Wissens und damit die Entstehung falscher Rechenprozeduren beschreibt.

3.1.1 Die induction hypothesis

Zentrale Annahme der Theorie des induktiven Lernens („induction hypothesis“, VanLehn, 1986) ist, dass arithmetisches Wissen wie das schriftliche Subtrahieren über induktive Prozesse erworben wird. Das heißt, ein Schüler erwirbt sein arithmetisches Wissen aus ausge-

arbeiteten Lösungsbeispielen, die im Unterricht dargeboten werden, und aus dem auf seine Lösungsversuche gegebenen Feedback. Instruktionale Unterrichtssequenzen sind dabei üblicherweise so aufgebaut, dass die zu erwerbenden Prozeduren sequentiell dargeboten werden, d.h. zuerst werden Subtraktionsaufgaben ohne Übertrag behandelt, bevor Subtraktionsaufgaben mit Übertrag eingeführt werden. Prototypisch wird der Lehrer die Aufgabe zuerst an der Tafel gemeinsam mit den Schülern lösen, bevor die Schüler aufgefordert sind, ähnliche Probleme selbständig zu bearbeiten. In der Forschung zu ausgearbeiteten Lösungsbeispielen („worked-out examples“) konnte zum einen die Lerneffektivität solcher Lösungsbeispiele empirisch belegt werden (Paas & van Merriënboer, 1994; Sweller & Cooper, 1985), zum anderen konnte auch die Präferenz der Lernenden für ausgearbeitete Lösungsbeispiele im Gegensatz zu anderen Instruktionsformen überzeugend demonstriert werden (LeFevre & Dixon, 1986).

Sehen sich Schüler nun mit einem neuen Problem konfrontiert, für welches noch kein adäquater Lösungsalgorithmus existiert, versuchen sie dieses neue Problem unter Anwendung bereits bekannter Prozeduren, die in vorhergehenden instruktionalen Einheiten erworben wurden, zu lösen. Fehlerhafte Problemlösungen resultieren nun aus einer Übergeneralisierung oder Überdiskriminierung solcher bereits erworbener Prozeduren. VanLehn (1990) beschrieb diese Entstehung fehlerhafter Aufgabenlösungen auf der Basis induktiven Schließens mit seiner repair theory, welche im nächsten Kapitel dezidiert vorgestellt wird.

3.1.2 Die repair theory

Die repair theory (Brown & VanLehn, 1980; VanLehn, 1990) postuliert, dass ein Schüler über ein bestimmtes prozedurales Wissen verfügt, welches dem momentanen Stand der instruktionalen Unterweisung entspricht. Wurde im Unterricht das Lösen von Aufgaben ohne Übertrag behandelt, so hat der Schüler im Laufe der instruktionalen Unterweisung die Kompetenz erworben, solche Aufgaben unter Anwendung des korrekten Lösungsalgorithmus erfolgreich zu bearbeiten. Bekommt er nun eine Aufgabe präsentiert, die einen Übertrag erfordert, findet er sich in einer Sackgasse („impasse“, VanLehn, 1990) wieder, da er für dieses neue Problem über keine adäquate Lösungsprozedur verfügt. Um die Sackgasse zu überwinden, wählt der Schüler eine Reparatur („repair“) aus einem Set existierender Heuristiken aus. Dabei rekuriert er auf bereits existierende Subprozeduren, die sich bisher bewährt haben. Ein systematischer Fehler resultiert nun aus der unangemessenen Übergeneralisierung einer sol-

chen Subprozedur auf ein neues Problem. So könnte er auf das neue Problem „Aufgabe mit Übertrag“ die Subprozedur „von der kleineren Ziffer zur größeren Ziffer ergänzen“ anwenden, die sich als adäquat beim Lösen von „Aufgaben ohne Übertrag“ erwiesen hat. Neben dem Phänomen der Übergeneralisierung können systematische Fehler auch aus der Überdiskriminierung einer Regel resultieren. Wurde beispielsweise der Übertrag am Beispiel zweistelliger Subtraktionsaufgaben eingeführt, macht der Schüler die Erfahrung, dass ein Übertrag nur an der Einerstelle ausgeführt werden muss. Sieht er sich nun mit dreistelligen Subtraktionsaufgaben konfrontiert, wird er die Regel unangemessen überdiskriminieren, indem er einen erforderlichen Übertrag nur an der Einerstelle ausführt und nicht an einer Zehnerstelle, auch wenn in der Zehnerspalte die obere Ziffer kleiner als die untere sein sollte.

Ein systematischer Fehler beschreibt demnach die Anwendung einer Fehlstrategie als Ergebnis einer fehlerhaften Induktion (Übergeneralisierung oder Überdiskriminierung) aus Lösungsbeispielen. In einer Studie zur empirischen Evidenz seiner Induktions-Hypothese konnte VanLehn (zitiert nach Ben-Zeev, 1995) 33% der beobachteten Fehler über fehlerhaftes induktives Schließen aus Lösungsbeispielen erklären. Ben-Zeev (1995) konnte über die Einführung eines neuen Zahlensystems („NewAbacus“) mit 67% Fehlererklärung einen stärkeren empirischen Beleg für die Induktions-Hypothese liefern.

Ein weiteres wichtiges Konstrukt der repair theory, welches vor allem im Hinblick auf die Intervention bei Schülerfehlern eine Rolle spielt, stellt das von Brown und VanLehn (1980) beschriebene „bug migration“ Phänomen dar. Dieses Phänomen beschreibt die Beobachtung, dass bei Erreichen einer Sackgasse (im Sinne eines neuen Problems) und ohne dass eine Intervention stattgefunden hat, *verschiedene* Repairstrategien zum Einsatz kommen können. Stabiles Fehlverhalten, im Sinne eines konsistenten Einsatzes einer bestimmten Fehlstrategie bei einer bestimmten Sackgasse, ist nach VanLehn dann zu beobachten, wenn:

- (a) der Lerner eine starke Präferenz für die Wahl einer bestimmten Repairstrategie hat, beispielsweise weil eine Präferenz für diese Strategie bereits vor dem Auftreten der Sackgasse vorlag, oder weil die Präferenz sich aus der Erfahrung mit verschiedenen anderen Repairstrategien entwickelt hat
- (b) der Lerner zwar verschiedene Repairstrategien ausprobiert hat, sich aber zum momentanen Zeitpunkt nur an eine erinnert und diese anwendet
- (c) die durch den Einsatz der Repairstrategie revidierte Lösungsprozedur so automatisiert wird, dass eine Sackgasse nicht mehr erreicht wird.

Liegt keine Präferenz für eine bestimmte Repairstrategie im Sinne einer Fehlstrategie vor, wird der Lerner bei mehrmaligem Erreichen einer Sackgasse verschiedene Fehlstrategien generieren, die sich beim Lösen bekannter Probleme bewährt haben. Das Phänomen der bug migration untersuchte VanLehn (1990) in einer eigenen Studie. Er analysierte die Fehler zu zwei Testzeitpunkten, zwischen denen keine Intervention stattfand und berichtet, dass 38% der beobachteten Fehler im Sinne der bug migration zu interpretieren waren.

Diese Beobachtung der Fehlerinstabilität diente Ben-Zeev (1996) als Grundlage, den Begriff des rationalen Fehlers (rational error) gegenüber dem Begriff des systematischen Fehlers zu präferieren, da der Begriff des rationalen Fehlers den regelbasierten Ursprung des Fehlers beschreibt, ohne auf die Fehlerstabilität zu referenzieren. Da sich der Begriff des systematischen Fehlers aber in der Forschung etabliert hat, wird dieser trotz seiner Ambiguität in dieser Arbeit unter der Prämisse verwendet, dass der Terminus „systematisch“ auf den regelbasierten Ursprung und nicht auf dessen konsistente Anwendung referenziert.

Allerdings impliziert der regelbasierte Ursprung systematischer Fehler nicht, dass diese Fehler auf konzeptionellen Missverständnissen oder Fehlkonzepten der schriftlichen Subtraktion beruhen, sondern dass systematische Fehler im Sinne der repair theory das Ergebnis rein syntaxbasierter Veränderungen der Prozeduren repräsentieren, denen kein teleologisches Verständnis (VanLehn, 1990) zugrunde liegt. Als Beleg dafür, dass die meisten Schüler kein teleologisches Verständnis des Subtraktionsalgorithmus besitzen, führt VanLehn (1990, S. 39) folgendes Beispiel an:

Unter der Voraussetzung, dass bisher nur das Lösen von Aufgaben, die ein Leihen von 0 verschiedenen Ziffern erforderte, gelehrt wurde, stellt eine häufige Sackgasse das Lösen von Subtraktionsaufgaben dar, die ein Leihen von 0 (borrow-from-zero problem, z.B. $508 - 28$) erfordern. Hätte der Schüler ein „teleologisches“ Verständnis der Prozedur aufgebaut, würde der Schüler jedoch nie eine solche Sackgasse erreichen, da beiden Typen von Aufgaben das gleiche Designprinzip zugrunde liegt.

Dass systematische Fehler in der schriftlichen Subtraktion häufig nicht Folge von Fehlkonzepten sind, scheint plausibel unter der Annahme, dass vor allem für sehr abstrakte bzw.

algorithmische Lehrinhalte, wie im Fall des schriftlichen Rechnens, Verstehensprozesse keine notwendige Voraussetzung für den Erwerb dieser kognitiven Fähigkeit darstellen, da diese Fähigkeit rein mechanisch erworben werden kann. Dass oftmals keine Einsicht in die Bedeutung des Ergebnisses vorliegt, wird ersichtlich, wenn noch so abenteuerliche Ergebnisse vorbehaltlos akzeptiert werden, z.B. wenn das Ergebnis nach der Subtraktion größer als der Minuend ist oder wenn Subtraktionsaufgaben bearbeitet werden, die nicht lösbar sind, z.B. die Subtraktion einer größeren Zahl von einer kleineren Zahl. So konnte auch Murray (1983) beobachten, dass gerade für sehr abstrakte bzw. algorithmische Lehrinhalte Widersprüche häufig entweder gar nicht wahrgenommen, als Paradoxa oder Mysterien akzeptiert oder Konflikte (Diskrepanzen) ganz einfach ignoriert werden.

3.2 Aufgabenanforderungen und systematische Fehler beim schriftlichen Subtrahieren

In diesem Kapitel werden Arbeiten von verschiedenen Autoren vorgestellt, welche in Abhängigkeit von ihrem wissenschaftstheoretischen Hintergrund eine unterschiedliche Terminologie für das Konzept der „Aufgabenanforderung“ präferieren, z.B. Schwierigkeitsmerkmale bei Gerster (1982) sowie Kühnhold und Padberg (1986), Sackgassen bei VanLehn (1990), Elementarkompetenzen bei Korossy (1996). Da dieses Kapitel vor allem auf die Beschreibung von Anforderungen, welche die Bearbeitung bestimmter Aufgaben an den Lerner stellt, fokussiert, wird im Sinne eines konsistenten Sprachgebrauchs der Begriff der Aufgabenanforderungen präferiert und im weiteren Verlauf synonym für die Begriffe Schwierigkeitsmerkmale, Sackgassen und Elementarkompetenzen verwendet.

Die Erkenntnis, dass systematische Fehler in der schriftlichen Subtraktion auf das Erreichen einer Sackgasse und der rein syntaxbasierten Generierung verschiedener Repairstrategien zurückzuführen sind, repräsentiert eine wesentliche Information für die Entwicklung informativer tutorieller Feedbackformen. Die Verwendung verschiedener Subtraktionsverfahren in den USA und in Deutschland erlaubt allerdings keine direkte Anwendung der von VanLehn (1990) beschriebenen Sackgassen und Repairstrategien. Ziel dieses Kapitels ist es, typische Aufgabenanforderungen und systematische Fehler für das seit 1958 in Deutschland gelehrt Subtraktionsverfahren „Ergänzen mit Erweitern bei Zehnerübergang“ zu beschreiben. Um ein Verständnis für die in den USA und in Deutschland gelehrt Subtrakti-

onsverfahren zu ermöglichen, welches notwendig ist, um das weitere Vorgehen nachvollziehen zu können, werden diese im Kapitel 3.2.1 mit ihren Vor- und Nachteilen kurz vorgestellt.

3.2.1 Abziehen mit Borgen vs. Ergänzen mit Erweitern

Wird in den USA das „Abziehverfahren mit Borgen bei Zehnerübergang“ (decomposition method) gelehrt, so wird an vielen Schulen in Deutschland immer noch das „Ergänzungsverfahren mit Erweitern bei Zehnerübergang“ gelehrt, welches auf einen KMK-Beschluss des Jahres 1958 zurückgeht und 1976 in einem weiteren Beschluss bestätigt wurde. Da aber der Bildungsbereich in der Länderhoheit liegt und in den letzten Jahren die Kritik an dem Ergänzungsverfahren zunahm (Padberg, 1994; Padberg, 1996), hält seit einigen Jahren das Abziehverfahren Einzug in deutsche Lehrpläne. So wurde in Nordrhein-Westfalen bereits 1996 das Verfahren der schriftlichen Subtraktion freigegeben. Nach dem Lehrplan für die Grundschulmathematik in Bayern kann seit dem Jahr 2000 sowohl das Abziehverfahren als auch das Ergänzungsverfahren gelehrt werden. In Sachsen gilt derzeit noch der Lehrplan von 1992, welcher das Ergänzungsverfahren festschreibt. Der neue Lehrplan, in welchem empfohlen wird, für die Erarbeitung eines inhaltlichen Verständnisses der Rechenverfahren das Abziehverfahren mit Borgen zu lehren, befindet sich momentan noch in der Erprobung, so dass davon ausgegangen werden kann, dass auch in absehbarer Zeit noch ein Großteil der GrundschullehrerInnen nach dem bekannten Ergänzungsverfahren mit Erweitern lehren wird.

Abzieh- und Ergänzungsverfahren unterscheiden sich durch ihre Sprechweise, während die Schreibweise für beide gleich ist. In Tabelle 4 sind die Sprechweisen beider Verfahren an einem Beispiel erläutert sowie die Vor- und Nachteile des Ergänzungsverfahrens im Vergleich zum Abziehverfahren kurz aufgelistet.

Ein weiterer Unterschied zwischen dem in den USA und dem in Deutschland gelehrt Verfahren wird manifest, wenn die Aufgabe einen Übertrag erfordert, welcher im Sinne der Borgetechnik oder der Erweiterungstechnik ausgeführt werden kann. In Tabelle 5 sind beide Verfahrensweisen mit ihren jeweiligen Vor- und Nachteilen nach Padberg (1996) kurz dargestellt.

Tabelle 4:

Gegenüberstellung des Abziehverfahrens und des Ergänzungsverfahrens mit den Vor- und Nachteilen des Ergänzungsverfahrens gegenüber dem Abziehverfahren nach Padberg (1996)

	Sprechweise beim Abziehverfahren	Sprechweise beim Ergänzungsverfahren
754	4 minus 2 gleich <u>2</u>	2 plus <u>2</u> gleich 4
- 342	5 minus 4 gleich <u>1</u>	4 plus <u>1</u> gleich 5
412	7 minus 3 gleich <u>4</u>	3 plus <u>4</u> gleich 7
Nachteile	<ul style="list-style-type: none"> • Das Abziehen wird als natürlichere Sinngebung der Subtraktion aufgefasst, da beim Abziehen subtrahiert, beim Ergänzen addiert wird. • Aufgrund von Verwechslungen mit der schriftlichen Addition unterlaufen beim Ergänzungsverfahren leichter Fehler. • Die Schreib- und Sprechweise klaffen beim Ergänzungsverfahren auseinander (Differenz steht in der Schreibweise unten, in der Sprechweise in der Mitte). 	
Vorteile	<ul style="list-style-type: none"> • Beim Ergänzungsverfahren wird nur das weniger fehleranfällige „Einsundeins“ im Gegensatz zum „Einsminuseins“ benötigt. • Das Vorwärtszählen wird besser beherrscht als das Rückwärtszählen. • Die Subtraktion mehrerer Subtrahenden ist beim Ergänzungsverfahren leichter und daher weniger fehleranfällig. • Der Zusammenhang zwischen Addition und Subtraktion als inverse Rechenoperationen wird beim Ergänzungsverfahren deutlicher. 	

Bei der Borgetechnik wird im Minuenden eine Einheit des nächsthöheren Stellenwertes (im Beispiel die Zehnerstelle des Minuenden) entbündelt, d.h. es wird eine Einheit „geborgt“. Im Unterschied zur Borgetechnik liegt der Erweiterungstechnik das Gesetz von der Konstanz der Differenz zugrunde. Im Sinne dieses Gesetzes bleibt die Differenz zweier Zahlen unverändert, wenn zum Minuenden bzw. Subtrahenden dieselbe Zahl addiert bzw. subtrahiert wird.

Cox (1975) fand in ihrer Studie, dass der am zweithäufigsten beobachtete systematische Fehler aus einem fehlerhaften Entbündeln des Minuenden resultierte, demnach auf das spezielle Verfahren der Borgetechnik zurückgeführt werden kann. Dieser Befund zeigt, dass eine direkte Übertragung der von VanLehn (1990) identifizierten Sackgassen und systematischen Fehler auf den deutschsprachigen Bereich nicht möglich ist. Von daher schien es notwendig, das in Deutschland gelehrt Verfahren der Ergänzungstechnik mit Erweitern dahingehend zu untersuchen, welche Aufgabenanforderungen im Sinne potentieller Sackgassen auftreten können und welche systematischen Fehler zu beobachten sind, um auf diesen Erkenntnissen aufbauend, fehlerspezifisches ITF zu entwickeln.

Tabelle 5:

Rechenvorschrift, Vor- und Nachteile für Borgetechnik und Erweiterungstechnik nach Padberg (1996)

Beispiel	Borgetechnik		Erweiterungstechnik	
	Enaktive/ikonische Realisierung	Symbolische Schreibweise	Enaktive/ikonische Realisierung	Symbolische Schreibweise
	<p>Zehner Einer</p>	$\begin{array}{r} 4 \\ 8^1 4 \\ - 26 \\ \hline 28 \end{array}$	<p>Zehner Einer</p>	$\begin{array}{r} 54^{10} \\ - 26 \\ \hline 28 \end{array}$
Vorteile	<ul style="list-style-type: none"> • Umformungen werden ausschließlich im Minuenden vorgenommen. • Diese Technik ist besonders unter Verwendung konkreten Materials für den Schüler gut nachvollziehbar. 		<ul style="list-style-type: none"> • Aufgaben mit Zwischennull im Minuenden sind problemlos zu lösen • Subtraktionsaufgaben mit mehreren Subtrahenden lassen sich ohne größere Probleme bearbeiten • Technik lässt sich leicht durch konkretes Material oder ikonisch veranschaulichen 	
Nachteile	<ul style="list-style-type: none"> • Das Lösen von Aufgaben mit mehreren Nullen im Minuenden ist kompliziert • Das Rechnen mit mehreren Subtrahenden kann die Entbündelung von zwei oder auch mehr Einheiten erfordern. 		<ul style="list-style-type: none"> • Starke Veränderung der ursprünglichen Aufgabenstellung durch Abänderung der Zahlen • Gesetz der Konstanz der Differenz schwer zu verstehen • Endform der Kurzschreibweise lässt entscheidende Idee des Erweiterns nicht mehr sichtbar werden, Gefahr der Mechanisierung • Grundidee der Erweiterungstechnik ist so kompliziert, dass nur wenige Kinder sie wirklich verstehen 	

3.2.2 Aufgabenanforderungen und systematische Fehler nach Gerster (1982)

Im deutschsprachigen Raum existieren nur wenige Arbeiten (Gerster, 1982; Kühnhold & Padberg, 1986), die sich mit der Beschreibung typischer Aufgabenanforderungen und systematischer Fehler beim Lösen schriftlicher Subtraktionsaufgaben beschäftigt haben. Beide mir bekannten Arbeiten sollen in diesem und im nächsten Kapitel ausführlich beschrieben werden, da deren Erkenntnisse als Grundlage für die Entwicklung des fehlerspezifischen ITFs dienen.

Gerster (1982) identifizierte im Verlauf einer vierjährigen Analyse von Schülerfehlern in Hausarbeiten, Klassenarbeiten, lernzielorientierten Tests und Erprobungsfassungen diagnostischer Tests folgende typischen Aufgabenanforderungen (bei Gerster „Schwierigkeitsmerkmale“), welche sich als fehlerauslösend erwiesen haben:

- Aufgaben mit Stellenunterschied (Leerstelle)
- Aufgaben mit einem Übertrag
- Aufgaben mit zwei Überträgen
- Aufgaben mit Null in der oberen Zahl
- Aufgaben mit Null in der unteren Zahl
- Keine Null in den gegebenen Zahlen (Anm. d. Verf.: d.h. Null im Ergebnis)
- Aufgaben mit Übertrag in leere Stelle
- Aufgaben mit Übertrag in leere Stelle und Ergänzen bis 10
- Aufgaben mit Übertrag zur Null
- Aufgaben mit Übertrag zur Null und Ergänzen bis 10
- Aufgaben mit Übertrag zur 9
- Aufgaben mit Übertrag zur 9 und Ergänzen bis 10
- Aufgaben mit gleichen Ziffern übereinander
- Aufgaben mit gleichen Ziffern übereinander beim höchsten Stellenwert
- Aufgaben mit gleichen Ziffern übereinander und dazu Übertrag
- Aufgaben mit gleichen Ziffern übereinander nach Berücksichtigung des Übertrags
- Aufgaben, in denen die untere Ziffer minus dem Übertrag gleich der oberen Ziffer
- Aufgaben, in denen die obere Zahl kleiner als die untere Zahl

Bei der Entwicklung der diagnostischen Tests wurden diese identifizierten Aufgabenanforderungen berücksichtigt, indem gezielt Aufgaben zu diesen Anforderungen konstruiert wurden. In Tabelle 6 ist der von Gerster auf Basis der empirischen Aufgaben- und Fehler-

analysen entwickelte diagnostische Test S1 dargestellt, welcher die in vorausgegangenen Fehleranalysen identifizierten Anforderungen enthält.

Tabelle 6:

Gersters diagnostischer Subtraktionstest S1 (1982)

Zahlenraum bis 1000 Aufgaben ohne bzw. mit Stellenunterschied wechseln zeilenweise				
	Keine Null	Null in der oberen Zahl	Null in der unteren Zahl	Keine Null in den gegebenen Zahlen
Kein Übertrag	1 45 <u>-32</u>	2 608 <u>-203</u> (7)	3 958 <u>-104</u>	4 496 <u>-136</u> (7)
	5 479 <u>-27</u>	6 309 <u>-4</u>	7 867 <u>-40</u>	8 687 <u>-82</u> (7)
Ein Übertrag	9 72 <u>-49</u>	10 704 <u>-262</u>	11 773 <u>-407</u> (3)	12 365 <u>-258</u> (8)(10)
	13 849 <u>-62</u> (1)	14 690 <u>-23</u>	15 657 <u>-8</u>	16 295 <u>-87</u> (10)
Zwei Überträge	17 821 <u>-788</u> (8)(10)	18 900 <u>-439</u>	19 506 <u>-207</u> (9)(4)	20 853 <u>-459</u> (9)
	21 613 <u>-25</u> (1)	22 604 <u>-8</u> (1)(2)	23 703 <u>-97</u> (1)(6)	24 994 <u>-98</u> (1)(5)(9)
SF ^a	25 593 <u>-739</u> (12)	26 1000 <u>-694</u> (8)(6)	27 43 <u>-102</u> (12)	28 951 <u>-67</u> (1)(11)

^aSonderfälle

	Zusätzliche Aufgabenanforderungen	In Aufgabe Nr.
(1)	Übertrag in leere Stelle	13, 15, 21, 22, 23, 24, 28
(2)	Übertrag in leere Stelle und Ergänzen bis 10	22
(3)	Übertrag zur Null	11
(4)	Übertrag zur Null und Ergänzen bis 10	19
(5)	Übertrag zur 9	24
(6)	Übertrag zur 9 und Ergänzen bis 10	23, 26
(7)	Gleiche Ziffern übereinander	2, 4, 8
(8)	... beim höchsten Stellenwert	12, 17, 26
(9)	... dazu Übertrag	19, 20, 24
(10)	... nach Berücksichtigung des Übertrags	12, 16, 17
(11)	Untere Ziffer minus Übertrag (gleich obere Ziffer)	28
(12)	Obere Zahl kleiner als untere Zahl	25, 27

Der von Gerster entwickelte diagnostische Test beinhaltet 28 Aufgaben, die sich alle in ihrem Anforderungsprofil unterscheiden. So enthält jede Aufgabe unterschiedliche Aufgabenanforderungen in unterschiedlicher Anzahl. Es ist zu vermuten, dass Aufgaben, die mehrere

Aufgabenanforderungen beinhalten, auch höhere Anforderungen an den Lerner stellen, als Aufgaben, die nur eine Aufgabenanforderung aufweisen (beispielsweise Aufgabe 5 und 7). Diagnostiziert wird mit diesem Test daher nicht die Beherrschung einzelner Aufgabenanforderungen, sondern die Beherrschung einer bestimmten Aufgabe mit einem spezifischen Anforderungsprofil. Auf die theoretische Begründung dieser Annahme wird im Kapitel 3.2.4 detailliert eingegangen. Als problematisch erweist sich zudem die Aufgabenanforderung „Keine Null in den gegebenen Zahlen“, nicht nur wegen seiner missverständlichen Bezeichnung, welche als „Null im Ergebnis“ zu interpretieren ist. Seine Problematik besteht primär in der doppelten Codierung dieser Aufgabenanforderung in diesem Test, einmal als 0 im Ergebnis und ein zweites mal als zusätzliche Aufgabenanforderung „gleiche Ziffern übereinander (ohne oder nach Übertrag)“. So resultiert eine 0 im Ergebnis nur, wenn gleiche Ziffern übereinander stehen. Ein weiterer Kritikpunkt an dem diagnostischen Test kann darin gesehen werden, dass die Aufgabenanforderung „Untere Ziffer minus Übertrag (gleich obere Ziffer)“ eine Fehlstrategie und keine Aufgabenanforderung beschreibt.

Neben den fehlerauslösenden Aufgabenanforderungen identifizierte Gerster auf der Basis vorliegender Schülerlösungen systematische Fehler, die den Einsatz einer bestimmten Fehlstrategie vermuten ließen. Die von Gerster klassifizierten systematischen Fehler sind mit Beispielen ausführlich in Tabelle 7 aufgelistet.

Betrachtet man die Fehlerklassifikation von Gerster vor dem Hintergrund der repair theory von VanLehn (1990), weist auch seine Fehlerklassifikation einige Probleme auf, auf die nachfolgend kurz eingegangen werden soll. Im Sinne der repair theory generiert der Lerner bei Erreichen einer Sackgasse eine Repairstrategie, die ihm ein Fortschreiten im Aufgabenlösungsprozess ermöglicht. Da die von Gerster beschriebenen Aufgabenanforderungen solche Sackgassen repräsentieren, sollte im Sinne der repair theory eine Fehlerklassifikation eher auf Grundlage dieser Aufgabenanforderungen erfolgen. Das heißt, es sollte analysiert werden, welche systematischen Fehler bei welchen Aufgabenanforderungen zu beobachten sind. Betrachtet man hingegen die Fehlerklassifikation von Gerster, so wird nicht transparent, nach welchen Kriterien seine Ober- und Unterkategorien gebildet wurden.

Tabelle 7:

Systematische Fehler in der schriftlichen Subtraktion nach Gerster (1982)

Oberkategorie: Fehler beim Übertrag (Ü)	Beispiele
Übertrag nicht berücksichtigt (Ü)	$\begin{array}{r} 72 \\ - 49 \\ \hline 33 \end{array}$
Übertrag in besonderen Fällen nicht berücksichtigt (Ü _b)	$\begin{array}{r} 773 \\ - 407 \\ \hline 376 \end{array}$
<ul style="list-style-type: none"> Zur Ziffer 0 (Ü₀) 	$\begin{array}{r} 118 \\ - 75 \\ \hline 143 \end{array}$
<ul style="list-style-type: none"> In leere Stelle (Ü_L) 	$\begin{array}{r} 651 \\ - 498 \\ \hline 163 \end{array}$
<ul style="list-style-type: none"> Zur Ziffer 9 (Ü₉) 	$\begin{array}{r} 853 \\ - 459 \\ \hline 404 \end{array}$
<ul style="list-style-type: none"> In Spalte mit gleichen Ziffern (Ü₌) 	$\begin{array}{r} 760 \\ - 426 \\ \hline 344 \end{array}$
<ul style="list-style-type: none"> Nach Ergänzen bis 10 (Ü¹⁰) 	$\begin{array}{r} 651 \\ - 498 \\ \hline 253 \end{array}$
<ul style="list-style-type: none"> Nach Ergänzen ab 9 + 1 (Ü₁₀) 	$\begin{array}{r} 687 \\ - 67 \\ \hline 610 \end{array}$
Übertragsziffer zu viel (Ü _z)	$\begin{array}{r} 687 \\ - 67 \\ \hline 610 \end{array}$
<ul style="list-style-type: none"> ... aus Spalte mit gleichen Ziffern (evtl. nach Addition eines Übertrags) (Ü_{z=}) 	
Falsches Operieren mit einer Übertragsziffer (Ü _O)	$\begin{array}{r} 704 \\ - 262 \\ \hline 642 \end{array}$
<ul style="list-style-type: none"> Übertrag zur oberen Ziffer addiert (Ü₊) 	
<ul style="list-style-type: none">ausgelöst durch Ziffer 0 (Ü₊₀) 	$\begin{array}{r} 773 \\ - 407 \\ \hline 386 \end{array}$
<ul style="list-style-type: none"> ...ausgelöst durch Leerstelle (Ü_{+L}) 	$\begin{array}{r} 849 \\ - 62 \\ \hline 987 \end{array}$
<ul style="list-style-type: none"> Übertrag von der unteren Ziffer subtrahiert (Ü₋) 	$\begin{array}{r} 751 \\ - 67 \\ \hline 704 \end{array}$
<ul style="list-style-type: none"> Übertrag in falscher Spalte (Ü_S) 	$\begin{array}{r} 295 \\ - 87 \\ \hline 118 \end{array}$
<ul style="list-style-type: none"> ... Ausweichen vor der Null (Ü_{S0}) 	$\begin{array}{r} 704 \\ - 106 \\ \hline 508 \end{array}$
<ul style="list-style-type: none"> Übertrag mehrmals berücksichtigt (Ü_m) 	$\begin{array}{r} 295 \\ - 87 \\ \hline 108 \end{array}$

Oberkategorie: Falsche Rechenrichtung (F)	Beispiele
• ... ohne Übertrag (F_{ou})	$\begin{array}{r} 72 \\ - 49 \\ \hline 37 \end{array}$
• ... mit Übertrag ($F_{mü}$)	$\begin{array}{r} 674 \\ - 227 \\ \hline 443 \end{array}$
• obwohl untere Ziffer kleiner als obere Ziffer ($F_{<}$)	$\begin{array}{r} 396 \\ - 52 \\ \hline 264 \end{array}$
Oberkategorie: Fehler mit der Null (N)	Beispiele
• Null durch Abziehen ($N_{.}$)	$\begin{array}{r} 772 \\ - 415 \\ \hline 360 \end{array}$
• Null durch Abziehen von der Null ($N_{0.}$)	$\begin{array}{r} 640 \\ - 116 \\ \hline 530 \end{array}$
• Falsche Rechenrichtung: Abziehen der oberen Null von der unteren Ziffer (N_R)	$\begin{array}{r} 704 \\ - 262 \\ \hline 562 \end{array}$
• Null dominiert (N_d)	$\begin{array}{r} 956 \\ - 105 \\ \hline 801 \end{array}$
• Null vermeiden (N_v)	$\begin{array}{r} 343 \\ - 142 \\ \hline 241 \end{array}$
Oberkategorie: Fehler durch unterschiedliche Stellenzahl (leere Stelle) (L)	Beispiele
• „Überstehende“ Stellen nicht berücksichtigt (L_b)	$\begin{array}{r} 682 \\ - 61 \\ \hline 21 \end{array}$
Oberkategorie: Addition statt Subtraktion (inverse Operation) IO	Beispiele
• Addition statt Subtraktion in einzelner Spalte (IO_c)	$\begin{array}{r} 687 \\ - 82 \\ \hline 609 \end{array}$

So werden bei Gerster alle systematischen Fehler, die bei den Aufgabenanforderungen „Null in der oberen Zahl“ und „Null in der unteren Zahl“ zu beobachten sind sowie Fehler, als deren fehlerhaftes Ergebnis eine 0 in der Lösung resultiert, unter die Oberkategorie „Fehler mit der Null“ subsummiert. Einen weiteren „Fehler mit der Null“ repräsentiert eine Aufgabe mit gleichen Ziffern übereinander, als deren Differenz diese Ziffer im Ergebnis steht (z.B. $4-4=4$). Andererseits repräsentieren in der Fehlerklassifikation von Gerster aber auch einzelne Aufgabenanforderungen, z.B. „Fehler durch unterschiedliche Stellenzahl“ als auch einzelne Fehlstrategien, z.B. „Addition statt Subtraktion“ oder „Falsche Rechenrichtung“ separate Oberkategorien. Ein weiteres Problem, welches noch einmal sehr schön den Unterschied zwischen beobachteten systematischen Fehler und vermuteter Fehlstrategie verdeutlicht, liegt in

der Beschreibung der beiden Fehlstrategien „Übertrag zur oberen Ziffer addiert“ bzw. „Übertrag von der unteren Ziffer subtrahiert“. Da beide Fehlstrategien in dem gleichen systematischen Fehler resultieren, lässt sich nicht eindeutig bestimmen, welche Fehlstrategie eingesetzt wurde. Eine Unterscheidung zwischen diesen beiden Fehlstrategien scheint daher nicht angemessen.

Nach Gerster (1982, S. 14) ließen sich 80% aller beobachteten Schülerfehler auf den Einsatz einer bestimmten Fehlstrategie zurückführen. Dieser Befund stellt eine Replikation des von Radatz bereits 1980 (S. 72) berichteten systematischen Fehleranteils von 70-90% an der Gesamtmenge fehlerhafter Aufgabenlösungen dar. Leider liegen keine Angaben über die beobachteten Fehlerhäufigkeiten der einzelnen systematischen Fehler vor.

3.2.3 Aufgabenanforderungen und systematische Fehler nach Kühnhold und Padberg (1986)

Kühnhold und Padberg (1986) führten wie Gerster (1982) eine empirische Untersuchung zu Aufgabenanforderungen und systematischen Fehlern in der schriftlichen Subtraktion in 31 Klassen des vierten Schuljahres an 16 Grundschulen durch. Auf der Grundlage der in der angloamerikanischen und deutschsprachigen Literatur beschriebenen Ergebnisse sowie angeregt durch die von Gerster (1982) konstruierten Tests erstellten Kühnhold und Padberg gleichfalls einen diagnostischen Test (siehe Tabelle 8) mit dem Ziel der Identifikation und Klassifikation systematischer Fehler.

Im Unterschied zu Gerster analysierten Kühnhold und Padberg die beobachteten systematischen Fehler entsprechend ihrer Häufigkeit. Dabei bewerteten sie in Anlehnung an die Vorgehensweise von Cox (1975) einen Fehler als systematisch, wenn dieser bei mindestens der Hälfte aller in Frage kommenden Aufgaben gemacht wurde, wobei Kühnhold und Padberg nicht explizieren, was unter „aller in Frage kommenden Aufgaben“ zu verstehen ist. Auf die Problematik der Definition eines systematischen Fehlers über seine Häufigkeit, im Sinne seines konsistenten Auftretens, wurde bereits im Kapitel 3.1.2 ausführlich eingegangen. Die Beobachtung des „intratest bug migration“ Phänomens (VanLehn, 1990, S. 55) zeigte, dass ein systematischer Fehler nicht über sein konsistentes Auftreten definiert werden sollte, sondern über seinen regelbasierten Ursprung.

Tabelle 8:

Diagnostischer Subtraktionstest nach Kühnhold und Padberg (1986)

	Zahlenraum bis 100 000		
	Keine Null	Null in den gegebenen Zahlen	Null im Ergebnis
Kein ÜT ^a	1 746 <u>-532</u> (7)	2 8067 <u>-4020</u> (7)	3 5738 <u>-717</u> (7)
Ein Übertrag	4 713 <u>-281</u> (1)	5 7705 <u>-462</u> (3)	6 3964 <u>-2558</u> (7)(8)(10)
	7 3279 <u>-628</u> (1)	8 5437 <u>-2091</u> (3)	9 5268 <u>-4838</u> (7)(8)(10)
Zwei Überträge	10 5643 <u>-4295</u> (5)	11 1503 <u>-396</u> (6)	12 74254 <u>-4156</u> (7)(9)
	13 9638 <u>-675</u> (1)(9)	14 8973 <u>-8085</u> (3)(8)	15 123781 <u>-3978</u> (1)(10)(10)
Mind. 3 Überträge	16 88555 <u>-33999</u> (5)(5)	17 60107 <u>-309</u> (1)(2)(4)	18 72184 <u>-3978</u> (1)(10)
	19 43362 <u>-42974</u> (5)(8)(10)	20 20010 <u>-420</u> (1)(2)	21 51365 <u>-9385</u> (1)(5)(7)(9)
SF ^b	22 6352 <u>-6413</u> (12)	23 1000 <u>-333</u> (1)	24 8345 <u>-37642</u> (7)(10)

Anmerkung: Die in Klammern angegebenen Ziffern wurden nachträglich von der Autorin hinzugefügt und beinhalten die Kennzeichnung zusätzlicher Aufgabenmerkmale nach Gerster (1982), siehe Kapitel 3.2.2

^akein Übertrag

^bSonderfälle

Entsprechend der von Kühnhold und Padberg verwendeten Definition konnten insgesamt 19 systematische Fehler identifiziert werden, wobei folgende drei systematischen Fehler, die zusammen über 50% des Gesamtfehleranteils ausmachen, am häufigsten beobachtet wurden:

- (1) Spaltenweise Unterschiedsbildung (im Sinne der Klassifikation nach Gerster handelt es sich dabei um die Fehlstrategie „Falsche Rechenrichtung“)
- (2) Keine Berücksichtigung des Übertrags
- (3) Kein Übertrag in Leerstelle

Unter Berücksichtigung der Anzahl der Aufgaben, bei denen die einzelnen Fehlerklassen (in Anlehnung an Gersters Klassifikationsschema) prinzipiell auftreten können, fand sich folgende Reihenfolge der am häufigsten beobachteten Fehlergruppen:

- (1) Übertragsfehler
- (2) Rechenrichtungsfehler
- (3) Perseverationsfehler
- (4) Fehler mit der Null
- (5) Fehler durch unterschiedliche Stellenzahl
- (6) Einsundeinsfehler
- (7) Anwendung der inversen Operation (Addition statt Subtraktion)

Wie die Fehlerklassifikation von Gerster lässt sich auch die Fehlerklassifikation von Kühnhold und Padberg (1986) dahingehend kritisieren, dass eine Fehlerklassifikation nicht auf der Grundlage potentieller Sackgassen, d.h. auf der Grundlage von Aufgabenanforderungen erfolgte, sondern dass Fehlerklassen über verschiedene Aufgabenanforderungen hinweg gebildet wurden.

3.2.4 Aufgabenanforderungen und systematische Fehler – Eine theoretische und empirische Reanalyse

Gerster entwickelte seine diagnostischen Tests mit dem Ziel, Lehrer und Lehramtsstudierende für Schülerfehler und deren Rolle im Lernprozess zu sensibilisieren. Er weist explizit darauf hin, dass es sich bei diesen Tests um informelle, nicht standardisierte Tests handelt, die im Hinblick auf die beschriebenen Aufgabenanforderungen und Fehlertypen keinen Anspruch auf Vollständigkeit erheben (1982, S.10).

Nach Narciss (2004) sollte die Entwicklung informativer tutorieller Feedbackformen auf der Basis einer kognitiven Anforderungsanalyse erfolgen, um definieren zu können, welche bereichsspezifischen Wissens Elemente notwendig sind, um bestimmte Aufgaben erfolgreich bearbeiten zu können. Dies schien auf Basis der diagnostischen Tests von Gerster sowie Kühnhold und Padberg nicht realisierbar, da diese im Sinne aktueller Ansätze der Wissensrepräsentation und Wissensdiagnostik keine systematische Beschreibung der für eine erfolgreiche Bearbeitung einer Aufgabe notwendigen Wissens Elemente liefern, bzw. eine reliable Diagnostik nicht beherrschter Subprozeduren auf Grundlage dieser diagnostischen Tests nicht

möglich schien. Um zu entscheiden, für welche Aufgaben mit welchen Aufgabenanforderungen fehlerspezifisches ITF entwickelt werden sollte, wurde daher in einem ersten Schritt eine kognitive Anforderungsanalyse auf der Grundlage wissenstheoretischer Ansätze durchgeführt, die im Kapitel 3.2.4.1 ausführlich beschrieben ist. Hierbei dienten die Arbeiten von VanLehn (1990), Gerster (1982) sowie Kühnhold und Padberg (1986) als wertvolle Grundlage.

Im Anschluss an die kognitive Anforderungsanalyse wurden für die definierten Aufgabenanforderungen systematische Fehler auf Basis einer Fehleranalyse vorliegender Daten einer Vorstudie identifiziert. Deren Ergebnisse sind im Kapitel 3.2.4.2 dargestellt.

3.2.4.1 Definition der Aufgabenanforderungen mit Hilfe einer kognitiven Anforderungsanalyse

Die Entwicklung von fehlerspezifischem ITF für computergestützte Lernaufgaben zum Themenbereich der schriftlichen Subtraktion erforderte entsprechend der Designrichtlinien für ITF nach Narciss (2004) eine kognitive Anforderungsanalyse, um die für eine erfolgreiche Aufgabenbearbeitung erforderlichen Wissensselemente zu identifizieren.

Für die schriftliche Subtraktion existieren in der Literatur eine Reihe von Arbeiten, in denen auf der Basis kognitiver Anforderungs- und Aufgabenanalysen das Vorgehen beim Lösen schriftlicher Subtraktionsaufgaben sehr detailliert, d.h. auf einem sehr elementaren Niveau beschrieben ist, z.B. das prozedurale Netzwerk von Brown und Burton (1978), das Produktionssystem von Young und O'Shea (1982), die Lernhierarchie von Gagné, Briggs und Wager (1992), der syntaktische Subtraktionsalgorithmus nach Resnick (1982).

Im Rahmen der vorliegenden Arbeit erwiesen sich diese Ansätze als wenig praktikabel, da eine vollständige Beschreibung aller möglichen deklarativen und prozeduralen Wissensselemente auf einem solch elementaren Niveau unter einem Kosten-Nutzen-Aspekt wenig effizient gewesen wäre. Aus diesem Grund wurde in dieser Arbeit auf aktuellere Ansätze der Wissensstrukturierung und Wissensdiagnose rekuriert (Doignon & Falmange, 1985; Korossy, 1996; Korossy & Held, 2001).

Diese Ansätze definieren das Wissen zu einem bestimmten Gegenstandsbereich über eine Menge von Aufgaben, d.h. der Wissenszustand einer Person lässt sich durch das Aufgabenlösungsverhalten dieser Person bestimmen. Entscheidend für den Ansatz der Wissensstruktur-Theorie ist die Erfahrungstatsache, dass Wissen Strukturmerkmale insofern aufweist, als die Lösung bestimmter Aufgaben die Lösung bestimmter anderer Aufgaben erwarten lässt.

Damit kann aus der Menge der kombinatorisch möglichen Lösungsmuster eine Menge von empirisch erwartbaren Lösungsmustern ausgezeichnet werden. Ein solches empirisch erwartbares Lösungsmuster heißt in der Wissensstruktur-Theorie Wissenszustand, die Menge aller Wissenszustände wird Wissensstruktur genannt. Formal wird eine Wissensstruktur dargestellt als ein geordnetes Paar $\{A, K\}$, bestehend aus einer endlichen, nicht-leeren Menge A von Aufgaben und einer Familie K von Teilmengen aus A , den Wissenszuständen. Die Erstellung der Wissensstruktur erfolgt üblicherweise über die Auswahl einer Stichprobe repräsentativer Aufgaben, die dann häufig auf der Basis von Expertenbefragungen in geeigneter Weise strukturiert werden, um schließlich eine Wissensstruktur als Familie empirisch erwartbarer Wissenszustände zu etablieren (Korossy, 1996). Eine Kritik an der Wissensstrukturtheorie, die sich auch als relevant für die vorliegende Arbeit erweisen sollte, formuliert Korossy (1996, S. 284) folgendermaßen:

„Ist eine solche Wissensstruktur erstellt und empirisch validiert, so lässt sich der Wissenszustand einer Person operationalisiert als eine bestimmte Menge gelöster bzw. lösbarer Aufgaben diagnostizieren. Was aber ‚weiß‘ diese Person nun ‚wirklich‘? Wird sie eine bestimmte, zusätzlich vorgelegte Aufgabe lösen können oder nicht? Welches ‚Wissen‘ müsste ihr vermittelt werden, damit sie einen bestimmten ‚höheren‘ Wissenszustand erreicht?“

Antworten auf diese Fragen gibt die Wissensstrukturtheorie nicht. Aus diesem Grund erweiterte Korossy (1996; Korossy & Held, 2001) die Wissensstrukturtheorie um seine „Kompetenz-Performanz-Konzeption“, wobei „Performanz als empirisch beobachtbares Verhalten (beim Lösen von Aufgaben) und Kompetenz als inhaltlich-qualitatives, nicht direkt beobachtbares, theoretisches Konstrukt zur Erklärung und Prognose von Performanz-Beobachtungen aufgefasst wird.“ (Korossy, 1996, S. 310). Dies bedeutet, dass auf der Basis beobachtbarer Lösungsmuster auf vorhandene/nicht vorhandene Kompetenzen geschlossen werden kann. Dieser Rückschluss von beobachtbaren Aufgabenlösungen auf zugrundeliegende Kompetenzen stellt eine wesentliche Bedingung für die Bereitstellung von ITF dar, da sich unter Verwendung des Kompetenz-Performanz-Ansatzes im Gegensatz zu den Annahmen der klassischen Wissensstrukturtheorie nun folgende Fragen beantworten lassen: (a) Welche Anforderungen beinhaltet eine Aufgabe? (b) Was muss der Lerner wissen bzw. können, um diese Anforderungen zu bewältigen und (c) Wie lassen sich bestimmte, vor allem inkorrekte Lösungen erklären? Die Beantwortung dieser Fragen ist notwendig, um bei Aufgaben, die bestimmte Anforderungen beinhalten und die vom Lerner noch nicht bewältigt werden, fehler-

spezifisches ITF anzubieten, welches für die Anforderungsbewältigung relevantes Wissen bereitstellt.

Im Gegensatz zu den bereits zitierten Arbeiten der Wissensrepräsentation zum Themenbereich der schriftlichen Subtraktion (z.B. Brown & Burton, 1978; Young & O'Shea, 1982), erhebt der Kompetenz-Performanz-Ansatz zudem keinen Anspruch auf eine auch nur annähernd so komplexe Form der Wissensrepräsentation, sondern repräsentiert einen eher pragmatischen Ansatz, Wissen zu modellieren und zu diagnostizieren (Korossy, 1996).

Im weiteren wird beschrieben, wie basierend auf dem Vorgehen von Korossy (1996) eine Anforderungsstruktur für den Gegenstandsbereich der schriftlichen Subtraktion über die Menge aller möglichen Aufgabentypen definiert und modelliert wurde. Ein Aufgabentyp ist über die Art und Anzahl der Aufgabenanforderungen sowie deren Kombinationsmöglichkeiten definiert. Entsprechend der Wissensstrukturtheorie lässt sich ein Wissenszustand als Menge der Aufgabentypen beschreiben, die nach unten abgeschlossen ist. Das heißt, wenn ein Aufgabentyp in der Menge der Aufgabentypen enthalten ist, dann sind auch seine untergeordneten Aufgabentypen in dieser Menge enthalten.

Die Erstellung der Anforderungsstruktur für den Bereich der schriftlichen Subtraktion beschränkte sich hierbei auf den Zahlenraum bis 1000. In einem ersten Schritt wurden folgende in Tabelle 9 aufgeführte elementare Aufgabenanforderungen basierend auf den bereits ausführlich dargestellten empirischen und kognitionspsychologischen Erkenntnissen zu Aufgabenanforderungen in der schriftlichen Subtraktion identifiziert (Gerster, 1982; Kühnhold & Padberg, 1986; VanLehn, 1990).

Mit der Menge A der elementaren Aufgabenanforderungen wurde die erste (unterste) Analyseebene und damit der Auflösungsgrad der Modellierung festgelegt. Die Kombinationen dieser elementaren Aufgabenanforderungen wiederum definieren die Art der Aufgabentypen als Teilmengen von A. So wurden in einem nächsten Schritt Aufgabentypen definiert, welche die Kombination von zwei der 5 elementaren Aufgabenanforderungen beinhalten. Diese Aufgabentypen repräsentieren die zweite (nächst höhere) Ebene der hierarchischen Anforderungsstruktur (siehe Abbildung 5). Dabei wird in Analogie zu Korossy (1996) dieser Art der Anforderungsmodellierung folgende Eigenschaft unterstellt:

Eine vollständige Anforderungsstruktur bildet nicht nur einzelne elementare Aufgabenanforderungen, sondern auch deren Kombinationen ab.

Tabelle 9:

Die Menge A der elementaren Aufgabenanforderungen

Elementare Aufgabenanforderungen	Abkürzung	Elementarkompetenz	Beispiele		
Leere Stelle	L	Wissen und Anwenden der Regeln für das Subtrahieren mit einer leeren Stelle	375 <u>- 43</u>	37 <u>- 4</u>	
Übertrag	Ü	Wissen und Anwenden der Regeln für das Subtrahieren mit einem Übertrag	375 <u>-147</u>	357 <u>- 174</u>	35 <u>- 17</u>
Null im Subtrahenden	N _u	Wissen und Anwenden der Regeln für das Subtrahieren mit einer Null im Subtrahenden	375 <u>-140</u>	357 <u>- 104</u>	35 <u>- 10</u>
Null im Minuenden (Ergänzen zur 10)	N ^o	Wissen und Anwenden der Regeln für das Subtrahieren mit einer Null im Minuenden	370 <u>-147</u>	307 <u>- 174</u>	30 <u>- 17</u>
Gleiche Ziffern übereinander	Z	Wissen und Anwenden der Regeln für das Subtrahieren mit gleichen Ziffern übereinander	375 <u>- 145</u>	357 <u>154</u>	537 <u>514</u>

Dieser Eigenschaft inhärent ist hierbei die Annahme, dass eine Kombination verschiedener elementarer Aufgabenanforderungen andere im Sinne höherer Anforderungen an den Lerner stellt, als die Bewältigung einzelner elementarer Aufgabenanforderungen. Für dieses Postulat lassen sich verschiedene kognitionspsychologische und gegenstandsspezifische Erklärungen anführen. Eine Erklärung kann aus der Forschungsliteratur zum kognitiven Fertigkeitserwerb („cognitive skill acquisition“, VanLehn, 1996) und zum Arbeitsgedächtnis gewonnen werden. Nach VanLehn (1996; Fitts, 1964) kann beim kognitiven Fertigkeitserwerb zwischen drei Phasen unterschieden werden: einer frühen, einer mittleren und einer späten Phase. (1) In der frühen Phase des kognitiven Fertigkeitserwerbs ist der Lerner bemüht, das zu erwerbende Domainwissen zu verstehen, ohne notwendigerweise seine Anwendung anzustreben. (2) Die mittlere Phase beginnt, wenn der Lerner seine Aufmerksamkeit auf das Lösen von Problemen aus diesem Wissensgebiet richtet. Ziel dieser Phase ist, durch das Lösen konkreter Probleme Wissenslücken zu schließen und/oder fehlerhaft erworbenes Wissen zu korrigieren. Wenige systematische Fehler und ein gelegentliches Auftreten unsystematischer Fehler signalisieren das Ende dieser Phase und den Beginn der dritten Phase. (3) In dieser dritten Phase werden die erworbenen Fertigkeiten durch häufiges Üben automatisiert, d.h. Geschwindigkeit und Akkuratheit werden verbessert.

Bearbeiten Lernende nun Aufgaben, die mehrere elementare Aufgabenanforderungen beinhalten, fehlerhaft, lösen hingegen aber Aufgaben mit nur einem dieser elementaren Anforderungen korrekt, kann dies vor dem Hintergrund der drei Phasen des kognitiven Fertigkeitserwerbs folgendermaßen erklärt werden: Der Lernende verfügt zwar über das Wissen, welches für die Bewältigung der einzelnen elementaren Anforderungen notwendig ist. Er muss aber auf dieses Wissen noch bewusst zugreifen, d.h. eine Automatisierung der auszuführenden Prozeduren ist noch nicht erfolgt, so dass eine Aufgabe mit mehreren elementaren Aufgabenanforderungen noch erhöhte kognitive Anforderungen stellt. Das heißt, dass Fehler, die bei Aufgaben mit mehreren elementaren Anforderungen auftreten, nicht darauf zurückgeführt werden müssen, dass dem Lernenden das für eine erfolgreiche Aufgabenbewältigung notwendige Wissen fehlt. Vielmehr können Fehler bei solchen Aufgaben aufgrund der Komplexität dieser Aufgabe ihre Ursache in einem cognitive overload haben. Das Lösen dieser Aufgaben bindet aufgrund einer noch nicht ausreichenden Automatisierung die begrenzten Ressourcen der mentalen Verarbeitungskapazität, was in einer erhöhten Fehleranzahl resultieren könnte.

Eine weiteres Argument für das Postulat, dass eine vollständige Anforderungsstruktur nicht nur einzelne elementare Aufgabenanforderungen, sondern auch deren Kombinationen abbildet, kann aus dem konkreten Gegenstandsbereich der schriftlichen Subtraktion gewonnen werden, da die Art der Kombination zweier elementarer Aufgabenanforderungen unterschiedliche Anforderungen an den Lerner zu stellen scheint. So zeigt ein direkter Vergleich der Lösungswahrscheinlichkeiten zweier Aufgaben², welche die Aufgabenanforderungen „gleiche Ziffern“ und „Übertrag“ beinhalten, dass die Aufgabe 348-149 (von 27 der 30 Schüler falsch gelöst) höhere Anforderungen an die Lernenden zu stellen scheint als die Aufgabe 787-297 (von 21 der 30 Schüler falsch gelöst). Diese Diskrepanz in den Lösungswahrscheinlichkeiten scheint darauf zurückführbar, dass die Aufgabe 348-149 durch den Übertrag zu gleichen Ziffern einen zweiten Übertrag erfordert, im Gegensatz zur Aufgabe 787-297. Von daher definiert nicht allein die Kombination zweier Aufgabenanforderungen einen neuen Aufgabenzustand, sondern zusätzlich die Art der Kombination. Für die Aufgabenanforderungen, von denen angenommen wird, dass die Art der Kombination Einfluss auf die kognitiven Anforderungen hat, wurden zwei Aufgabentypen definiert. In Tabelle 10 sind für alle Aufgabentypen dieser zweiten Ebene deren Abkürzungen, Beschreibungen und Beispiele dargestellt.

² Diese Aufgaben entstammen dem Vortest der 1. Studie.

Tabelle 10:

Aufgabentypen der Ebene 2 der Anforderungsstruktur, die über die Kombination von 2 der fünf elementaren Aufgabenanforderungen definiert sind

Abkürzung	Aufgabentyp	Beispiele		
LÜ	Übertrag in leere Stelle	17 - 9	176 - 92	
L+Ü	Leere Stelle und Übertrag	192 - 76		
L + N _u	Leere Stelle und Null im Subtrahenden	192 - 70		
LN ^o	Ergänzen zur 10 und Übertrag in leere Stelle	102 - 76		
L+N ^o	Leere Stelle und Ergänzen zur 10	190 - 76		
L+Z	Leere Stelle und gleiche Ziffern übereinander	196 - 92	169 - 29	
N _u Ü	Übertrag zur Null	424 - 206		
Ü+N _u	Übertrag und Null im Subtrahenden	442 - 260		
Ü+N ^o	Übertrag und Ergänzen zur 10	440 - 262	404 - 226	
ZÜ	Übertrag zu gleichen Ziffern	446 - 248		
Ü+Z	Gleiche Ziffern übereinander und Übertrag	464 - 284	446 - 428	
N ^o +N _u	Ergänzen zur 10 und Null im Subtrahenden	402 - 240		
N _u N ^o	Ergänzen zur 10 und Übertrag zur 0	420 - 204		
N _u +Z	Null im Subtrahenden und gleiche Ziffern übereinander	424 - 220	442 - 202	244 - 202
ZN ^o	Ergänzen zur 10 und Übertrag zu gleichen Ziffern übereinander	440 - 242		
N ^o +Z	Ergänzen zur 10 und gleiche Ziffern übereinander	404 - 224		

In einem nächsten Schritt wurden auf der dritten Anforderungsebene Aufgabentypen definiert, die sich aus Kombinationen der Aufgabentypen der 1. und der 2. Ebene zusammensetzen. Da die Anforderungsmodellierung für den Gegenstandsbereich der schriftlichen Subtraktion im Zahlenraum bis 1000 erfolgte, waren auf dieser dritten Ebene den Kombinationsmöglichkeiten sachimmanente Grenzen gesetzt. Alle noch verbleibenden Kombinationsmög-

lichkeiten sind auf der dritten Anforderungsebene abgebildet (siehe Abbildung 5). Der besseren Lesbarkeit wegen sind die Relationen zwischen dieser 3. Ebene und den Ebenen 2 und 1 in der Abbildung 5 gestrichelt dargestellt. In Tabelle 11 sind für jeden der Aufgabentypen auf Ebene 3 Abkürzung, Beschreibung und Beispiel aufgelistet.

Tabelle 11:

Aufgabentypen der Ebene 3 der Anforderungsstruktur, die über die Kombination der Aufgabentypen der 1. und 2. Ebene definiert sind

Abkürzung	Aufgabentyp	Beispiel
LÜ+N _u	Übertrag in leere Stelle und Null im Subtrahenden	346 - 60
LÜ+N ^o	Übertrag in leere Stelle und Ergänzen zur 10	340 - 66
LÜ+Z	Übertrag in leere Stelle und gleiche Ziffern übereinander	346 - 66
LN ^o +Ü	Ergänzen zur 10 und Übertrag in leere Stelle und Übertrag	304 - 68
LN ^o +Z	Ergänzen zur 10 und Übertrag in leere Stelle und gleiche Ziffern übereinander	307 - 67
L+Z+N _u	Leere Stelle und gleiche Ziffern und Null im Subtrahenden	376 - 70
Z+N _u Ü	Gleiche Ziffern und Übertrag zur Null	424 - 406
L+ZN ^o	Leere Stelle und Ergänzen zur 10 und Übertrag zu gleichen Ziffern übereinander	370 - 76
Z+N _u N ^o	Gleiche Ziffern übereinander und Ergänzen zur 10 und Übertrag zur 0	420 - 406

Diese drei Anforderungsebenen mit ihren 30 Aufgabentypen ergeben die in Abbildung 5 als Hassediagramm dargestellte Anforderungsstruktur. Im Sinne dieser Anforderungsstruktur sollten zuerst Aufgaben mit jeweils einer elementaren Aufgabenanforderung angeboten werden. Erst wenn diese vom Lerner sicher beherrscht werden, sollten Aufgaben präsentiert werden, welche kombinierte elementare Aufgabenanforderungen beinhalten.

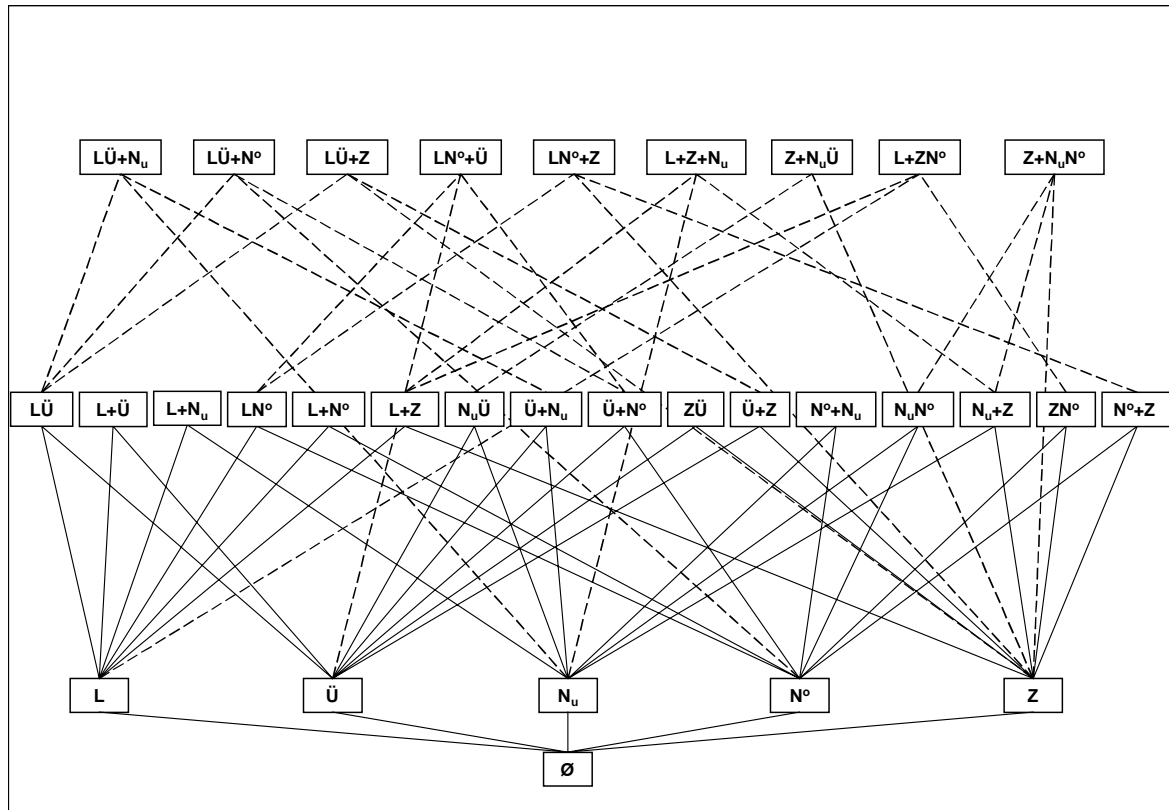


Abbildung 5: Hasse-Diagramm der Anforderungsstruktur zum Gegenstandsbereich der schriftlichen Subtraktion im Zahlenraum bis 1000

L... Leere Stelle, \ddot{U} ...Übertrag, N_u ...Null im Subtrahenden, N° ... Ergänzen zur 10, Z... gleiche Ziffern übereinander

Basierend auf dieser kognitiven Anforderungsanalyse und im Sinne des sequentiellen Vorgehens bei der Vermittlung von Fertigkeiten der schriftlichen Subtraktion setzte sich die vorliegende Arbeit die Entwicklung informativer tutorieller Feedbackformen für die fünf elementaren Aufgabenanforderungen als Ziel. Die Entwicklung von ITF Formen für die Anforderungsebenen 2 und 3 konnte im Rahmen dieser Arbeit nicht realisiert werden, stellt sachlogisch aber den nächsten Schritt dar.

Ein weiteres Argument, welches für die Entwicklung von ITF Formen für Aufgaben, welche jeweils nur eine Aufgabenanforderung enthalten, spricht, kann aus der repair theory von VanLehn (1990) gewonnen werden. Eine zentrale Voraussetzung für die Bereitstellung von computergestütztem ITF repräsentiert die Lokalisierung der fehlerhaft oder unvollständig beherrschten Subprozedur. Nun können aber Aufgaben, welche eine Kombination mehrerer Aufgabenanforderungen beinhalten, eine Diagnose der nicht beherrschten Subprozedur aufgrund der von VanLehn beschriebenen Back-up Repairstrategie (1990, S. 72) erschweren. Die Annahme einer solchen Back-up Repairstrategie basiert auf der Prämisse, dass das Bearbeiten

schriftlicher Subtraktionsaufgaben als hierarchisches Produktionssystem beschrieben werden kann, dessen Subziele im Sinne eines Zielstapels (goal stack) gemanagt werden (VanLehn, 1990, S. 71). Dies impliziert u.a., dass das momentan aktive Ziel von dem Zielstapel entfernt werden kann (pop action) und das darunter liegende Ziel reaktiviert wird. Diese Rekursivität als ein Merkmal des Zielstapels wird durch die Back-up Repairstrategie beschrieben, welche beinhaltet, dass bei Erreichen einer Sackgasse (d.h. einer bestimmten Aufgabenanforderung) der Lerner in seiner Prozedur so weit zurückgeht, bis sich ihm eine andere Lösungsalternative bietet. Die Back-up Repairstrategie soll nachfolgend an einem ausgewählten Beispiel beschrieben werden (siehe Tabelle 12), um daran anschließend kurz deren Konsequenzen für eine Diagnostik nicht beherrschter Subprozeduren zu diskutieren.

Tabelle 12:

Idealisiertes Lernerprotokoll nach VanLehn (1990, S. 73)

a)	$\begin{array}{r} 305 \\ - 167 \\ \hline \end{array}$	In the units column I can't take 7 from 5, so I'll have to borrow.
b)	$\begin{array}{r} 305 \\ - 167 \\ \hline \end{array}$	To borrow, I first have to decrement the next column's top digit. But I can't take 1 from 0!
c)	$\begin{array}{r} 305 \\ - 167 \\ \hline 2 \end{array}$	So I'll go back to doing the units column. I still can't take 7 from 5, so I'll take 5 from 7 instead.
d)	$\begin{array}{r} 21 \\ \cancel{3}05 \\ - 167 \\ \hline 2 \end{array}$	In the tens column I can't take 6 from 0, so I'll have to borrow. I decrement 3 to 2 and add 10 to 0. That's no problem.
e)	$\begin{array}{r} 21 \\ \cancel{3}05 \\ - 167 \\ \hline 142 \end{array}$	Six from 10 is 4. That finishes the tens. The hundreds is easy, there's no need to borrow, and 1 from 2 is 1.

Dieses Beispiel zeigt, dass der Lerner einen Schritt in seinem Lösungsprozess zurückgeht, als er feststellt, dass er über keine adäquate Subprozedur für die Aufgabenanforderung „Leihen von 0“ verfügt, so dass ein Fehler an einer Stelle des Lösungsalgorithmus auftritt, die er problemlos hätte bewältigen können, wenn in der Zehnerspalte keine 0 im Minuenden gestanden hätte. Aufgrund dieses Phänomens sollte computergestütztes ITF, welches fehlerspezifisches Feedback anbietet, vor allem dort effektiv sein, wo sich die nicht beherrschte Subprozedur sicher diagnostizieren lässt.

Da in der Studie von Kühnhold und Padberg (1986, siehe Kapitel 3.2.3) hohe Fehlerhäufigkeiten für den Aufgabentyp „Übertrag in leere Stelle“, welcher die elementaren Aufgabenanforderungen „Übertrag“ und „Leere Stelle“ enthält, gefunden werden konnten, wurde neben den 5 elementaren Aufgabenanforderungen auch für diesen Aufgabentyp fehlerspezifisches ITF entwickelt.

3.2.4.2 Klassifikation systematischer Fehler mit Hilfe einer empirischen Fehleranalyse

In einem nächsten Schritt wurde im Sinne des Designalgorithmus von Narciss (2004) für die identifizierten Aufgabentypen mit ihren Aufgabenanforderungen eine Fehleranalyse durchgeführt, um Typen und Häufigkeiten systematischer Fehler zu identifizieren. Da die Effektivität von Feedback primär über seine korrektive Funktion erklärt wird, sollten im Rahmen einer Fehleranalyse Fehlstrategien, die zu diesen Fehlern geführt haben können, identifiziert werden, um gezielt informatives Feedback bereitstellen zu können. Eine Klassifikation der Fehler und eine Auszählung der Fehlerhäufigkeiten erfolgte im Unterschied zu der Fehlerklassifikation von Gerster (1982) und Kühnhold und Padberg (1986) in Abhängigkeit der jeweiligen Aufgabenanforderungen.

Um Aussagen über systematische Fehler für die ausgewählten Aufgabentypen zu erhalten, wurde ein diagnostischer Test konzipiert, welcher zu jedem der sechs Aufgabentypen drei Aufgaben beinhaltete (siehe Tabelle 13).

Dieser Test wurde im November 1998 an 76 Schüler der 4. Klasse von zwei Dresdner Grundschulen verteilt. Die Schüler hatten den Test während einer Unterrichtsstunde zu bearbeiten, wobei sie gebeten wurden, möglichst alle Aufgaben zu bearbeiten. Insgesamt konnten 178 Fehler beobachtet werden. Dies bedeutet, dass 13% der Aufgaben fehlerhaft bearbeitet wurden. Von diesen 178 Fehlern wurden 14 Fehler (8%) als unsystematisch kategorisiert. Dabei wurde jeder Fehler als unsystematisch gewertet, der keine Fehlstrategie erkennen ließ (auch vermutliche Perseverationsfehler fallen darunter). Sechsmal wurde eine Aufgabe nicht bearbeitet (3%) und 12 Fehler (7%) konnten auf eine fehlerhafte Ausführung des Basisalgorithmus zurückgeführt werden. Hierbei konnte in sieben Fällen (58%) die Fehlstrategie „nicht notwendiger Übertrag“ beobachtet werden. In fünf Fällen (42%) wurde in allen Spalten addiert statt subtrahiert.

Tabelle 13:

Diagnostischer Test zur Klassifikation systematischer Fehler

1	852 <u>- 21</u>	(1)	2	756 <u>- 120</u>	(2)	3	359 <u>- 229</u>	(3)
4	74 <u>- 38</u>	(4)	5	348 <u>- 75</u>	(5)	6	450 <u>- 327</u>	(6)
7	148 <u>- 34</u>	(1)	8	957 <u>- 504</u>	(2)	9	660 <u>- 450</u>	(3)
10	345 <u>- 172</u>	(4)	11	437 <u>- 61</u>	(5)	12	706 <u>- 243</u>	(6)
13	539 <u>- 16</u>	(1)	14	473 <u>- 201</u>	(2)	15	362 <u>- 341</u>	(3)
16	638 <u>- 465</u>	(4)	17	673 <u>- 92</u>	(5)	18	280 <u>- 156</u>	(6)

	Aufgabenanforderungen	In Aufgabe Nr.:
(1)	Leere Stelle	1, 7, 13
(2)	Null im Subtrahenden	2, 8, 14
(3)	Gleiche Ziffern übereinander	3, 9, 15
(4)	Übertrag	4, 10, 16
(5)	Übertrag in leere Stelle	5, 11, 17
(6)	Null im Minuenden (Ergänzen zur 10)	6, 12, 18

82% aller Fehler lassen sich auf systematische Fehler bei den sechs Aufgabentypen zurückführen. Dieses Ergebnis repliziert die von Gerster (1982) und Radatz (1980) berichteten systematischen Fehleranteile zwischen 70% und 90% am Gesamtfehleranteil. In Tabelle 14 findet sich für jeden Aufgabentyp aufgelistet:

- wie viele der möglichen Aufgabenlösungen (bei drei Aufgaben und 76 Schülern entspricht dies 228 Aufgabenlösungen) sich durch den Einsatz einer Fehlstrategie auszeichneten (Anzahl systematischer Fehler / Anzahl Aufgabenlösungen)
- welche systematischen Fehler wie häufig beobachtet werden konnten, relativiert auf den Gesamtanteil systematischer Fehler, die bei diesem Aufgabentyp beobachtet werden konnten (Anzahl spezifischer systematischer Fehler / Anzahl systematischer Gesamtfehler)

Die Ergebnisse der Fehleranalyse zeigen, dass in Abhängigkeit des Aufgabentyps bei den Schülern unterschiedliche systematische Fehler in unterschiedlicher Auftretenshäufigkeit zu beobachten waren. So konnte bei dem Aufgabentyp „Übertrag in Leerstelle“ als häufigster systematischer Fehler der Fehler „Übertrag zur oberen Ziffer addiert“ beobachtet werden.

Dieser systematische Fehler zeigte sich bei den beiden anderen Aufgabentypen, die einen Übertrag erforderten, „Übertrag“ und „Ergänzen zur 10“, gar nicht bzw. nur einmal. Ein weiteres Ergebnis der Fehleranalyse ist, dass die verschiedenen Aufgabentypen unterschiedliche Schwierigkeitsgrade beinhalten. So scheint vor allem der Aufgabentyp „Ergänzen zur 10“ erhöhte Anforderungen an den Schüler zu stellen. Schüler, die Subtraktionsaufgaben mit der Aufgabenanforderung „Übertrag“ korrekt bearbeiteten, zeigten beispielsweise Probleme bei der Bearbeitung von Subtraktionsaufgaben mit der Aufgabenanforderung „Ergänzen zur 10“. Diese Ergebnisse sprechen für eine Fehlerklassifikation auf der Grundlage von Aufgabenanforderungen und unterstreichen somit noch einmal die Kritik an den Fehlerklassifikationen von Gerster (1982) und Kühnhold und Padberg (1986), deren Fehlerkategorien über alle Aufgabenanforderungen gebildet wurden.

Tabelle 14:

Typen systematischer Fehler und deren Auftretenshäufigkeit, dargestellt für jeden Aufgabentyp

Aufgabentyp	Beispiel	Anzahl syst. Fehler (%)	Spez. syst. Fehler	Anzahl spez. syst. Fehler (%)
Leere Stelle	852 - <u>21</u>	9 (3,95%)	= ?31 → Ziffer über leerer Stelle nicht berücksichtigt	• 9 (100%)
Null im Subtrahenden	756 - <u>120</u>	10 (4,39%)	= 630 → Null im Ergebnis = 624 → Falsche Rechenrichtung ^a mit Übertrag = 634 → Falsche Rechenrichtung ohne Übertrag	• 2 (20%) • 4 (40%) • 4 (40%)
Gleiche Ziffern übereinander (in erster Spalte)	359 - <u>229</u> 362 - <u>341</u>	15 (6,58%)	= 120 → Übertrag bei gleichen Ziffern = 128 → Addiert statt subtrahiert in Spalte mit gleichen Ziffern = 139 → Gleiche Ziffer bei gleichen Ziffern = 021 → Null in erster Spalte	• 7 (46,67%) • 3 (20%) • 2 (13,33%) • 3 (20%)
Übertrag	74 - <u>38</u>	28 (12,28%)	= 44 → Falsche Rechenrichtung ohne Übertrag = 46 → Übertrag nicht berücksichtigt = 34 → Falsche Rechenrichtung mit Übertrag = 32 → Addiert statt subtrahiert in Übertragsspalte	• 13 (46,43%) • 12 (42,86%) • 2 (7,14%) • 1 (3,57%)
Übertrag in Leerstelle	348 - <u>75</u>	36 (15,79%)	= 473 → Übertrag zur oberen Ziffer addiert = 373 → Übertrag nicht berücksichtigt = 333 → Falsche Rechenrichtung ohne Übertrag = 173 → Bei Übertrag in leere Stelle zwei Überträge abgezogen = ?73 → Ziffer über leerer Stelle nicht berücksichtigt	• 16 (44,44%) • 9 (25%) • 8 (22,22%) • 2 (5,55%) • 1 (2,78%)
Ergänzen zur 10	450 - <u>327</u>	48 (21,05%)	= 133 → Übertrag nicht berücksichtigt = 137 → Falsche Rechenrichtung ohne Übertrag = 130 → Null im Ergebnis = 143 → Übertrag zur oberen Ziffer addiert = 127 → Falsche Rechenrichtung mit Übertrag	• 24 (50%) • 14 (29,17%) • 8 (16,67%) • 1 (2,08%) • 1 (2,08%)

^aFalsche Rechenrichtung bedeutet das Ergänzen von der oberen zur unteren Ziffer

3.3 Entwicklung von fehlerspezifischem ITF für das schriftliche Subtrahieren

Mit dem Wissen über typische Aufgabenanforderungen und den bei diesen Aufgabenanforderungen auftretenden systematischen Fehlern konnten in einem nächsten Schritt relevante fehlerspezifische Feedbackinformationen ausgewählt werden. Welche fehlerspezifischen Informationen ausgewählt wurden, wird in Kapitel 3.3.1 beschrieben. Nach der Auswahl relevanter fehlerspezifischer Informationen wurde deren Präsentationsmodus festgelegt. Dies wird in Kapitel 3.3.2 beschrieben. In Kapitel 3.3.3 erfolgt eine detaillierte Vorstellung des letztendlich resultierenden fehlerspezifischen ITF Algorithmus.

3.3.1 Selektion der ITF-Inhalte

Um eine Fehlerkorrektur und den Erwerb korrekten algorithmischen Wissens zu unterstützen, wurden in einem nächsten Schritt die Informationen ausgewählt, die bei einer fehlerhaften Aufgabenbearbeitung angeboten werden sollen. Dazu wurde das Wissen über die Entstehung von Fehlern in der schriftlichen Subtraktion sowie die aus der Voruntersuchung gewonnenen Erkenntnisse zu Aufgabenanforderungen und den bei diesen Aufgabenanforderungen auftretenden systematischen Fehler genutzt.

Eine zentrale Rolle im Hinblick auf die Entscheidung, welche Feedbackinformationen den Schülern bei einer fehlerhaften Aufgabenlösung präsentiert werden, spielt der algorithmische Charakter der schriftlichen Subtraktion, welcher definiert werden kann als „a set of syntactic rules that prescribe how problems should be written, an order in which certain operations must be performed, and which kinds of symbols belong in which positions.“ (Resnick, 1982, S.137). Nach Resnick (1982) ist ein Verständnis der dem Algorithmus zugrundeliegenden Annahmen für ein korrektes Bearbeiten algorithmischer Aufgaben nicht zwingend erforderlich. Dass das Wissen darüber, warum so gerechnet wird, die Bearbeitung prozeduraler Aufgaben unterstützt, gilt allerdings als unbestritten (Rittle-Johnson, Siegler & Alibali, 2001). So konnte bei Schülern, die über eine Einsicht in den Algorithmus verfügten, eine niedrigere Fehlerrate beobachtet werden (Mosel-Göbel, 1988) - ein Befund, welcher auch in der Studie von Brownell (1947, zitiert nach Sherrill, 1979) gefunden werden konnte. Von daher war und ist es ein Ziel wissenschaftlicher Forschung, über das Fördern des Verstehens mathematischer Strukturen den Erwerb korrekten algorithmischen Wissens zu unterstützen. Ob Schüler die dem Subtraktionsalgorithmus zugrundeliegenden mathematischen Strukturen verstehen, hängt vor allem von dem gelehrt Verfahren ab. So berichtet Mosel-Göbel (1988) bezüglich des

Verständnisses des gelehrten Verfahrens eine deutliche Überlegenheit des Borgeverfahrens gegenüber dem Erweiterungsverfahren. Während bei dem Erweiterungsverfahren nur 10% die Stufe des „begründeten Verständnisses“ erreichten, waren dies bei dem Borgeverfahren rund 80% der Schüler. Wie bereits im Kapitel 3.2.1 dargestellt, ist vor allem das Gesetz der Konstanz der Differenz so kompliziert, dass nur wenige Kinder es wirklich verstehen. Auch hinsichtlich der Frage, ob Kinder das Verfahren selbständig entdecken können, konnte eine Überlegenheit des Borgeverfahrens gegenüber dem Erweiterungsverfahren nachgewiesen werden. So entdeckte jeder dritte Schüler die Borgetechnik selbständig, hingegen konnte keiner der untersuchten Schüler die Erweiterungstechnik selbständig entdecken.

Diese Erkenntnisse führten zu der Entscheidung, bei systematischen Fehlern, die auf ein Nichtbeherrschen der Erweiterungstechnik zurückgeführt werden können, auf die Bereitstellung fehlerspezifischer Informationen, die auf die Vermittlung eines Verständnisses der dem Algorithmus zugrundeliegenden Annahmen fokussieren, zu verzichten. Fehlerspezifisches ITF sollte hingegen Informationen bereitstellen, die über die Diagnose der nicht beherrschten Aufgabenanforderung und des systematischen Fehlers den Erwerb der korrekten Subprozedur unterstützen, ohne primär auf ein Verstehen der dem Algorithmus zugrundeliegenden Annahmen abzuzielen.

Als relevante Informationen, die eine solche Fehlerkorrektur unterstützen, wurden daher folgende Informationen identifiziert:

- Angabe des Fehlerortes über farbliche Markierung der Fehlerstelle
- Präsentation von Informationen, die den Erwerb konditionalen Wissens unterstützen, d.h. unter welchen Bedingungen ist die Subprozedur anzuwenden
- Präsentation von Informationen, die den Erwerb der einzelnen Prozedurschritte und deren Reihenfolge unterstützen, dazu gehören die Vermittlung der korrekten Regel und die Demonstration des korrekten Lösungsweges
- Präsentation von Informationen, die den Erwerb von Monitoringfertigkeiten bei der Prozedurausführung unterstützen, beispielsweise der Hinweis auf die Umkehrrechnung
- Präsentation von Informationen über ausgearbeitete Lösungsbeispiele
- Präsentation der richtigen Lösung

Diese Informationen wurden genutzt, um in Abhängigkeit der diagnostizierten nicht beherrschten Aufgabenanforderung (z.B. Aufgabe mit Übertrag) und des diagnostizierten sy-

stematischen Fehlers (z.B. Übertrag nicht berücksichtigt) fehlerspezifisches ITF zu entwickeln.

3.3.2 Selektion von Form und Modus der ITF-Präsentation

Da die widersprüchlichen Ergebnisse einer Reihe von Studien, in denen verschieden komplexe Feedbackformen auf ihre Wirksamkeit hin verglichen wurden, gezeigt haben, dass die Identifikation relevanter Feedbackinformationen zwar eine notwendige, aber keine hinreichende Bedingung für die Effektivität von Feedback darstellt, wurde in einem nächsten Schritt definiert, wie diese ausgewählten Informationen präsentiert werden sollen. Dafür wurden die von Narciss (2004; Narciss & Huth, 2004) abgeleiteten Präsentationsprinzipien für ITF, welche im Kapitel 2.4.3 ausführlich vorgestellt wurden, genutzt.

Für die Präsentation der fehlerspezifischen Informationen bedeutet dies, dass diese dem Lerner nach einer fehlerhaften Aufgabenlösung ohne das Angebot der korrekten Lösung bereitgestellt werden. Der Lernende soll sich auf diese Weise dazu aufgefordert sehen, seine Aufgabe unter Anwendung der bereitgestellten fehlerspezifischen Informationen eigenständig zu korrigieren. Die korrekte Lösung wird dem Lernenden erst nach einem dritten fehlerhaften Lösungsversuch präsentiert. Eine weitere Möglichkeit, eine aktive Verarbeitung der Informationen anzuregen, wird in der Definition eines mastery levels gesehen, welches über ein Lernkriterium operationalisiert ist, das dem Lernenden vor der Aufgabenbearbeitung mitgeteilt wird. Die Festsetzung eines Lernkriteriums verfolgt das Ziel, den Lernenden zu einer intensiveren Informationsverarbeitung zu motivieren, da ein Ende der Übungseinheit nur über eine erfolgreiche Aufgabenbearbeitung zu erreichen ist.

Eine weitere Entscheidung, welche in dem Kontext der Informationsbereitstellung getroffen werden musste, betraf den Aspekt der system- oder lernerinitiierten Informationsbereitstellung. Das heißt, sollten die fehlerspezifischen Informationen nach jedem fehlerhaften Versuch vom System initiiert bereitgestellt werden, oder sollte der Lerner die fehlerspezifischen Informationen im Sinne eines on-demand Feedback lernerkontrolliert anfordern können. Für die Entwicklung von ITF für die schriftliche Subtraktion wurde sich für ein systeminitiiertes Feedback, welches nach jeder fehlerhaften Aufgabenbearbeitung unaufgefordert präsentiert wird, aus folgenden Gründen entschieden:

- In Kapitel 2.3.2.2 wurden verschiedene Studien zitiert, die gezeigt haben, dass on-demand Feedback von vielen Lernern nicht effizient genutzt wird, d.h. dass Lerner on-demand Feedback nur in seltenen Fällen anfordern.

- Aus der Forschung zu help seeking liegen Befunde vor, die eine positive Korrelation zwischen help seeking Fertigkeiten und Alter belegen (Newman & Schwager, 1995). So zeigen Jugendliche aufgrund besser entwickelter metakognitiver Fertigkeiten ein angemesseneres Hilfesuchverhalten als Kinder, welche die Zielpopulation dieser Arbeit repräsentieren.

Alle Rückmeldungen werden sowohl auditiv als auch visuell präsentiert. Für die kombinierte auditive und visuelle Präsentation sprachlicher Rückmeldungen wurde sich neben den bereits aufgeführten Gründen (Mayer & Moreno, 2002, siehe Kapitel 2.4.3) aus folgendem weiteren Grund entschieden: Nach Paechter (1996) sollten vor allem Lerner mit geringen Lesefertigkeiten – z.B. Grundschüler - von einer auditiven Informationsdarbietung profitieren. Der offensichtliche Nachteil, dass auditive Texte flüchtige Informationsangebote repräsentieren, bei denen die Lernenden ihr Lerntempo nicht selbst bestimmen können, wird darüber zu kontrollieren versucht, dass die auditiven Texte in einer Kurzversion visuell präsentiert werden.

3.3.3 Resultierender fehlerspezifischer ITF Algorithmus

Die Implementierung der ausgewählten Feedbackinformationen unter Berücksichtigung der aufgestellten Feedback-Präsentationsprinzipien resultierte in dem in Abbildung 6 dargestellten dreistufigen fehlerspezifischen ITF Algorithmus, welcher an dieser Stelle ausführlich vorgestellt werden soll. Der ITF Algorithmus beginnt mit der Präsentation einer Aufgabe mit einer spezifischen Aufgabenanforderung, beispielweise mit einer Übertragsaufgabe. Der Lerner hat maximal drei Lösungsversuche, an deren Anschluss ihm jeweils unterschiedlich elaboriertes Feedback dargeboten wird.

(1) Nach dem ersten fehlerhaften Lösungsversuch erhält der Lerner *Knowledge of Response* (KR) Feedback mit der Aufforderung, es noch einmal zu versuchen. Die Entscheidung, nach einem ersten fehlerhaften Lösungsversuch noch keine zusätzlichen Informationen bereitzustellen, hatte folgende zwei Gründe: (a) das Auftreten unsystematischer Fehler wie Flüchtigkeitsfehler bedarf keiner weiteren korrektiven Informationen und kann über die Aufforderung „Versuchs noch einmal“ abgefangen werden, (b) das Auftreten systematischer Fehler impliziert nicht, dass die korrekte Lösung nicht doch noch selbständig gefunden werden kann (siehe Kapitel 3.1.2, „bug migration“ Phänomen).

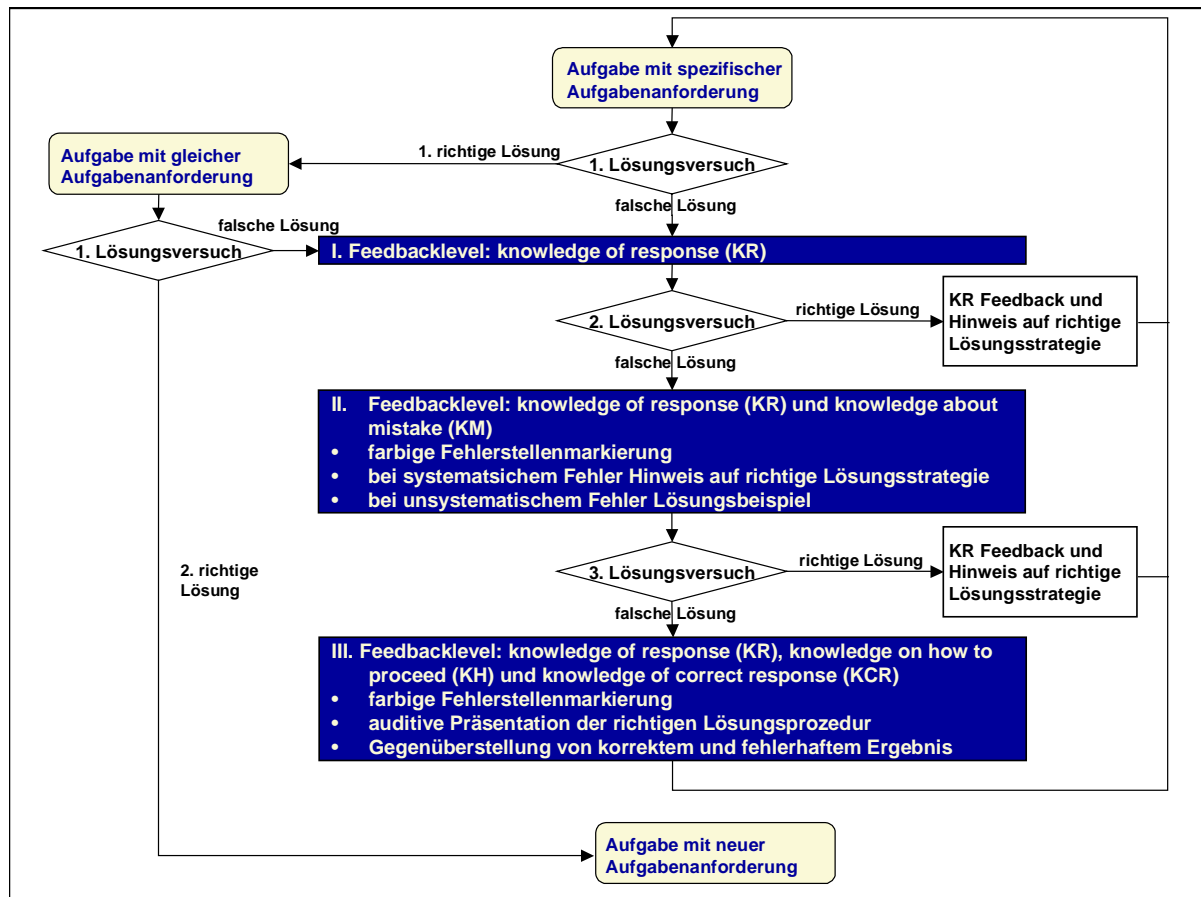


Abbildung 6: Fehlerspezifischer ITF Algorithmus für die schriftliche Subtraktion

(2) Nach dem zweiten fehlerhaften Lösungsversuch erhält der Lerner *Knowledge of Response* (KR) Feedback und *Knowledge about Mistake* (KM) Feedback. Das heißt, er bekommt mitgeteilt, dass seine Lösung noch einen Fehler enthält (KR). Zudem wird der Fehler-ort farblich markiert. Diagnostiziert das Programm einen systematischen Fehler, erfolgt in Abhängigkeit der diagnostizierten Fehlstrategie ein Hinweis auf die korrekte Lösungsstrategie. Dieser Hinweis beinhaltet zum einen konditionales Wissen - unter welchen Bedingungen ist die jeweilige Subprozedur anzuwenden - und zum anderen die für die Beherrschung der Subprozedur erforderlichen einzelnen Prozedurschritte. Handelt es sich um einen unsystematischen Fehler, d.h. kann das Programm die fehlerhafte Lösung auf keine der in der Vorstudie eruierten Fehlstrategien zurückführen, erhält der Lerner als Hinweis die Aufforderung, den korrekten Lösungsalgorithmus für diesen Aufgabentyp anhand eines ausgearbeiteten Lösungsbeispiels (worked-out example) zu erwerben.

(3) Nach einem dritten fehlerhaften Lösungsversuch erhält der Lerner *Knowledge of Response* (KR) Feedback, *Knowledge on how to proceed* (KH) Feedback und *Knowledge of*

Correct Response (KCR) Feedback. Der Fehlerort wird farblich markiert. Die korrekte Lösungsprozedur wird unter Berücksichtigung der nicht beherrschten Aufgabenanforderung, aber unabhängig von dem diagnostizierten Fehler (systematisch oder unsystematisch) an der korrekten Aufgabe visuell-auditiv demonstriert. Das korrekte Ergebnis wird sukzessiv präsentiert, ohne dass die fehlerhafte Schülerlösung gelöscht wird, so dass der Lerner seine Lösung mit der Programmlösung vergleichen kann.

Korrigiert der Lerner seine Aufgabe nach einem fehlerhaften Versuch erfolgreich, werden in Abhängigkeit des diagnostizierten systematischen Fehlers die korrekten Prozedurschritte für diese Aufgabenanforderung in der Rückmeldung noch einmal präsentiert. Die nochmalige Präsentation der korrekten Lösungsprozedur sollte eine Festigung derjenigen unterstützen. Eine Aufgabe mit einer neuen Aufgabenanforderung wird im Sinne des ziel erreichenden Lernens (mastery learning) präsentiert, wenn der Lerner zwei Aufgaben hintereinander im ersten Lösungsversuch korrekt gelöst hat.

3.4 Zusammenfassung

In diesem Kapitel wurde dargelegt, wie unter Anwendung der Gestaltungsprinzipien für ITF (Narciss, 2004; Narciss & Huth, 2004) für das Themengebiet der schriftlichen Subtraktion fehlerspezifisches ITF theoriegeleitet entwickelt wurde. Hierzu wurden in einem ersten Schritt auf der Grundlage von Ansätzen zur Wissensstrukturierung und Wissensdiagnose (Doignon & Falmange, 1985; Korossy, 1996) Anforderungen definiert, welche die Bearbeitung von schriftlichen Subtraktionsaufgaben an den Lerner stellt. Dabei wurde herausgearbeitet, dass unterschiedliche Anforderungsprofile nicht nur von der Art der in der Aufgabe enthaltenen Aufgabenanforderungen, sondern auch von der Anzahl der Aufgabenanforderungen und deren Kombination determiniert werden. In einem zweiten Schritt wurden für einen Ausschnitt dieser Anforderungen systematische Fehler identifiziert, die bei der Bearbeitung von Aufgaben mit diesen Anforderungen häufig zu beobachten sind. Im Unterschied zu den Arbeiten von Gerster (1982) und Kühnhold und Padberg (1986) wurden die beobachteten systematischen Fehler entsprechend den diesen Fehlern zugrundeliegenden nicht beherrschten Aufgabenanforderungen klassifiziert. Die auf der Grundlage dieser Anforderungs- und Fehleranalyse identifizierten Feedbackinformationen wurden abschließend unter Nutzung der formalen Gestaltungsprinzipien in einen Feedbackalgorithmus für die schriftliche Subtraktion implementiert.

4 EVALUATION DES FEHLERSPEZIFISCHEN ITFs IN SEINEM EINFLUSS AUF LERN- UND MOTIVATIONSPARAMETER – STUDIE 1

In der Forschungstradition zu den Wirkungen unterschiedlicher Feedbackarten wurden und werden in der Regel *entweder* kognitive *oder* motivationale Effekte untersucht. Darüber hinaus wurden in vielen dieser Studien Feedbackarten eingesetzt, die *eher* intuitiv *als* systematisch konstruiert wurden.

Mit der Entwicklung von fehlerspezifischem ITF für die schriftliche Subtraktion unter Anwendung der dargestellten Gestaltungs- und Evaluationsprinzipien bestand das zentrale Ziel der vorliegenden Studie darin, das fehlerspezifische ITF *sowohl* im Hinblick auf seine Lern- *als auch* seine motivationale Wirksamkeit zu untersuchen. Evaluationsergebnisse zu den in Kapitel 2 dargestellten Gestaltungs- und Evaluationsprinzipien liegen bisher nur im Rahmen eines experimentellen Konzeptlernparadigmas vor (Narciss, 1999, 2001, 2004). Die Ergebnisse dieser Konzepterwerbsstudien belegen, dass ITF nicht nur kognitive, sondern vor allem auch motivationale Effekte hat. Die vorliegende Studie liefert darüber hinaus Evaluationsergebnisse im Kontext einer schulischen Lernaufgabe – der schriftlichen Subtraktion.

4.1 Fragestellungen und Forschungshypothesen

Um hinsichtlich der Lern- und motivationalen Wirksamkeit des fehlerspezifischen ITFs valide Aussagen zu erhalten, wurden im Sinne des Multitrait-Multimethod Ansatzes nach Campbell und Fiske (1959) verschiedene beobachtbare und berichtbare Maße hypothesen-geleitet erhoben und überprüft. Dabei erfolgte die Formulierung der Fragestellungen und Hypothesen in Anlehnung an die von Narciss (2004) abgeleiteten Indikatoren für die Untersuchung von kognitiven, meta-kognitiven und motivationalen Feedbackeffekten (siehe Kapitel 2.5). Diese Überprüfung hinsichtlich des Einflusses des entwickelten ITFs erfolgte sowohl bezüglich beobachtbarer als auch berichtbarer Indikatoren einer Lern- und motivationalen Wirksamkeit von ITF.

Mit dem Ziel, Aussagen zu kognitiven Effekten des entwickelten ITF Algorithmus zu erhalten, wurde Fragestellung 1 formuliert:

Fragestellung 1: Unterstützt fehlerspezifisches ITF den Erwerb von Fertigkeiten der schriftlichen Subtraktion?

Die theoriegeleitete Entwicklung des fehlerspezifischen ITF für die schriftliche Subtraktion (siehe Kapitel 3) sollte die Identifikation korrekturrelevanter Informationen ermöglichen haben. Durch die Bereitstellung der als relevant identifizierten Informationen ohne dem Angebot der korrekten Lösung, verknüpft mit der Aufforderung einer nochmaligen Aufgabenbearbeitung, sollte eine aktive und vertiefte Auseinandersetzung mit diesen Informationen angeregt worden sein. Diese vertiefte Auseinandersetzung sollte zudem durch die Implementierung eines mastery levels weiter gefördert worden sein, so dass positive Feedbackeffekte auf alle erfassten Indikatoren zu erwarten sind.

Als beobachtbare Indikatoren wurden die Anzahl von Aufgaben, bei denen das mastery level erreicht wurde, die Anzahl im Nachtest korrekt gelöster Aufgaben sowie die Anzahl im Nachtest korrigierter Vortestaufgaben erhoben. Als berichtbarer Indikator wurde eine Einschätzung des wahrgenommenen Nutzens der angebotenen Feedbackinformationen erfasst.

Hypothese 1a: Da ITF von dem Lerner nach einer fehlerhaften Aufgabenbearbeitung eine nochmalige Bearbeitung einer Aufgabe mit einem identischen Anforderungsprofil verlangt, sollte sich ein lernförderlicher Einfluss bereits während der Aufgabenbearbeitung beobachten lassen. Die Lernenden mit fehlerspezifischem ITF sollten das gesetzte mastery level (zwei Aufgaben im 1. Lösungsversuch hintereinander korrekt bearbeiten) signifikant häufiger erreichen als Lernende, welche weniger informatives Feedback erhalten.

Hypothese 1b: Lernende, welche fehlerspezifisches ITF erhalten, sollten nicht nur während des Treatments fehlerhaft gelöste Aufgaben häufiger korrekt bearbeiten, sondern sollten über die Vermittlung von Lösungsstrategien Aufgaben zur schriftlichen Subtraktion auch nach dem Treatment signifikant häufiger korrekt lösen als Lernende, welche weniger informatives Feedback erhalten.

Hypothese 1c: Lernende, welche fehlerspezifisches ITF erhalten, sollten die im ITF bereitgestellten Informationen als hilfreich bewerten, da sie die Erfahrung machen, dass sie über das Angebot fehlerspezifischer Informationen sowohl bei der Korrektur fehlerhaft bearbeiteter Aufgaben als auch bei dem Erwerb von Fertigkeiten der schriftlichen Subtraktion unterstützt werden.

Neben der Untersuchung der Lernwirksamkeit des fehlerspezifischen ITFs besteht ein weiteres Anliegen dieser Studie darin, den Einfluss des ITFs auf post-aktionale motivationale

Parameter zu untersuchen. Zentral für die von Narciss (2004) formulierten motivationalen Funktionen von ITF (siehe Kapitel 2.5) ist die Annahme, dass ITF über die Bereitstellung korrekturrelevanter Informationen ohne der unmittelbaren Präsentation der richtigen Lösung Gelegenheiten bietet, Lernerfolge zu erleben, die intern attribuiert werden können. Das Erleben von Lernerfolg und dessen interne Attribution sollten zu einer positiven Kompetenzwahrnehmung führen, welche wiederum in einer positiven Selbstbewertung der eigenen Leistung zum Ausdruck kommen sollte. Narciss (2004) bezeichnet die aus dem ITF resultierenden motivationalen Effekte zusammengefasst als post-aktionales Kompetenzerleben. Indikatoren für das post-aktionale Kompetenzerleben sind nach Narciss (2004) die Zufriedenheit mit der eigenen Leistung, die Bewertung der eigenen Leistung und des Lernfortschritts sowie die wahrgenommene Aufgabenschwierigkeit. Da in vielen Motivationstheorien ein enger Zusammenhang zwischen dem Kompetenzerleben und intrinsischen Anreizwerten postuliert wird (z.B. die kognitive Evaluations-Theorie, Deci & Ryan, 1985; die Interessentheorie, Krapp, 1999; die Flow-Theorie, Csikszentmihalyi, 1975; siehe zusammenfassend Narciss, 2004), wird darüber hinaus erwartet, dass das Erleben von Kompetenz auch mit einer erhöhten intrinsischen Motivation einhergeht. Eine zweite Fragestellung lautete daher:

Fragestellung 2: Beeinflusst das Angebot von fehlerspezifischem ITF neben Lern- und Leistungsparametern auch post-aktionale motivationale Parameter?

Hypothese 2: Aufgrund der dargestellten motivationalen Funktionen des ITFs sollte das fehlerspezifische ITF in einem erhöhten post-aktionalen Kompetenzerleben und einer erhöhten intrinsischen Motivation resultieren. Das erhöhte Kompetenzerleben sollte darin zum Ausdruck kommen, dass die Lerner zufriedener mit ihrer Leistung sind, die Aufgaben als leichter und weniger anstrengend wahrnehmen, ihre Leistungen positiver bewerten, Lernfortschritte berichten und Freude beim Bearbeiten der Aufgaben haben.

Da die beobachtbaren Indikatoren für motivationale Effekte (z.B. Steigerung der Anstrengungsbereitschaft) nicht unabhängig von beobachtbaren Indikatoren für kognitive Effekte sind (Narciss, 2004) und Ausdauer aufgrund der festgelegten Treatmentzeit keinen sinnvollen Indikator repräsentiert, wurden zur Überprüfung der Fragestellung 2 nur berichtbare Indikatoren erhoben.

4.2 Methode

Zur Beantwortung der Fragestellungen und empirischen Prüfung der Forschungshypothesen wurde zu Schuljahresbeginn im September 2000 eine experimentelle Untersuchung mit Schülern der vierten Jahrgangsstufe aus vier Dresdner Grundschulen durchgeführt. Obwohl der Erwerb des Verfahrens der schriftlichen Subtraktion bereits im Lehrplan der dritten Jahrgangsstufe verankert ist, wurden für diese Untersuchung Schüler der vierten Jahrgangsstufe ausgewählt, da die Schüler bereits grundlegende Rechenkompetenzen erworben haben sollten. Diese Forderung resultierte aus der Überlegung, dass systematische Fehler vor allem in der zweiten Phase des kognitiven Fertigkeitserwerbs zu beobachten sind, wenn der Lerner bereits über grundlegendes Wissen verfügt (VanLehn, 1996, siehe Kapitel 3.2.4.1).

Um die Forschungshypothesen überprüfen zu können, wurde neben der Experimentalbedingung „fehlerspezifisches ITF“ eine Vergleichsbedingung realisiert, welche weniger informatives Feedback, sogenanntes „KR-KCR Feedback“, bereitstellt. Dieser Vergleichs-Feedbackalgorithmus wird im Kapitel 4.2.2 ausführlich beschrieben. Um eine experimentelle Kontrolle der unabhängigen Variablen zu garantieren und Versuchsleitereffekte möglichst auszuschließen, wurden beide Feedbackalgorithmen computergestützt implementiert. Eine Beschreibung der Software erfolgt im Kapitel 4.2.3.

4.2.1 Versuchspersonen

Die an der Evaluationsstudie teilnehmenden Schüler wurden auf der Basis eines Vortests, welcher zwei Wochen vor der eigentlichen Evaluation von 131 Schülern der vierten Jahrgangsstufe im Verlauf einer Unterrichtsstunde bearbeitet wurde, ausgewählt. Der Vortest beinhaltete 32 Aufgaben mit den sechs Aufgabenmerkmalen (siehe Tabelle 16 und Anhang A).

Da ITF vor allem über die Korrektur von Fehlern wirken soll, bestand ein Ziel dieses Vortestes darin, diejenigen Schüler auszuwählen, die das Verfahren der schriftlichen Subtraktion noch nicht sicher beherrschten und daher in der Aufgabenbearbeitung viele Fehler zeigten. Für die Teilnahme an der Studie wurden die Schüler ausgewählt, die mehr als 50% der im Vortest zu bearbeitenden Aufgaben fehlerhaft gelöst hatten. Auf der Basis dieses Selektionskriteriums konnten 38 Schüler ($w = 16$; $m = 22$) rekrutiert werden, die zum Zeitpunkt der Untersuchung im Durchschnitt 9.4 ($SD = 0.50$) Jahre alt waren.

4.2.2 Untersuchungsdesign

Im Hinblick auf die vorliegenden Fragestellungen wurde ein Vortest-Nachtest-Vergleichsgruppen-Plan mit folgenden unabhängigen Variablen, Kontrollvariablen und abhängigen Variablen realisiert:

- Die unabhängige Variable stellte die Feedbackbedingung dar, welche zweifach gestuft war.
 - Fehlerspezifisches Informatives Tutorielles Feedback (enthält KR-KM-KH-KCR Feedback)
(fehlerspezifisches ITF)
 - Knowledge of Response und Knowledge of Correct Response Feedback
(KR-KCR Feedback)

Die Beschreibung des KR-KCR Feedbackalgorithmus erfolgt mit der Abbildung 7 in diesem Kapitel.

- Als Kontrollvariablen wurden (1) das aufgabenspezifische Vorwissen, (2) das aufgabenspezifische Selbstkonzept sowie (3) die Computereinstellung und die Computer-/Lernsoftwareerfahrung über einen Vortest und eine Vorbefragung erfasst.
- Abhängige Variablen, die den Einfluss des Feedbacks auf kognitive Parameter erfassten, repräsentierten folgende vier Leistungsindikatoren: Treatmentleistung, Nachtestleistung, der Korrekturwahrscheinlichkeitsindex im Nachtest sowie die Einschätzung des Feedbacknutzens. Der Einfluss des Feedbacks auf die post-aktionalen motivationalen Parameter wurde mit Hilfe eines Fragebogens mit sechs Items erfasst.

Der dieser Untersuchung zugrundeliegende Vortest-Nachtest-Vergleichsgruppen-Plan lässt sich folgendermaßen (siehe Tabelle 15) darstellen:

Tabelle 15:

Vortest-Nachtest-Vergleichsgruppen-Plan

	Vortest	Treatment	Nachtest
Fehlerspezifisches ITF (Experimentalgruppe)	32 Subtraktions- aufgaben	Fehlerspezifisches ITF	32 Subtraktions- aufgaben
KR-KCR Feedback (Vergleichsgruppe)		KR-KCR Feedback	

Um intern valide Aussagen hinsichtlich der Wirksamkeit des fehlerspezifischen ITFs auf Leistungs- und motivationale Parameter treffen zu können, war eine Kontrolle bekannter Störfaktoren auf dem Wege der Konstanthaltung (z.B. instruktionaler Kontext) und der Kontrolle unbekannter Störfaktoren auf dem Wege der Randomisierung erforderlich (Hager, Patry & Brezing, 2000). Dieses Anliegen machte die Realisierung einer Vergleichsgruppe erforderlich, welche im wesentlichen die Kriterien einer vergleichenden Evaluation erfüllte. Danach sollte eine Vergleichsgruppe bis auf das spezifische Treatment die gleiche Intervention erfahren wie die Experimentalgruppe. Das Treatment der Vergleichsgruppe sollte dabei die gleichen Ziele wie das Treatment der Experimentalgruppe verfolgen, allerdings über den Einsatz anderer Vorgehensweisen, z.B. anderer didaktischer Methoden (Hager, Patry & Brezing, 2000). Eine weitere Anforderung an eine Vergleichsgruppe formuliert Hager dahingehend, dass das Treatment der Vergleichsgruppe seine isolierte Wirksamkeit bereits mehrfach und unter verschiedenen Randbedingungen unter Beweis gestellt haben sollte.

Diese Vergleichsbedingung sollte zudem dem Anspruch externer Validität genügen, um praktisch bedeutsame Aussagen bezüglich der Effektivität des entwickelten ITF zu ermöglichen. Eine Analyse kommerzieller Mathematiklernsoftware zum Thema „Schriftliches Subtrahieren“ hatte gezeigt, dass Feedback, welches im Anschluss an eine fehlerhafte Aufgabebearbeitung bereitgestellt wird, in aller Regel zweistufig ist. In einer ersten Feedbackstufe wird der Lerner darauf hingewiesen, dass seine Aufgabenlösung noch fehlerhaft ist (KR Feedback) und er erhält die Aufforderung, es noch einmal zu versuchen. Nach einem zweiten fehlerhaften Lösungsversuch wird ihm dann die richtige Lösung präsentiert (KCR Feedback). Eine Analyse der fehlerhaften Lösung, d.h. welche Fehlstrategie eventuell eingesetzt wurde bzw. welche Teilprozedur nicht beherrscht wird, erfolgt nicht. Der Lerner hat maximal die Option, Hilfe über das Aufrufen eines ausgearbeiteten Lösungsbeispiels zu erhalten. In Anlehnung an diese Analyse wurde als Vergleichsbedingung der in Abbildung 7 dargestellte KR-KCR Feedbackalgorithmus entwickelt.

Analog zum ITF Algorithmus erhält der Lerner nach seiner ersten fehlerhaften Aufgabenlösung die Rückmeldung, dass seine Antwort noch fehlerhaft sei (KR Feedback) und die Aufforderung, es noch einmal zu versuchen. Löst er die Aufgabe ein zweites Mal falsch, wird ihm die korrekte Lösung präsentiert (KCR Feedback) und er erhält eine neue Aufgabe mit der gleichen Aufgabenanforderung. Identisch zum fehlerspezifischen ITF Algorithmus war das mastery level darüber definiert, dass zwei Aufgaben mit derselben spezifischen Aufgabenan-

forderung in Folge beim ersten Versuch richtig gelöst werden müssen, ehe eine Aufgabe mit einer neuen Aufgabenanforderung präsentiert wird. In Anlehnung an kommerzielle Lernsoftware wird auch hier dem Lerner die Möglichkeit bereitgestellt, bei Problemen Hilfe in Form eines ausgearbeiteten Lösungsbeispiels anzufordern.

Die Entscheidung für diesen Feedbackalgorithmus hatte neben der Gewährleistung einer zufriedenstellend hohen ökologischen Validität als zweiten Grund die Intention, das ITF im Vergleich zu einer relativ komplexen Feedbackbedingung, in diesem Fall KCR Feedback, zu evaluieren, da empirische Evidenz hinsichtlich der Überlegenheit elaborierter Feedbackformen gegenüber keinem Feedback vorliegt (siehe Kapitel 2.2) und die isolierte Wirksamkeit von KCR Feedback in einer Reihe von Studien empirisch belegt werden konnte (Bangert-Drowns et al., 1991).

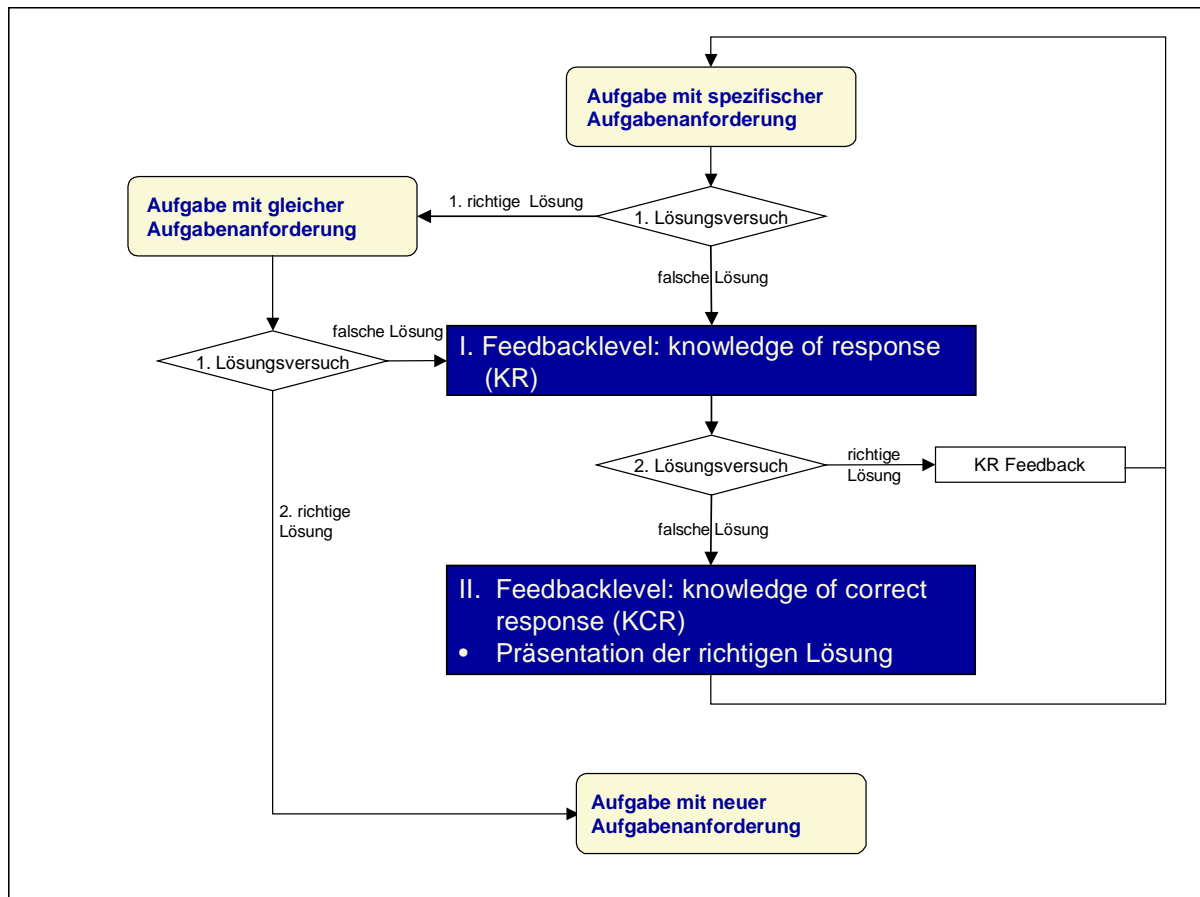


Abbildung 7: Darstellung des in der Vergleichsbedingung realisierten KR-KCR Feedbackalgorithmus

4.2.3 Programmbeschreibung der Übungssoftware „Subtratio“

Beide Feedbackalgorithmen wurden in ein computergestütztes Trainingsprogramm (Subtratio) implementiert. Dieses Trainingsprogramm wurde in Tcl/Tk, einer von John Ousterhout an der University of California in Berkeley entwickelten, universell einsetzbaren, interpretativen Programmiersprache, programmiert.

Die Benutzerführung von Subtratio war bis auf die experimentelle Feedbackvariation in beiden Programmen identisch. Zu Beginn des Treatments öffnete sich ein Fenster (siehe Abbildung 8), welches den Nutzer aufforderte, sich über seinen Vornamen und sein Geburtsdatum eindeutig zu identifizieren.

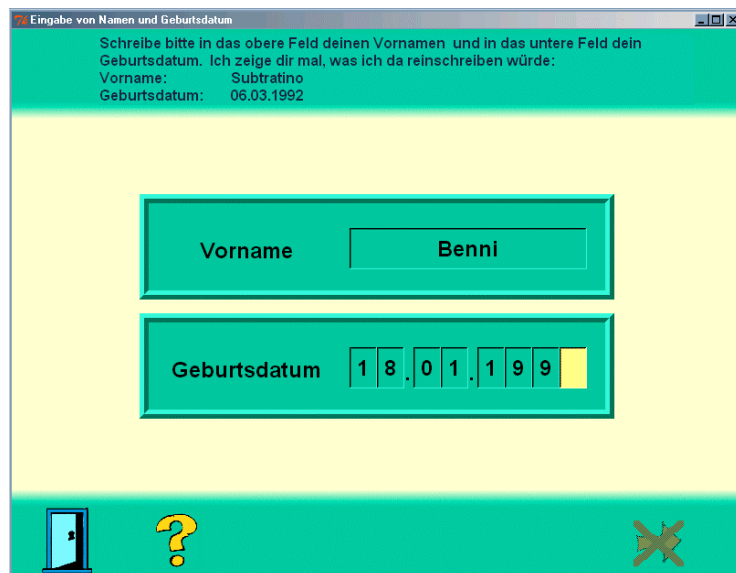


Abbildung 8: Eingabefenster zur Benutzeridentifikation

Mit diesen Informationen wurde für jeden Schüler ein Logfile angelegt, welches die Benutzeraktionen während der halbstündigen Arbeit mit dem Programm dokumentierte. Hatte der Schüler seinen Vornamen und sein Geburtsdatum eingetragen, wurde der Button mit dem roten Pfeil aktiv. Nach dessen Betätigung öffnete sich das Aufgabenmenüfenster mit den zur Bearbeitung zur Verfügung stehenden sechs Aufgabentypen (siehe Abbildung 9). Durch den Versuchsleiter wurden für jeden Schüler die Aufgabentypen ausgewählt (rot), von denen aufgrund des Vortests vermutet wurde, dass diese von dem jeweiligen Schüler nicht beherrscht werden.

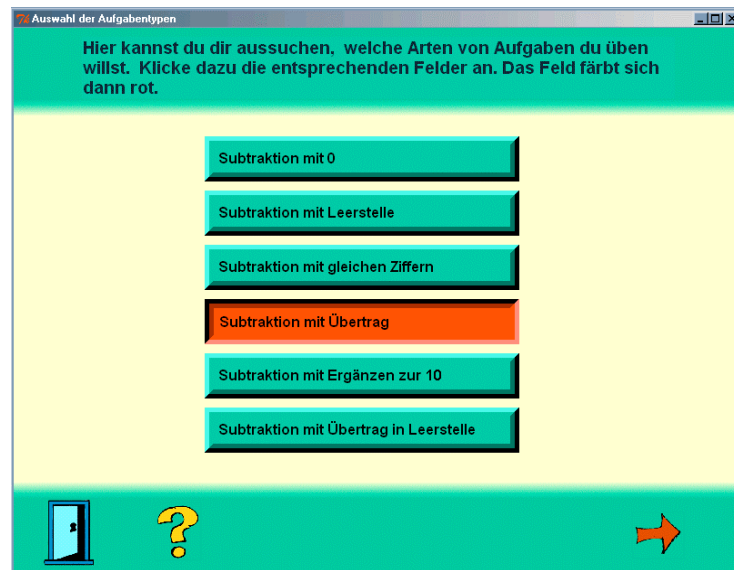


Abbildung 9: Aufgabenmenüfenster für die Auswahl der zu bearbeitenden Aufgabentypen

Nach der Auswahl und der Betätigung des roten Pfeils öffnete sich das Arbeitsfenster mit dem Rechenblatt (siehe Abbildung 10). Die dem Schüler auf diesem interaktiven Rechenblatt präsentierten Aufgaben wurden unter Berücksichtigung folgender Regeln zufällig generiert:

(1) ein Drittel der Aufgaben sind zweistellig, zwei Drittel der Aufgaben dreistellig, (2) jede Aufgabe enthält nur die für diesen Aufgabentyp typische Aufgabenanforderung, z.B. einen Übertrag, (3) die Aufgabenanforderung variiert systematisch in ihrer Lokation, z.B. Übertrag von der Einer- in die Zehnerspalte und Übertrag von der Zehner- in die Hunderter-spalte.

Der Schüler konnte seine Antwort per Tastatur oder per Mausklick über die links implementierten Buttons eingeben. Bei fehlerhafter Eingabe war eine Korrektur über die „Backspace“ Taste bzw. über den „Radiergummi“ Button möglich. War der Schüler der Meinung, die Aufgabe vollständig bearbeitet zu haben, konnte er über die Betätigung des Daumens ein Feedback anfordern.

Erster, korrekter Lösungsversuch

Bei einer korrekten Aufgabenlösung im ersten Lösungsversuch wurde ihm dies sowohl visuell-verbal über die grüne Rechteck-Fläche, visuell-nonverbal und auditiv-verbal von der Programmfigur „Subtratio“ mitgeteilt. Das Ergebnis wurde doppelt unterstrichen, der Daumen wurde inaktiv und der Schüler konnte eine neue Aufgabe mit der gleichen Aufgabenan-

forderung über die Betätigung des roten Pfeils aufrufen (siehe Abbildung 10). Hatte der Schüler auf diese Weise zwei Aufgaben des gleichen Aufgabentyps nacheinander im ersten Lösungsversuch korrekt bearbeitet, gelangte er über die Betätigung des roten Pfeils in das Aufgabenmenüfenster zurück, woraufhin ihm vom Versuchsleiter ein neuer Aufgabentyp ausgewählt wurde.

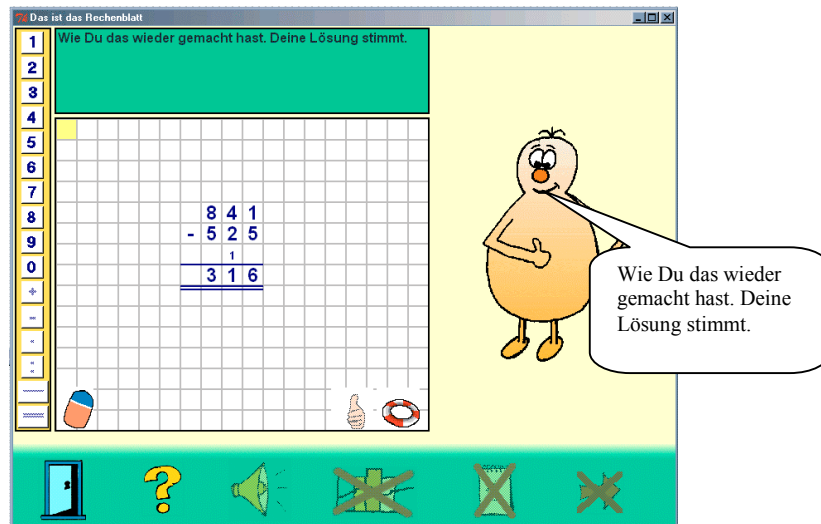


Abbildung 10: Rechenfenster mit korrekt bearbeiteter Übertrags-Aufgabe im ersten Versuch

Erster, fehlerhafter Lösungsversuch

Löste der Schüler die Aufgabe im ersten Versuch hingegen falsch, erhielt er nach Betätigung des Daumens über das grüne Rechteck und von Subtratino visuell und auditiv die Rückmeldung, dass das Ergebnis falsch ist, verbunden mit der Aufforderung, es noch einmal zu versuchen (Abbildung 11).

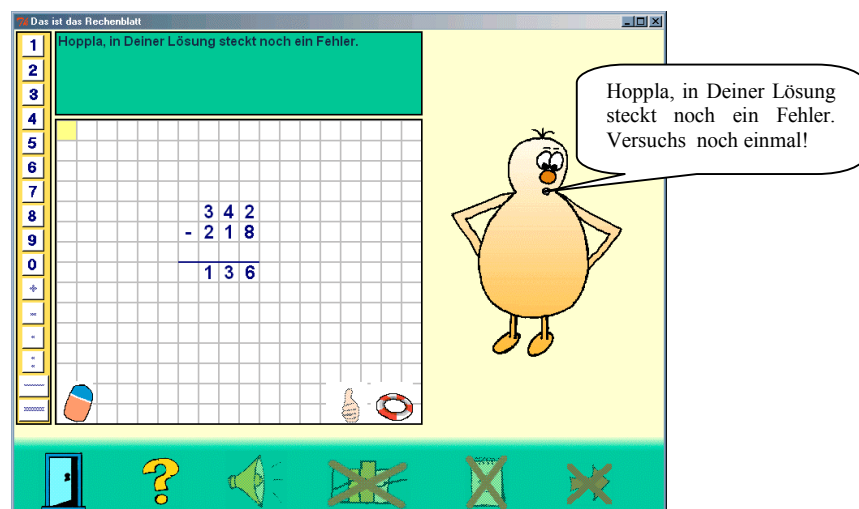


Abbildung 11: Rechenfenster mit fehlerhaft bearbeiteter Übertragsaufgabe im ersten Lösungsversuch

Das fehlerhafte Ergebnis wurde gelöscht, der rote Pfeil wurde inaktiv und der Daumen blieb bis zur nächsten Eingabe inaktiv. Bis zu dieser Stelle waren beide Programmversionen mit und ohne informativem Feedback völlig identisch. Erst mit dem zweiten Lösungsversuch unterschieden sich beide Feedbackbedingungen.

Zweiter, korrekter Lösungsversuch

Bearbeitete der Schüler unter der ITF Bedingung die Aufgabe im zweiten (bzw. im dritten) Lösungsversuch richtig, wurde ihm neben der Information, dass seine Lösung richtig ist, bei einem systematischen Fehler zusätzlich die korrekte Lösungsprozedur mitgeteilt, mit dem Ziel, diese durch die Wiederholung zu festigen. Wie in Abbildung 12 zu sehen, wurde gelegentlich auch ein strategischer Hinweis für nachfolgende Aufgabenbearbeitungen angeboten.

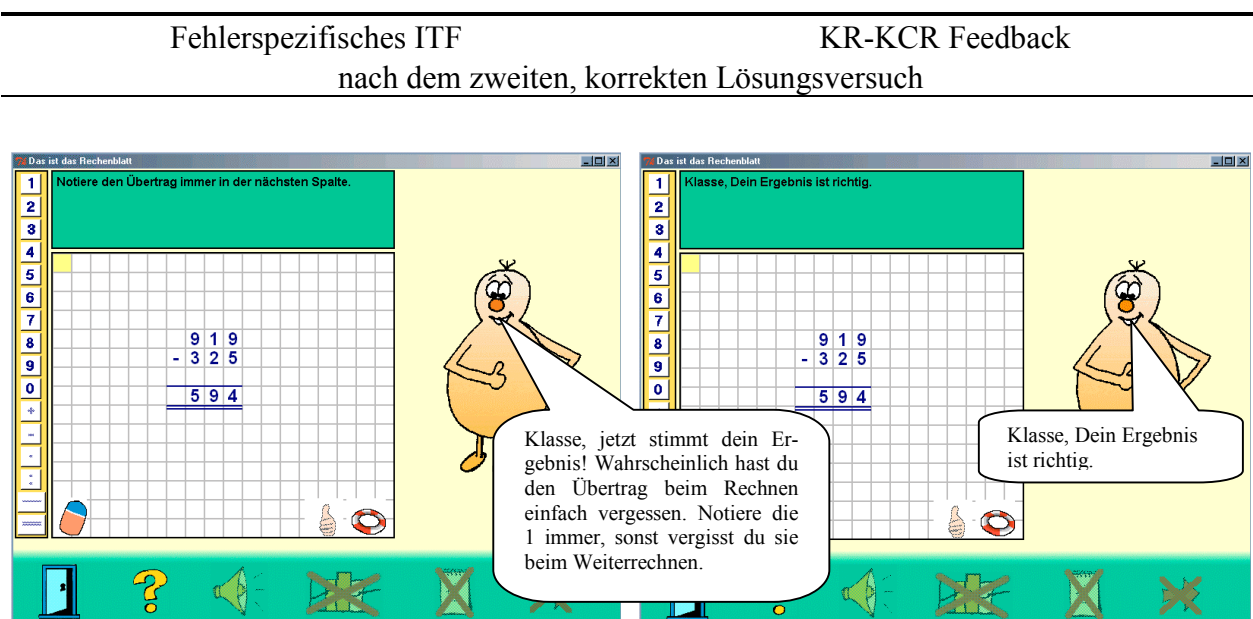


Abbildung 12: KR Feedback & Hinweis auf korrekte Lösungsstrategie

Abbildung 13: KR Feedback

Unter der KR-KCR Feedbackbedingung erhielt der Schüler lediglich die Bestätigung, dass seine Lösung richtig ist (siehe Abbildung 13).

Zweiter, fehlerhafter Lösungsversuch

Löste der Schüler die Aufgabe auch beim zweiten Lösungsversuch fehlerhaft, erhielt er unter der ITF Bedingung einen Hinweis auf den Ort und die Art des systematischen Fehlers sowie einen Hinweis auf die korrekte Lösungsstrategie, verbunden mit der Aufforderung, es noch einmal zu versuchen (siehe Abbildung 14). Unter der Vergleichsbedingung erhielt der Schüler nach einem zweiten fehlerhaften Lösungsversuch die korrekte Lösung präsentiert und er konnte sich die nächste Aufgabe mit der gleichen Aufgabenanforderung über die Betätigung des roten Pfeils aufrufen (siehe Abbildung 15).

Fehlerspezifisches ITF

nach dem zweiten, fehlerhaften Lösungsversuch

KR-KCR Feedback

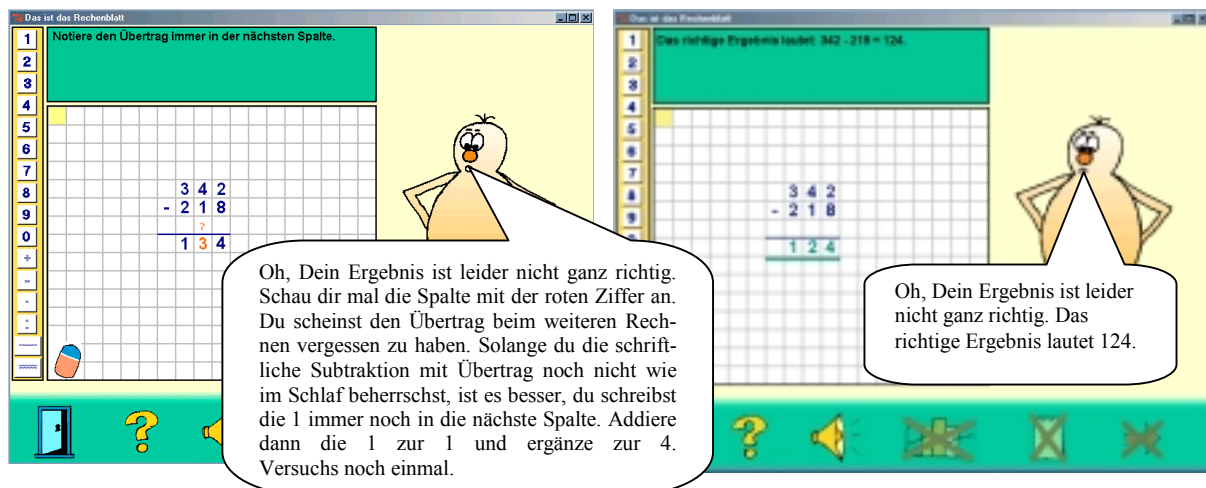


Abbildung 14: KM Feedback (Knowledge-about-Mistake Feedback) – Informationen zu Ort und Art des Fehlers und Hinweis auf korrekte Lösungsstrategie

Abbildung 15: KCR Feedback

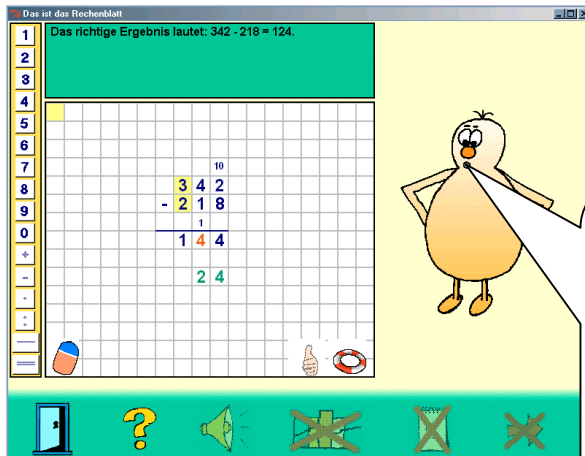
Dritter, fehlerhafter Lösungsversuch unter ITF Bedingung

Bearbeitete der Schüler unter der ITF Bedingung die Aufgabe zum dritten Mal fehlerhaft, wurde auf der dritten Rückmeldungsstufe die korrekte Lösungsprozedur sowohl auditiv als auch visuell präsentiert, d.h. dem Lerner wurde vorgerechnet, wie diese konkrete Aufgabe zu lösen ist. Dabei wurde konditionales Wissen in der Form vermittelt, dass der Lerner im Sinne einer Wenn-dann-Regel darüber informiert wurde, unter welcher Bedingung welche Subprozedur wie ausgeführt werden muss (Abbildung 16). Die korrekte Lösung erschien Schritt für Schritt unter der fehlerhaften Schülerantwort, so dass der Lerner seine fehlerhafte Aufgabenlösung mit der korrekten Aufgabenlösung vergleichen konnte. Im Anschluss konnte

sich der Lerner eine neue Aufgabe des gleichen Aufgabentyps über die Betätigung des roten Pfeils aufrufen.

Fehlerspezifisches ITF

KR-KCR Feedback



Der Lerner erhält eine neue Aufgabe des gleichen Aufgabentyps.

Oh schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der 8 bis zur 2 sind ... Huch, das geht ja gar nicht, denn ich kann nicht von einer größeren Ziffer zu einer kleineren ergänzen. Also muss ich die 2 um 10 erweitern und eine 1 in die nächste Spalte übertragen. 2 plus 10 sind 12. Jetzt kann ich weiterrechnen. Von der 8 bis zur 12 sind 4. Schreibe 4. Jetzt addiere ich die übertragene 1 zur 1. 1 plus 1 sind 2 und rechne weiter. Von der 2 zur 4 sind 2. Schreibe 2. Von der 2 bis zur 3 sind

Abbildung 16: KH Feedback (Knowledge on How to Proceed Feedback) - auditive und visuell-sequentielle Präsentation des korrekten Lösungsalgorithmus nach dem dritten, fehlerhaften Lösungsversuch

Unter beiden Feedbackbedingungen hatte der Schüler zudem jederzeit die Möglichkeit, sich ein ausgearbeitetes Lösungsbeispiel über die Betätigung des Rettungsringes anzeigen zu lassen. Ein solch ausgearbeitetes Lösungsbeispiel ist exemplarisch für den Aufgabentyp „Aufgaben mit Übertrag“ in Abbildung 17 dargestellt.

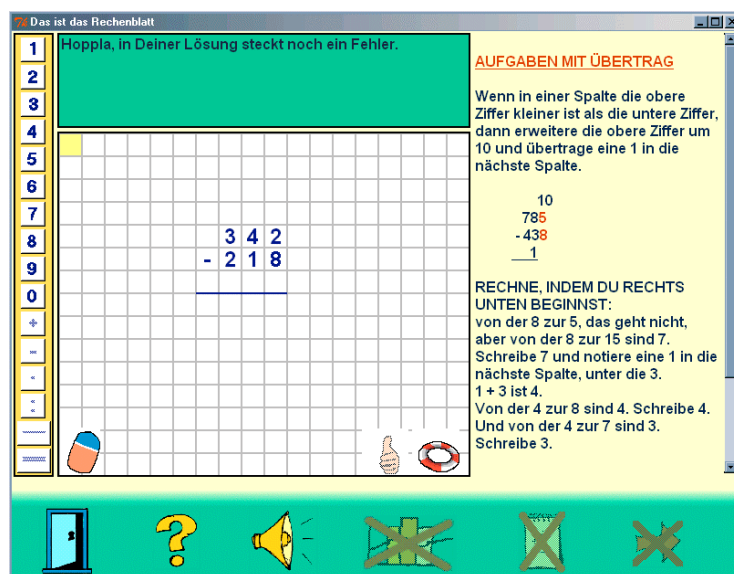


Abbildung 17: Rechenblatt mit ausgearbeitetem Lösungsbeispiel

4.2.4 Messinstrumente

Die im folgenden beschriebenen Messinstrumente dienten der Erfassung der Kontrollvariablen und abhängigen Variablen. Die zur Erfassung der Kontrollvariablen eingesetzten Fragebögen sind im Anhang A abgedruckt.

Kontrollvariable: Aufgabenspezifisches Vorwissen

Empirische Evidenz liegt dafür vor, dass das aufgabenspezifische Vorwissen ein bedeutsamer Prädiktor für Lernerfolg ist (Spada & Wichmann, 1996; Zielinski, 1998). In zahlreichen Studien hat sich das aufgabenspezifische Vorwissen sogar vielfach als vorhersagestärker für schulische Leistungen erwiesen als Intelligenz (vgl. Helmke & Schrader, 1998).

Darüber hinaus konnte im Rahmen der Feedbackforschung gezeigt werden, dass Vorwissen einen moderierenden Einfluss auf die Lernwirksamkeit von Feedback besitzt (siehe Kapitel 2.3.2.1). So analysierte Clariana (1993) in seiner Metaanalyse mögliche Wechselwirkungen zwischen dem Vorwissensniveau und der präsentierten Feedbackform. Er fand, dass primär vorwissensreiche Lerner im Vergleich zu vorwissensarmen Lernern beim Wissenserwerb von Multiple-Try Feedback (MTF) profitierten, wohingegen KCR Feedback sich als lernförderlicher gegenüber MTF bei Lernern mit geringem Vorwissen erwies.

Aus diesen Gründen wurde das aufgabenspezifische Vorwissen als Kontrollvariable in die Analyse aufgenommen und in seinem Einfluss auf die abhängigen Variablen kontrolliert.

Das aufgabenspezifische Vorwissen wurde über einen Vortest erfasst, welcher zwei Wochen vor der eigentlichen Untersuchung den Schülern dargereicht wurde. Dieser Vortest enthielt 32 Aufgaben zur schriftlichen Subtraktion mit den im Rahmen der kognitiven Anforderungsanalyse explorierten sechs Aufgabenanforderungen: Null im Subtrahenden, Leerstelle, gleiche Ziffern übereinander, Übertrag, Übertrag in Leerstelle, Ergänzen zur 10. Diese Aufgabenanforderungen wurden so kombiniert, dass in der Mehrzahl der Aufgaben jeweils 2 Aufgaben zwei dieser Aufgabenanforderungen aufwiesen (siehe Tabelle 16). Insgesamt war jede Aufgabenanforderung zehnmal in den Vortestaufgaben repräsentiert. In dieser Untersuchung wurde das aufgabenspezifische Vorwissen über die Anzahl der im Vortest korrekt bearbeiteten Subtraktionsaufgaben operationalisiert.

Tabelle 16:

Vortest mit 32 Subtraktionsaufgaben zur Erfassung des aufgabenspezifischen Vorwissens

1	847	2	562	3	673	4	374	5	460	6	948	7	316	8	459
	- 30		- 36		- 53		- 92		- 25		- 20		- 140		- 408
9	736	10	504	11	863	12	729	13	348	14	252	15	520	16	392
	- 70		330		- 49		- 390		- 149		- 77		- 367		- 91
17	651	18	787	19	236	20	704	21	613	22	328	23	741	24	344
	- 401		- 297		- 86		- 424		- 91		- 70		- 62		- 64
25	620	26	570	27	902	28	430	29	609	30	260	31	367	32	926
	- 82		- 37		- 760		- 272		- 349		- 81		- 35		- 12
Aufgabenanforderungen								Aufgaben							
Null im Subtrahenden								1, 6, 7, 8, 9, 10, 12, 17, 22, 27							
Leere Stelle								1, 2, 3, 5, 6, 11, 16, 26, 31, 32							
Gleiche Ziffern übereinander								3, 8, 13, 16, 17, 18, 20, 24, 29							
Übertrag								2, 7, 11, 12, 13, 14, 15, 18, 23, 28							
Ergänzen zur 10								5, 10, 15, 20, 25, 26, 27, 28, 29, 30							
Übertrag in leere Stelle								4, 9, 14, 19, 21, 22, 23, 24, 25, 30, 32							

Kontrollvariablen: Aufgabenspezifisches Selbstkonzept

In der Lehr-Lernforschung werden verschiedene Motivations-Konzepte in ihrem Einfluss auf zukünftiges Lernverhalten diskutiert, deren gemeinsame Annahme ist, dass jede Person bestimmte Einschätzungen bezüglich der eigenen Fähigkeiten bzw. Kompetenzen bei der Bearbeitung bestimmter Aufgaben vornimmt. Zu diesen Ansätzen zählen das akademische Selbstkonzept (z.B. Marsh, 1999, zitiert nach Narciss, 2004), die Self-Efficacy (Bandura, 1997, zitiert nach Narciss, 2004), das Selbstkonzept der eigenen Begabung (z.B. Meyer, 1973, zitiert nach Narciss, 2004), das Selbstvertrauen (z.B. Jerusalem, 1990, zitiert nach Narciss, 2004). Diese verschiedenen Forschungsansätze unterscheiden sich nach Narciss (2004) darin, inwiefern sie Einschätzungen bezüglich der eigenen Kompetenzen eine relative Generalität und Stabilität vs. eine relative Spezifität und Veränderlichkeit zuschreiben. Darüber hinaus differieren diese Ansätze auch hinsichtlich des Aspektes, ob das kompetenzbezogene Schema auch Wertkomponenten wie z.B. das Interesse am Lerngegenstand oder die Freude an einer Lernhandlung umfasst oder nicht (Narciss, 2004). Einig sind sich alle Autoren aber hinsichtlich des zentralen Einflusses individueller Kompetenzeinschätzungen auf zukünftiges Lern- und Leistungsverhalten, d.h. auf die Initiierung von Lernhandlungen und auch auf die Anstrengung, Ausdauer und Intensität, mit der diese Lernhandlungen verfolgt werden. So konnte in zahlreichen Studien zum akademischen Selbstkonzept und zur Self-Efficacy empirisch gezeigt werden, dass Personen, welche sich hinsichtlich einer bestimmten Aufgabenbearbeitung als kompetent einschätzten, diese Aufgaben mit einer höheren Ausdauer, Anstrengung und

Intensität verfolgten (Pajares & Schunk, 2001; Schunk & Pajares, 2001). Aus diesem Grund wurden für die vorliegende Studie sowohl Einschätzungen bezüglich der eigenen aufgabenspezifischen Kompetenz als auch Einschätzungen zu tätigkeitsbezogenen Anreizen über folgende vier Items auf einer vierstufigen Ratingskala als Kontrollvariable erfasst und in ihrem Einfluss auf die abhängigen Variablen kontrolliert:

- Das Lösen schriftlicher Subtraktionsaufgaben fällt mir normalerweise 1 = sehr schwer bis 4 = sehr leicht.
- Rechnen fällt mir normalerweise 1 = sehr schwer bis 4 = sehr leicht.
- Das Lösen schriftlicher Subtraktionsaufgaben macht mir 1 = sehr viel Spaß bis 4 = überhaupt keinen Spaß.
- Rechnen macht mir 1 = sehr viel Spaß bis 4 = überhaupt keinen Spaß.

Diese vier Items korrelierten mit $.36 \leq r \leq .71$ relativ hoch mit einander und wiesen - nach einer im Vorfeld erfolgten Umkodierung der beiden Spaß-Items - mit Cronbachs Alpha = .82 eine relativ hohe interne Konsistenz auf. Da Marsh (1999, zitiert nach Narciss, 2004) Wertkomponenten wie die Freude an einer Lernhandlung als Bestandteile des akademischen Selbstkonzeptes betrachtet, wurde über alle vier Items auch aufgrund ihrer hohen internen Konsistenz ein Summenscore gebildet und nur dieser bei der weiteren Auswertung berücksichtigt. Die entstandene Skala wird im weiteren Verlauf dieser Studie als „aufgabenspezifisches Selbstkonzept“ bezeichnet.

Kontrollvariablen: Computereinstellung und Computer-/Lernsoftwareerfahrung

In verschiedenen Studien zum computergestützten Lernen/Arbeiten (Heil & Agnew, 2000; Stoffers, Berendse, Deijen & Wolters, 2002) konnte gezeigt werden, dass Erfahrungen (z.B. mit der Maus) und Einstellungen zu Lernsoftware und zum Computer die am Rechner erzielten Leistungen dahingehend beeinflussen, dass Nutzer mit einer positiven Computereinstellung und höherer Computererfahrung bessere Leistungen am Computer erzielen. Für diese Untersuchung erscheinen diese Ergebnisse vor allem deswegen relevant, da zu vermuten ist, dass unterschiedliche Fertigkeiten in der Bedienung der Maus oder Tastatur aufgrund der unterschiedlich entwickelten feinmotorischen Fertigkeiten von Viertklässler die am Computer erbrachten Leistungen sogar noch stärker beeinflussen dürften. Um den Einfluss der Computererfahrung und -einstellung zu kontrollieren, wurden die in Tabelle 17 dargestellten Items in der Vorbefragung eingesetzt.

Tabelle 17:

Items zur Erfassung der Computer-/Lernprogrammerfahrung und Computereinstellung

Item	Itemtext	Antwortskala
Computerbesitz	Ich/meine Familie besitzt einen eigenen Computer.	Ja/Nein
Computererfahrung	Mit dem Computer arbeite/spiele ich ungefähr ...	täglich 1-2 mal in der Woche 1-2 mal im Monat seltener nie
Lernprogrammbesitz	Ich/meine Familie besitzen ein Computer-Lernprogramm.	Ja/Nein
Lernprogrammerfahrung	Mit diesem Computerlernprogramm übe ich ...	täglich 1-2 mal in der Woche 1-2 mal im Monat seltener nie
Computereinstellung	Das Arbeiten/Spielen am Computer macht mir ...	sehr viel Spaß (1) ein wenig Spaß (2) nicht so viel Spaß (3) überhaupt keinen Spaß (4)

Abhängige Variable: Treatmentleistung

Die Schüler arbeiteten unter beiden Feedbackbedingungen jeweils eine halbe Stunde mit dem computergestützten Programm. Die Treatmentleistung wurde darüber operationalisiert, wie viele Aufgabentypen in dieser halben Stunde erfolgreich bearbeitet werden konnten. Ein Aufgabentyp galt als erfolgreich bearbeitet, wenn der Schüler zwei Aufgaben dieses Aufgabentyps in Folge im ersten Lösungsversuch korrekt bearbeitet hatte. War dies der Fall, hatte der Lerner das mastery level erreicht und das Aufgabentypfenster öffnete sich, so dass ein neuer Aufgabentyp ausgewählt werden konnte. Hatte der Lerner alle sechs Aufgabentypen einmal erfolgreich bearbeitet, wurden diese Aufgabentypen dem Lerner erneut präsentiert, um eine Festigung der erworbenen Fertigkeiten zu unterstützen.

Die Treatmentleistung resultierte demnach aus der Anzahl erreichter mastery level Zustände, d.h. aus der Anzahl erfolgreich bearbeiteter Aufgabentypen während der Treatmentphase. Mit dem Maß der Treatmentleistung war es möglich, den lernwirksamen Einfluss beider Feedbackbedingungen bereits während der Treatmentphase zu erfassen.

Abhängige Variable: Nachttestleistung

Die Nachttestleistung wurde operationalisiert über die Anzahl korrekt bearbeiteter Aufgaben im Nachttest, welcher im Anschluss an die Treatmentphase als Paper and Pencil Test den Schülern ohne Zeitbegrenzung zur Bearbeitung vorgelegt wurde. Die im Nachttest enthaltenen Aufgaben, welcher die gleichen Aufgaben aus dem Vortest beinhaltet, waren nicht wie im Treatment nach Aufgabentypen geordnet, sondern wechselten sich in ihren Aufgabenmerkmalen ab. Das heißt, der Schüler hatte nicht hintereinander fünf Aufgaben mit einem Übertrag zu bearbeiten, sondern die Übertragungsaufgaben waren über den ganzen Nachttest verteilt.

Abhängige Variable: Korrekturwahrscheinlichkeitsindex

Viele Feedbackuntersuchungen erfassen die Wirksamkeit von Feedback allein über die Anzahl richtig gelöster Aufgaben im Nachttest, eine Strategie, die von einigen Feedbackforschern dahingehend kritisiert wurde, dass auf diese Weise das Potenzial der Fehler vernachlässigt werde. Forschergruppen um Phye (Phye, 1979; Phye & Bender, 1989) und Peeck (Peeck, van den Bosch & Kreupeling, 1985) plädieren deshalb für ein Vorgehen, welches auf der Basis von Vortest-Nachttest-Antwortmuster-Vergleichen Aussagen im Hinblick auf konkrete Fehlerkorrekturprozesse erlaubt. Ihr Vorgehen, welches primär die empirisch nachgewiesene korrigierende Funktion von Feedback gegenüber der bestätigenden Funktion von Feedback berücksichtigt (Bangert-Drowns et al., 1991), basiert auf der Unterscheidung folgender in Tabelle 18 dargestellter vier Vortest-Nachttest-Antwortmuster:

Tabelle 18:

Vortest-Nachttest-Antwortmuster und deren korrespondierenden Feedbackfunktionen

Feedbackfunktion	Aufgabenlösung im Vortest	Aufgabenlösung im Nachttest	Nomenklatur
Bestätigend	Richtig	Richtig	$r_1 \rightarrow r_2$
Korrigierend	Falsch	Richtig	$f_1 \rightarrow r_2$
Nicht funktional	Falsch	Falsch	$f_1 \rightarrow f_2$
Nicht funktional	richtig	Falsch	$r_1 \rightarrow f_2$

Feedback wirkt bestätigend, wenn die im Vortest korrekt bearbeitete Aufgabe auch im Nachttest richtig gelöst wurde und korrigierend, wenn eine im Vortest falsch bearbeitete Aufgabe im Nachttest korrekt gelöst wurde. Nicht funktional hingegen ist Feedback, wenn im Nachttest Fehler auftreten, unabhängig davon, ob die Aufgabe im Vortest korrekt oder falsch gelöst wurde.

Um die korrigierende Wirkung des Feedbacks präziser erfassen zu können, plädieren Phye und Bender (1989) für die Bildung eines Korrekturwahrscheinlichkeitsindex, welcher die bedingte Wahrscheinlichkeit darstellt, dass eine im Vortest fehlerhaft bearbeitete Aufgabe im Nachtest richtig gelöst wird. Zur Bestimmung des Index wird die Anzahl der im Vortest falsch, im Nachtest jedoch richtig gelösten Aufgaben (Antwortmuster $f_1 \rightarrow r_2$) zur Gesamtzahl der im Vortest falsch gelösten Aufgaben (Antwortmuster $f_1 \rightarrow r_2$ und $f_1 \rightarrow f_2$) ins Verhältnis gesetzt (siehe Tabelle 19).

Tabelle 19:

Definition des Korrekturwahrscheinlichkeitsindex als bedingte Wahrscheinlichkeit und dessen Operationalisierung in Form von Antwortmustern

	Bedingte Wahrscheinlichkeit	Antwortmuster
Korrekturwahrscheinlichkeitsindex	A – Aufgabe (x) im Vortest richtig B – Aufgabe (x) im Nachtest richtig $P(B \bar{A}) = P(\bar{A} \cap B)/P(\bar{A})$	$(f_1 \rightarrow r_2)$ $(f_1 \rightarrow r_2) \cap (f_1 \rightarrow f_2)$

Abhängige Variable: Einschätzung des wahrgenommenen Nutzens der angebotenen Feedbackinformationen

Die Einschätzung des wahrgenommenen Nutzens der angebotenen Feedbackinformationen wurde getrennt für die beiden Feedbackgruppen mit verschiedenen Items erfasst. In der ITF Gruppe wurden den Schülern im Anschluss an die Treatmentphase folgende in Tabelle 20 dargestellten Items zur Beantwortung vorgelegt:

Tabelle 20:

Items zur Erfassung des wahrgenommenen Nutzens des ITF

Item	Itemtext	Pole der 2 bzw. 4 stufigen Antwortskala
Nützlichkeit	Die Hinweise haben mir	sehr geholfen (1) gehoffen (2) nicht so geholfen (3) überhaupt nicht geholfen (4)
Hinweiswunsch	Ich würde solche Hinweise gerne öfter erhalten	ja (1) nein (2)
Selbständigkeit der Lösungsfindung	Subtratinos Hinweise haben mir dabei geholfen, die richtige Lösung doch noch allein zu finden	stimmt (1) stimmt ein bisschen (2) stimmt nicht (3) stimmt überhaupt nicht (4)

Der KR-KCR Feedbackgruppe wurden im Anschluss an die Treatmentphase folgende in Tabelle 21 aufgelisteten Items zur Beantwortung vorgelegt:

Tabelle 21:

Items zur Erfassung des Feedbackwunsches, der KR-KCR Feedbackgruppe vorgelegt

Item	Itemtext	Pole der 2 stufigen Antwortskala
Hinweiswunsch auf fehlerhafte Lösung	Es wäre hilfreich gewesen, wenn Subtratio mir gesagt hätte, was ich falsch gemacht habe.	ja (1) nein (2)
Hinweiswunsch auf korrekte Lösungsprozedur	Ich hätte gern einen Tipp von Subtratio bekommen, wie man richtig subtrahiert	ja (1) nein (2)

Abhängige Variable: Post-aktionale motivationale Parameter

Um den Einfluss des ITFs auf post-aktionale motivationale Parameter zu erfassen, wurde den Schülern im Anschluss an die Treatmentphase ein Fragebogen mit sechs Items zur Beantwortung vorgelegt (siehe Tabelle 22).

Auf einer 4 stufigen Ratingskala wurden die Schüler gebeten, einzuschätzen, wie sie die Arbeit mit dem Programm unter bestimmten motivationalen Aspekten wahrgenommen haben. Da vor allem Aspekte des Kompetenzerlebens und intrinsischer Anreize die Lernmotivation zukünftiger Lernaktivitäten beeinflussen, wurden die in Tabelle 22 zusammengestellten Items den Schülern im Anschluss an die Treatmentphase zur Beantwortung vorgelegt. Das post-aktionale Kompetenzerleben wurde in Anlehnung an Narciss (2004) darüber erfragt, wie die Schüler ihre eigene Leistung bewerten, wie zufrieden sie mit ihrer Leistung sind, wie viel sie dazugelernt haben und wie schwierig bzw. anstrengend sie die Bearbeitung der Aufgaben empfanden. Die intrinsische Motivation wurde darüber erfragt, wie viel Spaß ihnen das Bearbeiten der Aufgaben bereitet hat.

Vor einer weiteren Analyse wurden die Items 1, 3, 4, 5 und 6 umkodiert, so dass ein geringer Wert auf diesen Items einer geringen Ausprägung entsprach. Eine Faktorenanalyse über alle sechs Items führte zu einer einfaktoriellen Lösung, die mit einem Eigenwert von 4.40 73% der Varianz aufklärte. Die Berechnung des Konsistenzkoeffizienten (Cronbachs α) über alle sechs Items erbrachte mit einem alpha-Wert von .92 einen sehr hohen Reliabilitätskoeffizienten. Aufgrund dieser hohen internen Konsistenz der Items wurde über alle sechs Items ein Summenscore gebildet und dieser Gesamtscore für weitere Berechnungen genutzt. Da nach Ansicht verschiedener Autoren (z.B. Marsh, 1999; zitiert nach Narciss, 2004) kom-

petenzbezogene Schemata auch Wertkomponenten wie die Freude an einer Lernhandlung umfassen, wurde in Anlehnung an Narciss (2004) dieser Gesamtscore als „post-aktionales Kompetenzerleben“ bezeichnet.

Tabelle 22:

Items zur Erfassung post-aktionaler motivationaler Parameter

Item	Itemtext	Pole der 4stufigen Antwortskala
Zufriedenheit	Ich bin mit meiner Leistung in dieser halben Stunde ...	Sehr zufrieden (1) Zufrieden (2) Nicht zufrieden (3) Überhaupt nicht zufrieden (4)
Anstrengung	Ich fand das Lösen der Aufgaben sehr anstrengend.	Stimmt (1) Stimmt ein bisschen (2) Stimmt nicht (3) Stimmt überhaupt nicht (4)
Leistungseinschätzung	Ich fand meine Leistung in dieser halben Stunde ...	Sehr gut (1) Gut (2) Nicht gut (3) Überhaupt nicht gut (4)
Aufgabenschwierigkeit	Ich fand das Beantworten der Aufgaben ...	Sehr leicht (1) Leicht (2) Schwer (3) Sehr schwer (4)
Freude	Mir hat das Lösen der Aufgaben ...	Sehr viel Spaß gemacht (1) Spaß gemacht (2) Keinen Spaß gemacht (3) Überhaupt keinen Spaß gemacht (4)
Lerngewinn	Ich denke, ich habe in dieser halben Stunde ...	Sehr viel dazu gelernt (1) Viel dazu gelernt (2) Ein wenig dazu gelernt (3) Nichts dazu gelernt (4)

4.2.5 Untersuchungsablauf

Zwei Wochen vor der eigentlichen Untersuchung wurde den Schülern der als Paper and Pencil konzipierte Vortest, bestehend aus den in der Voruntersuchung als potentielle Fehlerquellen identifizierten Aufgabenanforderungen zur schriftlichen Subtraktion (siehe Kapitel 3.2.4), im Verlauf einer Unterrichtsstunde an der jeweiligen Grundschule durch den Klassenlehrer dargeboten. Für die Bearbeitung des Vortests gab es keine Zeitbegrenzung. Auf Basis der erhaltenen Daten wurden dann die Schüler, welche gravierende Probleme beim Lösen schriftlicher Subtraktionsaufgaben zeigten, für die eigentliche Untersuchung ausgewählt. Die experimentelle Evaluation wurde im Multimedia Labor der Professur für die Psychologie des Lehrens und Lernens an der TU Dresden durchgeführt. Die Schüler nahmen in einer einmaligen Sitzung in Gruppen von 6-7 Schülern an der Evaluation teil, welche von zwei Versuchsleitern betreut wurde. Die Untersuchung dauerte ca. 1 ½ Stunden. Die Schüler wurden zufällig einer der beiden Feedbackbedingungen zugewiesen. Zu Beginn der Untersuchung hatten die Schüler einen Vorbefragungsbogen zu bearbeiten, welcher Daten hinsichtlich folgender Aspekte erfasste: (a) vier Items zum aufgabenspezifischen Selbstkonzept, (b) zwei Items, welche die Häufigkeit der Computernutzung bzw. Lernprogrammnutzung erfassten, (c) ein Item, welches die Computereinstellung erfasste und (d) zwei Items, welche erfassten, ob der Schüler bzw. seine Familie einen eigenen Computer bzw. eine Lernsoftware besitzen. Neben den Angaben zum aufgabenspezifischen Fähigkeitskonzept und zum tätigkeitsspezifischen Anreiz sollten mit diesem Vorbefragungsbogen potenzielle Störgrößen wie Computererfahrung und -einstellung erfasst und gegebenenfalls kontrolliert werden.

Im Anschluss daran erhielten die Schüler eine Einweisung in die Programmbedienung durch einen der Versuchsleiter, welcher die Funktionalität der einzelnen Buttons und Möglichkeiten der Tastatureingabe am Beispiel des KR-KCR Feedbackalgorithmus erklärte. Danach hatten die Schüler über einen Zeitraum von 30 Minuten schriftliche Subtraktionsaufgaben am Rechner zu bearbeiten. Durch den Versuchsleiter wurden für jeden Schüler die Aufgabentypen ausgewählt, die im Vortest als problematisch für den Schüler identifiziert worden waren. Da die Mehrzahl der Schüler Schwierigkeiten mit Übertrags-Aufgaben hatte, begannen die Schüler am häufigsten mit der Bearbeitung von Aufgaben dieses Aufgabentyps. Zu Beginn des computergestützten Treatments wurde den Schülern mitgeteilt, dass ein Aufgabentyp als erfolgreich bearbeitet galt, wenn zwei Aufgaben mit einer bestimmten Aufgabenanforderung in Folge im ersten Lösungsversuch korrekt bearbeitet wurden. Hatte ein Schüler einen Aufgabentyp erfolgreich bearbeitet, wurde durch den Versuchsleiter der nächste Aufgabentyp aufgerufen. Um eine Festigung der erworbenen Fertigkeiten zu unterstützen, wurde

nach erfolgreicher Bearbeitung aller sechs Aufgabentypen wieder mit dem ersten Aufgabentyp begonnen. Nach 30 Minuten wurde die Aufgabenbearbeitung durch den Versuchsleiter abgebrochen.

Für jeden Aufgabentyp mit einer spezifischen Aufgabenanforderung wurden die zu präsentierenden Aufgaben zufällig generiert. In einem Logfile wurde jede präsentierte Aufgabe mit der korrekten Lösung, der Schülerlösung in jedem Versuch, der Fehlerart und mit der Häufigkeit des aufgerufenen Lösungsbeispiels dokumentiert. Im Anschluss an das Treatment hatten die Schüler den Nachbefragungsbogen mit den sechs Items zur Erfassung post-aktionaler motivationaler Parameter zu bearbeiten. In diesem Nachbefragungsbogen wurden auch in Abhängigkeit der Feedbackbedingung Informationen darüber erfragt, für wie hilfreich das ITF wahrgenommen wurde, bzw. ob sich die Vergleichsgruppe informatives Feedback im Sinne von Hinweisen gewünscht hätte. Abschließend wurde ihnen der Nachtest dargeboten, welcher die gleichen Aufgaben des Vortests enthielt. Der Untersuchungsablauf ist schematisch der Abbildung 18 zu entnehmen.

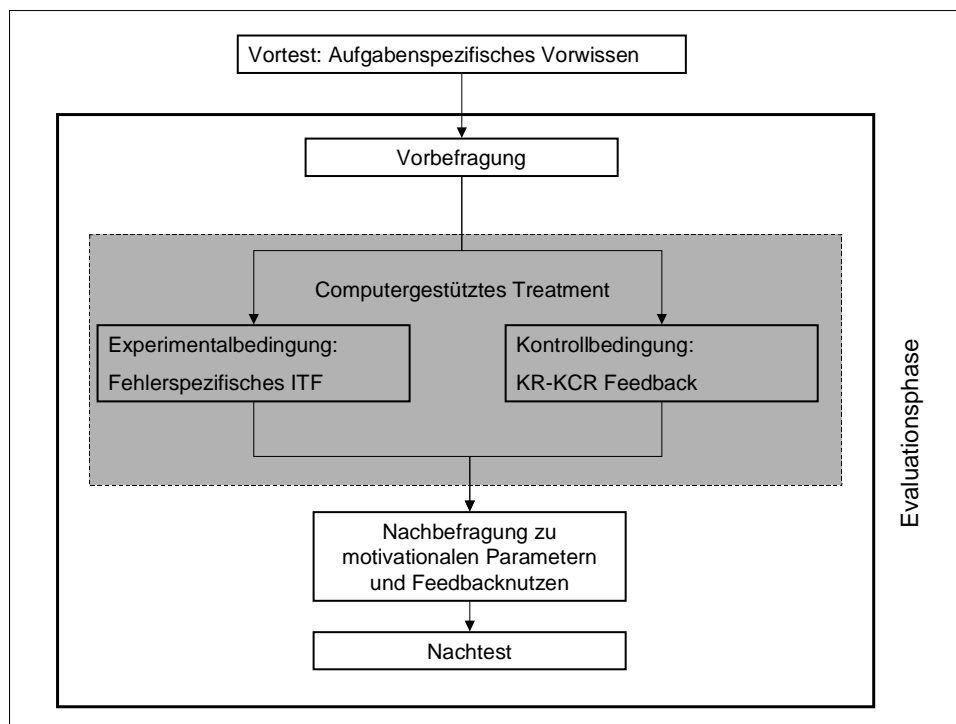


Abbildung 18: Beschreibung des Untersuchungsablaufs von Studie 1

4.2.6 Analysevorbereitung und Methoden der Datenanalyse

In diesem Kapitel werden die beiden Feedbackgruppen zuerst auf ihre Vergleichbarkeit untersucht, bevor die bei der Datenanalyse und Datenauswertung eingesetzten Methoden vorgestellt werden.

4.2.6.1 Vergleichbarkeit der Feedbackgruppen

Da eine zentrale Annahme zur Effektivität von ITF lautet, dass ITF vor allem bei der Korrektur auftretender Fehler wirksam wird, sollten an der Evaluationsstudie nur Schüler mit Problemen in der schriftlichen Subtraktion teilnehmen. Da aber die Vortestleistung zwei Wochen vor der eigentlichen Evaluation erhoben wurde und zu vermuten war, dass während dieses Zeitraums einige Schüler einen starken Lernfortschritt in der Beherrschung der schriftlichen Subtraktion erfahren, wurden im Vorfeld der eigentlichen Datenanalyse die vorliegenden Schülerdaten einer explorativen Datenanalyse unterzogen. Diese explorative Datenanalyse verfolgte das Ziel, diejenigen Schüler von der weiteren Datenauswertung auszuschließen, die zum Evaluationszeitpunkt ein hohes Leistungslevel erreicht hatten. Neben solchen inhaltlichen Überlegungen wurde auch aus methodischen Gesichtspunkten eine explorative Datenanalyse durchgeführt. So plädiert Sedlmeier (1996) vor allem bei kleinen Stichprobengrößen für den Einsatz einer explorativen Datenanalyse, da diese das Entdecken starker Asymmetrien und deutlicher Ausreißer ermöglicht, welche die anschließenden Analysen stark beeinflussen könnten.

Als Selektionsbasis wurde die Treatmentleistung gewählt, operationalisiert über die Anzahl erfolgreich bearbeiteter Aufgabentypen in der halbstündigen Treatmentphase. Ein Aufgabentyp galt als erfolgreich bearbeitet, wenn zwei Aufgaben mit einer bestimmten Aufgabenanforderung in Folge im ersten Versuch richtig bearbeitet worden waren (mastery level). Über eine explorative Datenanalyse wurden vier Ausreißer der Vergleichsgruppe identifiziert, welche zehn Aufgabentypen oder mehr in der Treatmentphase bearbeitet hatten. Da diese Schüler der Vergleichsgruppe kein informatives Feedback erhielten, war zu vermuten, dass diese Schüler mittlerweile über gute Kenntnisse der schriftlichen Subtraktion verfügten. Für die Experimentalgruppe konnten keine Ausreißer identifiziert werden. Eine Erklärung für dieses Ergebnis kann in einer Konfundierung der Wirksamkeit des fehlerspezifischen ITFs mit dem Vorliegen hoher Leistungsfähigkeit zum Treatmentzeitpunkt gesehen werden. Um die Vergleichbarkeit beider Feedbackgruppen zu gewährleisten, wurden diejenigen Schüler

der Experimentalgruppe von der weiteren Datenanalyse ausgeschlossen, welche gleichfalls zehn oder mehr Aufgabentypen in der Treatmentphase bearbeitet hatten. Dieses Selektionskriterium traf auf vier Schüler der Experimentalgruppe zu, so dass in die Datenauswertung 30 Datensätze (15 der Experimentalgruppe und 15 der Vergleichsgruppe) eingingen. Diese 30 Schüler ($w = 11$; $m = 19$) waren zum Zeitpunkt der Untersuchung im Durchschnitt 9.4 ($SD = .50$) Jahre alt.

Mit dem Ziel, die Vergleichbarkeit beider Feedbackgruppen zu überprüfen, wurde bei gegebenen Voraussetzungen für die in Tabelle 23 aufgeführten, mit dem Vorbefragungsbogen erfassten Variablen wie auch der im Vorfeld erfassten Vortestleistung der t -Test für unabhängige Stichproben berechnet. Bei nicht normalverteilten bzw. ordinal skalierten Daten wurde alternativ zum t -Test der nichtparametrische Mann-Whitney-U-Test benutzt. Der exakte Test nach Fisher wurde für die Analyse nominal skalierten Variablen berechnet. Deskriptive Maße der zentralen Tendenz, Dispersionsmaße, Häufigkeitsverteilungen und Ergebnisse der Signifikanztests sind für jede Variable in Tabelle 23 abgetragen.

Es wurden keine statistisch signifikanten Unterschiede zwischen beiden Feedbackgruppen für die in Tabelle 23 dargestellten Variablen gefunden, so dass auf eine Vergleichbarkeit der beiden Feedbackgruppen hinsichtlich dieser Variablen geschlossen werden kann.

Tabelle 23:

Deskriptive Maße der zentralen Tendenz, Dispersionsmaße und Häufigkeitsverteilungen der Kontrollvariablen, getrennt nach Feedbackgruppen dargestellt, sowie Ergebnisse der Signifikanztests

Kontrollvariablen	Feedbackbedingung		Signifikanztest	<i>p</i>	
	Fehlerspezifisches ITF (<i>n</i> = 15)	KR-KCR Feedback (<i>n</i> = 15)			
Alter	<i>M</i> (<i>SD</i>)	9.53 (.52)	9.33 .49	<i>U</i> = 90.00	.37
Aufgabenspezifisches Vorwissen	<i>M</i> (<i>SD</i>)	8.40 (6.48)	6.93 (4.15)	<i>t</i> = .738	.47
Aufgabenspezifisches Selbstkonzept	<i>M</i> (<i>SD</i>)	8.64 (2.21)	9.00 (2.86)	<i>t</i> = -.37	.71
Computereinstellung	<i>Me</i>	1.00	1.00	<i>U</i> = 91.00	.56
Computererfahrung	Täglich 1-2mal/Woche 1-2mal/Monat seltener nie	33.3% 20.0% 13.3% 26.7% 6.7%	33.3% 13.3% 20.0% 13.3% 20.0%	<i>U</i> = 105.00	.78
Lernprogrammerfahrung	Täglich 1-2mal/Woche 1-2mal/Monat seltener nie	6.7% 26.7% 20.0% 26.7% 20.0%	0.0% 20.0% 20.0% 6.7% 53.3%	<i>U</i> = 78.50	.16
Computerbesitz	Ja Nein	66.7% 33.3%	80.0% 20.0%	Exakter Test nach Fisher	.68
Programmbesitz	Ja Nein	86.7% 13.3%	53.3% 46.7%	Exakter Test nach Fisher	.11
Geschlecht	Weiblich Männlich	40% 60%	33.3% 66.7%	Exakter Test nach Fisher	1.00

4.2.6.2 Methoden der Datenanalyse

Das zentrale Anliegen dieser Studie lag in dem Nachweis einer höheren Lern- und motivationalen Wirksamkeit von fehlerspezifischem ITF gegenüber KR-KCR Feedback. Zur statistischen Prüfung dieser Wirksamkeitsunterschiedshypothesen waren zum einen kovarianzanalytische Verfahren indiziert, da im Rahmen korrelationsanalytischer Auswertungen vereinzelt signifikante Korrelationen zwischen den über den Vortest und die Vorbefragung erfassten Kontrollvariablen und den abhängigen Variablen gefunden werden konnten. In Tabelle 24 sind die Korrelationen zwischen den einzelnen Kovariaten und abhängigen Variablen dargestellt. Die einzelnen Items wurden im Vorfeld so umkodiert, dass geringe Werte eine geringe Ausprägung indizieren.

Tabelle 24:

Korrelationen zwischen Kovariaten und abhängigen Variablen

	2	3	4	5	6	7	8	9
1. Treatmentleistung	.22	.32	.39*	-.02	.15	.26	.57**	.29
2. Nachtestleistung	-	.95**	.33	.39*	.14	-.34	.13	.03
3. Korrekturwahrscheinlichkeit		-	.28	.20	.04	-.25	.26	.13
4. Post-aktionales Kompetenzerleben			-	.35	.57**	.04	.02	.22
5. Vortestleistung				-	.45*	-.23	-.28	-.06
6. Aufgabenspezifisches Selbstkonzept					-	.24	-.16	.30
7. Computererfahrung						-	.42*	.10
8. Lernprogrammerfahrung							-	-.12
9. Computereinstellung								-

* $p < .05$. ** $p < .01$.

Im Rahmen der kovarianzanalytischen Auswertungen wurde für die Bestimmung der zulässigen Kovariatenanzahl auf Formel 1 zurückgegriffen, wobei C die Anzahl der Kovariaten, J die Anzahl der Gruppen und N die totale Stichprobengröße repräsentiert (Huitema, 1980, S. 161). Bei einem Wert $> .10$ ist davon auszugehen, dass die Schätzung der angepassten Mittelwerte instabil sein wird.

$$\frac{C + (J - 1)}{N} > .10 \quad (\text{Formel 1})$$

Waren keine Kovariaten zu berücksichtigen, wurde der t -Test für unabhängige Stichproben als ein weiteres statistisches Auswertungsverfahren zur Überprüfung der Wirksamkeitsunterschiedshypothesen gewählt.

Neben dem Einsatz der genannten statistischen Verfahren, welche Aussagen darüber treffen, ob eine Wirkung aufgetreten ist, die unter probabilistischer oder signifikanzstatistischer Sichtweise größer ist als zufällige Schwankungen, wurde ergänzend hierzu auf die Berechnung von Effektstärkemaßen zurückgegriffen, welche angeben, wie intensiv die gegen Zufall abgesicherte statistische Wirkung ist, d.h. Aussagen hinsichtlich der praktischen Bedeutsamkeit ermöglichen. Die Effektstärkemaße wurden aus den Prüfgrößen der statistischen Signifikanztests abgeleitet. Bei vorliegenden t -Werten wurde der standardisierte Mittelwertsabstand d (Cohen, 1988) über Formel 2 bestimmt (Sedlmeier, 1996).

$$d = \frac{t\sqrt{n_A + n_B}}{\sqrt{n_A n_B}} \quad (\text{Formel 2})$$

Bei Verwendung der auf quadrierten Werten beruhenden parametrischen Teststatistik wurde als Effektstärkemaß das univariate ω^2 (Hays, 1973) berechnet, welches im Gegensatz zu dem von SPSS ausgegebenen Effektstärkemaß η^2 nicht auf der Grundlage von Stichprobeninformationen berechnet wird, sondern eine Populationsschätzung darstellt. Im Unterschied zu η^2 gewährleistet eine auf einer Populationsschätzung beruhende Effektstärkenberechnung eine „konservative“ Bestimmung der Effektstärke, d.h. die Effektstärke wird am wenigsten überschätzt (Wolf, 2001). Bei vorliegenden F -Werten wurde das Effektstärkemaß ω^2 über Formel 2 bestimmt, wobei n_h die Hypothesenfreiheitsgrade repräsentiert und n_e die Fehlerfreiheitsgrade (error) (Wolf, 2001).

$$\omega^2 = \frac{F - 1}{F + \frac{n_e + 1}{n_h}} \quad (\text{Formel 3})$$

Um Aussagen bezüglich der Interpretierbarkeit der erhaltenen Effektstärkemaße zu erhalten, bietet sich zum einen der Rückgriff auf die von Cohen (1988) vorgeschlagenen Konventionen an, die in Tabelle 25 dargestellt sind.

Tabelle 25:

Interpretation der Effektstärkemaße nach Cohen (1988)

	Kleiner Effekt	Mittlerer Effekt	Starker Effekt
Effektstärkemaß d	$\leq .35$	$.35 < d \leq .65$	$> .65$
Effektstärkemaß η^2 bzw. ω^2	$\leq .03$	$.03 < d \leq .10$	$> .10$

Obwohl diese Werte anfangs empirisch nicht begründet worden waren, scheinen doch die mittleren Effekte den in verschiedenen Bereichen der Psychologie zu findenden durchschnittlichen Effekten gut zu entsprechen (Sedlmeier, 1996). Trotzdem warnen verschiedene Autoren – wie auch Cohen selbst – davor, diese Konventionen als alleinige Interpretationsbasis zu nutzen, da eine Interpretation der Effektstärken sich immer an bisherigen Ergebnissen des Forschungsfeldes orientieren sollte (Cohen, 1988; Sedlmeier, 1996). So kann in bestimmten Forschungsbereichen bereits ein kleiner Effekt von Bedeutung sein, wenn in diesem Forschungsfeld bisher keine Effekte nachgewiesen werden konnten. Allerdings sollte auch bedacht werden, dass ein Effektstärkenvergleich eine Vergleichbarkeit der Untersuchungssituationen voraussetzt, welche häufig nicht gegeben ist, wie etwa hinsichtlich der Dauer der Intervention, der Art der Vergleichsbedingung, etc. (Hager, Patry & Brezing, 2000).

Für diese Studie wurden Studien als Vergleichsbasis herangezogen werden, die den Einfluss informativen Feedbacks auf Lern- oder motivationale Parameter untersucht haben. Die Studien zum Einfluss von ITF auf Lernparameter stammen vor allem aus dem Bereich der Forschung zu intelligenten tutoriellen Systemen, die über eine Diagnose des aktuellen Wissensstandes des Lerners (Generierung eines Lernermodells) Feedback anbieten, welches auf den diagnostizierten Wissensstand des Lerners zugeschnitten ist. Mit dem Ziel, menschliches Verhalten zu imitieren, wurden in intelligenten tutoriellen Systemen der 2. Generation aber weitere Elemente (z.B. natürlichsprachige Dialoge, Avatare) implementiert, die eine direkte Vergleichbarkeit erschweren. Ein weiteres Problem ist darin zu sehen, dass obwohl ziemlich viel Aufwand in die Entwicklung intelligenter tutorieller Systeme investiert wurde, im Vergleich hierzu nur wenige Evaluationsstudien vorliegen. Diese wenigen Evaluationsstudien berichten von Lernzuwachsen zugunsten der intelligenten tutoriellen Systeme im Vergleich zu

konventionellem Unterricht im Effektstärkenbereich von .3 und 1.0 (Corbett, Anderson, Graesser, Koedinger & VanLehn, 1999). Evaluationsstudien, die Lerngewinne von intelligenten tutoriellen Systemen mit einer computerbasierten Kontrollbedingung verglichen, zeigten Effektstärken zwischen .6 und .8 (Albacete & VanLehn, 2000; Heffernan, 2001; Nagata & Swisher, 1995). Da sich das Treatment dieser Studien im Vergleich zu der vorliegenden Studie über einen längeren Zeitraum erstreckte, wurden für die vorliegende Studie Effektstärken um $d = .6$ bzw. $\omega^2 = .06$ als Ausdruck praktischer Bedeutsamkeit definiert.

Den Einfluss informativen Feedbacks auf motivationale Parameter untersuchte Narciss (2004). Sie fand in ihrer Konzeptlernstudie, welche ITF mit zwei weniger informativen Feedbackformen verglich, eine hohe ITF-Wirksamkeit auf das post-aktionale Kompetenzerleben vor allem für Personen mit einer hohen situationsbezogenen Kompetenzeinschätzung ($\eta^2 = .08$). Vor diesem Hintergrund wurden für die vorliegende Studie Effektstärken um $d = .5$ bzw. $\omega^2 = .07$ als hohe Effektstärken definiert.

4.3 Ergebnisdarstellung

Die Ergebnisdarstellung erfolgt entsprechend der aufgestellten Hypothesen, beginnend mit der Analyse der Treatmentleistung, operationalisiert über das mastery level, gefolgt von der Analyse der Nachtestleistung, des Korrekturwahrscheinlichkeitsindex und der Einschätzung des Feedbacknutzens, und abschließend mit der Analyse der in der Nachbefragung erfassten motivationalen Parameter. Es wird berichtet, ob Unterschiede in diesen abhängigen Variablen zwischen den beiden Feedbackgruppen gefunden werden konnten.

4.3.1 Treatmentleistung

Die Treatmentleistung wurde operationalisiert über die Anzahl erfolgreich bearbeiteter Aufgabentypen während der halbstündigen Treatmentphase.

Da die abhängige Variable „Treatmentleistung“ hoch mit dem in der Vorbefragung erfasstem Item „Lernprogrammerfahrung“ korrelierte ($r = -.57$, $p = .00$), wurde das Item „Lernprogrammerfahrung“ als Kovariate in der Datenauswertung berücksichtigt. Für die unabhängige Variable „Feedbackbedingung“ konnte nach Auspartialisierung des Einflusses der Variable „Lernprogrammerfahrung“ ein hochsignifikanter Haupteffekt, $F(1, 23) = 12.53$, $p = .00$, $\omega^2 = .316$, nachgewiesen werden. Die Schüler unter der ITF Bedingung ($M = 3.64$, $SE = .44$) bearbeiteten während der halbstündigen Treatmentphase signifikant mehr Aufgabentypen

als die Schüler unter der KR-KCR Feedbackbedingung ($M = 1.34, SE = .42$). So können 32% der Varianz der Treatmentleistung über eine höhere Lernwirksamkeit des ITF gegenüber dem KR-KCR Feedback aufgeklärt werden (H1a).

4.3.2 Nachtestleistung

Die Nachtestleistung wurde über 32 Aufgaben zur schriftlichen Subtraktion erfasst, welche den Schülern im Anschluss an die Treatmentphase dargeboten wurden und die gleichen Aufgaben aus dem Vortest enthielt. Da die Vortestleistung mit der Nachtestleistung korrelierte ($r = .39, p = .03$), wurde diese als Kovariate in die weitere Analyse mit aufgenommen. Eine Kovarianzanalyse erbrachte für die Kovariate Vortestleistung einen Haupteffekt, $F(1, 27) = 4.31, p = .047, \omega^2 = .10$. So zeigten die Schüler beider Feedbackgruppen nach dem Treatment signifikant bessere Leistungen beim Lösen schriftlicher Subtraktionsaufgaben, $M_{\text{vorher}} = 7.67, SD = 5.40; M_{\text{nachher}} = 15.83, SD = 8.92$ (siehe hierzu auch Abbildung 19). Für die unabhängige Variable Feedback wurde das Signifikanzniveau nur knapp verfehlt, $F(1, 27) = 3.10, p = .09, \omega^2 = .07$. Die angepassten Mittelwerte der Nachtestleistung zeigen, dass die Schüler der ITF Bedingung, $M = 18.44, SE = 2.09$, hypothesenkonform (H1b), wenn auch nicht statistisch signifikant, im Nachtest mehr Aufgaben korrekt bearbeiteten als die Schüler der KR-KCR Feedbackbedingung, $M = 13.23, SE = 2.09$ (siehe hierzu auch Abbildung 19).

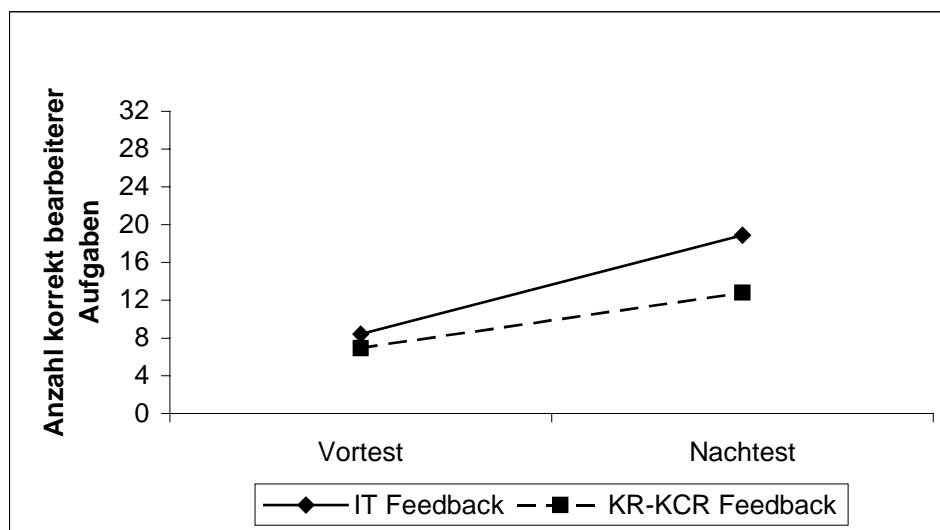


Abbildung 19: Vor-Nachtestleistungen der Feedbackgruppen

4.3.3 Korrekturwahrscheinlichkeitsindex

Analog zu der von Phye und Bender (1989) vorgeschlagenen Verfahrensweise wurden in einem nächsten Schritt die Vortest- und Nachtestlösungen nach folgenden vier Antwortmustern analysiert: $r_1 \rightarrow r_2$, $f_1 \rightarrow r_2$, $f_1 \rightarrow f_2$, $r_1 \rightarrow f_2$. Für jeden Schüler wurde der relative Anteil dieser vier Antwortmuster bestimmt. In Tabelle 26 sind für beide Feedbackbedingungen die durchschnittlichen Häufigkeiten der jeweiligen Antwortmuster in Form von Mittelwerten aufgeführt.

Tabelle 26:

Mittelwerte und Standardabweichungen der vier Vortest-Nachtest-Antwortmuster, getrennt für jede Feedbackgruppe dargestellt, Ergebnisse der t-Tests

Vortest-Nachtest-Antwortmuster	ITF	KR-KCR Feedback		
	($n = 15$) $M (SD)$	($n = 15$) $M (SD)$	$t (28)$	p
richtig \rightarrow richtig $r_1 \rightarrow r_2$	6.47 (5.34)	5.80 (3.91)	.39	.70
falsch \rightarrow richtig $f_1 \rightarrow r_2$	12.33 (7.41)	6.87 (7.36)	2.03	.05
falsch \rightarrow falsch $f_1 \rightarrow f_2$	11.27 (8.76)	18.20 (7.89)	-2.28	.03
richtig \rightarrow falsch $r_1 \rightarrow f_2$	1.93 (1.94)	1.13 (1.46)	1.28	.21

Mittelwertvergleiche zeigten für das Antwortmuster $f_1 \rightarrow r_2$ und für das Antwortmuster $f_1 \rightarrow f_2$ signifikante Unterschiede zwischen den beiden Feedbackgruppen (siehe Tabelle 26). Das heißt, dass Schüler unter der ITF Bedingung im Vergleich zu Schülern der KR-KCR Feedbackbedingung ihre im Vortest falsch gelösten Aufgaben häufiger im Nachtest korrekt bearbeiteten und vice versa Schüler der KR-KCR Feedbackbedingung ihre im Vortest fehlerhaft gelösten Aufgaben häufiger auch im Nachtest falsch bearbeiteten, im Vergleich zur ITF Gruppe.

Gemäß der Annahme, dass Feedback primär korrigierend wirkt, wurde in einem nächsten Schritt das von Phye und Bender (1989) vorgeschlagene Maß der Korrekturwahrscheinlichkeit (corrective probability) berechnet. Die Korrekturwahrscheinlichkeit kann beschrieben werden als die bedingte Wahrscheinlichkeit, dass eine Aufgabe im Nachtest richtig gelöst

wird, wenn sie im Vortest falsch gelöst wurde. Zur Bestimmung dieses Index wird die Anzahl der im Vortest falsch, im Nachtest jedoch richtig gelösten Aufgaben (Antwortmuster $f_1 \rightarrow r_2$) zur Gesamtzahl der im Vortest falsch gelösten Aufgaben (Antwortmuster $f_1 \rightarrow r_2$ und $f_1 \rightarrow f_2$) ins Verhältnis gesetzt. Mittelwerte und Standardabweichungen für diesen Index sind getrennt nach den Feedbackgruppen in Tabelle 27 dargestellt.

Tabelle 27:

Mittelwerte und Standardabweichungen für den Korrekturwahrscheinlichkeitsindex getrennt nach Feedbackgruppen

	IT Feedback ($n = 15$)	KR-KCR Feedback ($n = 15$)
	$M (SD)$	$M (SD)$
Korrekturwahrscheinlichkeitsindex	.54(.30)	.27 (.28)

Schüler, welche fehlerspezifisches ITF erhielten, lösten 54% der Aufgaben, die sie im Vortest fehlerhaft bearbeitet hatten, im Nachtest korrekt. Hingegen bearbeiteten Schüler, welche KR-KCR Feedback erhalten hatten, nur 27% der im Vortest fehlerhaft gelösten Aufgaben im Nachtest richtig. Bei gegebener Normalverteilung und Varianzhomogenität wurde der t -Test für unabhängige Stichproben berechnet, welcher einen signifikanten Unterschied für den Faktor Feedback zeigte, $t(28) = 2.52, p = .018, d = 0.92, 95\text{-KI} (0.05; 0.49)$.

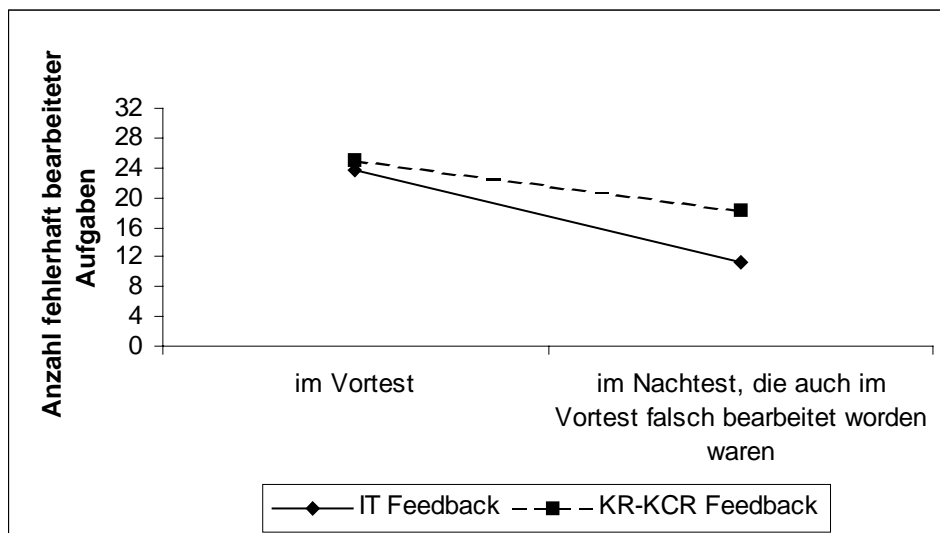


Abbildung 20: Darstellung der Mittelwerte der im Vortest fehlerhaft gelösten Aufgaben und der Mittelwerte der Aufgaben, die sowohl im Vortest als auch im Nachtest fehlerhaft bearbeitet worden waren, d.h. im Nachtest nicht korrigiert werden konnten

Schüler, welche ITF erhielten, waren bei der Korrektur ihrer im Vortest falsch gelösten Aufgaben demnach statistisch signifikant erfolgreicher als Schüler, welche KR-KCR Feedback erhalten hatten (H1b). Dieses Ergebnis veranschaulicht auch noch einmal Abbildung 20.

4.3.4 Einschätzung des wahrgenommenen Feedbacknutzen

Daten zum wahrgenommenen Feedbacknutzen wurden über drei Items im Anschluss an die Treatmentphase nur für die ITF Gruppe erhoben, da nur diese informatives Feedback erhalten hatte. Die KR-KCR Feedbackgruppe erhielt im Anschluss an die Treatmentphase zwei Items, welche Informationen zum Wunsch nach instruktionaler Unterstützung erfassten. In Tabelle 28 sind sowohl deskriptive Daten als auch Ergebnisse der Chi-Quadrat-Tests nach Pearson dargestellt.

Tabelle 28:

Häufigkeiten (%) und Ergebnisse der Chi-Quadrat-Tests zum wahrgenommenen Feedbacknutzen der ITF Gruppe bzw. zum Feedbackwunsch der KR-KCR Feedbackgruppe

Gruppe	Items	Beobachtete Häufigkeiten (%)	Erwartete Häufigkeiten (%)	Chi ² -Test	<i>p</i>	
ITF	Die Hinweise haben mir	sehr geholfen:	4 (26.7)	3.8 (25)	$\chi^2(3, N = 15) = 6.60$.09
		geholfen:	7 (46.7)	3.8 (25)		
		nicht so geholfen:	4 (26.7)	3.8 (25)		
		überhaupt nicht geholfen:	0 (0.00)	3.8 (25)		
	Ich würde solche Hinweise gerne öfter erhalten	ja:	13 (86.7)	7.5 (50)	$\chi^2(1, N = 15) = 8.07$.01
		nein:	2 (13.3)	7.5 (50)		
	Subtratino Hinweise haben mir dabei geholfen, die richtige Lösung doch noch allein zu finden	stimmt:	13 (86.7)	3.8 (25)	$\chi^2(3, N = 15) = 30.60$.00
		stimmt ein bisschen:	1 (6.7)	3.8 (25)		
		stimmt nicht:	1 (6.7)	3.8 (25)		
		stimmt überhaupt nicht:	0 (0.0)	3.8 (25)		
	KR-KCR	Es wäre hilfreich gewesen, wenn Subtratino mir gesagt hätte, was ich falsch gemacht habe.				

ja:	11 (73.3)	7.5 (50)	$\chi^2(1, N = 15) = .07$
nein:	4 (26.7)	7.5 (50)	3.27
Ich hätte gern einen Tipp von Subtratino bekommen, wie man richtig subtrahiert.			
ja:	10 (71.4)	7.0 (50)	$\chi^2(1, N = 14) = .11$
nein:	4 (28.6)	7.0 (50)	2.57

Die Ergebnisse der Befragung nach dem wahrgenommenen Nutzen der dargebotenen Feedbackinformationen zeigen, dass die Schüler, welche ITF erhielten, dieses in der Mehrzahl als sehr hilfreich einschätzten und sich solche Hinweise öfter wünschten (H1c). Die Schüler der KR-KCR Feedbackbedingung äußerten gleichfalls in der Mehrzahl den Wunsch nach einer instruktionalen Unterstützung bei auftretenden Problemen. Die Ergebnisse der Chi-Quadrat-Tests zeigen für fast alle Items signifikante Abweichungen der beobachteten Häufigkeiten von den erwarteten Häufigkeiten.

4.3.5 Post-aktionales Kompetenzerleben

Wie bereits bei der Darstellung der abhängigen Variablen ausgeführt, wurde auf grund der hohen internen Konsistenz der sechs motivationalen Items ($\alpha = .92$) über alle Items ein Summenscore gebildet und dieser Gesamtscore als post-aktionales Kompetenzerleben für weitere Berechnungen genutzt.

Die Skala „post-aktionales Kompetenzerleben“ korrelierte sowohl mit der Skala „aufgabenspezifisches Selbstkonzept“ signifikant zu $.57$, als auch mit der abhängigen Variablen Treatmentleistung zu $.39$. Um zu überprüfen, ob Lerner, die fehlerspezifisches ITF erhielten, ein höheres Kompetenzerleben berichten als die Schüler der KR-KCR Feedbackgruppe, wurde in einem ersten Auswertungsschritt eine Kovarianzanalyse mit der Kovariaten „aufgabenspezifisches Selbstkonzept“ berechnet. Signifikante Haupteffekte konnten sowohl für die Kovariate „aufgabenspezifisches Selbstkonzept“, $F(1, 25) = 14.15, p = .00, \omega^2 = .328$, als auch für die unabhängige Variable Feedback, $F(1, 25) = 8.35, p = .01, \omega^2 = .214$, nachgewiesen werden. Schüler mit einem hohen aufgabenspezifischen Fähigkeitskonzept hinsichtlich der Bearbeitung schriftlicher Subtraktionsaufgaben, denen zudem die Bearbeitung solcher Aufgaben Spaß bereitet, bewerteten ihre Leistungen während der Treatmentphase statistisch signifikant positiver. Hinsichtlich der motivationsförderlichen Wirkung des ITF zeigten sich hypo-

thesenkonforme Ergebnisse (H2). Lerner mit ITF ($M = 13.82$, $SE = 0.96$) berichteten ein höheres Kompetenzerleben als Lerner, die KR-KCR Feedback ($M = 9.90$, $SE = 0.96$) erhielten. Das heißt, diese Lerner waren zufriedener mit ihren Leistungen, bewerteten diese günstiger, empfanden die Aufgaben als weniger schwierig und anstrengend und erfuhren einen höheren Lerngewinn. Mit diesen kompetenzbezogenen Bewertungen einhergehend erlebten die Schüler der ITF Bedingung auch mehr Spaß beim Bearbeiten der computergestützten Subtraktionsaufgaben.

In einem zweiten Analyseschritt wurde die Treatmentleistung als eine zweite Kovariate mit berücksichtigt, obwohl nach Formel 1 durch Hinzunahme einer zweiten Kovariaten der definierte Grenzwert von .10 mit .11 überschritten wurde, dies allerdings nur knapp. Eine Kovarianzanalyse zeigte nur für die Skala „aufgabenspezifisches Selbstkonzept“ einen signifikanten Haupteffekt, $F(1, 23) = 12.54$, $p = .00$, $\omega^2 = .316$. Allerdings wurde für die unabhängige Variable Feedback das 5%-Niveau nur knapp verfehlt, $F(1, 23) = 3.43$, $p = .08$, $\omega^2 = .089$.

4.4 Diskussion

Die vorliegende Studie diente der empirischen Evaluation des entwickelten fehlerspezifischen ITFs für die schriftliche Subtraktion hinsichtlich seiner Lern- und motivationalen Wirksamkeit. Das in der vorliegenden Studie untersuchte fehlerspezifische ITF für schriftliche Subtraktionsaufgaben war auf der Grundlage der von Narciss (2004) vorgeschlagenen Gestaltungs- und Evaluationsprinzipien für informatives tutorielles Feedback entwickelt und evaluiert worden. Von daher repräsentiert diese Studie gleichzeitig auch eine empirische Überprüfung und damit Validierung dieser Gestaltungs- und Evaluationsprinzipien im Kontext einer schulischen Lernaufgabe.

Die Ergebnisse dieser Studie bestätigen die Annahme einer lern- und motivationsförderlichen Wirksamkeit des entwickelten ITFs für die schriftliche Subtraktion. So konnten für den lernwirksamen Einfluss des fehlerspezifischen ITFs die aufgestellten Hypothesen größtenteils bestätigt werden. Schüler, welche fehlerspezifisches ITF im Anschluss an eine fehlerhafte Aufgabenbearbeitung erhielten, zeigten sowohl während der Treatmentphase als auch bei der Bearbeitung von Nachtestaufgaben, die im Vortest fehlerhaft bearbeitet wurden, signifikant bessere Lernergebnisse als Schüler, die KR-KCR Feedback erhalten haben. Diese Ergebnisse bestätigen die Annahme, dass ITF die Korrektur von Fehlern unterstützt (Narciss, 2004). Die Berechnung von Gruppenunterschieden im Nachtest verfehlte nur knapp die Signi-

fikanzgrenze. Erklärbar ist dieses Ergebnis zum einen darüber, dass neben der korrektiven Funktion, welche primär bei der ITF Bedingung wirksam geworden sein dürfte, bei beiden Feedbackbedingungen korrekte Lösungen im Anschluss bestätigt wurden, so dass das Feedback beider Untersuchungsbedingungen auch bestätigend gewirkt haben dürfte. Einen empirischen Beleg für diese Annahme liefert das Ergebnis, dass für den Korrekturwahrscheinlichkeitsindex im Unterschied zur Nachtestleistung ein signifikanter Unterschied zugunsten der fehlerspezifischen ITF Gruppe nachgewiesen werden konnte. Hinsichtlich des signifikanten Unterschieds in der Treatmentleistung zugunsten der ITF Gruppe und des nicht gefundenen Nachtestunterschiedes kann als weitere Erklärung angeführt werden, dass die Lernenden bei der Bearbeitung des Nachtests im Unterschied zur Treatmentphase eine Transferleistung zu erbringen hatten. So hatten die Lernenden bei jeder Aufgabe eine Entscheidung im Hinblick auf die anzuwendende Subprozedur zu treffen. Eine erfolgreiche Aufgabenbeantwortung erforderte somit neben der Anwendung prozeduralen Wissens auch die Anwendung konditionalen Wissens, d.h. von Wissen, welche Bedingung, im Sinne welches Aufgabenmerkmal, welche Subprozedur erfordert. Ließ sich die Lernwirksamkeit des ITFs anhand beobachtbarer Indikatoren nachweisen, konnte dessen Lernwirksamkeit auch durch berichtbare Indikatoren bestätigt werden. So bewerteten die Schüler unter ITF die bereitgestellten Hinweise als sehr hilfreich und wünschten sich solche Hinweise öfter.

Auch hinsichtlich des motivationsförderlichen Einflusses von ITF konnte die aufgestellte Hypothese bestätigt werden. So bewerteten die Schüler unter der ITF Bedingung ihre während der Treatmentphase erbrachten Leistungen durchweg positiver als die Schüler unter der KR-KCR Feedbackbedingung. Die Schüler der ITF Gruppe zeigten demnach ein höheres post-aktionales Kompetenzerleben (Narciss, 2004) als die Schüler der Vergleichsbedingung. Dieses erhöhte Kompetenzerleben der Schüler kann zum einen darüber erklärt werden, dass die Schüler aufgrund einer besseren Treatmentleistung Erfahrungen zunehmender Kompetenz sammeln konnten, die darin zum Ausdruck kamen, dass sie (1) zufriedener mit ihren Leistungen waren, (2) diese positiver bewerteten, (3) angaben, mehr dazu gelernt zu haben, die Aufgaben als (4) weniger schwierig und (5) weniger anstrengend wahrnahmen und (6) ihnen das Bearbeiten der Aufgaben mehr Freude bereitet hat.

Zudem ist zu vermuten, dass über das Angebot zusätzlicher Informationen ohne gleichzeitige Bekanntgabe der korrekten Lösung, die Lerner die Erfahrung machten, ihre fehlerhafte Lösung eigenständig, unter Nutzung der strategischen Informationen, erfolgreich korrigieren

zu können. Lernerfolge können damit eher internal attribuiert werden (Narciss, 2004). Dies wiederum dürfte dazu geführt haben, dass die Lernenden der ITF Bedingung zufriedener mit ihrer Leistung waren, da bei ihnen Gefühle eigener Wirksamkeit hervorgerufen wurden. Eine Unterstützung dieser Ergebnisinterpretation liefert möglicherweise der Befund, dass, auch nach Ausparialisierung des Einflusses der Treatmentleistung auf das post-aktionale Kompetenzerleben, die Schüler der ITF Gruppe ihre Leistungen positiver bewerteten, auch wenn dieser Unterschied das Signifikanzniveau knapp verfehlte. Das heißt, dass das erhöhte Kompetenzerleben nicht nur über eine tatsächliche Leistungsverbesserung mediiert war. Möglicherweise unterstützte die tutorielle Komponente des ITF eine interne Attribuierung der Lernerfolge, da diese Erfolge der eigenen Anstrengung zugeschrieben werden konnten.

Die Berechnung der Effektstärken zeigt, dass die Unterschiede zwischen den Feedbackbedingungen nicht nur statistisch signifikant sondern auch praktisch bedeutsam sind. Die für die einzelnen beobachtbaren und berichtbaren Indikatoren berechneten Effektstärken fielen sogar stärker als erwartet aus. So konnten für die Treatmentleistung und die Selbstbewertung der Treatmentleistung über 30% bzw. 20% der Varianz über die Feedbackvariation erklärt werden, ein Effekt, welcher als groß interpretiert werden kann. Auch wenn die Feedbackgruppen sich in der Nachtestleistung nicht signifikant unterschieden, kann die berechnete Effektstärke von $\omega^2 = .07$ entsprechend der Effektstärkenklassifikation als mittlerer Effekt interpretiert werden und bestätigt damit Ergebnisse vergleichbarer Feedbackstudien. So fanden Evaluationsstudien, die Lerngewinne von intelligenten tutoriellen Systemen mit einer computerbasierten Kontrollbedingung verglichen (Albacete & VanLehn, 2000; Heffernan, 2001; Nagata & Swisher, 1995), mittlere Effektstärken zwischen $d = .6$ und $d = .8$. Auch diese Studien nutzten wesentliche Komponenten von informativem tutoriellem Feedback. So wurden den Lernenden im Rahmen eines computergestützten tutoriellen Dialogs bei einer fehlerhaften Aufgabenbearbeitung korrekturrelevante Informationen ohne gleichzeitigem Angebot der korrekten Lösung bereitgestellt.

Andererseits stehen die Ergebnisse der vorliegenden Studie in teilweisem Kontrast zu den häufig nicht signifikanten bzw. sogar widersprüchlichen Ergebnissen klassischer Feedbackstudien (Kulhavy et al., 1985; Kulhavy & Stock, 1989; Phye, 1979). Im Unterschied zu diesen Studien wurde das ITF in der vorliegenden Studie über einen sehr aufwendigen theoretisch und empirisch fundierten Designprozess entwickelt. So sollte die theoriegeleitete Entwicklung des fehlerspezifischen ITF die Identifikation korrekturrelevanter Informationen er-

möglichst haben. Durch die Bereitstellung der als relevant identifizierten Informationen ohne das Angebot der korrekten Lösung, verknüpft mit der Aufforderung einer nochmaligen Aufgabenbearbeitung, sollte eine aktive und vertiefte Auseinandersetzung mit diesen Informationen angeregt worden sein. Diese vertiefte Auseinandersetzung sollte zudem durch die Implementierung eines mastery levels weiter gefördert worden sein. Im Unterschied zu anderen Studien kann die nachgewiesene Lernwirksamkeit auch nicht auf eine höhere Studierzeit des ITF zurückgeführt werden. So war für beide Feedbackgruppen die Treatmentphase auf 30 Minuten begrenzt.

Dass sich die beiden Feedbackbedingungen in ihrer Anzahl an Korrekturversuchen unterscheiden, war der Überlegung geschuldet, ITF mit einer möglichst ökologisch validen Kontrollbedingung zu vergleichen. Der implementierte KR-KCR Feedbackalgorithmus repräsentiert das in den meisten mathematischen Lernprogrammen realisierte Feedback, welches dem Lerner nur einen Korrekturversuch bereitstellt, bevor die korrekte Lösung präsentiert wird. So hatte der Schüler unter der ITF Bedingung zwei Korrekturversuche, der Schüler unter der KR-KCR Feedbackbedingung nur einen. Das Argument, dass die Anzahl der Korrekturversuche und weniger die bereitgestellten Informationen bei einer Aufgabenbearbeitung die Schülerleistung moderiert haben könnte, lässt sich darüber entkräften, dass die präsentierten Aufgaben zu einem Aufgabentyp sich nur in den Ziffernkombinationen unterschieden, die Anforderungen aber innerhalb eines Aufgabentyps sonst identisch waren. So kann nach einer zweimaligen fehlerhaften Aufgabenbearbeitung unter der KR-KCR Feedbackbedingung die Bearbeitung einer neuen Aufgabe des gleichen Aufgabentyps im Sinne eines dritten Lösungsversuchs interpretiert werden.

Die Ergebnisse der vorliegenden Studie können als Bestätigung der postulierten Lern- und motivationalen Wirksamkeit von ITF gesehen werden, wenn dieses auf der Grundlage der von Narciss (2004) formulierten Gestaltungs- und Evaluationsprinzipien entwickelt und evaluiert wurde. Damit liefern die Ergebnisse auch eine empirische Validierung dieser Gestaltungs- und Evaluationsprinzipien für einen schulischen Aufgabenbereich.

5 EVALUATION DES FEHLERSPEZIFISCHEN ITFS IM VERGLEICH ZU UNTERSCHIEDLICH KOMPLEXEN FEEDBACKFORMEN – STUDIE 2

In der ersten Studie wurde informatives tutorielles Feedback (ITF) in seinem Einfluss auf verschiedene Leistungs- und Motivationsparameter mit dem wenig informativen KR-KCR Feedback verglichen. Für alle abhängigen Variablen waren hypothesenkonforme Unterschiede zu verzeichnen, die in den meisten Fällen statistisch abgesichert werden konnten.

Die vorliegende zweite Studie verfolgte aufbauend auf den Erkenntnissen der ersten Studie weiterführende Fragestellungen, welche Aussagen zur Lern- und motivationalen Wirksamkeit des fehlerspezifischen ITFs im Vergleich zu komplexeren Feedbackformen als das in der ersten Studie untersuchte KR-KCR Feedback ermöglichen sollten. Die Beantwortung folgender Fragen stand dabei im Fokus dieser zweiten Studie:

- (1) Lässt sich eine Lern- und motivationale Wirksamkeit des fehlerspezifischen ITFs im Vergleich zu Feedbackformen nachweisen, welche einmalig fehlerspezifische Informationen bereitstellen? Diese Fragestellung sollte klären, ob eine dem fehlerspezifischen ITF vergleichbare Lern- und motivationale Wirksamkeit auch auf einem ökonomischerem Weg zu erzielen ist.
- (2) Welche der einzelnen Feedbackstufen des fehlerspezifischen ITFs erweist sich am wirksamsten? Da im fehlerspezifischen ITF korrekturrelevante Informationen auf zwei verschiedenen Wegen bereitgestellt werden, sollte die Beantwortung dieser Frage dabei helfen, spezifische Aussagen zu der isolierten Wirksamkeit der beiden Feedbackstufen zu erhalten.

Um diese Fragen zu beantworten, wurden zwei weitere Feedbackformen aus dem vorliegenden ITF Algorithmus entwickelt, welche sich hinsichtlich der Komplexität ihres Informationsgehaltes und ihres Präsentationsmodus zwischen dem ITF und dem KR-KCR Feedback einordnen lassen. Diese beiden Feedbackformen werden im weiteren Verlauf der Arbeit als KR-KM Feedback (Knowledge of Response and Knowledge about Mistake Feedback) und als KR-KH Feedback (Knowledge of Response und Knowledge on How to proceed Feedback) bezeichnet. Beide Feedbackformen sind analog zu dem KR-KCR Feedback zweistufig und präsentieren daher im Unterschied zum ITF korrekturrelevante Informationen gemeinsam mit der korrekten Lösung. Eine unmittelbare Anwendung der präsentierten Informationen an der gleichen Aufgabe, welche den Lerner zu einer aktiveren Auseinandersetzung mit diesen

Informationen anregen soll, ist damit im Gegensatz zum ITF nicht gefordert. Das heißt, diesen beiden neuen Feedbackformen fehlt die tutorielle Komponente, welche das fehlerspezifische ITF kennzeichnet. Neben dem Präsentationsmodus unterscheiden sich beide neuen Feedbackformen auch hinsichtlich ihres Informationsgehaltes. Das KR-KM Feedback präsentiert hierbei die Informationen, die im ITF Algorithmus dem Lerner auf dem zweiten Feedbacklevel angeboten werden, wohingegen das KR-KH Feedback dem dritten Feedbacklevel des ITF Algorithmus entspricht. Obwohl beide neuen Feedbackformen korrekturrelevante Informationen bereitstellen, erfolgt dies aufgrund der gemeinsamen Präsentation dieser Informationen mit der korrekten Lösung nur einmalig bei jeder Aufgabenbearbeitung. Im Unterschied hierzu gewährleistet das zweistufige Angebot fehlerspezifischer Informationen durch das ITF den Erwerb korrektiver Informationen über zwei verschiedene Informationswege, die ergänzend wirksam werden können.

In die vorliegende Studie wurden diese beiden neuen Feedbackformen neben dem fehlerspezifischen ITF und dem KR-KCR Feedback aufgenommen, um vergleichende Aussagen dieser vier Feedbackformen hinsichtlich ihrer Lern- und motivationalen Wirksamkeit zu erhalten. Da bereits Kulhavy und Stock (1989) das Fehlen einer systematischen Validierung existierender Feedbackbefunde kritisiert haben, bestand ein weiteres Anliegen dieser 2. Studie darin, die in der ersten Studie erhaltenen Ergebnisse zu replizieren.

5.1 Fragestellungen und Forschungshypothesen

Nach Hager (2000, S. 212) sollte ein Vergleich zwischen verschiedenen Programmen nur durchgeführt werden, wenn die Vergleiche innerhalb der Programme zu dem vorhergesagtem Ausgang geführt haben, da andernfalls die Gefahr von Fehlinterpretationen zu groß ist. Ein erstes Ziel dieser Studie bestand demnach darin, die isolierte Wirksamkeit der vier Feedbackformen zu untersuchen, um abzusichern, dass signifikante Wirksamkeitsunterschiede in den abhängigen Variablen nicht auf gegenläufige Veränderungen in den einzelnen Feedbackgruppen zurückgeführt werden können, die Programmterfolge vortäuschen, obwohl keine Verbesserungen aufgetreten sind. Eine erste Fragestellung lautete daher:

Fragestellung 1: Lassen sich mit allen vier Feedbackbedingungen leistungsbezogene und motivationale Merkmale fördern?

Dass Lernaufgaben den Wissenserwerb unterstützen, gilt als ein gut replizierter Befund der Lehr-Lernforschung (Haynie, 1994; 1997). Ein weiterer gut replizierter Befund der Feedbackforschung zeigt aber auch, dass die alleinige Bearbeitung von Lernaufgaben ohne eine Rückmeldung über die Qualität der Aufgabenbearbeitung weniger lernwirksam ist, als das Bearbeiten von Lernaufgaben, an deren Anschluss der Lerner wenigstens darüber informiert wird, ob die Aufgabe fehlerhaft oder korrekt gelöst wurde, d.h. Knowledge of Result bereitgestellt wird (siehe Kapitel 2.2).

Hypothese 1: Da in der Feedbackbedingung mit dem geringsten Informationsgehalt dieser Studie – der KR-KCR Feedbackbedingung – die Lerner Aufgaben zu bearbeiten haben, an deren Anschluss ihnen nicht nur mitgeteilt wird, ob ihr Ergebnis fehlerhaft oder korrekt ist (KR Feedback), sondern ihnen nach einem zweiten fehlerhaften Lösungsversuch die richtige Lösung präsentiert wird (KCR Feedback), ist zu vermuten, dass bei Lernern unter allen vier Feedbackbedingungen signifikante Leistungszuwächse zu beobachten sind. Diese Leistungszuwächse sollten sich aufgrund eines erhöhten Kompetenzerlebens auch positiv auf die motivationale Bewertung der erfahrenen Leistungssituation auswirken.

Mit dem Nachweis einer isolierten Wirksamkeit aller vier Feedbackbedingungen ist es möglich, in einem nächsten Schritt Wirksamkeitsunterschiede zwischen den verschiedenen informativen Feedbackformen hinsichtlich lernbezogener und motivationaler Merkmale zu untersuchen. Konnte in der ersten Studie die Lern- und motivationale Wirksamkeit des fehlerspezifischen ITFs im Vergleich zu der wenig informativen KR-KCR Feedbackform nachgewiesen werden, bestand das zentrale Anliegen dieser zweiten Studie darin, das entwickelte ITF gegen Feedbackformen mit einem höheren Informationsgehalt als KR-KCR Feedback zu evaluieren. Gleichzeitig diene diese Studie aber auch einer Replikation der in Studie 1 erhaltenen Ergebnisse. Eine zweite Fragestellung lautete daher:

Fragestellung 2: Lässt sich eine höhere Lern- und motivationale Wirksamkeit von ITF auch im Vergleich zu komplexeren Feedbackformen als KR-KCR Feedback nachweisen bzw. lassen sich die in Studie 1 gewonnenen Ergebnisse replizieren, d.h. kann eine höhere Lernwirksamkeit von ITF gegenüber KR-KCR Feedback auch in dieser Studie gezeigt werden?

Mit dieser Fragestellung sollen Wirksamkeitsunterschiede zwischen diesen verschiedenen Feedbackformen untersucht werden. Beide in die zweite Studie zusätzlich aufgenommenen Feedbackformen – KR-KM Feedback und KR-KH Feedback - unterscheiden sich von ITF sowohl im Hinblick auf ihren Informationsgehalt als auch bezüglich ihres Präsentationsmodus. So weisen beide Feedbackformen einen geringeren Informationswert als ITF auf. Diese Informationen werden zudem dem Lerner zusammen mit der richtigen Lösung nach einem zweiten fehlerhaften Lösungsversuch präsentiert, so dass eine unmittelbare Anwendung der präsentierten Informationen an derselben fehlerhaft bearbeiteten Aufgabe nicht erfolgt, eine aktive Auseinandersetzung mit den präsentierten Informationen somit nicht unmittelbar gefordert wird.

Zur Untersuchung der Fragestellung wurden in dieser Studie die gleichen beobachtbaren Leistungsindikatoren wie in Studie 1 genutzt, d.h. (1) die Treatmentleistung, operationalisiert über die Anzahl erfolgreich bearbeiteter Aufgabentypen während des computergestützten Treatments (2) die Anzahl im Nachtest korrekt gelöster Aufgaben sowie (3) die Anzahl im Nachtest korrigierter Vortestaufgaben

Hypothese 2a: Unter folgender Annahme – Je informativer das Feedback und je stärker eine aktive Auseinandersetzung mit den präsentierten Informationen angeregt wird, desto stärker wird der Erwerb von Fertigkeiten der schriftlichen Subtraktion gefördert – sind folgende Ergebnisse zu erwarten:

H2a₁: ITF erweist sich als lernwirksamer als die drei weniger informativen Feedbackformen.

H2a₂: KR-KM Feedback und KR-KH Feedback, welche sich hinsichtlich ihres Informationsgehaltes zwischen dem ITF und dem KR-KCR Feedback einordnen lassen, erweisen sich als nicht so lernwirksam wie ITF, aber als lernwirksamer als KR-KCR Feedback.

Eine erhöhte Lernwirksamkeit sollte in einem erhöhten Kompetenzerleben resultieren. Dieses Kompetenzerleben sollte unter der ITF Bedingung verstärkt zu beobachten sein, da über die nochmalige Bearbeitung der fehlerhaft bearbeiteten Aufgabe auf der Grundlage der bereitgestellten korrekturrelevanten Informationen Lernerfolge der eigenen Anstrengung zugeschrieben werden sollten. Da in vielen Motivationstheorien ein enger Zusammenhang zwi-

schen dem Kompetenzerleben und intrinsischen Anreizwerten postuliert wird (z.B. die kognitive Evaluations-Theorie, Deci & Ryan, 1985; die Interessentheorie, Krapp, 1999; die Flow-Theorie, Csikszentmihaly, 1975; siehe zusammenfassend Narciss, 2004), wird darüber hinaus erwartet, dass das Erleben von Kompetenz auch mit einer erhöhten intrinsischen Motivation einhergeht.

Hypothese 2b: Unter der Annahme, dass Erfolgserlebnisse zu einem erhöhten Kompetenzerleben führen, vor allem wenn diese Erfolge noch internal attribuiert werden, sind folgende Ergebnisse zu erwarten:

H2b₁: Lernende mit ITF zeigen aufgrund häufigerer Erfolgserlebnisse und deren internalen Attribution ein höheres Kompetenzerleben, eine positivere Selbstbewertung ihrer Leistungen und eine höhere intrinsische Motivation im Vergleich zu den drei anderen Feedbackbedingungen.

H2b₂: Da Lernende, welche entweder KR-KM Feedback oder KR-KH Feedback dargeboten bekommen, einen höheren Lernfortschritt zeigen als Lernende, welche KR-KCR Feedback dargereicht bekommen, werden diese ihre Leistungen im Vergleich zu Lernern, welche KR-KCR Feedback erhalten, aufgrund eines erhöhten Kompetenzerlebens auch positiver bewerten.

5.2 Methode

Da diese Studie der ersten Studie in vielen Merkmalen gleicht, wird im weiteren Verlauf nur auf Aspekte eingegangen, die für diese Studie neu und relevant sind. Auf bereits bekannte Aspekte wird verwiesen.

Zur Beantwortung der Fragestellungen und Prüfung der Forschungshypothesen wurde zu Schuljahresbeginn im August/September 2001 eine experimentelle Untersuchung mit Schülern der vierten Jahrgangsstufe von zehn Dresdner Grundschulen und Schülern der sechsten Jahrgangsstufe von drei Förderschulen durchgeführt. Entsprechend der Fragestellungen waren eine Experimentalgruppe und drei Vergleichsgruppen realisiert. Im Unterschied zur ersten Studie beinhaltete die zweite Studie zwei Treatmentzeitpunkte, die eine Woche auseinander lagen. Diese Entscheidung resultierte aus dem Anliegen dieser Studie, isolierte Wirksamkeitsaussagen zu allen vier Feedbackbedingungen zu treffen.

5.2.1 Versuchspersonen

Mit dem Ziel, Schüler mit Problemen in der schriftlichen Subtraktion zu identifizieren, wurde 501 Schülern der 13 Schulen während des regulären Unterrichts ein Vortest mit 24 Aufgaben zur schriftlichen Subtraktion zur Bearbeitung vorgelegt. Für die Teilnahme an der Evaluationsstudie wurden die Schüler ausgewählt, die mindestens 1/3 aller Aufgaben des Vortests (8 von 24) fehlerhaft bearbeitet hatten. Dieses veränderte Selektionskriterium (50% fehlerhaft bearbeitete Aufgaben in Studie 1) resultierte aus den in diesem Vortest eingesetzten Aufgaben. Im Gegensatz zum Vortest der Studie 1 beinhalteten die Aufgaben der Studie 2 nur eine Aufgabenanforderung (in Studie 1 zwei Aufgabenanforderungen) und stellten somit geringere Anforderungen an den Schüler. Auf der Basis dieses Selektionskriteriums wurden 113 Schüler ausgewählt, 82 Grundschüler der vierten Jahrgangsstufe ($w = 36$, $m = 46$), die zum Zeitpunkt der Untersuchung im Durchschnitt 9.73 ($SD = .79$) Jahre alt waren und 31 Förderschüler der sechsten Jahrgangsstufe ($w = 10$, $m = 21$), die zum Zeitpunkt der Erhebung im Durchschnitt 12.0 ($SD = .56$) Jahre alt waren.

5.2.2 Untersuchungsdesign

Eine Beantwortung der Forschungsfragen dieser Studie erforderte wie auch in Studie 1 eine vergleichende Evaluation zur Prüfung der Wirksamkeitsunterschiedshypothesen. Im Hinblick auf die vorliegenden Fragestellungen wurde ein Vortest-Nachtest-Vergleichsgruppen-Plan über zwei Messzeitpunkte mit folgenden unabhängigen Variablen, Kontrollvariablen und abhängigen Variablen realisiert:

- Die unabhängige Variable stellte die Feedbackbedingung dar, welche vierfach gestuft war:
 - Fehlerspezifisches Informatives Tutorielles Feedback (enthält KR-KM-KH-KCR Feedback)
(fehlerspezifisches ITF)
 - Knowledge of Response und Knowledge about Mistake Feedback mit KCR (KR-KM Feedback)
 - Knowledge of Response und Knowledge on How to proceed Feedback mit KCR (KR-KH Feedback)
 - Knowledge of Response und Knowledge of Correct Response Feedback (KR-KCR Feedback)

Die Beschreibung der in dieser Studie neu implementierten Feedbackbedingungen KR-KM Feedback und KR-KH Feedback erfolgt mit den Abbildungen 21 und 22 in diesem Kapitel.

- Als Kontrollvariablen wurden (1) das aufgabenspezifische Vorwissen, (2) individuelle motivationale Voraussetzungen wie das aufgabenspezifische Fähigkeitskonzept, das Interesse und die Freude beim Lösen schriftlicher Subtraktionsaufgaben sowie (3) die Computererfahrung über einen Vortest und eine Vorbefragung erfasst.
- Abhängige Variablen, die den Einfluss des Feedbacks auf kognitive Parameter erfassten, repräsentierten wie auch in Studie 1 folgende drei Leistungsindikatoren: die Treatmentleistung, die Nachtestleistung sowie der Korrekturwahrscheinlichkeitsindex im Nachtest. Der Einfluss des Feedbacks auf post-aktionale motivationale Parameter wurde über einen Fragebogen mit sieben Items erfasst, der geringfügig von dem in Studie 1 eingesetzten Fragebogen differierte. Alle abhängigen Variablen wurden zu beiden Messzeitpunkten erhoben.

Der dieser Untersuchung zugrundeliegende Vortest-Nachtest-Vergleichsgruppen-Plan mit zwei Messzeitpunkten lässt sich folgendermaßen darstellen (Tabelle 29):

Tabelle 29:

Vortest-Nachtest-Vergleichsgruppen-Plan mit zwei Messzeitpunkten

	Vortest	Treatment MP1	Nachtest 1	Treatment MP2	Nachtest 2
ITF (EG ^a)		ITF		ITF	
KR-KM (VG1 ^b)	24 Subtraktions- aufgaben	KR-KM	24 Subtraktions- aufgaben	KR-KM	24 Subtraktions- aufgaben
KR-KH (VG2)		KR-KH		KR-KH	
KR-KCR (VG3)		KR-KCR		KR-KCR	

^aEG = Experimentalgruppe

^bVG = Vergleichsgruppe

Im folgenden erfolgt eine Beschreibung der unabhängigen Variablen Feedback. Der ITF Algorithmus wurde bereits ausführlich im Kapitel 3.3.3 und der KR-KCR Feedbackalgorithmus detailliert im Kapitel 4.2.2 beschrieben. Im Unterschied zur 1. Studie wurde für diese

Studie das Erreichen des mastery levels folgendermaßen definiert: Das Lernkriterium ist erreicht, wenn drei Aufgaben mit derselben Aufgabenanforderung in Folge beim ersten Versuch richtig gelöst werden. Das Lernkriterium von zwei Aufgaben auf drei Aufgaben anzuheben, erfolgte mit der Intention, die Wahrscheinlichkeit, zwei Aufgaben in Folge durch Kopfrechnen richtig zu lösen, zu verringern. So lag nach Auswertung der Log Files der 1. Studie die Vermutung nahe, dass einige Schüler das mastery level nur erreicht hatten, wenn in Folge der automatischen Aufgabengenerierung zwei zweistellige Aufgaben nacheinander präsentiert wurden, die von diesen Schülern korrekt im Kopf gelöst werden konnten.

Im weiteren sind der KR-KM Feedbackalgorithmus und der KR-KH Feedbackalgorithmus ausführlich beschrieben und in Abbildung 21 und Abbildung 22 dargestellt.

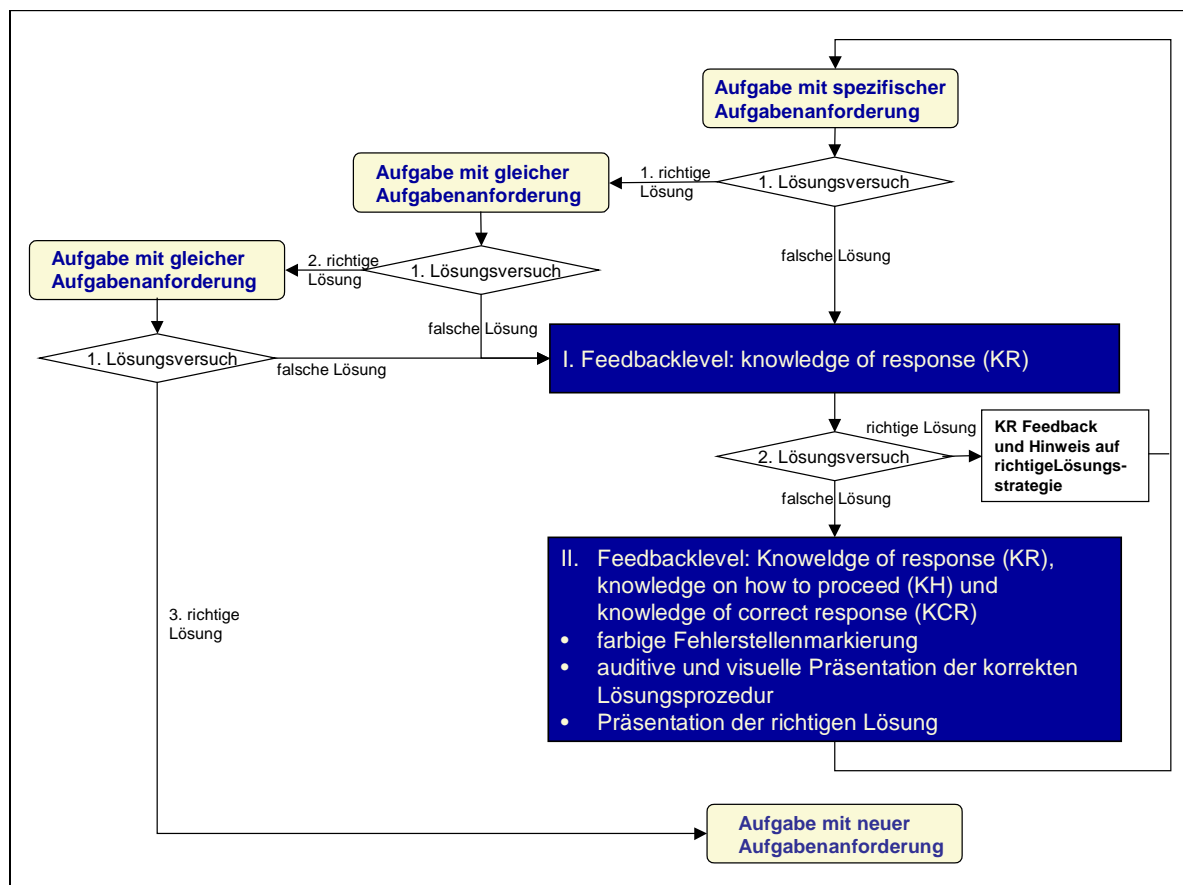


Abbildung 21: Darstellung des KR-KM Feedbackalgorithmus

Analog zum ITF und KR-KCR Feedback erhält der Lerner in beiden Feedbackvarianten nach einer fehlerhaften Aufgabenbearbeitung die Rückmeldung, dass seine Aufgabenlösung noch falsch sei, verbunden mit der Aufforderung, es noch einmal zu versuchen. Nach

einer erneuten fehlerhaften Aufgabenlösung wird die Fehlerstelle rot markiert. In Abhängigkeit des jeweiligen Feedbackalgorithmus erfolgt nun entweder ein auditiver Hinweis auf den diagnostizierten Fehler und die korrekte Lösungsstrategie (KR-KM Feedback) oder dem Lerner wird an der konkreten fehlerhaft gelösten Aufgabe modelliert, wie diese Aufgabe zu lösen ist, d.h. dem Lerner wird die korrekte Lösungsprozedur sowohl visuell als auch auditiv vorgeführt (KR-KH Feedback). Im Anschluss an diese korrektiven Informationen wird in beiden Feedbackvarianten das korrekte Ergebnis (KCR Feedback) präsentiert. Eine Aufgabe mit dem gleichen Aufgabenmerkmal wird so lange zur erneuten Bearbeitung vorgelegt, bis der Lerner drei Aufgaben in Folge im 1. Lösungsversuch korrekt bearbeiten konnte.

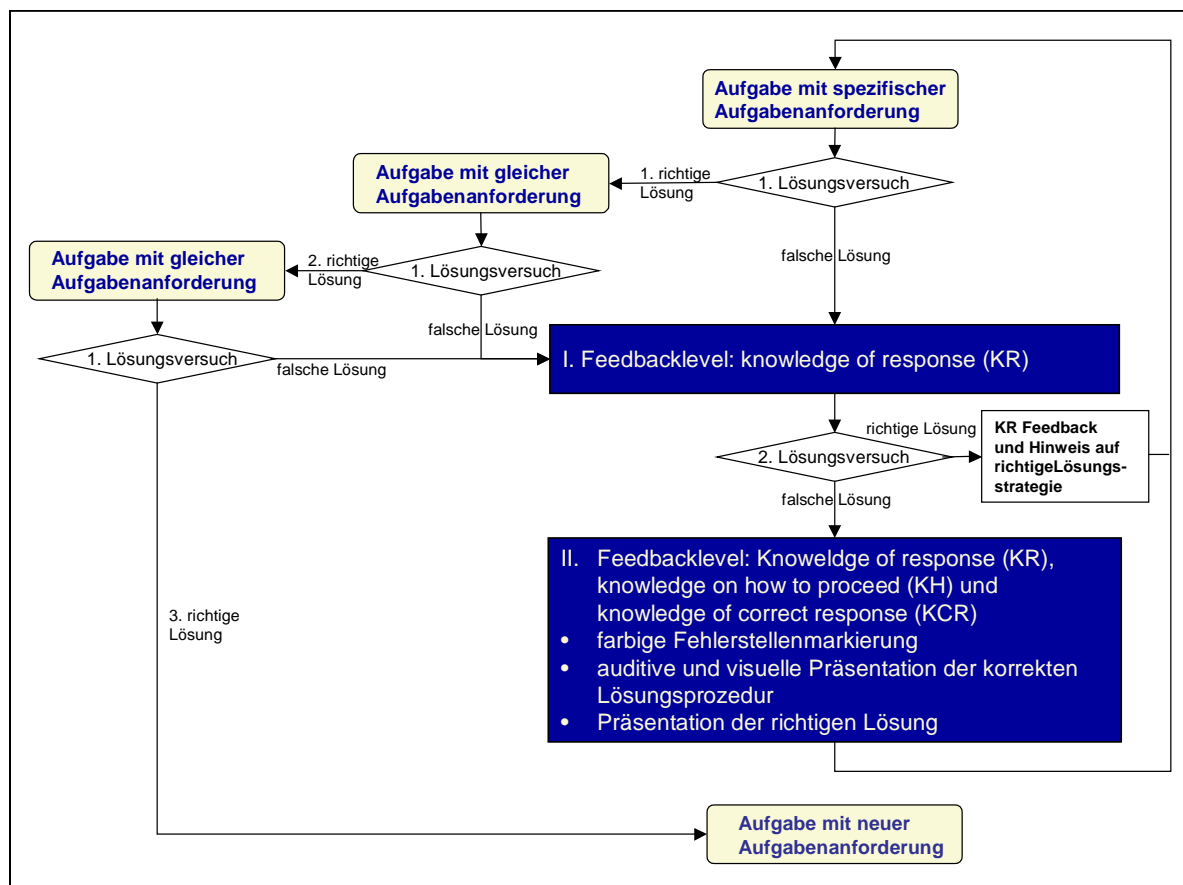


Abbildung 22: Darstellung des KR-KH Feedbackalgorithmus

Beide Feedbackalgorithmen, in ihrer Komplexität zwischen dem ITF und dem KR-KCR Feedbackalgorithmus anzusiedeln, sind - und da sind sie dem KR-KCR Feedbackalgorithmus vergleichbar - zweistufig. Nach einer try-again Aufforderung und einem zweiten

fehlerhaften Lösungsversuch wird entweder KM oder KH in Kombination mit KCR Feedback angeboten.

5.2.3 Programmbeschreibung der Übungssoftware „Subtrativo“

Die detaillierte Beschreibung der Übungssoftware „Subtrativo“ erfolgte bereits im Kapitel 4.2.3. Vergleichbar zum ITF Algorithmus und zum KR-KCR Feedbackalgorithmus melden der KR-KM Feedbackalgorithmus und der KR-KH Feedbackalgorithmus nach einem ersten fehlerhaften Lösungsversuch dem Schüler zurück, dass seine Lösung noch einen Fehler enthalte (KR Feedback), verbunden mit der Aufforderung, es noch einmal zu versuchen (siehe Abbildung 11). Bearbeitete der Schüler die Aufgabe auch im zweiten Lösungsversuch falsch, erhielt er entweder KM Feedback (siehe Abbildung 23) oder KH Feedback (siehe Abbildung 24), verbunden mit der Mitteilung der richtigen Lösung (KCR Feedback).

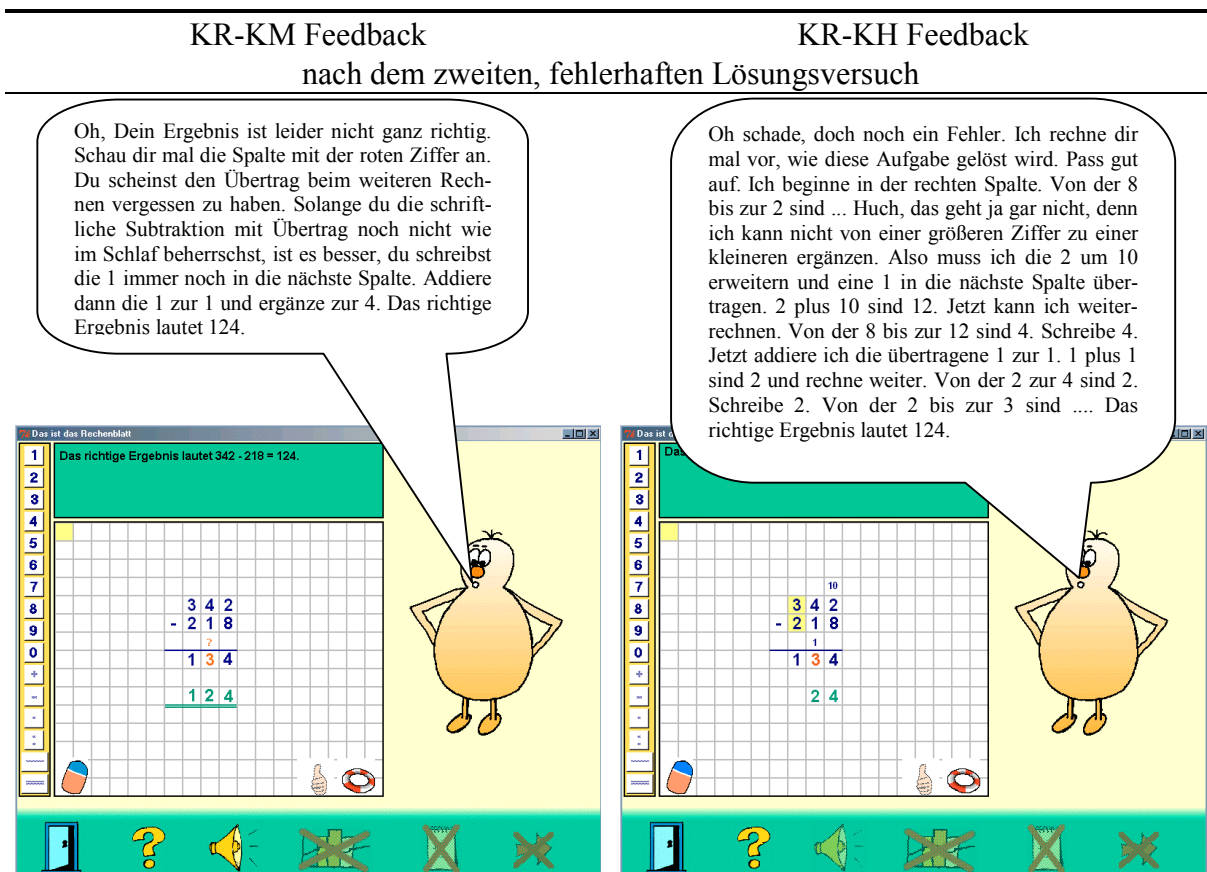


Abbildung 23: KM Feedback - Informationen zu Ort und Art des Fehlers und Hinweis auf korrekte Lösungsstrategie

Abbildung 24: KH Feedback - auditive und visuell-sequentielle Präsentation des korrekten Lösungsalgorithmus

5.2.4 Messinstrumente

Die im weiteren vorgestellten Untersuchungsmaterialien dienten der Erfassung der Kontrollvariablen und der abhängigen Variablen. Variablen, deren Erfassung bereits ausführlich in Studie 1 beschrieben wurde, sind über einen jeweiligen Vermerk indiziert.

Kontrollvariable: Aufgabenspezifisches Vorwissen

Auf die zentrale Rolle, die das aufgabenspezifische Vorwissen als Prädiktor für Lernerfolg spielt, wurde bereits im Kapitel 4.2.4 ausführlich eingegangen. In Studie 1 konnte diese Annahme empirisch bestätigt werden. So korrelierte die Vortestleistung mit der Nachtestleistung zu .39. Aus diesem Grund wurde auch in dieser Studie das aufgabenspezifische Vorwissen als Kontrollvariable über einen Vortest mit 24 Aufgaben zur schriftlichen Subtraktion erhoben. Um identische Aufgabenanforderungen der Vortestaufgaben und der Treatmentaufgaben zu gewährleisten, wurde für diese Untersuchung im Vergleich zur ersten Untersuchung folgendes Vorgehen bei der Konstruktion des Vortestes gewählt: Es wurden vergleichbar zur ersten Untersuchung wieder die in der Voruntersuchung identifizierten Aufgabenanforderungen als Konstruktionsgrundlage genutzt. Im Gegensatz zur 1. Studie, in welcher jede Vortestaufgabe in der Mehrzahl zwei Aufgabenanforderungen beinhaltete und somit an den Lerner höhere Anforderungen stellte als die Bearbeitung der computergestützt präsentierten Treatmentaufgaben, beinhaltete jede Vortestaufgabe der 2. Studie per definitionem nur eine Aufgabenanforderung. Auf dieser Konstruktionsgrundlage wurden für jede Aufgabenanforderung vier Aufgaben entwickelt (siehe Tabelle 30). Das aufgabenspezifische Vorwissen wurde analog zur 1. Studie über die Anzahl der im Vortest korrekt bearbeiteten Subtraktionsaufgaben operationalisiert.

Tabelle 30:

Vortest mit 24 Subtraktionsaufgaben zur Erfassung des aufgabenspezifischen Vorwissens

1	724	2	589	3	578	4	482	5	802	6	379	7	797	8	297
	<u>- 501</u>		<u>- 62</u>		<u>- 523</u>		<u>- 165</u>		<u>- 534</u>		<u>- 95</u>		<u>- 160</u>		<u>- 85</u>
9	483	10	925	11	690	12	256	13	562	14	856	15	269	16	758
	<u>- 381</u>		<u>- 359</u>		<u>- 267</u>		<u>- 73</u>		<u>- 201</u>		<u>- 34</u>		<u>- 165</u>		<u>- 462</u>
17	240	18	526	19	478	20	468	21	298	22	924	23	507	24	828
	<u>- 128</u>		<u>- 42</u>		<u>- 130</u>		<u>- 53</u>		<u>- 178</u>		<u>- 782</u>		<u>- 386</u>		<u>- 36</u>

Aufgabenanforderung	Aufgaben
Null im Subtrahenden	1, 7, 13, 19
Leere Stelle	2, 8, 14, 20
Gleiche Ziffern übereinander	3, 9, 15, 21
Übertrag	4, 10, 16, 22
Ergänzen zur 10	5, 11, 17, 23
Übertrag in leere Stelle	6, 12, 18, 24

Kontrollvariablen: Individuelle motivationale Voraussetzungen wie das aufgabenspezifische Fähigkeitskonzept, die Freude als tätigkeitsspezifischer Anreiz und das Interesse als gegenstandsspezifischer Anreiz

Auf die Bedeutung individueller motivationaler Voraussetzungen für das Lern- und Leistungsverhalten wurde bereits im Kapitel 4.2.4 eingegangen. In dieser Studie wurden vor dem Hintergrund der in Studie 1 dargelegten Erkenntnisse zum Einfluss kompetenzbezogener Schemata auf das Lern- und Leistungsverhalten das aufgabenspezifische Fähigkeitskonzept, die Freude als tätigkeitsspezifischer Anreiz und das Interesse als gegenstandsspezifischer Anreiz über einen Fragebogen mit fünf Items erfasst.

Im Unterschied zur ersten Studie wurde in der vorliegenden Studie das aufgabenspezifische Fähigkeitskonzept über drei Items operationalisiert, die sich in den Arbeiten von Narciss (1999, 2001, 2004) bewährt hatten. Neben diesen drei Items zum aufgabenspezifischen Fähigkeitskonzept (siehe Tabelle 31) wurden zwei weitere Items aufgenommen, welche Einschätzungen zu gegenstands- und tätigkeitsbezogenen Anreizen erfassten (siehe Tabelle 31).

Der besseren Lesbarkeit halber wurden alle fünf Items umgepolt, so dass ein hoher Wert auch eine hohe Ausprägung auf diesem Item bedeutet. Aufgrund der in der ersten Studie dargelegten unterschiedlichen Sichtweisen, ob kompetenzbezogene Schemata auch affektive Komponenten wie die Freude und das Interesse beinhalten oder nicht, wurde eine Entscheidung hinsichtlich der Zusammenfassung der Items auf der Grundlage empirischer Itemanalysen und exploratorischer Faktorenanalysen getroffen.

Tabelle 31:

Items zur Erfassung individueller motivationaler Voraussetzungen (vgl. Narciss, 2004)

Variable	Item	6stufige Ratingskala
Einschätzung der eigenen Begabung	Ich denke, ich bin für das Subtrahieren sehr begabt.	
Subjektive Aufgabenschwierigkeit	Ich finde das Lösen von Subtraktionsaufgaben sehr einfach.	Stimmt total (1)
Subjektive Leistungseinschätzung	Ich denke, ich kann Subtraktionsaufgaben sehr gut lösen.	Stimmt (2) Stimmt etwas (3) Stimmt eher nicht (4)
Gegenstandsspezifischer Anreiz (Interesse)	Ich finde das Lösen von Subtraktionsaufgaben sehr interessant.	Stimmt nicht (5) Stimmt überhaupt nicht (6)
Tätigkeitsspezifischer Anreiz (Freude)	Das Lösen von Subtraktionsaufgaben macht mir sehr viel Spaß.	

Eine Hauptkomponentenanalyse über alle fünf Items lieferte eine zweifaktorielle Faktorenstruktur mit Eigenwerten größer 1, die zusammen 76% der Varianz der fünf Items erklären (Eigenwerte von 2.73 und 1.06). Um die Faktoren sinnvoll interpretieren zu können, wurde im Anschluss eine orthogonale Rotation über die Faktorenladungsmatrix nach der Varimax-Methode durchgeführt. Diese lieferte für die drei Items zum aufgabenspezifischen Fähigkeitskonzept Faktorladungen größer .78 auf Faktor 1, so dass dieser Faktor als aufgabenspezifische Fähigkeit interpretiert werden kann. Auf Faktor 2 lud vor allem das Item Interesse mit .94. Das Item Freude lud auf beiden Faktoren ungefähr gleich hoch. Auf Grund der Ergebnisse dieser Faktorenanalyse wurde über die drei Items zum aufgabenspezifischen Fähigkeitskonzept der Konsistenzkoeffizient (Cronbachs α) berechnet, welcher mit $\alpha = .81$ auf eine hohe interne Konsistenz der Items hinweist, so dass über alle drei Items zum Fähigkeitskonzept ein Summenscore gebildet und dieser Wert bei der weiteren Auswertung berücksichtigt wurde. Fehlte nicht mehr als eine Itembeantwortung, so wurde dieser Missing-Wert durch den Mittelwert der Skala der übrigen Items ersetzt.

Kontrollvariable: Computererfahrung

Da sich in der 1. Studie die Computererfahrung als ein guter Prädiktor für die Treatmentleistung am Computer erwiesen hatte, wurde auch in dieser Studie die Computererfahrung als Kontrollvariable über folgendes Item auf einer 5 stufigen Ratingskala von „täglich“ [1] bis „nie“ [5] erfasst:

- Wie häufig arbeitest oder spielst du am Computer?

Auch im Fall der Computererfahrung wurden die Werte aus Gründen der Lesbarkeit umgepolt, so dass ein hoher Wert einer hohen Computererfahrung entspricht.

Abhängige Variable: Treatmentleistung

Die Schüler arbeiteten unter allen vier Feedbackbedingungen jeweils eine halbe Stunde mit dem computergestützten Programm. Die Treatmentleistung wurde darüber operationalisiert, wie viele Aufgabentypen in dieser halben Stunde erfolgreich bearbeitet werden konnten. Ein Aufgabentyp galt als erfolgreich bearbeitet, wenn der Schüler drei Aufgaben dieses Aufgabentyps in Folge im ersten Lösungsversuch korrekt bearbeitet hatte. War dies der Fall, hatte der Lerner das mastery level erreicht und das Aufgabentypfenster öffnete sich, so dass ein neuer Aufgabentyp ausgewählt werden konnte. Hatte der Lerner alle sechs Aufgabentypen einmal erfolgreich bearbeitet, wurden diese Aufgabentypen dem Lerner erneut präsentiert, um eine Festigung der erworbenen Fertigkeiten zu unterstützen.

Die Treatmentleistung resultierte demnach aus der Anzahl erreichter mastery level Zustände, d.h. aus der Anzahl erfolgreich bearbeiteter Aufgabentypen während der Treatmentphase. Mit dem Maß der Treatmentleistung war es möglich, den lernwirksamen Einfluss aller vier Feedbackbedingungen bereits während der Treatmentphase zu erfassen.

Abhängige Variable: Nachtestleistung

Die Nachtestleistung wurde analog zur Studie 1 operationalisiert über die Anzahl korrekt bearbeiteter Aufgaben im Nachtest, welcher im Anschluss an die Treatmentphase als Paper and Pencil Test den Schülern zur Bearbeitung vorgelegt wurde. Die im Nachtest präsentierten Aufgaben enthielten die gleichen 24 Aufgaben aus dem Vortest.

Abhängige Variable: Korrekturwahrscheinlichkeitsindex

Der Korrekturwahrscheinlichkeitsindex präsentierte analog zur Studie 1 die bedingte Wahrscheinlichkeit, dass eine im Vortest fehlerhaft gelöste Aufgabe im Nachtest richtig gelöst wird. Zur Bestimmung des Index wurde die Anzahl der im Vortest falsch, im Nachtest jedoch richtig gelösten Aufgaben (Antwortmuster $f1 \rightarrow r2$) zur Gesamtzahl der im Vortest falsch gelösten Aufgaben (Antwortmuster $f1 \rightarrow r2$ und $f1 \rightarrow f2$) ins Verhältnis gesetzt.

Abhängige Variable: Post-aktionale motivationale Parameter

Um den Einfluss der Feedbackbedingungen auf die post-aktionalen motivationalen Parameter zu erfassen, wurde den Schülern im Anschluss an jede Treatmentphase ein Fragebogen mit sieben Items dargeboten. Auf einer 6 stufigen Ratingskala von „stimmt total“ [1] bis „stimmt überhaupt nicht“ [6] wurden die Schüler gebeten, einzuschätzen, wie sie die Arbeit mit dem Programm unter bestimmten motivationalen Aspekten wahrgenommen haben.

Im Unterschied zur 1. Studie wurden zwei weitere Items, die den intrinsischen Wert der Aufgabe erfassen sollten, neben dem tätigkeitsspezifischen Anreiz Freude, mit erhoben: das Interesse als gegenstandsspezifischer Anreiz und Flow (Csikszentmihalyi, 1975) als Ausdruck des völligen Aufgehens in einer Tätigkeit. Neben diesen drei Items, die den intrinsischen Wert der Aufgabe erfassen sollten, wurden vier weitere Items zur Erfassung des Kompetenzerlebens und der Selbstbewertung in den Fragebogen aufgenommen, so dass die in Tabelle 32 dargestellten Items den Schülern zur Beantwortung vorgelegt wurden. Alle Items außer dem Item „Lernfortschritt“ wurden aus Gründen der Lesbarkeit umgepolt, so dass ein hoher Wert auf einem Item einer positiven Ausprägung entspricht.

Für den nach jeder Treatmentphase dargebotenen Fragebogen mit den motivationalen Items wurde jeweils eine Hauptkomponentenanalyse und eine Reliabilitätsanalyse über alle sieben Items durchgeführt. Für den Fragebogen zum 1. Messzeitpunkt konnten zwei Faktoren mit einem Eigenwert größer 1 extrahiert werden. Der erste Faktor mit einem Eigenwert von 3.72 klärte 53% der Varianz auf, der zweite Faktor mit einem Eigenwert von 1.07 klärte 15% der Varianz auf. Um die Faktoren sinnvoll interpretieren zu können, wurde im Anschluss eine orthogonale Rotation über die Faktorenladungsmatrix nach der Varimax-Methode durchgeführt. Diese lieferte für die Items Zufriedenheit, Leistungseinschätzung, Lerngewinn, Lernfortschritt und Freude Faktorladungen größer .61 auf Faktor 1. Auf Faktor 2 luden die Items Interesse und Flow mit Faktorladungen .76. Die Berechnung des Konsistenzkoeffizienten

(Cronbachs α) über alle sieben Items erbrachte mit einem alpha-Wert von .84 einen hohen Reliabilitätskoeffizienten. Da die Messqualität der Skala durch die Eliminierung des Flow-Items und des Interesse-Items nur minimal von .84 auf .86 verbessert werden konnte, wurde über alle sieben Items ein Summenscore gebildet und dieser Gesamtscore in die weiteren Berechnungen mit einbezogen. Dieser Gesamtscore wurde in Analogie zu Studie 1 als „post-aktionales Kompetenzerleben“ bezeichnet. Wie auch in Studie 1 war diese Bezeichnung dadurch motiviert, dass das Kompetenzerleben in einer positiven Selbstbewertung der eigenen Leistung, einem wahrgenommenen Lernfortschritt und einer erhöhten intrinsischen Motivation zum Ausdruck kommen sollte.

Tabelle 32:

Items zur Erfassung des post-aktionalen Kompetenzerlebens

Dimension	Itemtext	6 stufige Ratingskala
Zufriedenheit	Ich bin mit meiner Leistung in dieser halben Stunde sehr zufrieden.	
Leistungseinschätzung	Ich fand meine Leistung in dieser halben Stunde sehr gut.	
Lerngewinn	Ich denke, ich habe in dieser halben Stunde sehr viel dazugelernt.	Stimmt total (1) Stimmt (2) Stimmt etwas (3)
Lernfortschritt	In dieser halben Stunde bin ich nicht so vorwärtsgekommen wie ich wollte.	Stimmt eher nicht (4) Stimmt nicht (5) Stimmt überhaupt nicht (6)
Freude	Mir hat das Lösen der Subtraktionsaufgaben sehr viel Spaß gemacht.	
Interesse	Ich fand das Lösen der Subtraktionsaufgaben sehr interessant.	
Flow	Beim Lernen am Computer verging die Zeit wie im Fluge.	

Für den Fragebogen zum 2. Messzeitpunkt konnte ein Faktor mit einem Eigenwert größer 1 extrahiert werden. Dieser Faktor mit einem Eigenwert von 3.72 klärte 55% der Varianz auf. Die Berechnung des Konsistenzkoeffizienten (Cronbachs α) über alle sieben Items erbrachte mit einem alpha-Wert von .84 einen identischen, d.h. gleichfalls hohen, Reliabilitätskoeffizienten wie der Fragebogen zum 1. Messzeitpunkt. Aus diesem Grund wurde auch für diese Fragebogendaten über alle sieben Items ein Summenscore gebildet und dieser

Gesamtscore als „post-aktionales Kompetenzerleben“ zum Messzeitpunkt 2 in die weiteren Berechnungen mit einbezogen. Fehlte nicht mehr als eine Itembeantwortung, so wurde dieser Missing-Wert durch den Mittelwert der Skala der übrigen Items ersetzt.

5.2.5 Untersuchungsablauf

Zwei Wochen vor der eigentlichen Evaluation wurde in die Schulen der Vortest mit den 24 Aufgaben zur schriftlichen Subtraktion und der Vorbefragungsbogen zur Erfassung der Kontrollvariablen gereicht. Der Vorbefragungsbogen beinhaltete die Items zum aufgabenspezifischen Fähigkeitskonzept, zu den intrinsischen Werten schriftlicher Subtraktionsaufgaben und zur Computererfahrung. Die Schüler hatten beide Paper-und-Pencil Tests während einer Unterrichtsstunde zu bearbeiten. Auf der Basis des Vortestes wurden dann die Schüler an die TU Dresden eingeladen, die Probleme beim Bearbeiten schriftlicher Subtraktionsaufgaben zeigten. Die Evaluation wurde in dem Computerpool der Fachrichtung Psychologie durchgeführt. Die Schüler nahmen in Gruppen von 6-12 Schülern im Abstand von einer Woche an zwei Sitzungen teil, die jeweils ca. 1 ½ Stunden dauerten. In der ersten Sitzung wurden die Schüler zufällig einer der vier Feedbackbedingungen zugewiesen. Nach einer Einweisung in die Benutzerführung des Programms durch den Versuchsleiter arbeiteten die Schüler 30 Minuten mit der Software. Hatten die Schüler einen Aufgabentyp erfolgreich bearbeitet, wurde ein neuer Aufgabentyp durch den Versuchsleiter ausgewählt. Alle interessierenden Lerneraktionen wurden wie in Studie 1 in Log Files protokolliert. Im Anschluss an die computergestützte Treatmentphase hatten die Schüler den Fragebogen mit den sieben Items zur Erfassung der post-aktionalen motivationalen Parameter zu beantworten. Abschließend wurde ihnen der Nachtest dargeboten, welcher die gleichen Aufgaben des Vortests enthielt. Zur zweiten Sitzung wurde den Schülern die Feedbackbedingung zugewiesen, die sie eine Woche zuvor erhalten hatten. Ohne eine nochmalige Einweisung in die Programmbedienung wurden von dem/den Versuchsleiter(n) die als problematisch im Vortest identifizierten Aufgabentypen dem Schüler zur Bearbeitung ausgewählt. Nach einer halbstündigen Programmbearbeitung bekamen die Schüler den Fragebogen zur Erfassung der post-aktionalen motivationalen Parameter und den Nachtest dargeboten. Beide entsprachen dem zur 1. Sitzung eingesetzten Fragebogen und Nachtest. Der Untersuchungsablauf ist schematisch der Abbildung 25 zu entnehmen.

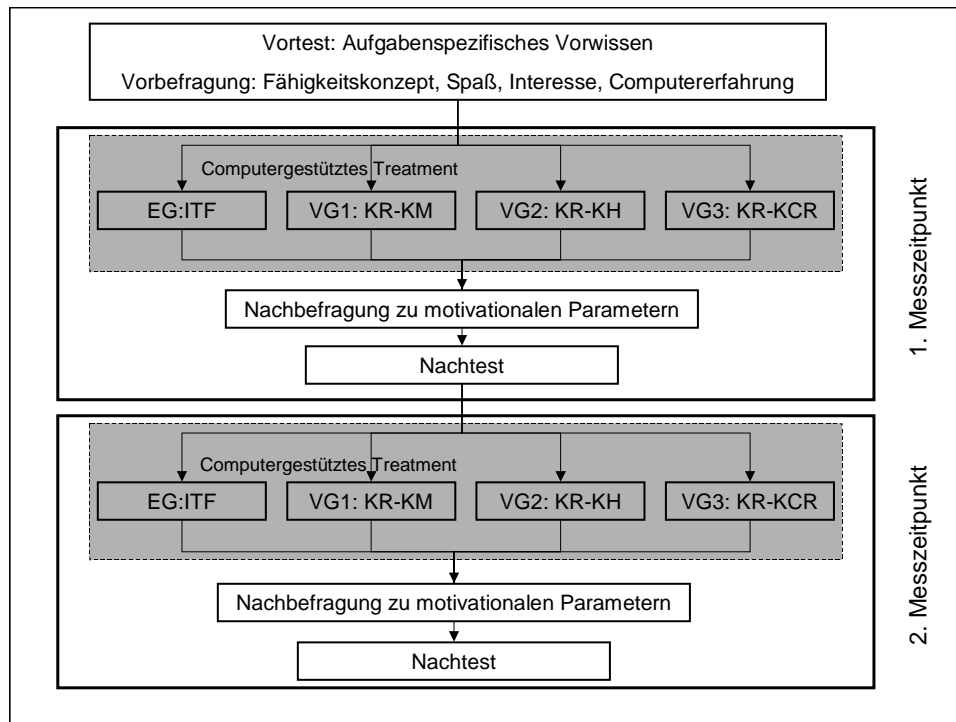


Abbildung 25: Beschreibung des Untersuchungsablaufs von Studie 2

5.2.6 Analysevorbereitung und Methoden der Datenanalyse

In diesem Kapitel werden die vier Feedbackgruppen zuerst auf ihre Vergleichbarkeit untersucht, bevor die bei der Datenanalyse und Datenauswertung eingesetzten Methoden vorgestellt werden.

5.2.6.1 Vergleichbarkeit der Feedbackgruppen

Analog zur Studie 1 wurde in einem ersten Schritt eine explorative Datenanalyse über die Treatmentleistung des 1. Messzeitpunktes durchgeführt, um die Lerner zu identifizieren, die zum 1. Messzeitpunkt eine hohe Fertigkeit beim Bearbeiten schriftlicher Subtraktionsaufgaben aufwiesen und von daher keine instruktionale Unterstützung bei der Bearbeitung der Aufgaben benötigten. Auf der Basis der explorativen Datenanalyse konnten drei Ausreißer der KR-KCR Feedbackgruppe identifiziert werden, die acht oder mehr Aufgabentypen während der 30minütigen Treatmentphase erfolgreich bearbeitet hatten. Für die anderen drei, informativeren Feedbackgruppen konnten keine Ausreißer identifiziert werden. Auch dieses Ergebnis kann möglicherweise wie in Studie 1 auf eine Konfundierung der Wirksamkeit der informativeren Feedbackbedingungen mit dem Vorliegen hoher Leistungsfähigkeit zum Treatmentzeitpunkt zurückgeführt werden. Um die Vergleichbarkeit der Feedbackgruppen zu

wahren, wurden wie in Studie 1 die Schülerdatensätze der anderen drei Feedbackgruppen von der weiteren Auswertung ausgeschlossen, die gleichfalls acht oder mehr Aufgabentypen während der computergestützten Treatmentphase bearbeitet hatten. Dies betraf einen Datensatz des ITF und jeweils zwei Datensätze der beiden Feedbackgruppen KR-KM und KR-KH, so dass in die eigentliche Datenanalyse zum 1. Messzeitpunkt 105 Schülerdatensätze eingingen, ITF: $n = 27$, KR-KCR: $n = 23$, KR-KM: $n = 31$ und KR-KH: $n = 24$.

Zum 2. Messzeitpunkt nahmen 92 der 105 Schüler an der Untersuchung teil. Die Ausfallquote betrug somit 12 %. Diese 92 Schüler waren wie folgt auf die einzelnen Feedbackgruppen verteilt: ITF: $n = 25$, KR-KCR: $n = 21$, KR-KM: $n = 23$ und KR-KH: $n = 23$. Starke Versuchspersonenverluste verzeichnete vor allem die KR-KM Feedbackgruppe. Da diese aber zum 1. Messzeitpunkt eine gewisse Überrepräsentation an Versuchspersonen aufwies und die Mittelwerte der 31 ($M = 12.26$) bzw. 23 Versuchspersonen ($M = 11.57$) in den Nachttestleistungen zum 1. Messzeitpunkt auf keinen signifikanten Unterschied hindeuten, scheint eine aus der Ausfallquote resultierende Verzerrung der Stichprobe nicht gegeben zu sein.

In Tabelle 33 sind für jede Feedbackgruppe Mittelwerte und Standardabweichungen bzw. Häufigkeitsverteilungen für jede Kontrollvariable dargestellt. Mit dem Ziel, die Vergleichbarkeit der vier Feedbackgruppen zu überprüfen, wurden über diese Kontrollvariablen verschiedene Prüfstatistiken berechnet. Eine einfaktorielle, multivariate Varianzanalyse über die Kontrollvariablen Fähigkeitskonzept, Interesse und Freude zeigte keinen statistisch signifikanten Unterschied zwischen den Feedbackgruppen, $F(9, 294) = 1.15$; $p = .33$. Mit dem Box- M -Test konnte die Gleichheit der Kovarianzmatrizen der Feedbackgruppen nachgewiesen werden, $F(18, 31381) = 1.18$, $p = .27$. Nachfolgend durchgeführte univariate Varianzanalysen zeigten für keine dieser Kontrollvariablen einen signifikanten Unterschied zwischen den Gruppen. Auch für nachfolgend aufgezählte Kontrollvariablen konnten keine signifikanten Unterschiede zwischen den Feedbackgruppen nachgewiesen werden: (a) aufgabenspezifisches Vorwissen, $F(3, 100) = .73$, $p = .53$, (b) Alter, $F(3, 101) = .47$, $p = .70$, (c) Geschlecht, $\chi^2(3, N = 105) = 3.81$, $p = .28$, (d) Computererfahrung, $H = 6.67$, $p = .08$.

Die Ergebnisse dieser Analysen zeigen, dass von einer Vergleichbarkeit der Gruppen hinsichtlich der erfassten Kontrollvariablen ausgegangen werden kann.

Tabelle 33:

Darstellung der Mittelwerte und Standardabweichungen bzw. Häufigkeitsverteilungen für jede Kontrollvariable getrennt nach Feedbackgruppe

Kontrollvariablen	Feedbackbedingung				
	ITF (n = 27)	KR-KM (n = 31)	KR-KH (n = 24)	KR-KCR (n = 23)	
Alter					
	<i>M</i>	10.48	10.39	10.13	10.52
	(<i>SD</i>)	(1.31)	(1.23)	(1.19)	(1.38)
Vorwissen					
	<i>M</i>	9.62	10.39	10.29	9.04
	(<i>SD</i>)	(4.57)	(3.62)	(2.82)	(3.93)
Fähigkeitskonzept					
	<i>M</i>	12.52	11.35	11.13	9.96
	(<i>SD</i>)	(3.32)	(3.00)	(2.92)	(3.78)
Freude					
	<i>M</i>	4.44	4.06	3.79	3.78
	(<i>SD</i>)	(1.47)	(1.46)	(1.32)	(1.48)
Interesse					
	<i>M</i>	4.40	3.97	4.04	3.70
	(<i>SD</i>)	(1.50)	(1.43)	(1.46)	(1.66)
Computererfahrung					
	Täglich	33.3%	9.7%	16.7%	30.4%
	1-2mal/Woche	25.9%	16.1%	33.3%	17.4%
	1-2mal/Monat	7.4%	0.0%	0.0%	0.0%
	seltener	14.8%	41.9%	33.3%	26.1%
	nie	7.4%	12.9%	12.5%	26.1%
Geschlecht					
	weiblich	12 (44.4 %)	17 (54.8%)	7 (29.2 %)	9 (39.1 %)

5.2.6.2 Methoden der Datenanalyse

Das zentrale Anliegen dieser Studie lag in der hypothesengeleiteten Untersuchung von Wirksamkeitsunterschieden zwischen den vier Feedbackgruppen zu zwei Messzeitpunkten. Zur statistischen Prüfung der Wirksamkeitsunterschiedshypothesen waren aufgrund des Messwiederholungsdesigns zum einen Varianzanalysen mit Messwiederholung indiziert. Da im Rahmen korrelationsanalytischer Auswertungen vereinzelt signifikante Korrelationen zwischen den über den Vortest und die Vorbefragung erfassten Kontrollvariablen und den abhängigen Variablen zu Messzeitpunkt 1 und zu Messzeitpunkt 2 (siehe Tabelle 34) gefunden werden konnten, wurden zum anderen kovarianzanalytische Verfahren genutzt.

Tabelle 34:

Korrelationen der Kontrollvariablen und abhängigen Variablen von Messzeitpunkt 1 und Messzeitpunkt 2 (85 < N < 105)

	V	FK	I	F	C	T1	N1	KW1	K1	T2	N2	KW2	K2
Vortest (V)	-	.18	-.21 *	-.02	-.04								
Fähigkeitskonzept (FK)		-	.53**	.22*	.09								
Interesse (I)			-	.48**	.08								
Freude (F)				-	.10								
Computererfahrung (C)					-								
Treatmentleistung1 (T1)	.24*	.07	-.24*	-.13	.23	-							
Nachtestleistung1 (N1)	.41*	.19	-.29**	-.08	.15	.68**	-						
Korrekturwahrscheinlichkeit1 (KW1)	.19	.17	-.24*	-.06	.21*	.68**	.89**	-					
Kompetenzerleben1 (K1)	.07	.01	-.10	.02	.32**	.51**	.45**	.48**	-				
Treatmentleistung2 (T2)	.20	.09	-.23**	-.11	.13	.63**	.62**	.27*	.27*	-			
Nachtestleistung2 (N2)	.44**	.24*	-.11	.09	.09	.50**	.65**	.19	.19	.73**	-		
Korrekturwahrscheinlichkeit2 (KW2)	.25*	.19	-.05	.09	.15	.44**	.46**	.47**	.12	.62**	.80**	-	
Kompetenzerleben2 (K2)	.26*	.11	.09	.26*	-.07	.18	.19	.20	.22*	.40**	.42**	.37**	-

Anmerkungen:

Hellgelb: Korrelationen der Kontrollvariablen; hellgrau: Korrelationen der abhängigen Variablen von Messzeitpunkt 1 und 2 mit den Kontrollvariablen, gelb: Korrelationen der abhängigen Variablen von Messzeitpunkt 1 und 2

* p < .05. ** p < .01.

Für die Bestimmung der zulässigen Kovariatenanzahl wurde Formel 1 (siehe S. 105) verwendet. Da Varianzanalysen mit Messwiederholung und kovarianzanalytische Verfahren aber nur einen globalen Vergleich ermöglichen und von daher nur Aussagen zulassen, ob zwischen den Gruppen ein signifikanter Mittelwertsunterschied vorliegt, aber nicht, zwischen welchen Gruppen dieser Mittelwertsunterschied beobachtet werden konnte, waren in einem weiteren Analyseschritt Einzelvergleiche erforderlich, um die aufgestellten Hypothesen beantworten zu können. Obwohl Hager (2000, S. 208) zu Recht die prinzipielle Brauchbarkeit globaler Prüfstatistiken in Frage stellt, werden Kennwerte von Globalvergleichen berichtet, da sich die Darstellung solcher globaler Prüfstatistiken in der Literatur und Forschung etabliert hat. Als ein Argument gegen globale Prüfstatistiken führt Hager an, dass nicht signifikante Ergebnisse solcher Globalanalysen in Detailanalysen signifikant werden können. Da die meisten Forscher zudem primär an Detailanalysen interessiert seien, plädiert Hager (2000) dafür, mit diesen zu beginnen.

Für die Detailanalysen wurde auf a-priori Einzelvergleiche entsprechend der aufgestellten gerichteten Hypothesen zurückgegriffen (Bortz, 1993; Hager, Patry & Brezing, 2000). Da die Anzahl aufgestellter Einzelvergleichshypothesen größer zwei war, wurden die a priori formulierten Hypothesen mit Korrektur des α -Fehler-Niveaus geprüft. Obwohl sowohl globale Prüfstatistiken als auch Statistiken der Einzelvergleiche berichtet werden, wird für die hypothesengeleitete Interpretation primär auf die Ergebnisse der Detailanalysen rekurriert.

Neben Vergleichen zwischen den einzelnen Feedbackbedingungen wurden im Vorfeld Vergleiche innerhalb der Feedbackbedingungen berechnet, um die vorhergesagte isolierte Wirksamkeit der einzelnen Feedbackbedingungen zu überprüfen. So können auch gegenläufige Veränderungen in den einzelnen Feedbackgruppen zu signifikanten Unterschieden in den abhängigen Variablen führen, die Programmterfolge vortäuschen, obwohl keine Verbesserungen aufgetreten sind. Intragruppenveränderungen wurden mit dem t -Test für abhängige Stichproben berechnet. Die deskriptiven Statistiken wurden für jede abhängige Variable aus Gründen der Nachvollziehbarkeit der Ergebnisse der Prüfstatistiken mit dargestellt.

Die Effektstärkemaße wurden sowohl aus den Prüfgrößen als auch aus den deskriptiven Maßen der statistischen Signifikanztests abgeleitet. Bei abhängigen Stichproben wurde bei vorliegendem t -Wert, der Korrelation r und der Anzahl n der Versuchspersonen die Effektstärke d über Formel 4 (Dunlap, Cortina, Vaslow & Burke, 1996, S.171) bestimmt:

$$d = t \sqrt{\frac{2(1-r)}{n}} \quad (\text{Formel 4})$$

Bei Verwendung der auf quadrierten Werten beruhenden parametrischen Teststatistik wurde wie in Studie 1 als Effektstärkemaß auf das univariate ω^2 (Hays, 1973) zurückgegriffen (siehe Formel 2, S.105).

Die Effektstärkenberechnung zwischen zwei Feedbackgruppen *A* und *B* zu einem Messzeitpunkt erfolgte unter Verwendung der adjustierten Mittelwerte \bar{X}^* nach Formel 5:

$$d = \frac{\bar{X}_A^* - \bar{X}_B^*}{\sqrt{\frac{s_A^2 + s_B^2}{2}}} \quad (\text{Formel 5})$$

Die Effektstärkenberechnung des Interaktionseffektes zwischen zwei Gruppen zu zwei Messzeitpunkten *i* und *j* wurde in Anlehnung an Dresel (2002) nach Formel 6 bestimmt:

$$d = \frac{(\bar{X}_{Ai}^* - \bar{X}_{Aj}^*) - (\bar{X}_{Bi}^* - \bar{X}_{Bj}^*)}{\sqrt{\frac{s_{Ai}^2 + s_{Aj}^2 + s_{Bi}^2 + s_{Bj}^2}{4}}} \quad (\text{Formel 6})$$

5.3 Ergebnisdarstellung

Die Ergebnisdarstellung beginnt mit der Analyse der Nachtestleistung, gefolgt von der Analyse des Korrekturwahrscheinlichkeitsindex und der Treatmentleistung, operationalisiert über das mastery level, und schließt mit der Analyse der in der Nachbefragung erfassten motivationalen Parameter ab. Es wird berichtet, ob Unterschiede in den abhängigen Variablen innerhalb der Feedbackgruppe und zwischen den vier Feedbackgruppen gefunden werden konnten.

5.3.1 Nachtestleistung

Die Nachtestleistung war operationalisiert über die Anzahl erfolgreich bearbeiteter Aufgaben im Nachtest. Für die Experimentalgruppe ITF und die drei Vergleichsgruppen KR-KM,

KR-KH und KR-KCR wird angenommen, dass über die drei Messzeitpunkte für alle vier Gruppen eine Verbesserung der Fertigkeiten der schriftlichen Subtraktion beobachtet werden kann. Diese Annahme beruht auf der in der Literatur nachgewiesenen Lernwirksamkeit jeglicher Form von Aufgabenbearbeitung und jeglicher Form von Feedback auf die Aufgabenbearbeitung. Die Verbesserungen sollten aber in Abhängigkeit der Komplexität der Feedbackbedingung stark variieren. Ob eine isolierte Wirksamkeit der Feedbackbedingungen in erwarteter Richtung zu beobachten ist, wurde daher in einem ersten Schritt geprüft. In einem zweiten Auswertungsschritt wurde untersucht, ob sich signifikante Unterschiede in der Nachtestleistung zwischen den Feedbackgruppen entsprechend den aufgestellten Hypothesen finden lassen.

5.3.1.1 Intragruppenvergleiche – Isolierte Wirksamkeit der Feedbackbedingungen

Um zu überprüfen, ob die Ergebnisse in erwarteter Richtung vorliegen, d.h. ob eine isolierte Wirksamkeit jeder Feedbackbedingung zu jedem der beiden Messzeitpunkte nachgewiesen werden kann, wurden in einem ersten Auswertungsschritt unter Nutzung des *t*-Testes für abhängige Stichproben die Intragruppenveränderungen analysiert. In Tabelle 35 sind neben den Mittelwerten und Standardabweichungen der korrekt gelösten Aufgaben im Vortest, Nachtest1 und Nachtest2 die *t*-Werte und Effektstärken der Intragruppenvergleiche für jede Feedbackgruppe aufgeführt.

Zum ersten Messzeitpunkt, d.h. nach der ersten Treatmentphase, zeigten alle Feedbackgruppen außer der Feedbackgruppe KR-KM hochsignifikante Verbesserungen in der Nachtestleistung im Vergleich zum Vortest. Dabei zeigte die ITF Gruppe den stärksten Leistungszuwachs. Zum zweiten Messzeitpunkt, d.h. nach der zweiten Treatmentphase kann nur noch ein hochsignifikanter Leistungszuwachs für die KR-KM Feedbackgruppe berichtet werden, welche zum ersten Messzeitpunkt keine signifikante Verbesserung aufwies. Die drei anderen Feedbackgruppen zeigten zwar auch noch einmal eine Leistungssteigerung, die aber unter keiner der drei Feedbackbedingungen statistisch signifikant wurde. Diese Intragruppenvergleiche bestätigen die vorhergesagte isolierte Wirksamkeit der einzelnen Feedbackbedingungen (H1). Unter jeder Feedbackbedingung bearbeiteten die Schüler im Anschluss an die Treatmentphasen Aufgaben zur schriftlichen Subtraktion erfolgreicher als vor der Treatmentphase. Dieses Ergebnis kann dahingehend interpretiert werden, dass die Bearbeitung von Aufgaben, an deren Anschluss die Schüler mindestens die korrekte Lösung mitgeteilt be-

kommen, sich lernförderlich auf die Leistungen der Schüler auswirkt. Die in nachfolgenden Auswertungen berichteten signifikanten Unterschiede zwischen den Feedbackgruppen können daher nicht auf gegenläufige Veränderungen in den einzelnen Feedbackgruppen zurückgeführt werden.

Tabelle 35:

Deskriptive Statistiken, Prüfstatistiken und Effektstärken für Intragruppenvergleiche hinsichtlich des Einflusses der Treatmentbedingungen auf die Nachttestleistung

Feedbackgruppen	<i>M</i> (<i>SD</i>)		<i>t</i> Vor-Nach1 (<i>df</i> _{Fehler})	<i>d</i> Vor-Nach1	<i>M</i> (<i>SD</i>)		<i>t</i> Nach1-Nach2 (<i>df</i> _{Fehler})	<i>d</i> Nach1-Nach2
	Vor	Nach1 ^a			Nach1 ^a	Nach2		
ITF	9.62 (4.57)	15.46 (6.83)	-6.14*** (25)	.93	15.84 (6.66)	16.32 (7.30)	-.42 (24)	.07
KR-KM	10.39 (3.26)	12.26 (4.92)	-1.99 ⁺ (30)	.44	11.57 (5.06)	15.87 (4.55)	-3.91** (22)	.89
KR-KH	10.29 (2.82)	14.67 (6.66)	-3.05** (23)	.85	14.61 (6.81)	15.48 (6.35)	-1.01 (22)	.13
KR-KCR	9.04 (3.93)	11.87 (5.07)	-3.11** (22)	.61	11.81 (5.25)	13.05 (5.58)	-1.21 (20)	.23

Anmerkungen:

^a Für den 1. Nachttest wurden für jede Feedbackgruppe zwei Mittelwerte und Standardabweichungen abgetragen, da zum 2. Messzeitpunkt nur 92 Datensätze vorliegen (105 zum 1. Messzeitpunkt).

⁺p<.10. *p<.05. **p<.01. ***p<.001.

5.3.1.2 Intergruppenvergleiche – Vergleichende Wirksamkeit der Feedbackbedingungen

Um die vergleichende Wirksamkeit der Feedbackbedingungen zu untersuchen, wurde in einem ersten Auswertungsschritt zu jedem Messzeitpunkt eine ANCOVA über die Nachttestleistung berechnet. Als Kovariate wurden die Kontrollvariablen aufgenommen, die mit der jeweiligen abhängigen Variable korrelierten. In der ANCOVA zum ersten Messzeitpunkt wurden als Kovariate die Vortestleistung und das Interesse berücksichtigt. In die ANCOVA zum zweiten Messzeitpunkt wurden die Vortestleistung und das Fähigkeitskonzept als Kovariate aufgenommen. In einem zweiten Analyseschritt wurde eine Varianzanalyse mit Messwiederholung mit folgenden abhängigen Variablen durchgeführt: (1) der Nachttestleistung zu Messzeitpunkt 1 und (2) der Nachttestleistung zu Messzeitpunkt 2. Als Kovariate wurden die Vor-

testleistung, das Interesse und das aufgabenspezifische Fähigkeitskonzept berücksichtigt, da diese mindestens mit einer der abhängigen Variablen korrelierten. Im Anschluss an diese globalen Prüfstatistiken wurden α -adjustierte Einzelvergleiche für die gefundenen Haupt- und Interaktionseffekte mit den unter Auspartialisierung der Kovariaten geschätzten Mittelwerte der einzelnen Feedbackgruppen berechnet. Da die Berechnung von Einzelvergleichen von Interaktionseffekten in SPSS nicht möglich ist, wurden Differenzen der Nachttestleistung zu Messzeitpunkt 1 und der Vortestleistung sowie der Nachttestleistung zu Messzeitpunkt 2 und der Nachttestleistung zu Messzeitpunkt 1 gebildet und für diese Differenzen jeweils eine ANCOVA mit anschließenden Einzelvergleichen durchgeführt. Die Ergebnisse der globalen Prüfstatistiken als auch der Einzelvergleiche sind in Tabelle 36 dargestellt.

Die Ergebnisse der Kovarianzanalysen zeigen für den Faktor Feedback einen Haupteffekt zum ersten Messzeitpunkt. Zum zweiten Messzeitpunkt lässt sich kein statistisch signifikanter Unterschied zwischen den Feedbackgruppen nachweisen. Dieses Ergebnis wird auch durch die nicht signifikanten Einzelvergleiche zum zweiten Messzeitpunkt unterstützt. Der signifikante Unterschied zwischen den Feedbackbedingungen zum Messzeitpunkt 1 lässt sich zum einen auf eine höhere Lernwirksamkeit der ITF Bedingung gegenüber der KR-KCR Feedbackbedingung zurückführen. Schüler unter der ITF Bedingung bearbeiteten hypothesenkonform im Anschluss an die erste Treatmentphase statistisch signifikant mehr Aufgaben im Nachttest korrekt als Schüler der KR-KCR Feedbackbedingung. Ein zweiter signifikanter Unterschied zum 1. Messzeitpunkt konnte zwischen der ITF Bedingung und der KR-KM Feedbackbedingung gefunden werden. Auch hier bearbeiteten die Schüler unter der ITF Bedingung im Vergleich zur KR-KM Feedbackbedingung statistisch signifikant mehr Aufgaben im Nachttest korrekt. So zeigten die Schüler unter der KR-KM Feedbackbedingung im Anschluss an die 1. Treatmentphase eine den Schülern der KR-KCR Feedbackbedingung vergleichbare Nachttestleistung. Wie aus Abbildung 26 ersichtlich, verbesserten sich die Schüler der KR-KM Feedbackbedingung allerdings zum zweiten Messzeitpunkt. So zeigten sie im Anschluss an die zweite Treatmentphase einen stärkeren Lernfortschritt als die Schüler der KR-KCR Feedbackbedingung, welcher aber nicht statistisch signifikant wurde. Allerdings verdeutlichen die Einzelvergleiche der Interaktionseffekte, dass Schüler mit ITF vom Vortest zum Nachttest1 einen statistisch signifikant größeren Lernzuwachs im Vergleich zu Schülern mit KR-KM Feedback erzielten. Dieser signifikante Interaktionseffekt trat mit umgekehrten Vorzeichen auf, wurden die Leistungen der ITF und KR-KM Feedbackgruppe von Nachttest1 zu Nachttest2 verglichen. Dies bedeutet, dass die Schüler mit KR-KM Feedback zum 2. Treatmentzeitpunkt dieses Lerndefizit aufholten. So zeigten sie einen statistisch signifikant höheren Lernzuwachs von Nachttest1 zu Nachttest2 im Vergleich zur ITF Gruppe.

Tabelle 36:

Statistische Kennwerte und Effektstärken der globalen Prüfstatistiken sowie Einzelvergleiche und Effektstärken für die Haupt- und Interaktionseffekte der Nachtestleistungen zu beiden Messzeitpunkten

	Globale Prüfstatistiken								
	ANCOVA zu MP1			ANCOVA zu MP2			Varianzanalyse mit Messwiederholung		
	$df(H, F)^a$	F	ω^2	$df(H, F)$	F	ω^2	$df(H, F)$	F	ω^2
Feedback (F)	3, 96	4.24**	.09	3, 84	.81	-.01	3, 82	2.23 ⁺	.04
Vortest (V)	1, 96	17.75***	.15	1, 84	18.65***	.17	1, 82	18.36***	.17
Interesse (I)	1, 96	8.85**	.07				1, 82	6.30*	.06
FK ^b				1, 84	1.48	.01	1, 82	2.05	.01
F x Z ^c							3, 82	3.02*	.07
V x Z							1, 82	.32	-.01
I x Z							1, 82	4.36*	.04
FK x Z							1, 82	.05	-.01
Einzelvergleiche mit Bonferroni-Korrektur									
	MP1		MP2		Feedback x Zeit				
	Mittlere Differenz (SE)	d	Mittlere Differenz (SE)	d	Mittlere Differenz (SE) MP1 ^d	d	Mittlere Differenz (SE) MP2 ^c	d	
ITF vs. KR-KCR	4.26* (1.51)	.71	2.65 (1.71)	.40	3.02 (1.56)	.66	-1.21 (1.51)	-.23	
ITF vs. KR-KM	4.43* (1.41)	.74	1.08 (1.62)	.18	3.98* (1.45)	.79	-4.16* (1.48)	-.74	
ITF vs. KR-KH	1.86 (1.48)	.27	1.40 (1.64)	.20	1.47 (1.54)	.24	-.68 (1.47)	-.13	
KR-KH vs. KR-KM	2.57 (1.41)	.44	-.31 (1.60)	-.06	2.50 (1.48)	.40	-3.49 (1.48)	-.62	
KR-KH vs. KR-KCR	2.41 (1.52)	.41	1.26 (1.66)	.20	1.55 (1.59)	.26	-.53 (1.51)	-.10	
KR-KCR vs. KM-KR	.16 (1.45)	.04	-1.57 (1.67)	-.30	.96 (1.50)	.20	-2.95 (1.53)	-.59	

Anmerkungen.

^aH = Hypothesenfreiheitsgrade, F = Fehlerfreiheitsgrade

^bFK = Fähigkeitskonzept

^cZ = Zeit

^dAbhängige Variablen zur Berechnung der Einzelvergleiche der Interaktionseffekte bildeten Differenzwerte zwischen Nachtest1 und Vortest sowie Nachtest2 und Nachtest1

⁺p<.10. *p<.05. **p<.01. ***p<.001.

In Abbildung 26 sind die unter Auspartialisierung der Kovariate Interesse und Fähigkeitskonzept geschätzten Mittelwerte des Vortests, Nachtests1 und Nachtests2 für die einzelnen Feedbackbedingungen dargestellt. Die Ergebnisse der globalen Prüfstatistiken und der Detailanalysen werden in dieser Abbildung noch einmal graphisch zusammengefasst präsentiert.

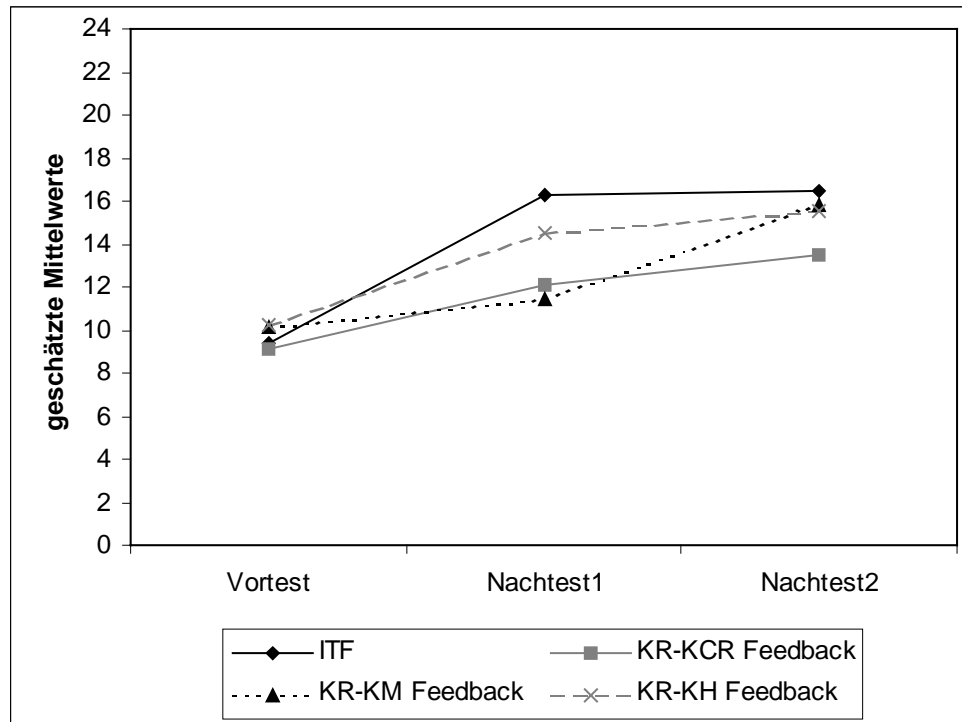


Abbildung 26: Darstellung der unter Auspartialisierung der Kovariaten Interesse und Fähigkeitskonzept geschätzten Mittelwerte des Vortests, Nachtests1 und Nachtests2 für die einzelnen Feedbackbedingungen

Zusammengefasst lässt sich festhalten, dass sich ITF zum 1. Messzeitpunkt hypothesenkonform als lernwirksamer erwiesen hat als die drei anderen Feedbackformen (H2a₁), auch wenn sich dieser Unterschied nur für die beiden Feedbackbedingungen KR-KM und KR-KCR auf dem 5%-Niveau absichern ließ. Zum zweiten Messzeitpunkt lässt sich die höhere Lernwirksamkeit von ITF zwar noch in der Tendenz erkennen, allerdings nicht mehr statistisch nachweisen (H2a₁). Möglicherweise ist dieses Ergebnis auf einen Deckeneffekt zurückzuführen. Schüler, die zum 1. Messzeitpunkt bereits mehr als 20 Aufgaben im Nachtest korrekt bearbeiteten hatten, konnten sich zum zweiten Messzeitpunkt nicht mehr so stark verbessern. Dies betraf 30% der Schüler der ITF Gruppe. Die Annahme, dass sich die beiden Feedbackformen KR-KH und KR-KM, welche sich hinsichtlich ihres Informationsgehaltes zwischen

dem ITF und dem KR-KCR Feedback einordnen lassen und im Gegensatz zum ITF keine unmittelbare Informationsanwendung erfordern, als nicht so lernwirksam wie ITF, aber als lernwirksamer als KR-KCR Feedback erweisen, ließ sich statistisch nicht absichern (H2a₂). Allerdings lässt Abbildung 26 eine Tendenz in die erwartete Richtung erkennen.

5.3.2 Korrekturwahrscheinlichkeitsindex

Da die Berechnung des Korrekturwahrscheinlichkeitsindex auf Basis der Vortest- und Nachtestdaten erfolgt, wurden in einem nächsten Auswertungsschritt die Feedbackgruppen auf Unterschiede im Korrekturwahrscheinlichkeitsindex hin untersucht. Der Korrekturwahrscheinlichkeitsindex beschreibt hierbei die bedingte Wahrscheinlichkeit, dass eine im Vortest fehlerhaft gelöste Aufgabe im Nachtest richtig gelöst wird. Zur Bestimmung des Korrekturwahrscheinlichkeitsindex zu Messzeitpunkt 1 wurde die Anzahl der im Vortest falsch, im Nachtest1 jedoch richtig gelösten Aufgaben (Antwortmuster $f1 \rightarrow r2$) zur Gesamtzahl der im Vortest falsch gelösten Aufgaben (Antwortmuster $f1 \rightarrow r2$ und $f1 \rightarrow f2$) ins Verhältnis gesetzt. Entsprechend wurde der Korrekturwahrscheinlichkeitsindex zu Messzeitpunkt 2 aus dem Verhältnis der im Nachtest1 falsch, im Nachtest2 jedoch richtig gelösten Aufgaben (Antwortmuster $f2 \rightarrow r3$) zur Gesamtzahl der im Nachtest1 falsch gelösten Aufgaben (Antwortmuster $f2 \rightarrow r3$ und $f2 \rightarrow f3$) berechnet.

In einem ersten Auswertungsschritt wurde überprüft, ob eine isolierte Wirksamkeit der Feedbackbedingungen dahingehend zu beobachten ist, dass die Schüler aller Feedbackbedingungen ihre im Vortest (bzw. Nachtest1) fehlerhaft gelösten Aufgaben im Nachtest1 (bzw. Nachtest2) häufiger korrigierten. In einem zweiten Auswertungsschritt wurde untersucht, ob sich signifikante Unterschiede zwischen den Feedbackgruppen dahingehend zeigen, dass Schüler ihre im Vortest (bzw. Nachtest1) fehlerhaft bearbeiteten Aufgaben im Nachtest1 (bzw. Nachtest2) um so häufiger korrigieren, je komplexer das bereitgestellte Feedback war.

5.3.2.1 Intragruppenvergleiche – Isolierte Wirksamkeit der Feedbackbedingungen

In Tabelle 37 sind für beide Messzeitpunkte zum einen die Mittelwerte und Standardabweichungen der sowohl im Vortest als auch im Nachtest1 fehlerhaft bearbeiteten Aufgaben sowie die Mittelwerte und Standardabweichungen der sowohl im Nachtest1 als auch im Nachtest2 fehlerhaft bearbeiteten Aufgaben, welche auch im Vortest fehlerhaft bearbeitet worden waren, für jede Feedbackgruppe dargestellt. Der für jede Feedbackgruppe zu jedem

Messzeitpunkt berechnete t -Test für abhängige Stichproben gibt darüber Auskunft, ob die Schüler statistisch signifikant ihre vor dem Treatment falsch gelösten Aufgaben nach dem Treatment erfolgreich korrigieren konnten. Der in der Tabelle 37 dargestellte Korrekturwahrscheinlichkeitsindex gibt an, wie viel Prozent der fehlerhaft bearbeiteten Aufgaben aus dem Vortest (bzw. aus dem Nachtest1) im Nachtest1 (bzw. im Nachtest2) korrigiert werden konnten.

Tabelle 37:

Deskriptive Statistiken, Prüfstatistiken und Effektstärken für Intragruppenvergleiche hinsichtlich des Einflusses der Treatmentbedingungen auf die Korrektur vor der Treatmentphase fehlerhaft bearbeiteter Aufgaben, Angabe des Korrekturwahrscheinlichkeitsindex

Feed- back- grup- pen	M (SD)		$t_{\text{Vor-Nach1}}$ (df_{Fehler})	$d_{\text{Vor-}}\text{Nach1}$	KWI ^c	M (SD)		$t_{\text{Nach1-Nach2}}$ (df_{Fehler})	$d_{\text{Nach1-}}\text{Nach2}$	KWI ^d
	Vor ^a	Nach1 ^b				Nach1 ^a	Nach2 ^b			
ITF	14.38 (4.57)	7.23 (6.31)	9.24*** (25)	1.20	50%	8.16 (6.66)	5.08 (6.07)	5.59*** (24)	.47	38%
KR- KM	13.61 (3.26)	9.16 (3.24)	7.48*** (31)	1.37	33%	12.43 (5.06)	6.70 (4.23)	6.25*** (22)	1.22	46%
KR-KH	13.71 (2.82)	7.21 (5.00)	6.08*** (23)	1.57	47%	9.39 (6.81)	6.00 (6.09)	5.25*** (22)	.51	36%
KR- KCR	14.96 (3.93)	10.13 (4.84)	6.85*** (22)	1.07	32%	12.24 (5.21)	8.24 (4.66)	6.56*** (20)	.78	33%

Anmerkungen:

^aMittelwert der im Vortest bzw. Nachtest1 fehlerhaft bearbeiteten Aufgaben für jede Feedbackgruppe

^bMittelwert der im Nachtest1 bzw. Nachtest2 fehlerhaft bearbeiteten Aufgaben, die auch im Vortest bzw. im Nachtest1 fehlerhaft bearbeitet worden waren, d.h. die Anzahl der nicht korrigierten Aufgaben

^cKWI: Korrekturwahrscheinlichkeitsindex (Anzahl korrigierter Aufgaben im Nachtest1 durch die Anzahl im Vortest fehlerhaft bearbeiteter Aufgaben) in Prozent

^dKWI: Korrekturwahrscheinlichkeitsindex (Anzahl korrigierter Aufgaben im Nachtest2 durch die Anzahl im Nachtest1 fehlerhaft bearbeiteter Aufgaben) in Prozent

*** $p < .001$.

Für jeden Messzeitpunkt konnte über alle Feedbackbedingungen ein statistisch signifikanter Unterschied nachgewiesen werden (H1). Das heißt, die Schüler jeder Feedbackbedingung bearbeiteten im Anschluss an das Treatment ihre vor dem Treatment fehlerhaft gelösten Aufgaben statistisch signifikant häufiger korrekt. Zu beiden Messzeitpunkten konnten von den Schülern der vier Feedbackbedingungen zwischen 32% und 50% ihrer fehlerhaft bearbeiteten Aufgaben erfolgreich korrigiert werden. Dabei zeigte die ITF Gruppe die höchste Korrekturrate. Auch dieses Ergebnis weist darauf hin, dass die Bearbeitung von Lernaufgaben, an deren Anschluss der Schüler mindestens KCR Feedback erhält, den Erwerb von Fertigkeiten in der schriftlichen Subtraktion unterstützt. Die in nachfolgenden Auswertungen berichteten signifikanten Unterschiede zwischen den Feedbackgruppen können daher nicht auf gegenläufige Veränderungen in den einzelnen Feedbackgruppen zurückgeführt werden.

5.3.2.2 Intergruppenvergleiche - Vergleichende Wirksamkeit der Feedbackbedingungen

Um die vergleichende Wirksamkeit der Feedbackbedingungen zu untersuchen, wurden in einem ersten Auswertungsschritt globale Prüfstatistiken berechnet. In Tabelle 38 sind die Ergebnisse der ANCOVA mit der abhängigen Variablen „Korrekturwahrscheinlichkeitsindex“ zu jedem Messzeitpunkt abgetragen. In der Kovarianzanalyse zu Messzeitpunkt 1 wurden als Kovariate das Interesse und die Computererfahrung berücksichtigt, in der Kovarianzanalyse zu Messzeitpunkt 2 die Kovariate Vortestleistung. In die Varianzanalyse mit Messwiederholung gingen als abhängige Variablen ein (1) der Korrekturwahrscheinlichkeitsindex zu Messzeitpunkt 1 (2) der Korrekturwahrscheinlichkeitsindex zu Messzeitpunkt 2. Als Kovariate wurden die Vortestleistung, das Interesse und die Computererfahrung berücksichtigt, da diese mindestens mit einer der abhängigen Variablen korrelierten. Die Anwendung von Formel 1 zur Bestimmung der zulässigen Kovariatenanzahl erbrachte mit .07 einen Wert kleiner .10, so dass eine Varianzanalyse mit Messwiederholung indiziert war. Die Ergebnisse der globalen Prüfstatistiken als auch der Einzelvergleiche sind in Tabelle 38 dargestellt. Da die Berechnung von Einzelvergleichen von Interaktionseffekten in SPSS nicht möglich ist, wurden Differenzwerte zwischen fehlerhaft bearbeiteten Aufgaben im Vortest und den gleichen fehlerhaft bearbeiteten Aufgaben im Nachtest1 sowie Differenzwerte zwischen fehlerhaft bearbeiteten Aufgaben im Nachtest1 und den gleichen fehlerhaft bearbeiteten Aufgaben im Nachtest2 gebildet und für diese Differenzen jeweils eine ANCOVA mit anschließenden Einzelvergleichen durchgeführt.

Tabelle 38:

Statistische Kennwerte und Effektstärken der globalen Prüfstatistiken und Einzelvergleiche für die Haupt- und Interaktionseffekte des Korrekturwahrscheinlichkeitsindex zu beiden Messzeitpunkten

	Globale Prüfstatistiken								
	ANCOVA zu MP1			ANCOVA zu MP2			Varianzanalyse mit Messwiederholung		
	<i>df</i> (H,F) ^a	<i>F</i>	ω^2	<i>df</i> (H,F)	<i>F</i>	ω^2	<i>df</i> (H,F)	<i>F</i>	ω^2
Feedback (F)	3, 89	4.95**	.11	3, 85	.80	-.01	3, 79	2.63 ⁺	.06
Vortest (V)				1, 85	13.49***	.03	1, 79	9.47**	.09
Interesse (I)	1, 89	9.32*	.08				1, 79	2.29	.02
Computer (C)	1, 89	2.03	.01				1, 79	2.63 ⁺	.02
F x Z ^b							3, 79	2.65 ⁺	.06
V x Z							1, 79	4.14*	.04
I x Z							1, 79	3.08 ⁺	.03
C x Z							1, 79	.01	-.01

	Einzelvergleiche mit Bonferroni-Korrektur							
	MP1		MP2		Feedback x Zeit			
	Mittlere Differenz (SE)	<i>d</i>	Mittlere Differenz (SE)	<i>d</i>	Mittlere Differenz (SE) MP1 ^c	<i>d</i>	Mittlere Differenz (SE) MP2 ^c	<i>d</i>
ITF vs. KR-KCR	.26** (.08)	1.04	.14 (.09)	.43	2.33 (1.15)	.63	-.22 (.86)	-.16
ITF vs. KR-KM	.26** (.08)	.99	.06 (.09)	.18	2.70 ⁺ (1.07)	.74	-.91 (.84)	-.37
ITF vs. KR-KH	.13 (.08)	.33	.02 (.09)	.08	.65 (1.13)	.14	.61 (.84)	.20
KR-KH vs. KR-KM	.13 (.08)	.44	.03 (.09)	.18	2.05 (1.09)	.47	-1.52 (.85)	-.48
KR-KH vs. KR-KCR	.14 (.08)	.46	.11 (.09)	.43	1.67 (1.17)	.38	-.83 (.87)	-.26
KR-KCR vs. KR-KM	.00 (.08)	-.04	-.08 (.09)	-.30	.37 (1.10)	.11	-.69 (.87)	-.25

Anmerkungen.

^aH = Hypothesenfreiheitsgrade, F = Fehlerfreiheitsgrade

^bZ = Zeit

^cAbhängige Variablen zur Berechnung der Einzelvergleiche der Interaktionseffekte bildeten Differenzwerte zwischen fehlerhaft bearbeiteten Aufgaben im Vortest und den gleichen fehlerhaft bearbeiteten Aufgaben im Nachtest1 sowie Differenzwerte zwischen fehlerhaft bearbeiteten Aufgaben im Nachtest1 und den gleichen fehlerhaft bearbeiteten Aufgaben im Nachtest2

⁺p<.10. *p<.05. **p<.01. ***p<.001.

Die Ergebnisse replizieren im wesentlichen die Ergebnisse der Nachttestleistung, die bereits im Kapitel 5.3.1 berichtet wurden. So konnte nur zum 1. Messzeitpunkt ein statistisch signifikanter Unterschied in der Korrektur fehlerhaft bearbeiteter Aufgaben zwischen den Feedbackgruppen nachgewiesen werden. Nachfolgend durchgeführte Einzelvergleiche hinsichtlich des Korrekturwahrscheinlichkeitsindex zu Messzeitpunkt 1 zeigen, dass Schüler der ITF Bedingung ihre im Vortest fehlerhaft bearbeiteten Aufgaben im Nachttest1 häufiger korrekt bearbeiteten als Schüler der KR-KCR Feedbackgruppe. Dieser hypothesenkonforme Effekt ließ sich auf dem 1%-Niveau absichern. Eine vergleichbare, statistisch signifikante Überlegenheit des ITF konnte auch gegenüber der KR-KM Feedbackbedingung nachgewiesen werden (H2a₁).

Wie aus Abbildung 27 ersichtlich, verbesserten die Schüler der KR-KM Feedbackbedingung ihre Korrekturleistungen aber nach der zweiten Treatmentphase. Die zum 2. Messzeitpunkt nicht mehr nachweisbare Überlegenheit von ITF kann wie auch schon bei der Nachttestleistung über einen Deckeneffekt erklärt werden.

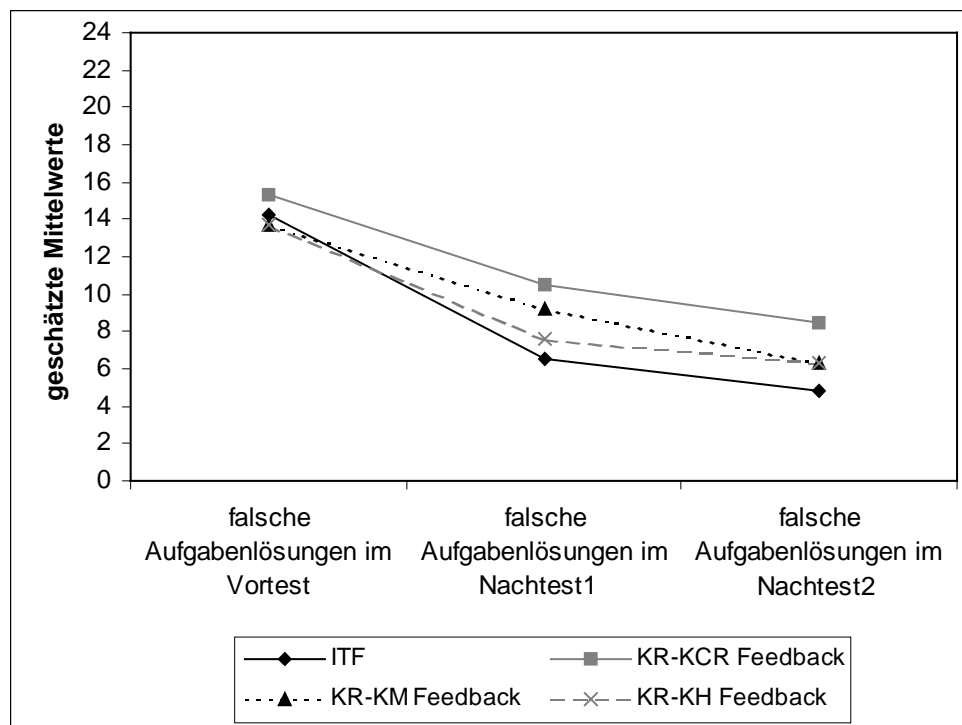


Abbildung 27: Darstellung der unter Auspartialisierung der Kovariaten Interesse und Computererfahrung geschätzten Mittelwerte der fehlerhaft bearbeiteten Aufgaben im Vortest, im Nachttest1 und im Nachttest2 für die einzelnen Feedbackbedingungen

Hinsichtlich der erwarteten mittleren Wirksamkeit der beiden Feedbackformen KR-KH und KR-KM ließ sich diese statistisch zwar nicht absichern, aber auch hier ist aus Abbildung 27 eine erwartungskonforme Tendenz erkennbar (H2a₂).

Diese Ergebnisse, welche im wesentlichen den in Kapitel 5.3.1 berichteten Ergebnissen der Nachtestleistung entsprechen, bestätigen die Annahme, dass die Wirksamkeit von Feedback vor allem über seine korrektive Funktion erklärt werden kann, das heißt, Feedback führt dann zu Lernerfolgen, wenn es Informationen bereitstellt, mit denen Lerner ihre Fehler korrigieren können.

5.3.3 Treatmentleistung

Die Lernwirksamkeit von Feedback lässt sich zeitlich sowohl nach der Treatmentphase, als auch bereits während der Treatmentphase untersuchen. In der Literatur werden Feedbackeffekte in der Regel über die Leistung in einem Nachtest beurteilt (z.B. Collins, Carnine & Gersten, 1987; Kulhavy et al., 1985; Nagata, 1993). Da anzunehmen ist, dass Feedback bereits kurzfristig wirkt und Einfluss auf die Bearbeitung der Aufgaben während des Treatments hat, wurde in einem nächsten Auswertungsschritt die Treatmentleistung analysiert. Diese war über das mastery level operationalisiert, welches die Anzahl erfolgreich bearbeiteter Aufgabentypen während der Treatmentphase beinhaltet. Ein Aufgabentyp galt dann als erfolgreich bearbeitet, wenn drei Aufgaben mit dem gleichen Aufgabenmerkmal in Folge im ersten Versuch richtig bearbeitet worden waren.

Auch hier wurde wieder in einem ersten Auswertungsschritt überprüft, ob eine isolierte Wirksamkeit der Feedbackbedingungen in erwarteter Richtung zu beobachten ist. Da die Treatmentleistung nur zu den beiden Treatmentzeitpunkten erhoben werden konnte, können Aussagen hinsichtlich der isolierten Wirksamkeit der Feedbackbedingungen nur im Vergleich der Leistungen vom 1. zum 2. Messzeitpunkt getroffen werden. In einem zweiten Auswertungsschritt wurde untersucht, ob sich signifikante Unterschiede in der Treatmentleistung zwischen den Feedbackgruppen entsprechend den aufgestellten Hypothesen finden lassen.

5.3.3.1 Intragruppenvergleiche – Isolierte Wirksamkeit der Feedbackbedingungen

In Tabelle 39 sind für beide Messzeitpunkte die Mittelwerte und Standardabweichungen der während der Treatmentphase erfolgreich bearbeiteten Aufgabentypen für jede Feedbackgruppe abgetragen. Der für jede Feedbackgruppe berechnete *t*-Test für abhängige Stichproben

gibt darüber Auskunft, ob die Schüler zum Messzeitpunkt 2 statistisch signifikant mehr Aufgabentypen bearbeiteten als zu Messzeitpunkt 1.

Die Ergebnisse der Intragruppenvergleiche zeigen, dass unter allen Feedbackbedingungen die Schüler während der zweiten Treatmentphase signifikant mehr Aufgabentypen erfolgreich bearbeiteten als während der ersten Treatmentphase (H1). Das heißt, sie erreichten das mastery level zum zweiten Messzeitpunkt häufiger als zum ersten Messzeitpunkt. Dieser Wirksamkeitsunterschied vom ersten zum zweiten Messzeitpunkt ließ sich mindestens auf dem 1%-Niveau absichern. Die in nachfolgenden Auswertungen berichteten signifikanten Unterschiede zwischen den Feedbackgruppen können daher nicht auf gegenläufige Veränderungen in den einzelnen Feedbackgruppen zurückgeführt werden.

Tabelle 39:

Deskriptive Statistiken, Prüfstatistiken und Effektstärken für Intragruppenvergleiche hinsichtlich des Einflusses der Feedbackbedingungen auf die während der Treatmentphase bearbeiteten Aufgabentypen

Feedbackgruppen	<i>M (SD)</i>		<i>t (df_{Fehler})</i>	<i>d</i>
	Treatmentleistung MP1	Treatmentleistung MP2		
ITF	2.83 (2.50)	6.21 (4.11)	-4.83*** (23)	.93
KR-KM	1.52 (1.93)	4.26 (3.77)	-4.44*** (22)	.56
KR-KH	2.59 (2.44)	5.14 (4.37)	-3.40** (21)	.80
KR-KCR	1.33 (1.98)	3.14 (3.75)	-2.96** (20)	.65

** $p < .01$. *** $p < .001$

5.3.3.2 Intergruppenvergleiche - Vergleichende Wirksamkeit der Feedbackbedingungen

Um die vergleichende Wirksamkeit der Feedbackbedingungen zu untersuchen, wurden in einem ersten Auswertungsschritt globale Prüfstatistiken berechnet. In Tabelle 40 sind die Ergebnisse der ANCOVA mit der abhängigen Variablen „Treatmentleistung“ zu jedem Messzeitpunkt abgetragen. In der Kovarianzanalyse zum Messzeitpunkt 1 wurden als Kovariate die Vortestleistung, das Interesse und die Computererfahrung berücksichtigt, in der Kovarianzanalyse zu Messzeitpunkt 2 die Kovariate Interesse. Die Anwendung von Formel 1 zur Berechnung der zulässigen Kovariatenanzahl erbrachte für den Messzeitpunkt 1 mit 3 Kovaria-

ten mit .06 einen Wert kleiner .10, so dass eine Kovarianzanalyse berechnet werden konnte. In die Varianzanalyse mit Messwiederholung gingen als abhängige Variablen (1) die Treatmentleistung zu Messzeitpunkt 1 und (2) die Treatmentleistung zu Messzeitpunkt 2 ein, operationalisiert über die Anzahl erfolgreich bearbeiteter Aufgabentypen. Als Kovariate wurden die Vortestleistung, das Interesse und die Computererfahrung berücksichtigt.

Tabelle 40:

Statistische Kennwerte und Effektstärken der globalen Prüfstatistiken und Einzelvergleiche für die Haupt- und Interaktionseffekte der Treatmentleistung zu beiden Messzeitpunkten

	Globale Prüfstatistiken								
	ANCOVA zu MP1			ANCOVA zu MP2			Varianzanalyse mit Messwiederholung		
	<i>df</i> (H,F) ^a	<i>F</i>	ω^2	<i>df</i> (H,F)	<i>F</i>	ω^2	<i>df</i> (H,F)	<i>F</i>	ω^2
Feedback (F)	3, 88	2.81*	.06	3, 82	3.13*	.07	3, 78	3.13*	.07
Vortest (V)	1, 88	4.51*	.04				1, 78	2.71	.02
Interesse (I)	1, 88	7.73**	.07	1, 82	6.24*	.06	1, 78	7.98**	.08
Computer (C)	1, 88	4.00*	.03				1, 78	2.21	.01
F x Z ^b							3, 78	.99	.00
V x Z							1, 78	.72	.00
I x Z							1, 78	.74	.00
C x Z							1, 78	.23	-.01
Einzelvergleiche mit Bonferroni-Korrektur									
	MP1		MP2		Feedback x Zeit				
	Mittlere Differenz (SE)	<i>d</i>	Mittlere Differenz (SE)	<i>d</i>	Mittlere Differenz ^c (SE)	<i>d</i>			
ITF vs. KR-KCR	1.68* (.61)	1.18	3.43* (1.20)	.86	1.66 (.99)	.52			
ITF vs. KR-KM	1.30 (.61)	.49	2.50 (1.18)	.62	.77 (.98)	.12			
ITF vs. KR-KH	.83 (.60)	.34	1.59 (1.18)	.37	1.03 (.98)	.29			
KR-KH vs. KR-KM	.48 (.59)	.23	.91 (1.18)	.22	-.26 (.98)	-.08			
KR-KH vs. KR-KCR	.85 (.61)	.41	1.84 (1.20)	.59	.63 (.99)	.20			
KR-KM vs. KR-KCR	.37 (.60)	.19	.92 (1.18)	.25	.89 (1.00)	.31			

Anmerkungen.

^aH = Hypothesenfreiheitsgrade, F = Fehlerfreiheitsgrade

^bZ = Zeit

^cAbhängige Variable zur Berechnung der Einzelvergleiche der Interaktionseffekte bildete die Differenz der Treatmentleistung von Messzeitpunkt 2 zu Messzeitpunkt 1

* $p < .05$. ** $p < .01$.

Auch hier erbrachte die Anwendung von Formel 1 mit .07 einen Wert kleiner .10, so dass eine Varianzanalyse mit Messwiederholung indiziert war. Da die Berechnung von Einzelvergleichen von Interaktionseffekten in SPSS nicht möglich ist, wurde eine ANCOVA mit Einzelvergleichen über die Differenz der Treatmentleistung von Messzeitpunkt 2 zu Messzeitpunkt 1 berechnet. Die Ergebnisse der globalen Prüfstatistiken als auch der Einzelvergleiche sind in Tabelle 40 dargestellt.

Für die Treatmentleistung ließ sich zu beiden Messzeitpunkten ein statistisch signifikanter Unterschied zwischen den Feedbackbedingungen auf dem 5%-Niveau absichern, welcher hypothesenkonform für beide Treatments darauf zurückzuführen ist, dass Schüler, welche ITF erhielten, das mastery level signifikant häufiger erreichten als Schüler mit KR-KCR Feedback (H2a₁). Dieser Befund wird auch durch den signifikanten Haupteffekt der Varianzanalyse mit Messwiederholung bestätigt. Auch wenn sich für die anderen Einzelvergleiche keine signifikanten Effekte nachweisen lassen, deutet Abbildung 28 doch darauf hin, dass die Feedbackkomplexität die Bearbeitung von Aufgaben während der Treatmentphase im erwartungskonformen Sinne moderiert.

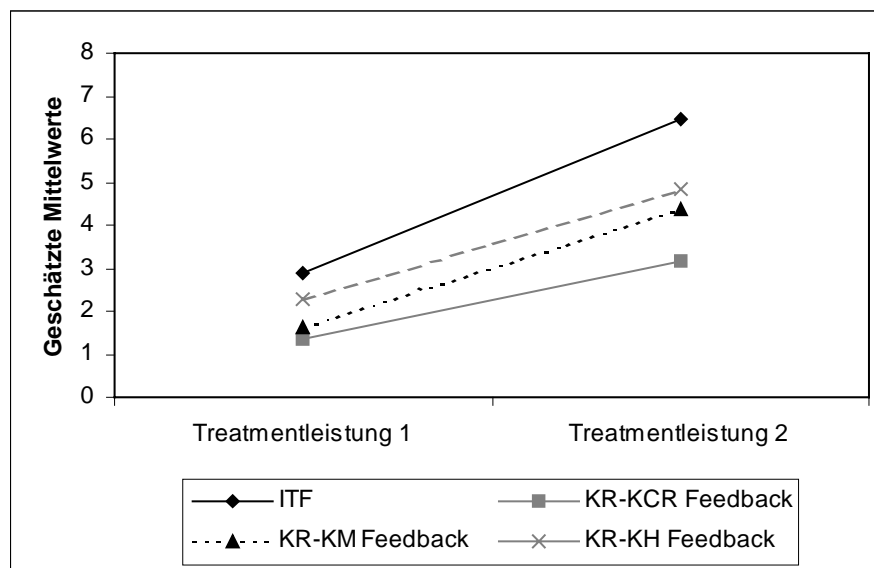


Abbildung 28: Darstellung der unter Auspartialisierung der Kovariaten Vortest, Interesse, und Computererfahrung geschätzten Mittelwerte der Treatmentleistung zu Messzeitpunkt 1 und 2 für die einzelnen Feedbackbedingungen

Je informativer das Feedback und je stärker eine aktive Auseinandersetzung mit den präsentierten Informationen angeregt wird, desto erfolgreicher waren die Schüler während beider Treatmentphasen. So lassen sich die beiden Feedbackformen KR-KM und KR-KH, von denen angenommen wird, dass sie bezüglich ihres Informationsgehaltes zwischen dem

ITF und dem KR-KCR Feedback liegen und welche zudem keine aktive Auseinandersetzung über eine nochmalige Aufgabenbearbeitung anregen, auch in ihrer Treatmentleistung zwischen dem hoch informativen und dem gering informativen Feedback wiederfinden, auch wenn sich dieser Befund statistisch nicht absichern ließ (H2a₂).

5.3.4 Post-aktionales Kompetenzerleben

Neben der Untersuchung von Wirksamkeitsunterschieden bezüglich leistungsbezogener Merkmale bestand ein weiteres Ziel dieser Studie darin, den Einfluss verschieden komplexer Feedbackformen auf post-aktionale motivationale Merkmale, im speziellen den Einfluss auf das post-aktionale Kompetenzerleben und die damit eng zusammenhängende intrinsische Motivation zu untersuchen. Zu diesem Zweck wurde in einem ersten Auswertungsschritt überprüft, ob eine isolierte Wirksamkeit der Feedbackbedingungen in erwarteter Richtung zu beobachten ist. Da die post-aktionalen motivationalen Merkmale nur zu den beiden Treatmentzeitpunkten erhoben werden konnten, können Aussagen hinsichtlich der isolierten Wirksamkeit der Feedbackbedingungen nur im Vergleich der motivationalen Einschätzungen vom 1. zum 2. Messzeitpunkt getroffen werden. In einem zweiten Auswertungsschritt wurde untersucht, ob sich signifikante Unterschiede in den post-aktionalen motivationalen Parametern zwischen den Feedbackgruppen entsprechend den aufgestellten Hypothesen finden lassen.

5.3.4.1 Intragruppenvergleiche – Isolierte Wirksamkeit der Feedbackbedingungen

In Tabelle 41 sind für beide Messzeitpunkte die Mittelwerte und Standardabweichungen der über die sieben Fragebogenitems im Anschluss an die Programmbearbeitung erfassten und zu einem Gesamtscore aggregierten post-aktionalen motivationalen Einschätzungen für jede Feedbackgruppe abgetragen. Der für jede Feedbackgruppe berechnete *t*-Test für abhängige Stichproben gibt darüber Auskunft, ob die Schüler zum Messzeitpunkt 2 einen größeren Kompetenzzuwachs und damit einhergehend eine erhöhte intrinsische Motivation bei der Aufgabenbearbeitung erfuhren als zu Messzeitpunkt 1.

Die Ergebnisse der *t*-Tests für abhängige Stichproben zeigen für die beiden Feedbackbedingungen KR-KM und KR-KH einen statistisch signifikanten Unterschied von Messzeitpunkt 1 zu Messzeitpunkt 2. Das heißt, die Schüler der KR-KM und KR-KH Feedbackbedingung berichteten zu Messzeitpunkt 2 ein höheres Kompetenzerleben und eine höhere intrinsische Motivation als zu Messzeitpunkt 1.

Tabelle 41:

Deskriptive Statistiken, Prüfstatistiken und Effektstärken für Intragruppenvergleiche hinsichtlich des Einflusses der Feedbackbedingungen auf das post-aktionale Kompetenzerleben

Feedbackgruppen	<i>M (SD)</i>		<i>t (df_{Fehler})</i>	<i>d</i>
	Kompetenzerleben MP1	Kompetenzerleben MP2		
ITF	35.92 (4.47)	36.63 (5.89)	-.75 (23)	.13
KR-KM	29.55 (8.72)	36.50 (5.15)	-2.68* (21)	.99
KR-KH	33.39 (7.88)	38.22 (5.07)	-2.88** (22)	.72
KR-KCR	31.60 (7.42)	31.30 (9.82)	1.53 (19)	-.02

* $p < .05$. ** $p < .01$.

Dass für die Feedbackbedingung ITF kein signifikanter Unterschied gefunden werden konnte, mag daran liegen, dass die Schüler unter der ITF Bedingung schon zu Messzeitpunkt 1 ($M = 35.92$) über ein hohes Kompetenzerleben und eine hohe intrinsische Motivation berichten, Steigerungen so nur begrenzt möglich waren (maximale Punktzahl 42). Im Gegensatz hierzu bewerteten die Schüler der Feedbackbedingung KR-KCR ihre Treatmentleistung zu beiden Messzeitpunkten fast durchweg negativer als die anderen drei Feedbackbedingungen. Die in nachfolgenden Auswertungen berichteten signifikanten Unterschiede zwischen den Feedbackgruppen können daher nicht auf gegenläufige Veränderungen in den einzelnen Feedbackgruppen zurückgeführt werden (H1).

5.3.4.2 Intergruppenvergleiche - Vergleichende Wirksamkeit der Feedbackbedingungen

Um die vergleichende Wirksamkeit der Feedbackbedingungen zu untersuchen, wurden in einem ersten Auswertungsschritt globale Prüfstatistiken berechnet. In Tabelle 42 sind die Ergebnisse der ANCOVA mit der abhängigen Variablen „post-aktionales Kompetenzerleben“ zu jedem Messzeitpunkt abgetragen. In der Kovarianzanalyse zum Messzeitpunkt 1 wurde als Kovariate die Computererfahrung berücksichtigt, in der Kovarianzanalyse zu Messzeitpunkt 2 die Kovariaten Vortestleistung und Freude. In die Varianzanalyse mit Messwiederholung gingen als abhängige Variablen (1) das post-aktionale Kompetenzerleben zu Messzeitpunkt 1 und (2) das post-aktionale Kompetenzerleben zu Messzeitpunkt 2 ein. Als Kovariate wurden die Vortestleistung, die Freude und die Computererfahrung berücksichtigt. Auch hier er-

brachte die Anwendung von Formel 1 mit .07 einen Wert kleiner .10, so dass eine Varianzanalyse mit Messwiederholung indiziert war. Zur Berechnung von Einzelvergleichen von Interaktionseffekten wurde eine ANCOVA mit Einzelvergleichen über die Differenz des post-aktionalen Kompetenzerlebens von Messzeitpunkt 2 zu Messzeitpunkt 1 berechnet. Die Ergebnisse der globalen Prüfstatistiken als auch der Einzelvergleiche sind in Tabelle 42 dargestellt.

Tabelle 42:

Statistische Kennwerte und Effektstärken der globalen Prüfstatistiken und Einzelvergleiche für die Haupt- und Interaktionseffekte der Selbstbewertung der Treatmentleistung zu beiden Messzeitpunkten

	Globale Prüfstatistiken								
	ANCOVA zu MP1			ANCOVA zu MP2			Varianzanalyse mit Messwiederholung		
	<i>df</i> (H,F) ^a	<i>F</i>	ω^2	<i>df</i> (H,F)	<i>F</i>	ω^2	<i>df</i> (H,F)	<i>F</i>	ω^2
Feedback (F)	3, 90	2.02	.03	3, 81	4.05**	.10	3, 77	3.52*	.09
Vortest (V)				1, 81	5.36*	.05	1, 77	4.01*	.04
Freude (F)				1, 81	7.39**	.07	1, 77	2.91 ⁺	.04
Computer (C)	1, 90	6.49*	.06				1, 77	1.40	.01
F x Z ^b							3, 77	2.25 ⁺	.04
V x Z							1, 77	.85	.00
F x Z							1, 77	3.27 ⁺	.03
C x Z							1, 77	6.90*	.01
Einzelvergleiche mit Bonferroni-Korrektur									
	MP1		MP2		Feedback x Zeit				
	Mittlere Differenz (SE)	<i>d</i>	Mittlere Differenz (SE)	<i>d</i>	Mittlere Differenz ^c (SE)	<i>d</i>			
ITF vs. KR-KCR	3.30 (2.10)	.53	4.64 (1.97)	.57	1.36 (2.59)	.19			
ITF vs. KR-KM	4.99 (2.10)	.81	.65 (1.92)	.11	-4.84 (2.64)	-.51			
ITF vs. KR-KH	1.89 (2.09)	.29	-2.03 (1.92)	-.31	-3.95 (2.56)	-.59			
KR-KH vs. KR-KM	3.10 (2.06)	.38	2.68 (1.92)	.43	-.89 (2.63)	-.08			
KR-KH vs. KR-KCR	1.41 (2.09)	.18	6.67** (1.95)	.78	5.31 (2.56)	.63			
KR-KM vs. KR-KCR	-1.69 (2.06)	-.21	3.99 (1.96)	.51	6.20 (2.67)	.58			

Anmerkungen.

^aH = Hypothesenfreiheitsgrade, F = Fehlerfreiheitsgrade

^bZ = Zeit

^cAbhängige Variable zur Berechnung der Einzelvergleiche der Interaktionseffekte bildete die Differenz des post-aktionalen Kompetenzerlebens von Messzeitpunkt 2 zu Messzeitpunkt 1

⁺p<.10. *p<.05. **p<.01.

Konnte zum 1. Messzeitpunkt zwischen den vier Feedbackgruppen kein signifikanter Unterschied hinsichtlich des post-aktionalen Kompetenzerlebens gefunden werden, zeigt sich ein solcher zum 2. Messzeitpunkt. Post-hoc Vergleiche zum zweiten Messzeitpunkt zeigen einen statistisch signifikanten Unterschied zwischen den beiden Feedbackbedingungen KR-KH und KR-KCR. Schüler, welche KR-KH Feedback erhielten, berichteten ein höheres Kompetenzerleben und eine höhere intrinsische Motivation als Schüler, welche KR-KCR Feedback erhielten. Auch wenn zwischen den anderen Feedbackbedingungen keine signifikanten Effekte gefunden werden konnten, bestätigt Abbildung 29 die aufgestellte Hypothese dahingehend, dass zum 1. Messzeitpunkt die ITF Gruppe eine höhere post-aktionale Motivation berichtet als die drei anderen Feedbackgruppen (H2b₁). Ein solch hypothesenkonformes Bild zeigt sich allerdings nicht hinsichtlich der Einschätzung des post-aktionalen Kompetenzerlebens der KR-KM Feedbackgruppe. So bewerteten die Schüler der KR-KM Feedbackgruppe ihre Treatmentleistung zum 1. Messzeitpunkt schlechter als die Schüler der KR-KCR Feedbackgruppe, obwohl die Schüler der KR-KM Feedbackgruppe zum ersten Messzeitpunkt geringfügig bessere Treatmentleistungen als die Schüler der KR-KCR Feedbackgruppe zeigten.

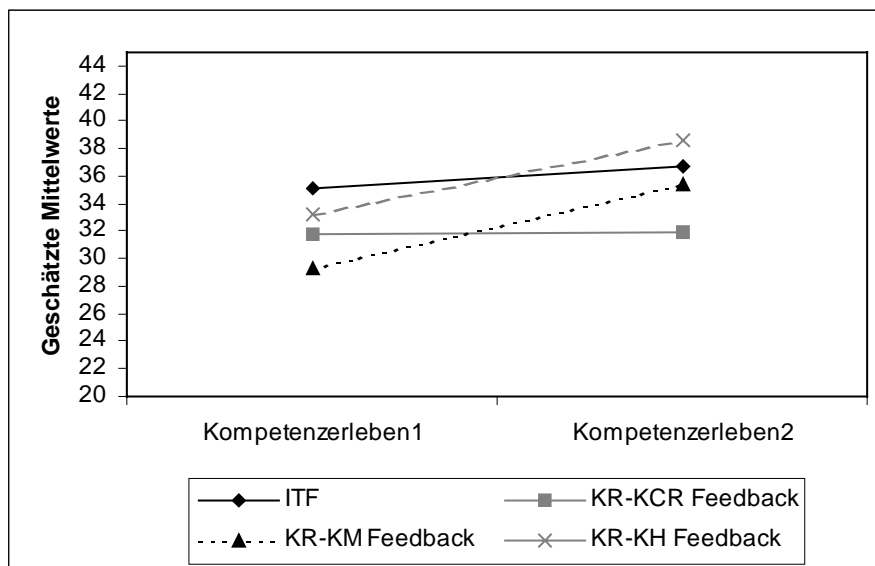


Abbildung 29: Darstellung der unter Auspartialisierung der Kovariaten Vortestleistung, Freude und Computererfahrung geschätzten Mittelwerte des post-aktionalen Kompetenzerlebens zu Messzeitpunkt 1 und 2 für die einzelnen Feedbackbedingungen

Dieses Bild änderte sich zum zweiten Messzeitpunkt dahingehend, dass jetzt die KR-KCR Feedbackgruppe ihre Leistung während des Treatments am negativsten bewertete und die KR-KH Feedbackgruppe ihre Treatmentleistung am positivsten einschätzte. Dass zum 2. Messzeitpunkt die Schüler, welche KR-KH Feedback erhielten, ein höheres Kompetenzerleben und eine höhere intrinsische Motivation berichten als die Schüler, welche ITF erhielten, kann möglicherweise darüber erklärt werden, dass die Schüler mit ITF zum 2. Messzeitpunkt nur einen sehr geringen Lernfortschritt zeigten. Ursächlich für diesen geringen Lernfortschritt der ITF Gruppe mag der schon häufiger zitierte Deckeneffekt gewesen sein.

In einem nächsten Auswertungsschritt wurde für jeden Messzeitpunkt eine erneute Kovarianzanalyse diesmal unter Hinzunahme der Kovariaten „Treatmentleistung zu Messzeitpunkt 1“ und „Treatmentleistung zu Messzeitpunkt 2“ berechnet. Im Vorfeld durchgeführte Korrelationsanalysen zeigten erwartungskonform hoch signifikante Zusammenhänge zwischen der Treatmentleistung und dem post-aktionalen Kompetenzerleben zu Messzeitpunkt 1 ($r = .51^{**}$) und zu Messzeitpunkt 2 ($r = .40^{**}$). Die Anwendung von Formel 1 zur Berechnung der zulässigen Kovariatenanzahl erbrachte für die Varianzanalyse mit Messwiederholung bei 5 Kovariaten mit .096 einen Wert kleiner .10, so dass eine Varianzanalyse mit Messwiederholung berechnet werden konnte. Die Ergebnisse der Kovarianzanalysen mit den dazugehörigen Detailanalysen über beide Messzeitpunkte sowie die Ergebnisse der Varianzanalyse mit Messwiederholung und den entsprechenden Einzelvergleichen sind in Tabelle 43 dargestellt.

Auch unter Auspartialisierung der Treatmentleistung bleibt der signifikante Unterschied im post-aktionalen Kompetenzerleben zu Messzeitpunkt 2 zwischen der KR-KH Feedbackgruppe und der KR-KCR Feedbackgruppe erhalten. Der Unterschied in der abhängigen Variablen zu Messzeitpunkt 1 zwischen der KR-KM Feedbackgruppe und den anderen drei Feedbackgruppen wurde nach Auspartialisierung der Treatmentleistung noch deutlicher.

Tabelle 43:

Statistische Kennwerte und Effektstärken der globalen Prüfstatistiken und Einzelvergleiche für die Haupt- und Interaktionseffekte des post-aktionalen Kompetenzerlebens unter Auspar-tialisierung des Einflusses der Treatmentleistung zu beiden Messzeitpunkten

	Globale Prüfstatistiken								
	ANCOVA zu MP1			ANCOVA zu MP2			Varianzanalyse mit Messwiederholung		
	<i>df</i> (H,F) ^a	<i>F</i>	ω^2	<i>df</i> (H,F)	<i>F</i>	ω^2	<i>df</i> (H,F)	<i>F</i>	ω^2
Feedback (F)	3,89	1.18	.01	3,79	3.12**	.10	3,74	1.77	.03
Treatment1 ^b (T1)	1,89	21.33***	.18				1,74	4.76*	.05
Treatment2 (T2)				1,79	15.13***	.11	1,74	3.16 ⁺	.03
Vortest (V)				1,81	2.68	.00	1,74	1.02	.00
Freude (F)				1,81	12.01**	.15	1,74	6.42*	.07
Computer (C)	1,89	6.49*	.03				1,74	.17	-.01
F x Z ^c							3,74	2.67 ⁺	.06
T1 x Z							1,74	13.85***	.14
T2 x Z							1,74	10.97**	.12
V x Z							1,74	1.45	.01
F x Z							1,74	3.01 ⁺	.03
C x Z							1,74	3.79 ⁺	.04

	Einzelvergleiche mit Bonferroni-Korrektur					
	MP1		MP2		Feedback x Zeit	
	Mittlere Differenz (SE)	<i>d</i>	Mittlere Differenz (SE)	<i>d</i>	Mittlere Diffe- renz ^d (SE)	<i>d</i>
ITF vs. KR-KCR	1.14 (1.95)	.18	2.52 (1.90)	.31	2.30 (2.70)	.20
ITF vs. KR-KM	3.46 (1.93)	.50	-.87 (1.82)	-.15	-4.19 (2.69)	-.44
ITF vs. KR-KH	1.01 (1.90)	.16	-3.03 (1.82)	-.54	-3.62 (2.56)	-.54
KR-KH vs. KR-KM	2.45 (1.87)	.30	2.16 (1.80)	.42	-.57 (2.64)	-.05
KR-KH vs. KR-KCR	.14 (1.91)	.02	5.56*(1.85)	.71	5.92 (2.60)	.70
KR-KM vs. KR-KCR	-2.32 (1.86)	-.29	3.39 (1.85)	.43	6.49 (2.67)	.72

Anmerkungen.

^aH = Hypothesenfreiheitsgrade, F = Fehlerfreiheitsgrade

^bTreatment1 = Treatmentleistung zu Messzeitpunkt 1

^cZ = Zeit

^dAbhängige Variable zur Berechnung der Einzelvergleiche der Interaktionseffekte bildete die Differenz des post-aktionalen Kompetenzerlebens von Messzeitpunkt 2 zu Messzeitpunkt 1

⁺p<.10. *p<.05. **p<.01. ***p<.001.

Zum 2. Messzeitpunkt zeigte die KR-KH Gruppe auch unter Kontrolle der Treatmentleistung die positivste Selbstbewertungsbilanz (siehe Abbildung 30).

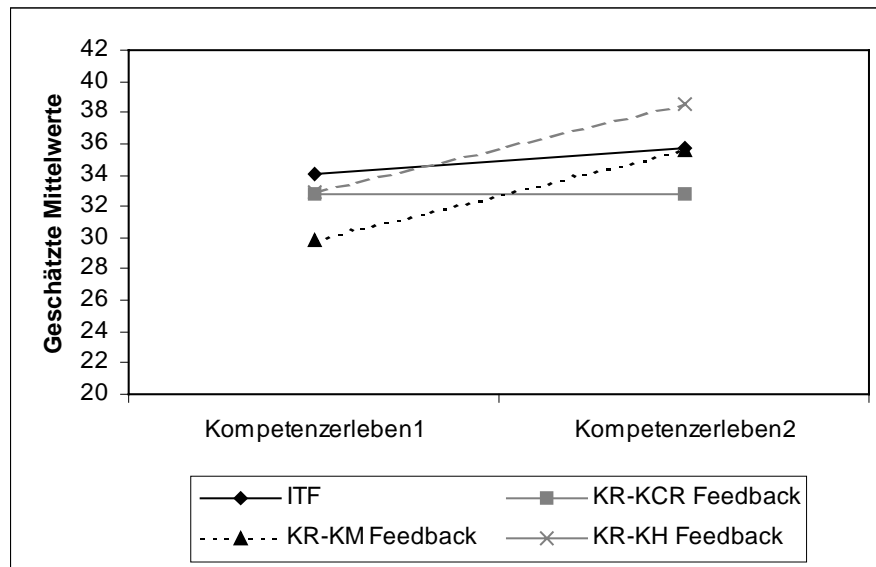


Abbildung 30: Darstellung der unter Auspartialisierung der Kovariaten Treatmentleistung, Vortestleistung, Freude und Computererfahrung geschätzten Mittelwerte des post-aktionalen Kompetenzerlebens zu Messzeitpunkt 1 und 2 für die einzelnen Feedbackbedingungen

5.4 Diskussion

Ziel der vorliegenden Studie bestand zum einen in einer Replikation der Ergebnisse der ersten Studie. In dieser ersten Studie konnte hypothesenkonform eine Überlegenheit des entwickelten fehlerspezifischen ITFs für die schriftliche Subtraktion gegenüber dem wenig informativen KR-KCR Feedback hinsichtlich Lern- und Motivationseffekte nachgewiesen werden. Neben dieser Replikation bestand ein weiteres zentrales Anliegen der vorliegenden Studie in einer empirischen Evaluation des fehlerspezifischen ITFs im Vergleich zu zwei komplexeren Feedbackformen als das in der ersten Studie implementierte KR-KCR Feedback. Die höhere Komplexität der beiden Feedbackformen „Knowledge-About-Mistake“ (KR-KM) Feedback und „Knowledge-on-How-to-Proceed“ (KR-KH) Feedback gegenüber dem KR-KCR Feedback zeichnete sich durch das einmalige Bereitstellen korrekturrelevanter Informationen mit gleichzeitiger Präsentation der korrekten Lösung aus. Dabei präsentierte das KR-KM Feedback dem Lerner in Abhängigkeit des diagnostizierten Fehlers einen auditiven Hinweis auf den diagnostizierten Fehler und die korrekte Lösungsstrategie. Das KR-KH Feedback hingegen modellierte an der konkreten, fehlerhaft bearbeiteten Aufgabe, wie diese Aufgabe gelöst wird, d.h. dem Lerner wurde die korrekte Lösungsprozedur sowohl auditiv als auch visuell demonstriert. Entsprechend dem KR-KM Feedback der zweiten Stufe des ITF Algo-

rithmus, dann repräsentierte das KR-KH Feedback die dritte Stufe des ITF Algorithmus. Neben Aussagen zur prinzipiellen Wirksamkeit der unterschiedlich komplexen Feedbackformen können damit auch Aussagen zur isolierten Wirksamkeit der im ITF implementierten Feedbackstufen getroffen werden.

Das fehlerspezifische ITF wurde in der vorliegenden Studie hinsichtlich seiner Lern- und motivationalen Wirksamkeit mit den beiden Feedbackformen KR-KM und KR-KH sowie mit dem KR-KCR Feedback verglichen. Im weiteren Verlauf werden zuerst Ergebnisse der vorliegenden Studie im Hinblick auf die Lernwirksamkeit der untersuchten Feedbackformen diskutiert (Kapitel 5.4.1), anschließend deren motivationale Effekte (Kapitel 5.4.2).

5.4.1 Lernwirksamkeit der untersuchten Feedbackformen

ITF sollte nach einer fehlerhaften Aufgabenlösung aufgrund der Präsentation korrektur-relevanter Informationen ohne unmittelbare Darbietung der richtigen Lösung zu einer aktiven Auseinandersetzung mit diesen Informationen anregen. Eine solch intensive Verarbeitung der Feedbackinformationen sollte in einer erhöhten Lernwirksamkeit resultieren. Diese Annahme konnte durch die vorliegenden Ergebnisse dahingehend bestätigt werden, dass Schüler, die ITF erhielten, sowohl während der Treatmentphase als auch bei der Bearbeitung des Nachttestes bessere Leistungen erzielten als Schüler, welche weniger komplexes Feedback präsentiert bekamen. Das heißt, Schüler mit ITF bearbeiteten mehr Aufgabentypen am Computer vollständig und korrigierten ihre im Vortest fehlerhaft gelösten Aufgaben auch häufiger im Nachttest als die Schüler der drei anderen Feedbackbedingungen. Eine solche Überlegenheit der ITF Gruppe konnte im Vergleich zur KR-KCR Feedbackgruppe und zur KR-KM Feedbackgruppe für den ersten Messzeitpunkt mindestens auf dem 5%-Niveau für die Nachttestleistung und den Korrekturwahrscheinlichkeitsindex der Nachttestleistung statistisch abgesichert werden. Hinsichtlich der Treatmentleistung konnte eine statistisch signifikante Überlegenheit des ITF gegenüber der KR-KCR Feedbackgruppe zu beiden Messzeitpunkten auf dem 5%-Niveau gefunden werden. Im Vergleich zum KR-KH Feedback ließ sich über beide Messzeitpunkte kein statistisch signifikanter Vorteil des ITF für die drei Leistungsindikatoren finden.

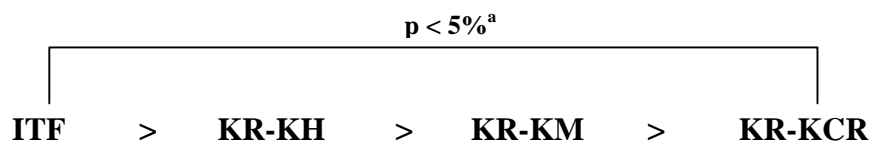
Zusammenfassend lässt sich zum Messzeitpunkt 1 über alle Leistungsmaße eine Wirksamkeitsreihenfolge der vier Feedbackformen in folgender Weise konsistent wiederfinden:



^afür Nachttestleistung, Korrekturwahrscheinlichkeitsindex und Treatmentleistung

^bfür Nachttestleistung und Korrekturwahrscheinlichkeitsindex

Dass zum 2. Messzeitpunkt von Studie 2 fast keine statistisch signifikanten Wirksamkeitsunterschiede zwischen den Feedbackbedingungen mehr beobachtet werden konnten, ist möglicherweise über einen Deckeneffekt der ITF Gruppe zu erklären. So lösten mehr als 50% der Schüler der ITF Gruppe im Nachttest1 mindestens 17 der 24 Aufgaben korrekt. Bei der KR-KCR Feedbackgruppe waren dies nur 13%, bei der KR-KM Feedbackgruppe nur 10% und bei der KR-KH Feedbackgruppe 33%, die mindestens 17 der 24 Aufgaben des Nachttestes korrekt bearbeiteten. Die Schüler dieser drei Feedbackbedingungen hatten demnach zum 2. Messzeitpunkt ein höheres Verbesserungspotenzial, welches sie auch nutzten. Auch wenn zum 2. Messzeitpunkt mit Ausnahme der Treatmentleistung zwischen ITF und KR-KCR Feedback keine signifikanten Unterschiede gefunden werden konnten, zeigte sich analog zum Messzeitpunkt 1 der Studie 2 folgende Wirksamkeitsreihenfolge der vier Feedbackbedingungen über fast alle Leistungsmaße:



^afür Treatmentleistung

Mit den Ergebnissen der Studie 2 konnten die Ergebnisse der 1. Studie repliziert und damit ein zentrales Ziel dieser Studie erreicht werden. Mit dieser Replikation erhöht sich die Reliabilität der in Studie 1 und Studie 2 gefundenen Ergebnisse, welches wiederum als ein Beleg für die Gültigkeit der entwickelten Gestaltungs- und Evaluationsprinzipien für ITF interpretiert werden kann. Im folgenden wird detailliert auf Wirksamkeitsunterschiede zwischen dem ITF und den drei Vergleichsbedingungen eingegangen, bevor Erklärungsansätze präsentiert

werden, welche die geringere Lernwirksamkeit des KR-KM Feedbacks im Vergleich zum KR-KH Feedbacks zum 1. Messzeitpunkt zu erklären versuchen.

5.4.1.1 Wirksamkeitsunterschiede zwischen ITF und den drei weniger komplexen Feedbackformen

Erwartungskonform, wenn auch nicht immer statistisch signifikant, konnte eine Überlegenheit des ITF gegenüber den weniger komplexen Feedbackformen bezüglich seiner Lernwirksamkeit nachgewiesen werden. Da sich das ITF sowohl hinsichtlich seiner informativen als auch seiner tutoriellen Komponente von den drei anderen Feedbackformen unterschied, lässt sich allerdings nicht eindeutig feststellen, auf welche Komponente diese Überlegenheit zurückgeführt werden kann. So wurden den Schülern unter der KR-KM und der KR-KH Feedbackbedingung korrekturrelevante Informationen mit der richtigen Lösung nach einem zweiten fehlerhaften Lösungsversuch präsentiert. Aufgrund der synchronen Präsentation der korrekturrelevanten Informationen mit der korrekten Lösung war das Angebot zusätzlicher Informationen nicht möglich. Das heißt, eine systematische Manipulation der tutoriellen Komponente resultierte gleichzeitig auch in einer Manipulation der informativen Komponente. Isolierte Wirksamkeitsaussagen hinsichtlich der tutoriellen bzw. informativen Komponente des ITF können daher nicht getroffen werden. Das heißt, die Wirksamkeit des ITF kann nur in seiner Gesamtheit mit den anderen Feedbackformen verglichen werden. Dass sich ITF auch den komplexeren Feedbackformen gegenüber als lernwirksamer erwiesen hat, wurde bereits erwähnt. Dies zeigt sich vor allem bei der Betrachtung der Treatmentleistung der Feedbackgruppen über beide Messzeitpunkte. So konnte sogar im Vergleich zum KR-KH Feedback – dem effektivsten der drei Feedbackvergleichsbedingungen - zu beiden Messzeitpunkten eine mittlere Effektstärke ($d = .34$ bzw. $d = .37$) zugunsten des ITF gefunden werden. Dieses Ergebnis kann als ein deutlicher Hinweis darauf interpretiert werden, dass ITF vor allem die aktuelle Bearbeitung von Aufgaben unterstützt.

Obwohl die Wirksamkeitsunterschiede hypothesenkonform in den erwarteten Richtungen zu beobachten waren, konnten sie nicht in allen Fällen statistisch abgesichert werden. Signifikante Unterschiede zeigten sich vor allem für den Vergleich der beiden extremen Feedbackbedingungen, dem ITF und dem KR-KCR Feedback. Dass es sich hierbei auch um praktisch bedeutsame Effekte handelt, zeigen die Effektstärken, welche teilweise mit $d = 1.18$ sogar höher als erwartet ausfielen. Dass die anderen Vergleiche nicht immer signifikant wur-

den, mag zum einen an den geringen Stichprobengrößen liegen, welche sich aber im Vorfeld aufgrund der Versuchspersonenrekrutierung über die Vortestleistung nicht präzise festlegen ließen. Eine alternative Erklärung kann aber auch in der Konstruktion der vier Feedbackbedingungen gesehen werden. So war in allen vier Feedbackbedingungen ein mastery level im Sinne des zielerreichenden Lernens (mastery learning) implementiert. Das heißt, unter allen vier Feedbackbedingungen mussten die Schüler drei Aufgaben hintereinander im ersten Versuch korrekt gelöst haben, um den nächsten Aufgabentyp präsentiert zu bekommen. Im Kapitel 2.4.3 wurde die Implementierung eines mastery levels als eine Möglichkeit diskutiert, eine aktive Auseinandersetzung mit den präsentierten Informationen bzw. mit der zu bearbeitenden Aufgabe anzuregen. Erklärt wurde diese Annahme darüber, dass die Lerner in dem Wissen, dass das Ende der Lehrinheit nur zu erreichen ist, wenn die Aufgaben eigenständig korrekt bearbeitet werden, sich bei der Aufgabenbearbeitung mehr anstrengen. Das heißt, die Lerner der KR-KM und KR-KH Feedbackgruppe mussten die bereitgestellten korrekturrelevanten Informationen nutzen, um die nächste Aufgabe korrekt bearbeiten zu können. Das bereits berichtete Problem vieler Feedbackstudien, in denen elaborierte Feedbackinformationen gemeinsam mit der richtigen Lösung angeboten wurden (z.B. Kulhavy et al., 1985), erweist sich demnach für die Feedbackformen KR-KM und KR-KH der vorliegenden Studie als nicht relevant, da die Darbietung der korrekten Lösung weder eine notwendige noch eine hinreichende Information präsentierte, um die nächste – d.h. neue - Aufgabe korrekt lösen zu können. Möglicherweise stellte die Implementierung eines Lernkriteriums aber auch einen Leistungsanreiz für die Schüler im Rahmen eines sozialen Vergleichs dar. Da die Schüler sich melden sollten, wenn sie einen Aufgabentyp erfolgreich am Computer bearbeitet hatten, war es ihnen möglich, ihren Mitschülern zu demonstrieren, wie schnell und erfolgreich sie bei der Aufgabenbearbeitung sind. Das heißt, eine aktive Auseinandersetzung wurde nicht nur über das adaptive Angebot korrekturrelevanter Informationen angeregt (nur für ITF zutreffend), sondern auch über die Implementierung eines mastery levels unter alle vier Feedbackbedingungen.

Insgesamt konnte die zentrale Bedeutung einer theoriegeleiteten Konstruktion informativen tutoriellen Feedbacks durch die Ergebnisse der vorliegenden Studie bestätigt werden. Unterstützung erhalten diese Ergebnisse auch durch Studien, in welchen gleichfalls computergestütztes ITF im Sinne eines systematisch entwickelten fehlerspezifischen Feedbacks den Lernenden angeboten wurde. In einer dieser Studien evaluierten Nagata und Swisher (1995)

die Effektivität zweier Feedbackformen in einer computergestützten Lernumgebung zum japanischen Fremdsprachenerwerb. Das von ihnen als traditionell bezeichnete Feedback verwies auf fehlende oder unerwartete Wörter in des Lerners Antwort, das „intelligente“ Feedback stellte hingegen weitere Informationen über die Natur des Fehlers in Form metalinguistischer Regeln bereit. Unter beiden Feedbackbedingungen hatten die Lernenden drei Korrekturversuche. Nagata und Swisher fanden, dass das fehlerspezifische Feedback dahingehend erfolgreicher war, dass es die Kompetenz des Lerners in der Verwendung komplexer grammatikalischer Strukturen im Vergleich zum traditionellen Feedback signifikant verbesserte. In einer anderen Studie untersuchten Collins, Carnine und Gersten (1987) die Effektivität zweier Feedbackformen in einer computergestützten Lernumgebung zum Ziehen korrekter Schlüsse aus Syllogismen. Die eine Feedbackbedingung präsentierte nach einem fehlerhaften Lösungsversuch die richtige Lösung (KCR), die andere Feedbackbedingung stellte fehlerspezifisches Feedback in Form von konkreten Hinweisen und Regeln zusammen mit der korrekten Lösung bereit. Eine aktive Auseinandersetzung mit den präsentierten Informationen wurde durch die Implementierung eines mastery levels induziert. Dieses mastery level war darüber definiert, dass die Lernenden eine fehlerhaft bearbeitete Aufgabe am Ende der Lehreinheit so lange erneut bearbeiten mussten, bis sie diese korrekt bearbeitet hatten. Auch in dieser Studie erwies sich das fehlerspezifische Feedback dem KCR Feedback statistisch signifikant überlegen.

Da in der vorliegenden Studie keine Follow-up Erhebung durchgeführt wurde, können keine Aussagen hinsichtlich der längerfristigen Wirksamkeit des ITF getroffen werden. Das Treatment wurde zu Beginn des vierten Schuljahres durchgeführt und die Schüler hatten im Anschluss an die Interventionsphase im Rahmen des regulären Unterrichts Aufgaben der schriftlichen Subtraktion im Sinne einer Festigungsphase zu bearbeiten. Aus diesem Grund wurde eine Konfundierung zwischen der im Treatment erworbenen Fertigkeiten und den im Unterricht erworbenen Fertigkeiten erwartet und deshalb auf eine Follow-up Untersuchung verzichtet.

5.4.1.2 Wirksamkeitsunterschiede zwischen KR-KM Feedback und KR-KH Feedback

Erwartungswidrig - und von daher auch überraschend – war die geringe Lernwirksamkeit der KR-KM Feedbackbedingung zum ersten Messzeitpunkt der zweiten Studie. Zur Erinnerung sei noch einmal angeführt, dass Knowledge about mistake (KR-KM) Feedback dem

Lerner in Abhängigkeit des diagnostizierten Fehlers korrekturrelevante Hinweise präsentierte. Diagnostizierbare, d.h. systematische Fehler waren im Vorfeld der Studien basierend auf einer empirischen Fehleranalyse identifiziert und in die Übungssoftware implementiert worden. Konnte das Programm einen solchen systematischen Fehler identifizieren, wurde dem Lerner neben einem Hinweis auf die vermutete Fehlstrategie ein Hinweis auf die korrekte Lösungsstrategie auditiv bereitgestellt. Machte der Schüler einen Fehler, welcher nicht in der Fehlerbibliothek des Programms enthalten war, erhielt der Lerner den Hinweis, sich ein ausgearbeitetes Lösungsbeispiel zu diesem Aufgabentyp über den Rettungsring-Button aufzurufen.

Eine Erklärung für die geringe Effektivität des KR-KM Feedbacks gegenüber dem KR-KH Feedback zum 1. Messzeitpunkt kann darin gesehen werden, dass Lerner gehäuft Fehler bei der Aufgabenbearbeitung zeigten, die nicht in der Fehlerbibliothek der Software enthalten waren und somit nicht diagnostiziert werden konnten. Der Lerner erhielt dann nur den Tipp, das ausgearbeitete Lösungsbeispiel aufzurufen. Im Gegensatz zum KR-KM Feedback wurde den Lernern der KR-KH Feedbackgruppe auch bei einem nicht diagnostizierten Fehler die korrekte Lösungsprozedur sowohl visuell als auch auditiv an der konkreten fehlerhaft bearbeiteten Aufgabe demonstriert. Um die Annahme zu überprüfen, dass korrekturrelevante Informationen unter der KR-KM Feedbackbedingung im Falle eines nicht diagnostizierbaren Fehlers dem Lerner nicht bereitgestellt wurden und daher nicht wirksam werden konnten, wurden die KR-KM und die KR-KH Feedbackgruppe aufgrund sonst gleicher Feedbackgestaltung hinsichtlich folgender Faktoren miteinander verglichen:

- (1) In einem ersten Schritt wurde analysiert, wie häufig nicht diagnostizierbare Fehler beobachtet werden konnten.
- (2) In einem zweiten Schritt wurde analysiert, wie häufig die Lerner der KR-KM bzw. KR-KH Feedbackbedingung, die im zweiten Lösungsversuch einen nicht diagnostizierbaren Fehler zeigten, die nächste Aufgabe im 1. Lösungsversuch korrekt bearbeiteten. Auch hier sei noch einmal daran erinnert, dass die Lerner der KR-KM und KR-KH Feedbackgruppe nur zwei Lösungsversuche hatten, und dass ihnen die korrekturrelevanten Informationen gemeinsam mit der korrekten Lösung im Anschluss an einen zweiten fehlerhaften Lösungsversuch präsentiert wurden.

Die Analyse zeigt, dass die Lerner beider Feedbackgruppen 25% bzw. 27% nicht diagnostizierbarer Fehler im 2. Lösungsversuch machten. Dieses Ergebnis ist konsistent mit den von Gerster (1982) und Radatz (1980) berichteten Anteilen unsystematischer Fehler am Ge-

samtfehleranteil und kann als Ausdruck für die gelungene Implementierung der im Vorfeld auf Basis einer Aufgaben- und Fehleranalyse identifizierten systematischen Fehler interpretiert werden. Zeigten die Lerner der KR-KM Feedbackgruppe einen solch nicht diagnostizierbaren Fehler im 2. Lösungsversuch, lösten sie in 24% der Fälle die nächste Aufgabe im 1. Lösungsversuch korrekt. Lerner der KR-KH Feedbackgruppe hingegen bearbeiteten in 41% der Fälle die nächste Aufgabe im 1. Lösungsversuch richtig. Obwohl der Mann-Whitney-Test für unabhängige Stichproben keinen signifikanten Unterschied zwischen beiden Gruppen zeigte ($U = 78.5, p = .30$), deutet dieses Ergebnis in der Tendenz doch darauf hin, dass Lerner der KR-KM Feedbackgruppe bei einem nicht diagnostizierbaren Fehler keine Hilfe erhielten, die sie dazu befähigte, die nächste Aufgabe im 1. Lösungsversuch korrekt zu bearbeiten.

Eine weitere Erklärung für die geringe Lernwirksamkeit der KR-KM Feedbackbedingung zum 1. Messzeitpunkt kann in der Präsentationsform der korrekturrelevanten Informationen gesehen werden. Wurden unter der KR-KM Feedbackbedingung die korrekturrelevanten Informationen dem Lerner nur über den auditiven Kanal bereitgestellt, bekamen die Lerner der KR-KH Feedbackbedingung die korrekturrelevanten Informationen sowohl über den auditiven als auch über den visuellen Kanal präsentiert. Bei der Ableitung der Gestaltungsprinzipien (Kapitel 2.4.3) wurde bereits auf die Überlegenheit einer bimodalen gegenüber einer monomodalen Informationspräsentation hingewiesen (Mayer & Moreno, 2002). So reduziert die synchrone Präsentation auditiv-verbaler und visuell-piktoraler Informationen im Sinne der cognitive load Theorie (Sweller, 1994) die Arbeitsgedächtnisbelastung des Lernenden aufgrund der Nutzung des sowohl verbalen Informationsverarbeitungssystems („phonologische Schleife“, Baddeley, 1986) als auch des visuellen Informationsverarbeitungssystems („visuell-räumlicher Notizblock“, Baddeley, 1986). Unter Berücksichtigung der Multimedia Theorie (Mayer & Moreno, 2002) und der cognitive load Theorie (Sweller, 1994) ist für die KR-KM Feedbackbedingung daher zu vermuten, dass die alleinige auditive Präsentation der korrekturrelevanten Informationen die Lernenden zu Messzeitpunkt 1 überforderte, da die Lernenden aufgrund der Fülle an auditiv präsentierten Informationen diese nicht angemessen aufnehmen und verarbeiten konnten. Diese Überlastung, welche vor allem bei der sukzessiven Verarbeitung von Informationen eine große Rolle spielt, tritt auf, wenn das Informationsangebot die Kapazität der Verarbeitungsprozesse übersteigt, wie es möglicherweise für die KR-KM Feedbackbedingung zutraf. So fielen in Abhängigkeit des systematischen Fehlers einige auditive Rückmeldungen verhältnismäßig lang aus. Möglicherweise verfügten die Lernenden

aber auch über zu wenig Wissen und Übung, um die korrekturrelevanten Informationen adäquat verstehen und anwenden zu können. Dafür spricht der starke Lernzuwachs der KR-KM Feedbackgruppe zu Messzeitpunkt 2. So könnte die wiederholte auditive Präsentation der fehlerspezifischen Feedbackinformationen zu Messzeitpunkt 1 eine Schemakonstruktion und Automatisierung im Sinne der cognitive-load Theorie (Sweller, 1994) unterstützt haben. Der angeregte Schemaerwerb der schriftlichen Subtraktion zu Messzeitpunkt 1 sollte die kognitive Belastung der Schüler, welche KR-KM Feedback erhielten, zum Messzeitpunkt 2 reduziert haben, so dass sie aufgrund einer geringeren Arbeitsgedächtnisbelastung die auditiv präsentierten Informationen dann effektiver verarbeiten konnten.

Eine weitere Erklärung für die höhere Lernwirksamkeit von KR-KH Feedback im Vergleich zu KR-KM Feedback kann auch darin gesehen werden, dass die Visualisierung der Lösungsprozedur für den Erwerb prozeduralen Wissens eine angemessene, d.h. effektive Präsentationsform darstellt, da es sich bei diesem Lehrziel um die Darstellung dynamischer, prozeduraler Inhalte handelt, deren Erwerb besonders durch animierte Visualisierungen unterstützt werden sollte.

Obwohl sich sowohl das KR-KM Feedback als auch das KR-KH Feedback zum 2. Messzeitpunkt als gleichermaßen lernwirksam erwiesen, weisen doch vor allem die Ergebnisse zum 1. Messzeitpunkt darauf hin, dass das KR-KM Feedback möglicherweise aufgrund der rein auditiven Präsentation der korrekturrelevanten Informationen den Lerner überforderte. So setzt das KR-KM Feedback scheinbar ein höheres aufgabenspezifisches Vorwissen voraus, um die rein auditiv präsentierten Informationen angemessen aufnehmen und verarbeiten zu können. Vielleicht boten sich den Schülern der KR-KM Feedbackbedingung aber auch weniger Korrekturgelegenheiten als den Schülern der KR-KH Feedbackbedingung, da diesen Schülern bei einem nicht diagnostizierbaren Fehler keine korrekturrelevanten Informationen bereitgestellt wurden. So bearbeiteten die Schüler der KR-KH Feedbackbedingung nach einem nicht diagnostizierbaren Fehler im zweiten Lösungsversuch die nächste Aufgabe häufiger korrekt als die Schüler der KR-KM Feedbackbedingung.

Da das KR-KM Feedback die zweite und das KR-KH Feedback die dritte Stufe des ITF Algorithmus repräsentieren, ist davon auszugehen, dass die hier diskutierten Erklärungsansätze zur Wirksamkeit der einzelnen Feedbackbedingungen auf das ITF übertragbar sind. Vor dem Hintergrund der geringeren Lernwirksamkeit des KR-KM Feedback besteht weiterer

Forschungsbedarf dahingehend, zu untersuchen, wie die zweite Stufe des ITF Algorithmus verbessert werden kann.

5.4.2 Motivationale Wirksamkeit der untersuchten Feedbackformen

Neben dem Nachweis der Lernwirksamkeit des ITF bestand ein zweites zentrales Anliegen der vorliegenden Arbeit darin, den Einfluss des ITF auf post-aktionale motivationale Parameter im Vergleich zu den drei anderen Feedbackformen zu untersuchen. Diese motivationalen Parameter beinhalteten in Anlehnung an die Arbeiten von Narciss (1999, 2001, 2004) Items zum post-aktionalen Kompetenzerleben und dem damit verknüpften intrinsischen Anreizerleben. Hinsichtlich der motivationsförderlichen Wirkung des ITFs wurde von der Annahme ausgegangen, dass ITF, welches korrekturrelevante Informationen bereitstellt, die Chance auf intern attribuibare Erfolgserlebnisse erhöht. Sowohl gehäufte Erfolgserlebnisse als auch deren interne Attribution sollten zu einem positiven Erleben der eigenen Kompetenz geführt haben. Ein erhöhtes Kompetenzerleben sollte sich darin äußern, dass Lernende zufriedener mit ihren Leistungen sind, diese positiver bewerten und einen erhöhten Lerngewinn/Lernfortschritt erfahren (Narciss, 2004). Da in vielen Motivationstheorien ein enger Zusammenhang zwischen dem Kompetenzerleben und intrinsischen Anreizwerten postuliert wird (z.B. die kognitive Evaluations-Theorie, Deci & Ryan, 1975; die Interessentheorie, Krapp, 1999; die Flow-Theorie, Csikszentmihaly, 1975; siehe zusammenfassend Narciss, 2004), wurde darüber hinaus erwartet, dass das Erleben von Kompetenz auch mit einer erhöhten intrinsischen Motivation einhergeht.

Das Ergebnis der ersten Studie, in welcher ein signifikanter Unterschied hinsichtlich der motivationalen Effekte zugunsten der ITF Gruppe gefunden wurde, konnte mit den Ergebnissen der zweiten Studie teilweise bestätigt werden. So berichteten Schüler der ITF Bedingung über ein höheres Kompetenzerleben und eine höhere intrinsische Motivation als Schüler der KR-KCR Feedbackbedingung. Dieser Unterschied zwischen der ITF Gruppe und der KR-KCR Feedbackgruppe konnte statistisch zwar nicht abgesichert werden, doch weisen Effektstärken von $d = .53$ zum ersten Messzeitpunkt und $d = .57$ zum zweiten Messzeitpunkt auf einen mittleren Effekt hin. Von daher stellen diese Ergebnisse eine partielle Replikation der Befunde der ersten Studie dar. Betrachtet man die beiden anderen, hinsichtlich ihrer Komplexität zwischen dem ITF und dem KR-KCR Feedback liegenden Feedbackformen, bewerteten Lerner, welche KR-KH Feedback erhielten, ihre Treatmentleistung zum 1. Messzeitpunkt

erwartungskonform nicht so positiv wie die ITF Gruppe, aber positiver als die KR-KCR Feedbackgruppe und die KR-KM Feedbackgruppe. Allerdings war keiner der beobachteten Unterschiede zum 1. Messzeitpunkt statistisch signifikant. Interessant in diesem Zusammenhang erscheint die motivationale Einschätzung der KR-KM Feedbackgruppe zu Messzeitpunkt 1. Obwohl Schüler der KR-KM Feedbackgruppe eine geringfügig bessere Treatmentleistung zum 1. Messzeitpunkt im Vergleich zur KR-KCR Feedbackgruppe zeigten, berichten sie über ein geringeres Kompetenzerleben und eine geringere intrinsische Motivation als die Schüler, welche KR-KCR Feedback erhielten. Dieser Unterschied wurde nach Ausparialisierung der Treatmentleistung noch größer. Auch wenn dieses Ergebnis statistisch nicht abgesichert werden konnte, weist es möglicherweise darauf hin, dass Schüler der KR-KM Feedbackgruppe ihre Leistungen deswegen negativer bewerteten, weil sie der Annahme waren, dass sie die rechnergestützten Aufgaben durch die Bereitstellung korrekturrelevanter Informationen besser hätten lösen müssen. Da sie die bereitgestellten Informationen nicht in dem intendierten Sinne zu einer Korrektur ihrer fehlerhaft bearbeiteten Aufgaben nutzen konnten, aber deren Korrekturpotenzial erkannten, waren sie eventuell unzufriedener mit ihren Leistungen als Schüler der KR-KCR Feedbackbedingung. Diese Interpretation bedarf allerdings weiterer empirischer Forschung und stellt von daher nur eine mögliche Sichtweise der erhaltenen Daten dar.

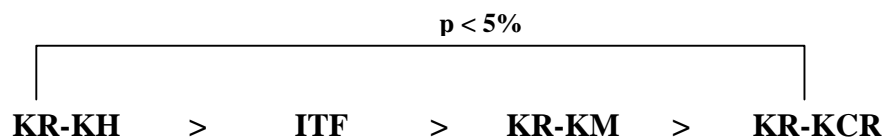
Zusammenfassend ließ sich folgende Wirksamkeitsreihenfolge der vier Feedbackformen auf die in der Studie erfassten motivationalen Parameter zum 1. Messzeitpunkt finden, wobei noch einmal betont wird, dass die Unterschiede nicht signifikant waren:

ITF > KR-KH > KR-KCR > KR-KM

Zum 2. Messzeitpunkt konnte ein statistisch signifikanter Unterschied in den motivationalen Parametern zwischen der KR-KH Feedbackgruppe und der KR-KCR Feedbackgruppe auf dem 5%-Niveau gefunden werden. Dieser Unterschied ist darauf zurückzuführen, dass die KR-KH Feedbackgruppe zum 2. Messzeitpunkt im Vergleich zum 1. Messzeitpunkt ein erhöhtes Kompetenzerleben und eine erhöhte intrinsische Motivation berichtet, im Gegensatz zur KR-KCR Feedbackgruppe, welche ihre Leistungen vergleichbar wie zu Messzeitpunkt 1 bewertete. Auch die KR-KM Feedbackgruppe berichtet im Vergleich zum 1. Messzeitpunkt ein deutlich erhöhtes Kompetenzerleben und eine erhöhte intrinsische Motivation. Dieses Er-

gebnis ist sicherlich über den Leistungszuwachs erklärbar, welchen die KR-KM Feedbackgruppe vom 1. zum 2. Messzeitpunkt erfuhr.

Im Gegensatz zu den beiden Feedbackbedingungen KR-KH und KR-KM nahmen die Schüler der ITF Bedingung ihre Leistungen nur geringfügig positiver wahr als zu Messzeitpunkt 1. Da die Schüler der ITF Bedingung aufgrund des vermuteten Deckeneffektes zu Messzeitpunkt 1 nur einen geringen Leistungszuwachs zum 2. Messzeitpunkt zeigten, ist es nicht verwunderlich, dass ihr Kompetenzerleben zu Messzeitpunkt 2 demjenigen zu Messzeitpunkt 1 vergleichbar ist. Zusammenfassend ließ sich folgende Wirksamkeitsreihenfolge der vier Feedbackformen auf die motivationalen Parameter zum 2. Messzeitpunkt finden, wobei noch einmal betont wird, dass nur der Unterschied zwischen KR-KH Feedback und KR-KCR Feedback statistisch signifikant war:



Im Kapitel 2.5 wurde bereits darauf hingewiesen, dass nur wenige Feedbackstudien vorliegen, welche die motivationale Wirksamkeit von informativem Feedback zum Untersuchungsgegenstand erklärt haben. In einer dieser wenigen Arbeiten konnte Narciss (1999; 2001; 2004) zeigen, dass sich informatives tutorielles Feedback im Vergleich zu ergebnisorientiertem KR Feedback günstig auf motivationale Variablen, wie das post-aktionale Kompetenzerleben, das Engagement, die Anstrengung und die Persistenz auswirkt. Diese positiven Effekte waren allerdings nur dann zu beobachten, wenn sich die Lerner wiederholt und ernsthaft auf die Bearbeitung der Konzeptlernaufgaben einließen. So fand Narciss (2004), dass insbesondere Personen mit niedriger situationsbezogener Kompetenz-Einschätzung (in der vorliegenden Studie als aufgabenspezifisches Fähigkeitskonzept erfasst) sich seltener auf die Aufgaben- und Feedbackbearbeitung einließen und diese auch häufiger abbrachen. Sie hatten damit seltener Gelegenheit, die Erfahrung zu machen, dass ITF sie bei ihrer Aufgabenbearbeitung unterstützt. Fehlende Erfolgserlebnisse konnten somit auch nicht in einem erhöhten Kompetenzerleben resultieren. Narciss (2004) leitete aus ihren Ergebnissen die Schlussfolgerung ab, dass ITF erst dann seine Wirkung entfalten kann, wenn Vermeidungsstrategien der Lernenden zumindest teilweise kontrolliert werden, beispielsweise über die Implementierung eines festes Zeitrahmens und/oder eines mastery levels. Im Gegensatz zu den Ergebnissen von

Narciss (2004) konnte in der vorliegenden Studie kein Zusammenhang zwischen der situationsbezogenen Kompetenzeinschätzung und dem post-aktionalen Kompetenzerleben gefunden werden. Das heißt, Schüler, welche ein positives aufgabenspezifisches Fähigkeitskonzept im Hinblick auf die Bearbeitung schriftlicher Subtraktionsaufgaben hatten, zeigten kein höheres Kompetenzerleben und keine höhere intrinsische Motivation. So erwies sich in der vorliegenden Studie das aufgabenspezifische Fähigkeitskonzept überhaupt als ein schlechter Prädiktor für die abhängigen Variablen. Da das aufgabenspezifische Fähigkeitskonzept noch vor der Bearbeitung des Vortestes erfragt wurde, fiel es den Schülern möglicherweise schwer, valide Aussagen darüber zu treffen, wie leicht ihnen die Bearbeitung solcher Aufgaben fällt, wie gut sie solche Aufgaben lösen können bzw. für wie begabt sie sich bezüglich der Bearbeitung schriftlicher Subtraktionsaufgaben halten.

Dass die gefundenen Wirksamkeitsunterschiede hinsichtlich der post-aktionalen motivationalen Parameter zwischen den Feedbackbedingungen geringer als erwartet ausfielen, kann möglicherweise über die Implementierung des mastery levels unter alle vier Feedbackbedingungen erklärt werden. So konnte in vielen Studien zum zielerreichenden Lernen gezeigt werden, dass allein schon die Implementierung eines mastery levels sich dahingehend motivierend auf die Lerner auswirkte, dass diese ein höheres Motivationslevel und eine positivere Fähigkeitswahrnehmung zeigten (Mory, 1996). Infolge der Implementierung eines mastery levels unter alle vier Feedbackbedingungen wurde zudem möglicherweise eine interne Attribution von Erfolgen auch bei der KR-KM und KR-KH Feedbackbedingung unterstützt. Da die präsentierten Aufgaben eines Aufgabentyps sich nur in ihren Ziffernkombinationen unterschieden, war es den Lernenden der KR-KM und KR-KH Feedbackbedingung möglich, die bereitgestellten korrekturrelevanten Feedbackinformationen unmittelbar auf die Bearbeitung der nächsten Aufgabe anzuwenden. Lösten sie diese Aufgabe erfolgreich, attribuierten sie diesen Erfolg möglicherweise internal, da sie die Erfahrung gemacht haben, diese Aufgabe aufgrund eigener Anstrengung/Fähigkeit unter Nutzung der korrekturrelevanten Informationen gelöst zu haben. Ob aufgrund der bereitgestellten korrekturrelevanten Informationen und deren Anwendung an der gleichen oder der nächsten Aufgabe bei Erfolg ein internaler Attributionsprozess unterstützt wird, ist in einer weiterführenden Studie zu klären, in welcher explizit erfasst werden sollte, wie die Lernenden Erfolg bzw. Misserfolg attribuieren.

Eine weitere Erklärung für die moderaten Wirksamkeitsunterschiede hinsichtlich der post-aktionalen motivationalen Parameter zwischen den Feedbackgruppen bietet möglicher-

weise wieder die Annahme eines Deckeneffektes. So zeigt eine Betrachtung der Häufigkeitsverteilungen der Einzelitems, dass für alle sieben motivationalen Items der zweiten Studie mit Ausnahme des Items „In dieser halben Stunde bin ich nicht so vorwärts gekommen wie ich wollte.“ weniger als 20% aller Schüler negative Einschätzungen vornahmen, so dass das Finden von signifikanten Unterschieden aufgrund der geringen Variabilität in den Aussagen erschwert worden sein könnte. Eine Erklärung für die geringe Bandbreite der Itembeantwortungen könnte in dem in der Literatur berichteten Befund gesehen werden, dass vor allem Grundschüler häufig übertrieben optimistische Selbsteinschätzungen ihrer Leistungen vornehmen (Helmke, 1991).

6 ZUSAMMENFASSUNG UND ABSCHLIEßENDE DISKUSSION

Informatives Feedback wird als ein bedeutendes instruktionales Mittel gesehen, um Lernende in ihrem Lernprozess zu unterstützen. Dabei wird informatives Feedback als jede Art von Rückmeldung definiert, die dem Lernenden im Anschluss an eine Aufgabebearbeitung durch die Lernumgebung bereitgestellt wird und den Lernenden über die Qualität seiner Antwort informiert (Mory, 1996). Obwohl Konsens im Hinblick auf die zentrale Bedeutung von informativem Feedback für das Lernen besteht, hat dieses Bewusstsein bisher wenig Resonanz in der Gestaltung multimedialer Lernumgebungen gefunden. Informatives Feedback in multimedialen Lernumgebungen gestaltet sich häufig als Richtig-/Falsch Rückmeldung bzw. als Bekanntgabe der richtigen Lösung. Werden Hilfen von dem Programm im Anschluss an eine fehlerhafte Aufgabebearbeitung angeboten, beinhalten diese häufig einen Verweis auf die relevante Lehrtextstelle oder die Bereitstellung eines ausgearbeiteten Lösungsbeispiels. Hinweise auf den Ort und die Art des Fehlers werden selten angeboten. Prototypisch für diese Form der Rückmeldung ist die Feedbackgestaltung mathematischer Grundschulsoftware. Eine Analyse gängiger kommerzieller Mathematiklernsoftware für den Grundschulbereich hat gezeigt, dass dem Lernenden in aller Regel nach dem ersten fehlerhaften Lösungsversuch mitgeteilt wird, dass seine Lösung noch einen Fehler beinhaltet, verbunden mit der Aufforderung, es noch einmal zu versuchen. Nach einem zweiten fehlerhaften Lösungsversuch wird dem Lernenden die richtige Lösung mitgeteilt und er bekommt die nächste Aufgabe präsentiert. Die in solcher Mathematiklernsoftware bereitgestellten Lösungsbeispiele erfordern von dem Lernenden neben einer hohen Lesekompetenz auch den Einsatz bestimmter (meta-)kognitiver Strategien bei der Verarbeitung der angebotenen Lösungsbeispiele (z.B. Generierung von Selbsterklärungen). Auf die aus einer oberflächlichen Verarbeitung ausgearbeiteter Lösungsbeispiele resultierende geringe Lerneffektivität solcher Lösungsbeispiele wurde bereits in den Kapiteln 2.3.1.2 und 2.3.2.2 verwiesen (Renkl, 2000). Aufgrund dieser Diskrepanz zwischen dem Wissen um die zentrale Bedeutung von Feedback für den Lernprozess und dessen suboptimaler Umsetzung in kommerzieller Lernsoftware bestand das primäre Ziel dieser Arbeit darin, theoriegeleitet informatives Feedback für den Bereich der schriftlichen Subtraktion zu entwickeln und hinsichtlich seiner Lern- und motivationalen Wirksamkeit zu evaluieren.

Ein Überblick über die Ergebnisse der Feedbackforschung zeigt allerdings, dass der Forschungsbereich zu Feedback zwar eine Vielzahl an Studien mit unterschiedlichsten Befunden aufweist, ein allgemeines Feedbackmodell, welches diese Ergebnisse integrierend darstellt und erklärt, bisher aber gefehlt hat. Aufgrund des Fehlens eines allgemeinen Feedbackmodells verwirrt die isolierte Betrachtung vieler, häufig auch widersprüchlicher Feedbackergebnisse eher, als dass sie klare Aussagen hinsichtlich der Gestaltung von Feedback erlaubt. Von daher erforderte eine systematische Entwicklung von informativem Feedback unter Nutzung vorliegender Erkenntnisse der Feedbackforschung von einem Feedbackdesigner bisher ein sehr breites und gleichzeitig sehr spezifisches Detailwissen, um der Multidimensionalität von Feedback gerecht zu werden. Mit dem heuristischen Modell zur Analyse und Untersuchung von Feedbackeffekten von Narciss (2004) wird erstmals ein Modell vorgelegt, welches die Multidimensionalität von Feedback angemessen berücksichtigt und die Bedeutung lerner- und systemseitiger Faktoren, welche die Feedbackeffektivität beeinflussen, herausarbeitet (siehe Kapitel 2.3). Auf der Grundlage dieses heuristischen Modells war es möglich, konkrete Gestaltungs- und Evaluationsprinzipien zu formulieren, welche eine theoriegeleitete Entwicklung und Evaluation von informativem tutoriellem Feedback ermöglichen (Narciss, 2004; Narciss & Huth, 2004). Diese Gestaltungs- und Evaluationsprinzipien wurden im Rahmen der vorliegenden Arbeit genutzt, um fehlerspezifisches ITF für die schriftliche Subtraktion systematisch zu entwickeln und zu evaluieren.

ITF zeichnet sich hierbei dadurch aus, dass bei Fehlern korrekturrelevante Informationen angeboten werden, ohne dem Lernenden die korrekte Lösung unmittelbar zu präsentieren. Der Lerner sieht sich auf diese Weise dazu aufgefordert, die Aufgabe unter Anwendung dieser strategischen Informationen bei einem sich anschließenden Lösungsversuch eigenständig zu lösen. Um die Lern- und motivationale Wirksamkeit von fehlerspezifischem ITF für die schriftliche Subtraktion zu untersuchen, wurde zu diesem Zweck basierend auf den Gestaltungsprinzipien systematisch fehlerspezifisches ITF für die schriftliche Subtraktion entwickelt und computergestützt implementiert. Hierfür wurde in einem ersten Schritt auf der Grundlage einer kognitiven Anforderungsanalyse für die schriftliche Subtraktion eine Anforderungsstruktur in Form eines Hasse-Diagramms erstellt. Diese Anforderungsstruktur verfolgt das Ziel, möglichst vollständig Aufgabentypen im Zahlenraum bis 1000 zu beschreiben, welche in Abhängigkeit der Art und Anzahl von Aufgabenanforderungen und deren Kombinationsmöglichkeiten unterschiedliche Anforderungen an den Lernenden stellen. Auf der Grundlage

dieser Anforderungsstruktur wurde in einem nächsten Schritt eine empirische Vorstudie durchgeführt, um mittels einer Fehleranalyse systematische Fehler für die in dieser Vorstudie enthaltenen sechs Aufgabentypen zu identifizieren. Systematische Fehler beschreiben hierbei Fehler, die auf eine fehlerhafte oder unvollständige Regelanwendung zurückgeführt werden können. Das Wissen über diese systematischen Fehler wurde in einem nächsten Schritt genutzt, um korrekturrelevante Informationen zu identifizieren, die dann entsprechend der konkreten Gestaltungsprinzipien für die Präsentation dieser Informationen in den ITF-Algorithmus implementiert wurden.

In zwei experimentellen Studien mit Viertklässlern verschiedener Dresdner Grundschulen und Sechstklässlern verschiedener Dresdner Förderschulen wurde dieses computergestützte ITF evaluiert. In der ersten Studie erfolgte diese Evaluation gegen eine Feedbackbedingung, welche das üblicherweise in mathematischer Lernsoftware realisierte KR-KCR Feedback präsentierte. In der zweiten Studie wurden zwei zusätzliche Feedbackbedingungen – Knowledge about mistake (KR-KM) und Knowledge on how to proceed (KR-KH) - aufgenommen, die in ihrem Informationsgehalt zwischen dem ITF und dem KR-KCR Feedbackalgorithmus anzusiedeln sind und der zweiten bzw. dritten Feedbackstufe des ITF Algorithmus entsprechen. Ziel dieser zweiten Studie war zum einen, isolierte Wirksamkeitsaussagen zu den einzelnen Feedbackstufen des ITFs zu erhalten. Zum anderen wurde mit dieser zweiten Studie der Frage nachgegangen, ob vergleichbare Lern- und Motivationseffekte des ITFs auch auf eine ökonomischere Vorgehensweise, d.h. über einen zweistufigen Feedbackalgorithmus, welcher korrekturrelevante Informationen beinhaltet, zu erhalten sind. Zur Überprüfung der Forschungshypothesen wurde in beiden Studien ein Vortest-Nachtest-Vergleichsgruppen-Design realisiert.

Die Ergebnisse dieser Studien bestätigen die postulierte Lern- und motivationale Wirksamkeit des entwickelten fehlerspezifischen ITFs für die schriftliche Subtraktion. Besonders im Vergleich zu der wenig informativen KR-KCR Feedbackbedingung konnten in beiden Studien signifikante Wirksamkeitsunterschiede nachgewiesen werden. So zeigten die Schüler unter der ITF Bedingung sowohl während des Treatments als auch im Nachtest signifikant bessere Leistungen. Diese besseren Lernleistungen kamen in einem positiven Kompetenzerleben und einer erhöhten intrinsischen Motivation zum Ausdruck. Auch im Vergleich zu den komplexeren Feedbackformen Knowledge about mistake (KR-KM) und Knowledge on how to proceed (KR-KH) konnten für die Schüler der ITF Bedingung bessere Leistungen beob-

achtet werden, auch wenn diese Unterschiede nur partiell signifikant wurden. Dass das bereitgestellte Feedback hoch wirksam war, belegt vor allem die kurze Interventionsphase von einer halben Stunde. So zeigten die Schüler der ITF Bedingung zum zweiten Messzeitpunkt der zweiten Studie keinen bedeutsamen Lernfortschritt mehr, da bereits zum ersten Messzeitpunkt ein hohes Leistungsniveau erreicht war.

Die Ergebnisse der vorliegenden Untersuchungen zeigen, dass es möglich ist, auf der Grundlage des von Narciss (2004) vorgestellten Feedbackmodells und der aus diesem Feedbackmodell abgeleiteten Gestaltungsprinzipien lern- und motivationsförderliches ITF zu entwickeln. So liefern die Ergebnisse beider Studien neben den Arbeiten von Narciss (1999, 2001, 2004) erste empirische Belege für die Annahme, dass informatives Feedback nicht nur den Lernprozess unterstützt, sondern auch motivationsförderlich wirkt.

Die Ergebnisse der vorliegenden Studien unterstreichen damit die Bedeutung einer kognitiven Anforderungs- und Fehleranalyse für die Feedbackeffektivität. So konnten in der vorliegenden Arbeit auf der Grundlage einer kognitiven Anforderungs- und Fehleranalyse korrekturrelevante Feedbackinformationen systematisch identifiziert werden. Im Gegensatz zu vielen Feedbackstudien, in denen die Feedbackinhalte eher intuitiv ausgewählt wurden (Kulhavy et al., 1985; Kulhavy & Stock, 1989), gewährleistet ein solches Vorgehen, dass dem Lerner nur Informationen bereitgestellt werden, die für eine Fehlerkorrektur notwendige Wissensselemente beinhalten.

Neben der systematischen Auswahl der Feedbackinhalte spielt aber auch die Art und Weise der Präsentation der als relevant identifizierten Feedbackinformationen eine zentrale Rolle für die Lern- und motivationale Wirksamkeit des ITFs. Die Ergebnisse der vorliegenden Studien unterstreichen somit auch die Forderung, dass Feedbackinhalte dem Lerner so präsentiert werden sollten, dass eine aktive Auseinandersetzung mit diesen angeregt und eine kognitive Überforderung des Lernenden vermieden wird. Für die Ableitung dieser Forderung wurde hierbei unter anderem auf Erkenntnisse aus der Forschung zu menschlichen Tutoren und intelligenten tutoriellen Systemen zurückgegriffen. So wird die empirisch belegte Effektivität menschlicher Tutoren (Chi et al., 2001; Merrill et al., 1995) darüber erklärt, dass diese erst intervenieren und Hilfe anbieten, wenn Lerner Probleme beim Bearbeiten einer Aufgabe zeigen (Wood & Wood, 1999). In diesem Fall bieten Tutoren adaptiv sequentielle Hilfe entsprechend den Bedürfnissen des Lernenden an, d.h. der Tutor gibt progressiv immer detailliertere und explizitere Hilfe, bis der Lerner in der Lage ist, das Problem selbständig zu lösen

(Graesser, Person & Magliano, 1995). Dass ein solches Vorgehen sinnvoll ist, belegen nicht nur die positiven Erkenntnisse zum menschlichen Tutoring, sondern zeigt auch die negative Befundlage der Feedbackforschung, in deren Studien die Lernenden sich häufig nicht dazu aufgefordert sahen, die richtige Lösung eigenständig, in Folge einer aktiven Auseinandersetzung mit den präsentierten Informationen, zu finden. So waren Lernenden die Informationen häufig schon vor der eigentlichen Aufgabenbearbeitung zugänglich (presearch availability, Anderson et al., 1971) bzw. wurden diese Informationen gemeinsam mit der korrekten Lösung präsentiert (Kulhavy et al., 1985). Eine Auseinandersetzung mit den bereitgestellten elaborierten Informationen war nicht gefordert, wenn für eine erfolgreiche Bearbeitung des Nachtstes nur das Wissen der korrekten Lösung erforderlich war.

Für das in der vorliegenden Arbeit entwickelte ITF wurde diese aktive Auseinandersetzung mit den präsentierten Feedbackinformationen darüber induziert, dass dem Lernenden die korrekturrelevanten Informationen ohne der richtigen Lösung präsentiert wurden, so dass der Lerner diese Informationen nutzen musste, um die Aufgabe in einem nächsten Lösungsversuch korrekt zu lösen. Neben der intendierten aktiven Auseinandersetzung birgt das sequentielle Angebot korrekturrelevanter Informationen darüber hinaus den Vorteil, dass eine Passung zwischen dem Könnensniveau des Lernenden und dem Feedbackinhalt hergestellt werden kann, ohne dabei auf eine aufwendige Lernermodellierung zurückgreifen zu müssen (z.B. Model-Tracing, Graesser, VanLehn, Rosé, Jordan & Harter, 2001). Je nach Fähigkeitsniveau des Lernenden werden diesem durch den abgestuften Algorithmus verschiedene Feedbacklevel mit zunehmendem Informationsgehalt angeboten. So benötigen Lerner mit einem hohen Fähigkeitsniveau häufig nur KR Feedback, um die Aufgabe in einem zweiten Lösungsversuch korrekt zu lösen. Für Lerner mit einem mittleren Fähigkeitsniveau genügt eventuell ein Hinweis auf die korrekte Lösungsprozedur, ohne diese an der konkreten Aufgabe zu demonstrieren (KR-KM Feedback). Lerner mit einem geringen Fähigkeitsniveau benötigen hingegen die Demonstration der korrekten Lösungsprozedur an der konkreten Aufgabe (KR-KH Feedback).

Ein weiterer Vorteil des adaptiven ITF-Algorithmus kann darin gesehen werden, dass dem Lernenden sequentiell nur solche Informationen präsentiert werden, die für eine korrekte Aufgabenlösung als relevant identifiziert worden waren. Die Vermittlung redundanter Informationen - z.B. von Informationen zu Teilfertigkeiten, die der Lerner bereits beherrscht - und damit die Gefahr des „cognitive overload“ werden durch das Angebot fehlerspezifischer Kor-

rektorhinweise vermieden. Dieses Potenzial des fehlerspezifischen ITFs sollte vor allem dann relevant werden, wenn dem Lerner komplexere Aufgabentypen, d.h. Aufgabentypen, die eine Kombination verschiedener Anforderungen beinhalten, präsentiert werden (siehe Hasse-Diagramm S. 63). Eine Diagnostik der nicht beherrschten Teilprozedur sollte dazu führen, dass korrekturrelevante Informationen nur für die Anforderung bereitgestellt werden, die vom Lerner nicht sicher beherrscht wird.

Neben seiner Adaptivität zeichnet sich das fehlerspezifische ITF durch die Implementierung eines mastery levels aus, welches gleichfalls das Ziel verfolgt, eine aktive Auseinandersetzung mit den präsentierten Feedbackinformationen anzuregen. Diese aktive Auseinandersetzung sollte darüber induziert werden, dass der Lerner realisiert, dass das Ende der Lehrereinheit nur zu erreichen ist, wenn die Aufgaben eigenständig korrekt bearbeitet werden. Dieses Erkenntnis wiederum motiviere den Lerner, sich intensiver mit den elaborierten Hinweisen auseinander zu setzen (Jacobs, 2002). Die Implementierung eines mastery levels unter allen Feedbackbedingungen wurde bereits als Grund diskutiert, dass die Unterschiede in den abhängigen Variablen zwischen dem ITF und den Feedbackformen KR-KM und KR-KH nicht immer signifikant wurden.

Ein weitere Erklärung für die nur partiell signifikanten Unterschiede zwischen dem ITF und dem KR-KM und KR-KH Feedback kann darin gesehen werden, dass es sich bei dem KR-KM und KR-KH Feedback im Vergleich zu klassischen Feedbackstudien um relativ komplexe Feedbackformen handelt. So wurde in klassischen Feedbackstudien in der Regel elaboriertes Feedback mit KCR bereits nach dem ersten Lösungsversuch angeboten, wohingegen das KR-KM und KR-KH Feedback der vorliegenden Studie einen zweistufigen Feedbackalgorithmus repräsentiert. Es ist zu vermuten, dass eindeutigere Ergebnisse in erwarteter Richtung zu beobachten gewesen wären, wenn den Lernern komplexere Aufgabentypen, d.h. Aufgaben mit höheren Anforderungen, präsentiert worden wären, so dass der bereits zum ersten Messzeitpunkt beobachtete Deckeneffekt hätte vermieden werden können.

Die in den vorliegenden Studien gefundenen Ergebnisse bestätigen im wesentlichen Befunde der Feedbackforschung, welche gezeigt haben, dass eine wiederholte Aufgabenbearbeitung unter Bereitstellung korrekturrelevanter Informationen (Graesser, Person & Magliano, 1995; Nagata & Swisher, 1995) und die Implementierung eines mastery levels (Collins, Carline & Gersten, 1987) wirkungsvolle Maßnahmen darstellen, um die Feedbackeffektivität zu erhöhen.

6.1 Grundlagen- und anwendungsbezogene Implikationen

Die detaillierte Darstellung der Forschungsliteratur zu Feedback und damit einhergehend das Aufzeigen wesentlicher Faktoren, welche die Wirksamkeit von Feedback beeinflussen, sollte verdeutlicht haben, dass eine effektive Feedbackkonstruktion theoriegeleitet erfolgen sollte. Für die Konstruktion von Feedback kann hierbei auf das heuristische Modell zur Analyse und Untersuchung von Feedbackeffekten nach Narciss (2004) rekurriert werden. Mit diesem Modell lassen sich zum einen system- und lernerseitige Faktoren beschreiben, welche bei der Feedbackentwicklung beachtet werden sollten. Zum anderen diente dieses Modell zur Ableitung konkreter Richtlinien, welche im Sinne von Handlungsanleitungen die theoriegeleitete Entwicklung und Evaluation von Feedback unterstützen. Diese Richtlinien beinhalten neben Empfehlungen zur Selektion relevanter Feedbackinhalte und deren Präsentation auch Empfehlungen hinsichtlich der Evaluation des entwickelten ITFs.

Designer von computergestützten Lernumgebungen sollte das heuristische Feedbackmodell zum einen für die Komplexität der die Feedbackeffektivität beeinflussenden Faktoren sensibilisieren. Solche Faktoren sind z. B. die aus den Lehr-Lernzielen, den Lehrinhalten und den Lernaufgaben resultierenden Anforderungen und die (meta-) kognitiven und motivationalen Lernerdispositionen. Diese system- und lernerseitigen Faktoren sollten wiederum die Entscheidung über Inhalt und Form des Feedbacks determinieren. Zum anderen sollte das Modell und die daraus abgeleiteten Gestaltungsprinzipien dazu dienen, der Multidimensionalität von Feedback bei dessen Konstruktion gerecht zu werden, indem es den Designer dabei unterstützt, bewusste Entscheidungen hinsichtlich folgender Aspekte zu treffen: dem Feedbackinhalt, dem Zeitpunkt der Feedbackpräsentation, der Lerner- oder Systemkontrolle der Feedbackpräsentation, den zu verwendenden Codes und Modi der Feedbackpräsentation, der sukzessiven oder einmaligen Präsentation der Feedbackinhalte, der Implementierung eines mastery levels.

Vor allem vor dem Hintergrund, dass computergestütztes Feedback zum gegenwärtigen Zeitpunkt selten über die Verwendung von Multiple-Choice Aufgaben hinausgeht, in deren Anschluss die fehlerhafte durch die richtige Antwort ersetzt wird, bieten das heuristische Feedbackmodell und die daraus abgeleiteten Gestaltungsprinzipien die Chance, informatives tutorielles Feedback zu konstruieren, welches den Lerner in seinem individuellen und selbständigen Lernprozess effektiv unterstützt und einen aktiven Wissenserwerb ermöglicht.

Neben diesen Implikationen, welche aus dem Feedbackmodell und den Gestaltungsprinzipien für den zukünftigen Feedbackdesigner resultieren, ergeben sich aus den Erkenntnissen und empirischen Befunden der vorliegenden Studien auch anwendungsbezogene Implikationen. So lassen sich hinsichtlich des Einsatzes von ITF aus den vorliegenden Erkenntnissen eine Reihe praktischer Maßnahmen ableiten. Eine zentrale Forderung, welche aus den Ergebnissen der vorliegenden Studien resultiert, besteht darin, bei der Konzeption und Entwicklung mathematischer Lern- und Übungssoftware die Erkenntnisse zur Konstruktion fehlerspezifischen ITFs zu nutzen. So zeigen die Ergebnisse der Studien, dass über die Bereitstellung fehlerspezifischen ITFs bereits im Rahmen einer kurzfristigen Intervention (30 Minuten) Fertigkeiten in der schriftlichen Subtraktion effektiv gefördert werden können. Vor allem bei Lernern mit großen Problemen in der schriftlichen Subtraktion kann über das Angebot fehlerspezifischen ITFs häufigen Misserfolgserlebnissen vorgebeugt und negative Emotionen wie Frustration und Scham vermieden werden. Lern- und Übungssoftware, welche fehlerspezifisches ITF nutzt, kann so beim individuellen Wissenserwerb zu Erfolgserlebnissen führen, die Freude am Lernen erhöhen und das Interesse am Lerngegenstand wecken. Aufgrund des sequentiellen Angebots der Feedbackinformationen und der darüber erzielten Passung zwischen Kenntnisniveau und Feedbackinhalt können auch Lerner mit unterschiedlichen kognitiven Lernvoraussetzungen von dem fehlerspezifischen ITF profitieren. Allerdings sollte fehlerspezifisches ITF im Sinne des kognitiven Fertigkeitserwerbs (VanLehn, 1996) frühestens zu Beginn der 2. Phase des Fertigkeitserwerbs eingesetzt werden, d. h. wenn die Lernenden bereits deklaratives Wissen, beispielsweise über die Einführung des Subtraktionsalgorithmus anhand ausgearbeiteter Lösungsbeispiele im Unterricht, erworben haben. Vor allem in der 2. Phase des Fertigkeitserwerbs, welche durch das Lösen konkreter Probleme gekennzeichnet ist, können Wissenslücken über das Angebot fehlerspezifischen ITF geschlossen und/oder fehlerhaft erworbenes Wissen korrigiert werden, d.h. sollte sich ITF als besonders effektiv erweisen.

Da fehlerspezifisches ITF bisher nur für die elementaren Aufgabentypen entwickelt wurde, steht in einem nächsten Schritt die Entwicklung fehlerspezifischen ITFs auch für die zweite und dritte Ebene der Anforderungsstruktur zur schriftlichen Subtraktion an (siehe Hasse-Diagramm). Voraussetzung hierfür ist allerdings eine im Vorfeld durchzuführende empirische Validierung der vorliegenden Anforderungsstruktur der schriftlichen Subtraktion. Mit der Entwicklung fehlerspezifischen ITFs auch für die zweite und dritte Anforderungsebene

wird letztendlich eine Übungssoftware vorliegen, deren Aufgabenstruktur die im Hasse-Diagramm dargestellte Anforderungsstruktur der schriftlichen Subtraktion im Zahlenraum bis 1000 widerspiegelt. Bevor der Lerner mit der Bearbeitung konkreter Aufgabentypen beginnt, wird eine Diagnostik des aktuellen Wissenszustandes erfolgen, in welcher dem Lerner Aufgaben aus der Anforderungsstruktur präsentiert werden. Dabei wird im Sinne der split-half Methode (Albert & Lukas, 1999) mit der Präsentation von Aufgaben aus der zweiten Anforderungsebene begonnen. In Abhängigkeit der vom Lerner korrekt bearbeiteten Aufgaben werden diesem als nächstes Aufgaben der ersten bzw. dritten Anforderungsebene präsentiert. Entsprechend der Wissensstrukturtheorie (Doignon & Falmagne, 1985) wird hierbei davon ausgegangen, dass die korrekte Bearbeitung eines Aufgabentyps der zweiten Anforderungsebene die korrekte Bearbeitung der mit diesem Aufgabentyp verbundenen Aufgabentypen der ersten Anforderungsebene impliziert. Nach abgeschlossener Diagnostik kann der Lerner mit der Aufgabenbearbeitung beginnen. Hierbei werden ihm zuerst Aufgaben präsentiert, die von ihm noch nicht beherrscht werden und die geringste Komplexität, d.h. geringsten Anforderungen, aufweisen. Über dieses Vorgehen sollte sowohl einer Über- als auch einer Unterforderung des Lerners vorgebeugt werden. Zudem garantiert ein solches systematisches Vorgehen, dass das fehlerspezifische ITF den Lerner kognitiv nicht überfordert. In einem nächsten Schritt bietet sich die Entwicklung von fehlerspezifischem ITF auch für die drei anderen Grundrechenarten (Addition, Multiplikation, Division) an, so dass letztendlich eine Übungssoftware resultiert, welches fehlerspezifisches ITF für den Bereich der schriftlichen Grundrechenarten anbietet.

6.2 Forschungsdesiderata

In der vorliegenden Arbeit wurde fehlerspezifisches ITF für den Gegenstandsbereich der schriftlichen Subtraktion entwickelt, indem in Anlehnung an die Gestaltungsprinzipien von Narciss (2004; Narciss & Huth, 2004) fehlerspezifische Informationen auf der Grundlage einer Anforderungs- und Fehleranalyse identifiziert wurden. Die Durchführung solcher Fehleranalysen empfiehlt sich dabei besonders für gut-strukturierte Wissensgebiete wie das Themengebiet der schriftlichen Subtraktion, welche eine detaillierte Anforderungs- und Fehlerbeschreibung gestatten und somit eine Diagnose nichtbeherrschter bzw. fehlerhaft beherrschter Teilfertigkeiten ermöglichen (Aleven, Stahl, Schworm, Fischer & Wallace, 2003). Von daher wurde solch fehlerspezifisches Feedback im Rahmen der Forschung zu intelligenten tutoriel-

len Systemen auch primär für gut-strukturierte Wissensdomänen entwickelt, z.B. für formale Themengebiete aus der Mathematik, der Informatik und der Physik.

Die Entwicklung von ITF auf der Basis von Fehleranalysen für schlecht-strukturierte Wissensgebiete scheint vor dem Hintergrund von Kosten-Nutzen-Erwägungen sehr aufwendig und wenig erfolgversprechend. So zeichnen sich schlecht-strukturierte Aufgaben dadurch aus, dass der Zielzustand nicht präzise beschrieben werden kann, verschiedene Lösungsstrategien zur korrekten Lösung führen können, das zu repräsentierende Wissen zu komplex ist, um es in einem Expertenmodell abzubilden bzw. die Lernerantwort zu komplex ist, um diese angemessen auswerten zu können. Eine Identifikation systematischer Fehler, in deren Folge fehlerspezifisches ITF bereitgestellt werden kann, scheint von daher für schlecht-strukturierte Aufgaben keine erfolgversprechende Strategie darzustellen. So plädiert auch Connelly (2001) im Hinblick auf Kosten-Nutzen-Erwägungen gegen ein antwort-spezifisches Feedback in Form fehlerspezifischer Hinweise oder kontextsensitiver Erklärungen, da die Erstellung solcher Feedbackinformationen anhand von Wissensrepräsentationen sehr zeit- und kostenintensiv wäre und Erfolg nicht garantiert werden kann.

Informatives tutorielles Feedback für schlecht-strukturierte Aufgaben, für die eine Fehleranalyse nicht möglich oder zu aufwendig wäre, kann andere Strategien nutzen, um den Lerner bei seinem Aufgabenlösungsprozess zu unterstützen. Eine Möglichkeit, wie sie in der Studie von Reimann und Narciss (2003) realisiert wurde, stellt das Angebot von Fallbeispielen dar, die dem Lernenden bei einer fehlerhaften Aufgabenbearbeitung angeboten werden. Auch in dieser Studie hatten die Lernenden im Sinne des ITF die Möglichkeit, sich intensiv mit dem Fallbeispiel auseinander zu setzen, um in einem zweiten Lösungsversuch die Aufgabe doch noch korrekt zu lösen. Eine zweite ökonomische Feedbackstrategie für schlecht-strukturierte Aufgaben stellt das Angebot von Musterlösungen dar. Über den Vergleich solcher Musterlösungen mit der eigenen Aufgabenlösung können „self-monitoring“ Strategien angeregt werden, die den Lerner dazu befähigen, individuelle Fehler zu entdecken und einen Revisionsprozess einzuleiten. Auf die Gefahr einer oberflächlichen Verarbeitung solcher Musterlösungen aufgrund suboptimaler (meta)-kognitiver Strategien wurde bereits im Kapitel 2.3.2.2 ausführlich hingewiesen. Eine dritte Möglichkeit ITF zu implementieren, welches strategische Informationen anbietet, die nicht auf einer im Vorfeld durchgeführten Fehleranalyse basieren, repräsentieren allgemeine, antwort-unspezifische Fragetechniken. Solche antwort-unspezifischen Fragen sollten den Lerner im Sinne von Prompts auf wichtige (meta)-

kognitive und heuristische Strategien hinweisen, die für eine erfolgreiche Bearbeitung der Aufgabe eine wichtige Voraussetzung darstellen. In einer der wenigen Untersuchungen zu schlecht-strukturierten Aufgaben nutzte Connelly (2001) ein solches Feedback, welches er als minimalistisches „state-based“ Feedback im Gegensatz zu einem „knowledge-based“ Feedback bezeichnete. Das in seinem Belvedere-System implementierte „state-based“ Feedback stellte dem Lerner domainunabhängige Hinweise zu der vom Lerner erstellten Argumentationsstruktur bereit, ohne dabei spezifische Inhalte der Statements zu tangieren, da diese eine Wissensrepräsentation und eine Lernermodellierung erforderlich gemacht hätten.

Ob diese Formen informativen tutoriellen Feedbacks, welche auf keiner Fehleranalyse basieren und von daher auch keine fehlerspezifischen Feedbackinformationen bereitstellen, den Lerner bei seiner Aufgabenbearbeitung unterstützen, bedarf allerdings weiterer Forschungsaktivitäten. So liegen zum momentanen Zeitpunkt wenig Arbeiten vor, welche die Bereitstellung von informativen Feedback für schlecht-strukturierte Aufgaben gezielt untersucht haben (Alevén & Ashley, 1995; Connelly, 2001; Reimann & Narciss, 2003).

Weiterer Forschungsbedarf besteht hinsichtlich der von Narciss (2004) postulierten motivierenden Funktionen von ITF. So besitzt ITF, welches strategische Informationen anbietet, ohne gleichzeitig die korrekte Lösung zu präsentieren, nach Narciss (2004) ein hohes Motivationspotenzial, da es dem Lerner Gelegenheiten gibt, Lernerfolge zu erleben, die intern attribuiert werden können. Dass Lernerfolge in einem erhöhtem Kompetenzerleben und einer erhöhten intrinsischen Motivation resultieren, konnte durch die Ergebnisse der vorliegenden Studien bestätigt werden. Da aber keine Erfolgs- und Misserfolgsattributionen im Anschluss an die Aufgabenbearbeitung erhoben wurden, können keine Aussagen darüber getroffen werden, ob über die unmittelbare erfolgreiche Anwendung der bereitgestellten Informationen motivationsgünstige Attributionsprozesse unterstützt wurden. Dies sollte Forschungsthema weiterführender Studien sein. Darüber hinaus sollte untersucht werden, wie sich ein erhöhtes Kompetenzerleben und eine erhöhte intrinsische Motivation auf die Anstrengung und Ausdauer bei der Bearbeitung nachfolgender Aufgaben auswirkt.

Vor allem im Hinblick auf die vorliegenden Ergebnisse der Anforderungsanalyse zum Themengebiet der schriftlichen Subtraktion steht in einem nächsten Schritt eine empirische Validierung der theoretisch gewonnenen Anforderungsstruktur an. Im Falle einer erfolgreichen empirischen Validierung sollte fehlerspezifisches ITF für alle Aufgabentypen systematisch entwickelt werden.

Da Evaluationsergebnisse zu den in Kapitel 2 dargestellten Gestaltungs- und Evaluationsprinzipien bisher nur im Rahmen eines experimentellen Konzeptlernparadigmas (Narciss, 1999, 2001, 2004) und zum Themenbereich der schriftlichen Subtraktion vorliegen, welche beide gut-strukturierte Wissensdomänen repräsentieren, sollten die Gestaltungs- und Evaluationsprinzipien auch auf andere gut- bzw. weniger gut-strukturierte Wissensbereiche angewandt und überprüft werden.

7 VERZEICHNISSE

7.1 Literaturverzeichnis

- Albacete, P. & VanLehn, K. (2000). Evaluating the effectiveness of a cognitive tutor for fundamental physics concepts [Electronic version]. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 22. Mahwah, NJ: Erlbaum.
- Albert, D. & Lukas, J. (1999). *Knowledge Spaces: Theories, Empirical Research, Applications*. Mahwah, NJ: Erlbaum.
- Aleven, V. & Ashley K. D. (1995). Using a well-structured model to teach in an Ill-structured domain [Electronic version]. *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*, 419-424. Mahwah, NJ: Erlbaum.
- Aleven, V. & Koedinger, K. R. (2000). Limitations of student control: Do students know when they need help [Electronic version]? In C. F. G. Gauthier & K. VanLehn (Eds.), *Proceedings of the 5th International Conference on Intelligent Tutoring Systems* (pp. 292 – 303). Berlin: Springer.
- Aleven V. & Koedinger, K. R. (2001). Investigations into help seeking and learning with a cognitive tutor [Electronic version]. In R. Luckin (Ed.), *Papers of the AIED-2001 Workshop on Help Provision and Help Seeking in Interactive Learning Environments* (pp. 47-58).
- Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research*, 73, 277-320.
- Anderson, R. C., Kulhavy, R. W. & Andre, T. (1971). Feedback procedures in programmed instruction. *Journal of Educational Psychology*, 62 (2), 148-156.
- Azevedo, R. & Bernard, R.M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research*, 13(2), 111-127.
- Baddeley, A. D. (1986). *Working memory*. London: Oxford University Press.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84 (2), 191-215.
- Bangert-Drowns, R. L., Kulik, C., Kulik, J. A. & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61, 213-238.

- Ben-Zeev, T. (1995). The nature and origin of rational errors in arithmetic thinking: Induction from examples and prior knowledge. *Cognitive Science*, 19, 341-376.
- Ben-Zeev, T. (1996). When erroneous mathematical thinking is just as "correct:" The oxymoron of rational errors. In R. J. Sternberg & T. Ben-Zeev (Eds.), *The nature of mathematical thinking* (pp. 55-79). Mahwah, NJ: Erlbaum.
- Bloom, B. (1968). Learning for mastery. *Evaluation Comment*, 1(2), 1-5.
- Bloom, B. (1972). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim: Beltz.
- Boekaerts, M. (1996). Self-regulated learning at the junction of cognition and motivation. *European Psychologist*, 1(2), 100-112.
- Bortz, J. (1993). *Statistik für Sozialwissenschaftler*. Berlin: Springer.
- Brown, J. S. & Burton, R. R. (1978) Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, 2, 155-191.
- Brown, J. S. & VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science*, 4, 379-426.
- Burton, R.R. (1982). Diagnosing bugs in a simple procedural skill. In D. Sleeman & J. S. Brown (Eds.), *Intelligent Tutoring Systems* (pp.157-183). New York: Academic Press.
- Butler, D. & Winne, P. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245-281.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multi- method matrix. *Psychological Bulletin*, 56, 81-105.
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P. & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13, 145-182.
- Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T. & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, 25, 471-533.
- Clariana, R. B. (1993). A review of multiple-try feedback in traditional and computer-based instruction. *Journal of Computer-Based Instruction*, 20 (3), 67-74.
- Clariana, R. B. *Feedback in computer-assisted learning*. NETg, University of Limerick Lecture Series. Verfügbar unter: <http://www.netg.com/research/whitepapers/clarianawp.asp> [März, 2003].

- Clariana, R. B. & Lee, D. (2001). The effects of recognition and recall study tasks with feedback in a computer-based vocabulary lesson. *Educational Technology Research and Development*, 49(3), 1042-1629.
- Clariana, R. B. & Smith, L. J. (1989). *The effects of AUC and KCR feedback on learners of different ability*. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Little Rock, Arkansas. (ERIC Document Reproduction Service No. ED 313 387)
- Cohen, A. D. (1987). Student processing of feedback on their compositions. In A. Wenden & Rubin, J. (Eds.), *Learner strategies in language learning* (pp. 57-69). Englewood Cliffs, NJ: Prentice/Hall International.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press.
- Collins, M., Carnine, D. & Gersten, R. (1987). Elaborated corrective feedback and the acquisition of reasoning skills: A study of computer-assisted instruction. *Exceptional Children*, 54, 254-262.
- Conati, C. & Van Lehn, K. (1999). Teaching meta-cognitive skills: Implementation and evaluation of a tutoring system to guide self-explanation while learning from examples. In S.P. Lajoie & M. Vivet (Eds.), *Artificial Intelligence in Education* (pp. 297 – 304). Amsterdam: IOS Press.
- Connelly III, J.W. (2001). *Assessing the impact of minimally intelligent, computer-generated, immediate feedback on an ill-defined problem solving task* [Electronic version]. Unveröffentlichte Dissertation. USA: University of Pittsburgh, Department of Psychology.
- Corbett, A. T., Anderson, J. R., Graesser, A. C., Koedinger, K., & VanLehn, K. (1999). Third generation computer tutors: Learn from or ignore human tutors? *Proceedings of the Conference of Computer-Human Interaction*, 85-86. New York: ACM Press.
- Cox, L. (1975). Systematic errors in the four vertical algorithms in normal and handicapped populations. *Journal for Research in Mathematics Education*, 6(4), 202-220.
- Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety*. San Francisco: Jossey-Bass.
- Day, J. D. & Córdón, L. A. (1993). Static and dynamic measures of ability: An experimental comparison. *Journal of Educational Psychology*, 85 (1), 75-82.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

- Dempsey, J. V. & Litchfield, B. C. (1993). Feedback, retention, discrimination error, and feedback study time. *Journal of Research on Computing in Education*, 25(3), 303-326.
- Doignon, J.-P. and Falmagne, J.-C. (1985). Spaces for the assessment of knowledge. *International Journal of Man-Machine Studies*, 23, 175-196.
- Dresel, M. (2002). *Motivationsförderung im schulischen Kontext: Effekte der Inhaltsvariation und Sequenzierung attributionalen Feedbacks*. Unveröffentlichte Dissertation, München.
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B. & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1, 170-177.
- Fitts, P. M. (1964). Perceptual-motor skills learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 243-285). New York: Academic Press.
- Gagné, R. M. (1985). *The conditions of learning (4th ed.)*. New York: Holt, Rinehart & Winston.
- Gagné, R. M., Briggs, L. J. & Wager, W. W. (1992). *Principles of instructional design*. Fort Worth: Harcourt Brace Jovanovich College Publication.
- Gerster, H.-D. (1982). *Schülerfehler bei schriftlichen Rechenverfahren*. Freiburg: Herder.
- Gertner, A. S., Conati, C. & VanLehn, K. (1998). Procedural help in Andes: Generating hints using a Bayesian network student model. *Proceedings of the 15th national Conference on Artificial Intelligence* (pp. 106-111). Cambridge, MA: MIT Press.
- Graesser, A. C., Person, N. K. & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied cognitive psychology*, 9, 495-522.
- Graesser, A. C., VanLehn, K., Rosé, C., Jordan, P. & Harter, D. (2001). Intelligent Tutoring Systems with Conversational Dialogue. *AI Magazine*, 22, 39-51.
- Hager, W. (2000). Planung von Untersuchungen zur Prüfung von Wirksamkeits- und Wirksamkeitsunterschiedshypothesen. In W. Hager, J.-L. Patry & H. Brezing (Hrsg.), *Evaluation psychologischer Interventionsmaßnahmen* (S. 202-239). Bern: Huber.
- Hager, W., Patry, J.-L. & Brezing, H. (2000). *Evaluation psychologischer Interventionsmaßnahmen*. Bern: Huber.
- Hancock, T. E., Thurman, R. A. & Hubbard, D. C. (1995). An expanded control model for the use of instructional feedback. *Contemporary Educational Psychology*, 20, 410-425.

- Haynie, W. J. (1994). Effects of multiple-choice and short-answer tests on delayed retention learning. *Journal of Technology Education*, 6(1), 32-44.
- Haynie, W. J. (1997). Effects of anticipation of tests on delayed retention learning. *Journal of Technology Education*, 9(1), 1-11.
- Hays, W. L. (1973). *Statistics for the social sciences*. New York: Holt, Rinehart, and Winston.
- Heffernan, N. T. (2001). *Intelligent tutoring systems are forgotten the tutor: Adding a cognitive model of human tutors* (Tech. Rep. CMU-CS-01-127) [Electronic version]. Unveröffentlichte Dissertation. USA, Pittsburgh: Carnegie Mellon University, School of Computer Science, Computer Science Department.
- Heift, T. (2001). Error-specific and individualized feedback in a web-based language tutoring system: Do they read it? *ReCALL*, 13(1), 99-109.
- Heil, M. C. & Agnew, B. O. (2000). *The effects of previous computer experience on Air Traffic-Selection and Training (AT-SAT) test performance* [Electronic version]. National Technical Information Service, Springfield, Virginia 22161.
- Helmke, A. (1991). Entwicklung des Fähigkeitsselbstbildes vom Kindergarten bis zur dritten Klasse. In R. Pekrun & H. Fend (Hrsg.), *Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung* (S. 83-99). Stuttgart: Enke.
- Helmke, A. & Schrader, F.-W. (1998). Determinanten der Schulleistung. In D.H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 60-67). Weinheim: PVU.
- Hoska, D. (1993). Motivating learners through CBI feedback: Developing a positive learner perspective. In J. V. Dempsey & G. C. Sales (Eds.), *Interactive instruction and feedback* (pp. 105-132). Englewood Cliffs, NJ: Educational Technology.
- Huitema, B.E. (1980) *The analysis of covariance and alternatives*. New York: Wiley.
- Jacobs, B. (2002) *Aufgaben stellen und Feedback geben*. Verfügbar unter: <http://www.phil.uni-sb.de/~jakobs/wwwartikel/feedback/projekt.htm> [November, 2003].
- Jonassen, D. H., Tessmer, M. & Hannum, W. H. (1999). *Task analysis methods for instructional design*. Mahwah, NJ: Erlbaum.
- Kline, F. M., Schumaker, J. B. & Deshler, D. D. (1991). Development and validation of feedback routines for instructing students with learning disabilities. *Learning Disability Quarterly*, 14 (3), 191-207.

- Kluger, A. N. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 199, 254-284.
- Korossy, K. (1996). Kompetenz und Performanz beim Lösen von Geometrie-Aufgaben. *Zeitschrift für Experimentelle Psychologie*, 11(3), 279-318.
- Korossy, K. & Held, T. (2001). Theory-based knowledge modelling in a subdomain of elementary algebra. *Zeitschrift für Psychologie*, 209(3), 277-315.
- Krapp, A. (1999). Intrinsische Lernmotivation und Interesse. Forschungsansätze und konzeptuelle Überlegungen. *Zeitschrift für Pädagogik*, 45, Heft 4.
- Kühnhold, K. & Padberg, F. (1986). Über typische Fehler bei der schriftlichen Subtraktion natürlicher Zahlen. *Mathematikunterricht*, 3, 6-16.
- Kulhavy, R.W. (1977). Feedback in written instruction. *Review of Educational Research*, 47, (1), 211-232.
- Kulhavy, R.W. & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, 1(4), 279-308.
- Kulhavy, R. W., White, M. T., Topp, B. W., Chan, A. L. & Adams, J. (1985). Feedback complexity and corrective efficiency. *Contemporary Educational Psychology*, 10, 285-291.
- Langley, P. & Ohlsson, S. (1984). Automated cognitive modelling. In *Proceedings of AAAI-84*. Los Altos, CA: Morgan Kaufmann.
- LeFevre, J. & Dixon, P. (1986). Do written instructions need examples? *Cognition and Instruction*, 3, 1-30.
- Mayer, R. & Anderson, R. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, 83 (4), 484-490.
- Mayer, R. & Anderson, R. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84 (4), 444-452.
- Mayer, R. E. & Moreno, R. (1998). A split-attention effect in multimedia learning: evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90 (29), 312-320.
- Mayer, R. E. & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, 12, 107-119.

- Marzano, R. J. & Arthur, S. (1977). *Teacher comments on student essays: It doesn't matter what you say*. (ERIC Document Reproduction Service No ED147864).
- Merrill, D. C., Reiser, B. J., Merrill, S. K. & Landes, S. (1995). Tutoring: Guided learning by doing. *Cognition and Instruction*, 13(3), 315-372.
- Merrill, M. D. (1983). Component Display Theory. In C. M. Reigeluth (Ed.). *Instructional-design theories and models. An overview of their current status* (pp. 384-434). Hillsdale, NJ: Erlbaum.
- Morrison, G. R., Ross, S. M., Gopalakrishnan, M. & Casey, J. (1995). The effects of feedback and incentives on achievement in computer-based instruction. *Contemporary Educational Psychology*, 20, 32-50.
- Mory, E. H. (1996). Feedback research. In D.H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 919-956). New York: Simon & Schuster Macmillan.
- Mosel-Göbel, D. (1988). Algorithmusverständnis am Beispiel ausgewählter Verfahren der schriftlichen Subtraktion. Eine Fallstudienanalyse bei Grundschullehrern. *Sachunterricht und Mathematik in der Grundschule*, 12, 554-559.
- Mousavi, S. Y., Low, R. & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87, 319-334.
- Murray, F. B. (1983) Equilibration as cognitive conflict. *Developmental Review*, 3, 54-61.
- Nagata, N. (1993). Intelligent computer feedback for second language instruction. *The Modern Language Journal*, 77(3), 330-339.
- Nagata, N. & Swisher, M.V. (1995). A study of consciousness-raising by computer: The effect of metalinguistic feedback. *Foreign Language Annals*, 28 (3), 337-347.
- Narciss, S. (April, 1999). *Motivational Effects of the Informativeness of Feedback*. Paper presented at the Conference of the American Educational Research Association, Montreal (available at ERIC Document Data Base ED430034).
- Narciss, S. (August, 2001). *Informative Feedback as a bridge from instruction to learning in computer-based instruction*. Vortrag auf der 9. Konferenz der der European Association for Research in Learning and Instruction, Fribourg, Switzerland.
- Narciss, S. (2004). *Entwicklung und Evaluation von informativem tutoriellem Feedback für komplexe Lernaufgaben*. Habilitationsmanuskript Technische Universität Dresden.

- Narciss, S. & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. In H. M. Niegemann, R. Brünken, & D. Leutner (Eds.), *Instructional design for multimedia learning* (S. 181-195). Münster: Waxmann.
- Newman, R. S. & Schwager, M. T. (1995). Students' help seeking during problem solving: Effects of grade, goal, and prior achievement. *American Educational Research Journal*, 32(2), 352-376.
- Norman, D.A. (1981). Categorization of action slips. *Psychological Review*, 88(1), 1-15.
- Paas, F. G. W. C. & Van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, 86 (1), 122-133.
- Padberg, F. (1994). Schriftliche Subtraktion – Änderungen erforderlich! *Mathematische Unterrichtspraxis*, 2, 24-34.
- Padberg, F. (1996). *Didaktik der Arithmetik*. Heidelberg: Spektrum Akademischer Verlag.
- Paechter, M. (1996). Auditive und visuelle Texte in Lernsoftware - Herleitung und empirische Prüfung eines didaktischen Konzepts zum Einsatz auditiver und visueller Texte in Lernsoftware. *Internationale Hochschulschriften*, 212. Münster: Waxmann.
- Pajares, F. & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept and school achievement. In R. J. Riding & S. G. Rayner (Eds.), *International perspectives on individual differences: Vol 2. Self-perception* (pp. 239-266). London: Ablex Publishing.
- Peeck, J., van den Bosch, A. B. & Kreupeling, W. J. (1985). Effects of informative feedback in relation to retention of initial responses. *Contemporary Educational Psychology*, 10 (4), 303-313.
- Phe, G. D. (1979). The processing of informative feedback about multiple-choice test performance. *Contemporary Educational Psychology*, 4, 381-394.
- Phe, G. D., & Bender, T. (1989). Feedback complexity and practice: Response pattern analysis in retention and transfer. *Contemporary Educational Psychology*, 14, 97-110.
- Phe, G. D., Gugliemella, J. & Sola, J. (1976). Effects of delayed retention on multiple-choice test performance. *Contemporary Educational Psychology*, 1(1), 26-36.
- Radatz, H. (1980). *Fehleranalysen im Mathematikunterricht*. Braunschweig: Vieweg.

- Reimann, G. & Narciss, S. (September 2003). *Individuelle Nutzung informativen tutoriellen Feedbacks in einer webbasierten Lernumgebung*. Vortrag auf der 9. Fachtagung der Pädagogischen Psychologie in Bielefeld.
- Renkl, A. (2000). *Worked-out examples: Instructional explanation support learning by self-explanations* (Research Report No. 139). Universität Freiburg, Institut für Psychologie.
- Resnick, L. B. (1982). Syntax and semantics in learning to subtract. In T. P. Carpenter, J. M. Moser & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (136-155). Hillsdale, NJ: Erlbaum.
- Rittle-Johnson, B., Siegler, R. S. & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*, 346-362.
- Schimmel, B. J. (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).
- Schimmel, B. J. (1988). Providing meaningful feedback in courseware. In D. H. Jonassen (Ed.), *Instructional designs for microcomputer courseware* (pp. 183-195). Hillsdale, NJ: Erlbaum.
- Schunk, D. H. & Pajares, F. (2001). The development of academic self-efficacy. In A. Wigfield & J. Eccles (Eds.), *Development of achievement motivation*. San Diego: American Press.
- Schunk, D. H. & Rice, J. M. (1993). Strategy fading and progress feedback: Effects on self-efficacy and comprehension among students receiving remedial reading services. *Journal of Special Education, 27*(3), 257-276.
- Sedlmeier, P. (1996). Jenseits des Signifikanztest-Rituals: Ergänzungen und Alternativen. *Methods of Psychological Research - online, 1*. [Internet: <http://www.mpr-online.de/>].
- Seligman, M. E. P. (1975). *Helplessness. On depression, development, and death*. San Francisco: Freeman.
- Siegel, M. A. & Misselt, L. A. (1984). Adaptive feedback and review paradigm for computer-based drills. *Journal of Educational Psychology, 76*(2), 310-317.
- Sherrill, J. M. (1979). Subtraction: Decomposition versus equal addends. *Arithmetic Teacher, September*, 16-17.

- Smith, P. L. & Ragan, T. J. (1993). Designing instructional feedback for different learning outcomes. In J. V. Dempsey & G. C. Sales (Eds.), *Interactive instruction and feedback* (pp. 75-103). Englewood Cliffs, NJ: Educational Technology.
- Spada, H. & Wichmann, S. (1996). Kognitive Determinanten der Lernleistung. In F. E. Weinert (Hrsg.), *Psychologie des Lernens und der Instruktion* (S. 119-153). Göttingen: Hogrefe.
- Stark, R. (1999). *Lernen mit Lösungsbeispielen*. Göttingen: Hogrefe.
- Steinberg, E.R. (1989). Cognition and learner control: A literatur review, 1977-1988. *Journal of Computer Based Instruction*, 16, 117-121.
- Stoffers, D., Berendse, H. W., Deijen, J. B. & Wolters, E. C. (2002). The influence of computer experience on visuo-motor control: Implications for visuo-motor testing in Parkinson's disease. *Neuropsychologia*, 40(11), 1779-1785.
- Suthers, D., Connelly, J., Lesgold, A., Paolucci, M., Toth, E., Toth, J. & Weiner, A. (2001). Representational and advisory guidance for students learning scientific inquiry. In K.D. Forbus & P.J. Feltovich (Eds.), *Smart machines in education: The coming revolution in educational technology* (pp. 7-35). Menlo Park, CA: AAAI/Mit Press.
- Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and Instruction*, 4, 295-312.
- Sweller, J. & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, 2 (1), 59-89.
- Sweller, J., Van Merriënboer, J. J. G. & Paas F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251-296.
- VanLehn, K. (1986). Arithmetic procedures are induced from examples. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 133-179). Hillsdale, NJ: Erlbaum.
- VanLehn, K. (1990). *Mind Bugs: The Origins of Procedural Misconceptions*. Cambridge, MA: MIT Press.
- VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology*, 47, 513-537.
- Wolf, B. (2001). Effektstärkemaße. In Rost, D. H. (Hrsg.) *Handwörterbuch Pädagogische Psychologie* (S. 96-102). Weinheim: Beltz.
- Wood, H., & Wood, D. (1999). Help seeking, learning and contingent tutoring. *Computers and Education*, 33, 153-169.

Young, R. M. & O'Shea, T. (1981). Errors in children's subtraction. *Cognitive Science*, 5, 153-177.

Zielinski, W. (1998). *Lernschwierigkeiten: Ursachen – Diagnostik – Intervention*. Stuttgart: Kohlhammer.

7.2 Abbildungsverzeichnis

<i>Abbildung 1:</i>	Heuristisches Modell zur Analyse und Untersuchung von Feedbackeffekten in Lehr-Lernsituationen nach Narciss (2004)	11
<i>Abbildung 2:</i>	Determinanten des Informationswertes von Feedback nach Narciss (2004).....	26
<i>Abbildung 3:</i>	Auswahl und Spezifikation von ITF-Komponenten nach Narciss (2004; Narciss und Huth, 2004).....	31
<i>Abbildung 4:</i>	Bedingungen und Wirkungen von Feedback in Lehr-Lernsituationen (nach Narciss, 2004).....	39
<i>Abbildung 5:</i>	Hasse-Diagramm der Anforderungsstruktur zum Gegenstandsbereich der schriftlichen Subtraktion im Zahlenraum bis 1000.....	66
<i>Abbildung 6:</i>	Fehlerspezifischer ITF Algorithmus für die schriftliche Subtraktion.....	76
<i>Abbildung 7:</i>	Darstellung des in der Vergleichsbedingung realisierten KR-KCR Feedbackalgorithmus.....	84
<i>Abbildung 8:</i>	Eingabefenster zur Benutzeridentifikation.....	85
<i>Abbildung 9:</i>	Aufgabenmenüfenster für die Auswahl der zu bearbeitenden Aufgabentypen.....	86
<i>Abbildung 10:</i>	Rechenfenster mit korrekt bearbeiteter Übertrags-Aufgabe im ersten Versuch.....	87
<i>Abbildung 11:</i>	Rechenfenster mit fehlerhaft bearbeiteter Übertragsaufgabe im ersten Lösungsversuch.....	87
<i>Abbildung 12:</i>	KR Feedback & Hinweis auf korrekte Lösungsstrategie.....	88
<i>Abbildung 13:</i>	KR Feedback.....	88
<i>Abbildung 14:</i>	KM Feedback (Knowledge-about-Mistake Feedback) – Informationen zu Ort und Art des Fehlers und Hinweis auf korrekte Lösungsstrategie.....	89
<i>Abbildung 15:</i>	KCR Feedback.....	89
<i>Abbildung 16:</i>	KH Feedback (Knowledge on How to Proceed Feedback) - auditive und visuell-sequentielle Präsentation des korrekten Lösungsalgorithmus nach dem dritten, fehlerhaften Lösungsversuch.....	90
<i>Abbildung 17:</i>	Rechenblatt mit ausgearbeitetem Lösungsbeispiel.....	90
<i>Abbildung 18:</i>	Beschreibung des Untersuchungsablaufs von Studie 1.....	100
<i>Abbildung 19:</i>	Vor-Nachtestleistungen der Feedbackgruppen.....	108
<i>Abbildung 20:</i>	Darstellung der Mittelwerte der im Vortest fehlerhaft gelösten Aufgaben und der Mittelwerte der Aufgaben, die sowohl im Vortest als auch im Nachtest fehlerhaft bearbeitet worden waren, d.h. im Nachtest nicht korrigiert werden konnten.....	110
<i>Abbildung 21:</i>	Darstellung des KR-KM Feedbackalgorithmus.....	124
<i>Abbildung 22:</i>	Darstellung des KR-KH Feedbackalgorithmus.....	125
<i>Abbildung 23:</i>	KM Feedback – Informationen zu Ort und Art des Fehlers und Hinweis auf korrekte Lösungsstrategie.....	126
<i>Abbildung 24:</i>	KH Feedback – auditive und visuell-sequentielle Präsentation des korrekten Lösungsalgorithmus.....	126
<i>Abbildung 25:</i>	Beschreibung des Untersuchungsablaufs von Studie 2.....	134
<i>Abbildung 26:</i>	Darstellung der unter Auspartialisierung der Kovariaten Interesse und Fähigkeitskonzept geschätzten Mittelwerte des Vortests, Nachtests1 und Nachtests2 für die einzelnen Feedbackbedingungen.....	144
<i>Abbildung 27:</i>	Darstellung der unter Auspartialisierung der Kovariaten Interesse und Computererfahrung	

	geschätzten Mittelwerte der fehlerhaft bearbeiteten Aufgaben im Vortest, im Nachtest1 und im Nachtest2 für die einzelnen Feedbackbedingungen.....	149
<i>Abbildung 28:</i>	Darstellung der unter Auspartialisierung der Kovariaten Vortest, Interesse, und Computererfahrung geschätzten Mittelwerte der Treatmentleistung zu Messzeitpunkt 1 und 2 für die einzelnen Feedbackbedingungen.....	153
<i>Abbildung 29:</i>	Darstellung der unter Auspartialisierung der Kovariaten Vortestleistung, Freude und Computererfahrung geschätzten Mittelwerte des post-aktionalen Kompetenzerlebens zu Messzeitpunkt 1 und 2 für die einzelnen Feedbackbedingungen.....	157
<i>Abbildung 30:</i>	Darstellung der unter Auspartialisierung der Kovariaten Treatmentleistung, Vortestleistung, Freude und Computererfahrung geschätzten Mittelwerte des post-aktionalen Kompetenzerlebens zu Messzeitpunkt 1 und 2 für die einzelnen Feedbackbedingungen.....	160

7.3 Tabellenverzeichnis

Tabelle 1:	<i>Effektstärken verschiedener Feedbackarten im Vergleich</i>	7
Tabelle 2:	<i>Kognitive, meta-kognitive und motivationale Funktionen von ITF nach Narciss (2004)</i>	38
Tabelle 3:	<i>Ableitung von Indikatoren für die Untersuchung von kognitiven, meta-kognitiven und motivationalen Feedbackeffekten nach Narciss (2004)</i>	40
Tabelle 4:	<i>Gegenüberstellung des Abziehverfahrens und des Ergänzungsverfahrens mit den Vor- und Nachteilen des Ergänzungsverfahrens gegenüber dem Abziehverfahren nach Padberg (1996)</i>	49
Tabelle 5:	<i>Rechenvorschrift, Vor- und Nachteile für Borgetechnik und Erweiterungstechnik nach Padberg (1996)</i>	50
Tabelle 6:	<i>Gersters diagnostischer Subtraktionstest SI (1982)</i>	52
Tabelle 7:	<i>Systematische Fehler in der schriftlichen Subtraktion nach Gerster (1982)</i>	54
Tabelle 8:	<i>Diagnostischer Subtraktionstest nach Kühnhold und Padberg (1986)</i>	57
Tabelle 9:	<i>Die Menge A der elementaren Aufgabenanforderungen</i>	62
Tabelle 10:	<i>Aufgabentypen der Ebene 2 der Anforderungsstruktur, die über die Kombination von 2 der fünf elementaren Aufgabenanforderungen definiert sind</i>	64
Tabelle 11:	<i>Aufgabentypen der Ebene 3 der Anforderungsstruktur, die über die Kombination der Aufgabentypen der 1. und 2. Ebene definiert sind</i>	65
Tabelle 12:	<i>Idealisiertes Lernerprotokoll nach VanLehn (1990, S. 73)</i>	67
Tabelle 13:	<i>Diagnostischer Test zur Klassifikation systematischer Fehler</i>	69
Tabelle 14:	<i>Typen systematischer Fehler und deren Auftretenshäufigkeit, dargestellt für jeden Aufgabentyp</i>	71
Tabelle 15:	<i>Vortest-Nachtest-Vergleichsgruppen-Plan</i>	82
Tabelle 16:	<i>Vortest mit 32 Subtraktionsaufgaben zur Erfassung des aufgabenspezifischen Vorwissens</i>	92
Tabelle 17:	<i>Items zur Erfassung der Computer-/Lernprogrammerfahrung und Computereinstellung</i>	94
Tabelle 18:	<i>Vortest-Nachtest-Antwortmuster und deren korrespondierenden Feedbackfunktionen</i>	95
Tabelle 19:	<i>Definition des Korrekturwahrscheinlichkeitsindex als bedingte Wahrscheinlichkeit und dessen Operationalisierung in Form von Antwortmustern</i>	96
Tabelle 20:	<i>Items zur Erfassung des wahrgenommenen Nutzens des ITF</i>	96
Tabelle 21:	<i>Items zur Erfassung des Feedbackwunsches, der KR-KCR Feedbackgruppe vorgelegt</i>	97
Tabelle 22:	<i>Items zur Erfassung post-aktionaler motivationaler Parameter</i>	98
Tabelle 23:	<i>Deskriptive Maße der zentralen Tendenz, Dispersionsmaße und Häufigkeitsverteilungen der Kontrollvariablen, getrennt nach Feedbackgruppen dargestellt, sowie Ergebnisse der Signifikanztests</i>	103
Tabelle 24:	<i>Korrelationen zwischen Kovariaten und abhängigen Variablen</i>	104
Tabelle 25:	<i>Interpretation der Effektstärkemaße nach Cohen (1988)</i>	106
Tabelle 26:	<i>Mittelwerte und Standardabweichungen der vier Vortest-Nachtest-Antwortmuster, getrennt für jede Feedbackgruppe dargestellt, Ergebnisse der t-Tests</i>	109
Tabelle 27:	<i>Mittelwerte und Standardabweichungen für den Korrekturwahrscheinlichkeitsindex getrennt nach Feedbackgruppen</i>	110
Tabelle 28:	<i>Häufigkeiten (%) und Ergebnisse der Chi-Quadrat-Tests zum wahrgenommenen Feedbacknutzen der ITF Gruppe bzw. zum Feedbackwunsch der KR-KCR Feedbackgruppe</i>	111
Tabelle 29:	<i>Vortest-Nachtest-Vergleichsgruppen-Plan mit zwei Messzeitpunkten</i>	123

Tabelle 30: <i>Vortest mit 24 Subtraktionsaufgaben zur Erfassung des aufgabenspezifischen Vorwissens</i>	128
Tabelle 31: <i>Items zur Erfassung individueller motivationaler Voraussetzungen (vgl. Narciss, 2004)</i>	129
Tabelle 32: <i>Items zur Erfassung des post-aktionalen Kompetenzerlebens</i>	132
Tabelle 33: <i>Darstellung der Mittelwerte und Standardabweichungen bzw. Häufigkeitsverteilungen für jede Kontrollvariable getrennt nach Feedbackgruppe</i>	136
Tabelle 34: <i>Korrelationen der Kontrollvariablen und abhängigen Variablen von Messzeitpunkt 1 und Messzeitpunkt 2 (85 < N < 105)</i>	137
Tabelle 35: <i>Deskriptive Statistiken, Prüfstatistiken und Effektstärken für Intragruppenvergleiche hinsichtlich des Einflusses der Treatmentbedingungen auf die Nachtestleistung</i>	141
Tabelle 36: <i>Statistische Kennwerte und Effektstärken der globalen Prüfstatistiken sowie Einzelvergleiche und Effektstärken für die Haupt- und Interaktionseffekte der Nachtestleistungen zu beiden Messzeitpunkten</i>	143
Tabelle 37: <i>Deskriptive Statistiken, Prüfstatistiken und Effektstärken für Intragruppenvergleiche hinsichtlich des Einflusses der Treatmentbedingungen auf die Korrektur vor der Treatmentphase fehlerhaft bearbeiteter Aufgaben, Angabe des Korrekturwahrscheinlichkeitsindex</i>	146
Tabelle 38: <i>Statistische Kennwerte und Effektstärken der globalen Prüfstatistiken und Einzelvergleiche für die Haupt- und Interaktionseffekte des Korrekturwahrscheinlichkeitsindex zu beiden Messzeitpunkten</i>	148
Tabelle 39: <i>Deskriptive Statistiken, Prüfstatistiken und Effektstärken für Intragruppenvergleiche hinsichtlich des Einflusses der Feedbackbedingungen auf die während der Treatmentphase bearbeiteten Aufgabentypen</i>	151
Tabelle 40: <i>Statistische Kennwerte und Effektstärken der globalen Prüfstatistiken und Einzelvergleiche für die Haupt- und Interaktionseffekte der Treatmentleistung zu beiden Messzeitpunkten</i>	152
Tabelle 41: <i>Deskriptive Statistiken, Prüfstatistiken und Effektstärken für Intragruppenvergleiche hinsichtlich des Einflusses der Feedbackbedingungen auf das post-aktionale Kompetenzerleben</i>	155
Tabelle 42: <i>Statistische Kennwerte und Effektstärken der globalen Prüfstatistiken und Einzelvergleiche für die Haupt- und Interaktionseffekte der Selbstbewertung der Treatmentleistung zu beiden Messzeitpunkten</i>	156
Tabelle 43: <i>Statistische Kennwerte und Effektstärken der globalen Prüfstatistiken und Einzelvergleiche für die Haupt- und Interaktionseffekte des post-aktionalen Kompetenzerlebens unter Auspartialisierung des Einflusses der Treatmentleistung zu beiden Messzeitpunkten</i>	159

7.4 Abkürzungsverzeichnis

ITF	...	Informatives Tutorielles Feedback
KR	...	Knowledge of Result
KCR	...	Knowledge of Correct Response
KM	...	Knowledge about Mistake
KH	...	Knowledge on How to proceed
MTF	...	Multiple Try Feedback
AUC	...	Answer Until Correct
DF	...	Delayed Feedback
EF	...	Elaborated Feedback

8 ANHÄNGE

8.1 Anhang A: Fragebögen

Studie 1

1. Vor-/Nachtest mit 32 Subtraktionsaufgaben
2. Vorbefragungsbogen mit Selbstkonzeptitems und Items zur Computereinstellung und -erfahrung
3. Nachbefragungsbogen mit 6 motivationalen Items und mit 2 Items zur Einschätzung des Feedbackwunsches für KR-KCR-Feedbackgruppe
4. Nachbefragungsbogen mit 6 motivationalen Items und mit 3 Items zur Einschätzung des wahrgenommenen Feedbacknutzens für ITF-Gruppe

Studie 2

1. Vortest, Nachtest1 und Nachtest2 mit 24 Subtraktionsaufgaben
2. Vorbefragungsbogen mit Items zum aufgabenspezifischen Fähigkeitskonzept, Freude, Interesse und Computererfahrung
3. Nachbefragungsbogen mit 7 motivationalen Items

Vortest / Nachtest der Studie 1

32 Aufgaben zur schriftlichen Subtraktion



Name:

Klasse:

Schule:

Datum:

1					2					3					4						
		8	4	7			5	6	2			6	7	3			3	7	4		
	-			3	0		-		3	6		-		5	3		-		9	2	
5						6					7					8					
		4	6	0			9	4	8			3	1	6			4	5	9		
	-			2	5		-		2	0		-		1	4	0	-		4	0	8
9						10					11					12					
		7	3	6			5	0	4			8	6	3			7	2	9		
	-			7	0		-		3	3	0	-		4	9		-		3	9	0
13						14					15					16					
		3	4	8			2	5	2			5	2	0			3	9	2		
	-			1	4	9	-		7	7		-		3	6	7	-		9	1	





1 7					1 8					1 9					2 0				
		6 5 1				7 8 7					2 3 6					7 0 4			
-		4 0 1			-	2 9 7				-	8 6				-	4 2 4			
2 1					2 2					2 3					2 4				
		6 1 3				3 2 8					7 4 1					3 4 4			
-		9 1			-	7 0				-	6 2				-	6 4			
2 5					2 6					2 7					2 8				
		6 2 0				5 7 0					9 0 2					4 3 0			
-		8 2			-	3 7				-	7 6 0				-	2 7 2			
2 9					3 0					3 1					3 2				
		6 0 9				2 6 0					3 6 7					9 2 6			
-		3 4 9			-	8 1				-	3 5				-	1 2			



Vorbefragungsbogen der Studie 1

Name: _____ Datum: _____

Geschlecht: Junge Klasse: _____
 Mädchen Schule: _____

Alter: _____

Beantworte bitte die folgenden Fragen so gut du kannst, in dem du ein Kreuz in das entsprechende Feld machst!

Hier siehst du ein Beispiel:

Ich sehe die Seamstraße

sehr
gern

ziemlich
gern

nicht so
gern

überhaupt
nicht gern

1) Das Lösen schriftlicher Subtraktionsaufgaben fällt mir normalerweise

sehr schwer

schwer

leicht

sehr leicht

2) Das Lösen schriftlicher Subtraktionsaufgaben macht mir

sehr viel
Spaß

ein wenig
Spaß

nicht so viel
Spaß

überhaupt keinen
Spaß

3) Rechnen fällt mir normalerweise

sehr schwer

schwer

leicht

sehr leicht

4) Rechnen macht mir

sehr viel
Spaß

ein wenig
Spaß

nicht so viel
Spaß

überhaupt keinen
Spaß

5) Ich/meine Familie besitzt einen eigenen Computer. Ja Nein

- 6) Mit dem Computer arbeite/spiele ich ungefähr
- täglich
 - 1-2mal in der Woche
 - 1-2mal im Monat
 - seltener
 - nie.
-

7) Das Arbeiten/Spielen am Computer macht mir

sehr viel
Spaß

ein wenig
Spaß

nicht so viel
Spaß

überhaupt keinen
Spaß

8) Ich/meine Familie besitzt ein Computer-Lernprogramm.

- Ja Nein

Wenn ja:
Wie heißt es? _____

- 9) Mit diesem Computer-Lernprogramm übe ich
- täglich
 - 1-2mal in der Woche
 - 1-2mal im Monat
 - seltener
 - nie.



So, jetzt kann es endlich richtig losgehen!

Nachbefragungsbogen KR-KCR-Feedbackgruppe der Studie 1

Name: _____ Datum: _____

Geschlecht: Junge Mädchen Klasse: _____

Alter: _____ Schule: _____

Beantworte bitte die folgenden Fragen so gut du kannst, in dem du ein Kreuz in das entsprechende Feld machst!

Hier siehst du ein Beispiel:

Ich sehe die Sesamstraße

sehr
gern

ziemlich
gern

nicht so
gern

überhaupt
nicht gern

ACHTUNG:

Alle Fragen beziehen sich nur auf die letzte halbe Stunde, die du an dem Matheprogramm gegessen hast!!!

1) Ich bin mit meiner Leistung in dieser halben Stunde

sehr
zufrieden

zufrieden

nicht
zufrieden

überhaupt
nicht zufrieden

2) Ich fand das Lösen der Aufgaben sehr anstrengend.

stimmt

stimmt
ein bisschen

stimmt
nicht

stimmt
überhaupt nicht

3) Ich fand meine Leistung in dieser halben Stunde

sehr gut

gut

nicht gut

überhaupt
nicht gut

4) Ich fand das Beantworten der Aufgaben

sehr
leicht

leicht

schwer

sehr
schwer

5) Mir hat das Lösen der Aufgaben

sehr viel Spaß
gemacht

Spaß
gemacht

keinen Spaß
gemacht

überhaupt keinen
Spaß gemacht

6) Ich denke, ich habe in dieser halben Stunde

sehr viel
dazu gelernt

viel
dazu gelernt

ein wenig
dazu gelernt

nichts
dazu gelernt

7) Es wäre hilfreich gewesen, wenn Subtratino mir gesagt hätte, was ich falsch gemacht habe.

ja

nein

8) Ich hätte gern einen Tipp von Subtratino bekommen, wie ich man richtig subtrahiert.

ja

nein



Vielen Dank für deine Hilfe, Subtratino.

Nachbefragungsbogen ITF-Gruppe der Studie 1

Name: _____ Datum: _____

Geschlecht: Junge Klasse: _____
 Mädchen Schule: _____

Alter: _____

Beantworte bitte die folgenden Fragen so gut du kannst, in dem du ein Kreuz in das entsprechende Feld machst!

Hier siehst du ein Beispiel:

Ich sehe die Sesamstraße

sehr
gern

 ziemlich
gern

 ~~nicht~~ so
gern

 überhaupt
nicht gern

ACHTUNG:

Alle Fragen beziehen sich nur auf die letzte halbe Stunde, die du an dem Matheprogramm gegessen hast!!!

1) Ich bin mit meiner Leistung in dieser halben Stunde

sehr
zufrieden

 zufrieden

 nicht
zufrieden

 überhaupt nicht
zufrieden

2) Ich fand das Lösen der Aufgaben sehr anstrengend.

stimmt

 stimmt
ein bisschen

 stimmt
nicht

 stimmt
überhaupt nicht

3) Ich fand meine Leistung in dieser halben Stunde

sehr gut

 gut

 nicht gut

 überhaupt
nicht gut

4) Ich fand das Beantworten der Aufgaben

sehr leicht

leicht

schwer

sehr schwer

5) Mir hat das Lösen der Aufgaben

sehr viel Spaß gemacht

Spaß gemacht

kein Spaß gemacht

überhaupt keinen Spaß gemacht

6) Ich denke, ich habe in dieser halben Stunde

sehr viel dazu gelernt

viel dazu gelernt

ein wenig dazu gelernt

nichts dazu gelernt

7) Die Hinweise haben mir

sehr geholfen

geholfen

nicht so geholfen

überhaupt nicht geholfen

8) Ich würde solche Hinweise gerne öfter erhalten.

ja

nein

9) Subtratinos Hinweise haben mir dabei geholfen, die richtige Lösung doch noch allein zu finden.

stimmt

stimmt ein bisschen

stimmt nicht

stimmt überhaupt nicht



Vielen Dank für deine Hilfe, Subtratino

Vorbefragungsbogen der Studie 2

Name: _____

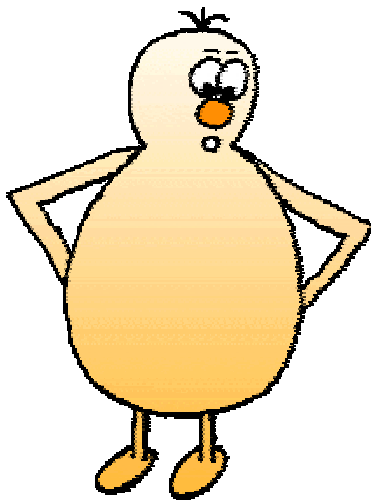
Datum: _____

- Junge
 Mädchen

Alter: _____

Klasse: _____

Schule: _____



Hallo, mein Name ist Subtratino.

Ich bin die Hauptperson in einem Computer lernprogramm zur schriftlichen Subtraktion.

Ich möchte heute von Dir erfahren, wie gut du bereits rechnen kannst.

Aus diesem Grund habe ich Dir ein Arbeitsblatt mitgebracht mit einigen Aufgaben zur schriftlichen Subtraktion. Ich bitte Dich, diese Aufgaben so gut wie möglich zu lösen.

Vorher würde ich Dich aber gern noch ein wenig besser kennen lernen. Es kann nämlich sein, dass wir uns bald wiedersehen.

Aus diesem Grund werde ich dir ein paar Fragen stellen, die du bitte so gewissenhaft wie möglich beantworten sollst.

Bei jeder vorgegebenen Aussage überlege genau, inwieweit diese auf dich zutrifft.

Mache dann bitte ein Kreuz auf die entsprechende Ziffer.

Hier siehst du ein Beispiel, welches ich bereits bearbeitet habe:

Bei der Aussage "Ich sehe die Sesamstraße sehr gern" habe ich die Ziffer 3 angekreuzt, "stimmt etwas".

Mach doch auch noch dein Kreuz über die Ziffer, die für dich zutrifft, und gehe dann zur nächsten Seite. Denn dann wird es ernst.

	stimmt total	stimmt	stimmt etwas	stimmt eher nicht	stimmt nicht	stimmt überhaupt nicht
Ich sehe die Sesamstraße sehr gern.	1	2	X	4	5	6

	stimmt total	stimmt	stimmt etwas	stimmt eher nicht	stimmt nicht	stimmt über- haupt nicht
1. Ich finde das Lösen von Subtraktionsaufgaben sehr einfach.	1	2	3	4	5	6
2. Ich denke, ich bin für das Subtrahieren sehr begabt.	1	2	3	4	5	6
3. Ich denke, ich kann Subtraktionsaufgaben sehr gut lösen.	1	2	3	4	5	6
4. Ich finde das Lösen von Subtraktionsaufgaben sehr interessant.	1	2	3	4	5	6
5. Das Lösen von Subtraktionsaufgaben macht mir sehr viel Spaß.	1	2	3	4	5	6

6. Wie häufig arbeitest oder spielst du am Computer?

- täglich
- 1-2 mal in der Woche
- 1-2 mal im Monat
- seltener
- nie.

Vielen Dank für Deine Mitarbeit!
Jetzt kannst du endlich loslegen mit dem Rechnen!
Viel Glück!

Nachbefragungsbogen Messzeitpunkt 1 und 2 der Studie 2

Name: _____

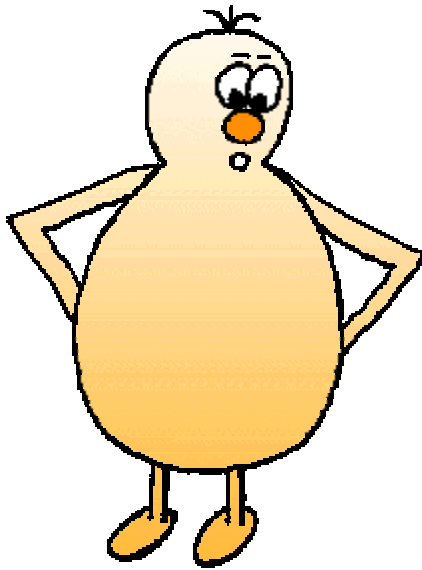
Datum: _____

- Junge
 Mädchen

Alter: _____

Klasse: _____

Schule: _____



Hallo, ich hoffe, du bist noch nicht zu müde, um noch einige Fragen zu beantworten. Ich möchte nämlich von Dir wissen, wie Du die Arbeit mit dem Lernprogramm empfunden hast.

Mache Deine Kreuze bitte sehr gewissenhaft. Deine Antworten sind mir sehr wichtig!

Wie immer noch ein Beispiel zur Erinnerung, obwohl ich natürlich weiß, dass du das schon längst kannst.

Bei der Aussage "Subtratino finde ich einfach süß." habe ich die Ziffer 2 angekreuzt, "stimmt".

Mach doch auch noch dein Kreuz über die Ziffer, die für dich zutrifft, und gehe dann zur nächsten Seite. Denn du weißt, dann wird es ernst.

	stimmt total	stimmt	stimmt etwas	stimmt eher nicht	stimmt nicht	stimmt überhaupt nicht
Subtratino finde ich einfach süß.	1	2	X	4	5	6

	stimmt total	stimmt	stimmt etwas	stimmt eher nicht	stimmt nicht	stimmt über- haupt nicht
1. Ich bin mit meiner Leistung in dieser halben Stunde sehr zufrieden.	1	2	3	4	5	6
2. Ich fand meine Leistung in dieser halben Stunde sehr gut.	1	2	3	4	5	6
3. Ich denke, ich habe in dieser halben Stunde sehr viel dazu-gelernt.	1	2	3	4	5	6
4. Ich fand das Lösen der Subtraktionsaufgaben sehr interes-sant.	1	2	3	4	5	6
5. Mir hat das Lösen der Subtraktionsaufgaben sehr viel Spaß gemacht.	1	2	3	4	5	6
6. Beim Lernen am Computer verging die Zeit wie im Fluge.	1	2	3	4	5	6
7. In dieser halben Stunde bin ich nicht so vorwärtsgekomen wie ich wollte.	1	2	3	4	5	6

Vielen Dank! Bald hast du es geschafft. Streng dich beim Rechnen jetzt noch einmal so richtig an!

8.2 Anhang B: Feedbackinhalte für ITF, KR-KCR, KR-KM und KR-KH Feedback

Fehlerspezifisches ITF

0. Lösungsversuch - auf Fertig-Button gehen, ohne vorher gerechnet zu haben

- Hey, hey erst rechnen, bevor Du diesen Knopf drückst. Wenn du nicht genau weißt wie, lass es dir doch von mir zeigen. Drücke den Rettungsring.

1. Lösungsversuch

- **Aufgabe richtig gelöst**
 - Spitze, richtig gerechnet!
 - Klasse, Dein Ergebnis ist richtig.
 - Sehr schön, Du hast richtig subtrahiert.
 - Wie Du das wieder gemacht hast. Deine Lösung stimmt.
- **Aufgabe falsch gelöst: Feedbacklevel 1: Knowledge of Response (KR)**
 - Oh, oh dein Ergebnis ist leider falsch. Versuchs noch einmal.
 - Oh, Dein Ergebnis ist leider nicht ganz richtig. Versuchs noch einmal.
 - Upps, hier ist was falsch. Versuchs noch einmal.
 - Oh, da hat sich ein Fehler in Dein Ergebnis eingeschlichen. Versuchs noch einmal.
 - Hoppla, in Deiner Lösung steckt noch ein Fehler. Versuchs noch einmal.
 - Oh, ich habe einen Fehler in deiner Lösung entdeckt. Versuchs noch einmal.

2. Lösungsversuch

- **Aufgabe richtig gelöst: KR Feedback und Hinweis auf richtige Lösungsstrategie**
 - ↳ **nach systematischem Fehler und nach kombiniertem systematischen/unsystematischen Fehler**
 - **mit der Null**
Klasse, jetzt stimmt dein Ergebnis! Merk dir, wenn du von der 0 zur oberen Ziffer ergänzt, kommt immer die obere Ziffer heraus. Wenn ich meine leere Bonbonschachtel mit [x] Bonbons auffülle, dann habe ich natürlich am Ende [x] Bonbons in meiner Schachtel.
 - **mit der Leerstelle**
Klasse, jetzt stimmt dein Ergebnis! Merk dir für die nächste Aufgabe: Denke dir für die leere Stelle eine 0 und ergänze von dieser 0 zur oberen Ziffer.
 - **mit gleichen Ziffern**
 - **Übertrag bei gleichen Ziffern**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch eine kleine Erinnerung. Wenn zwei gleiche Ziffern übereinanderstehen, ist ein Übertrag nicht nötig.
 - **Null erste Ziffer vom Ergebnis**
Klasse, jetzt stimmt dein Ergebnis! Denke daran, dass wenn in der ganz linken Spalte des Ergebnisses eine 0 rauskommen würde, diese 0 überhaupt nicht geschrieben wird. Du würdest ja auch nicht schreiben, du bist 08 Jahre alt, sondern du schreibst, du bist 8 Jahre alt.
 - **mit Übertrag**
 - **Übertrag nicht berücksichtigt**
Klasse, jetzt stimmt dein Ergebnis! Wahrscheinlich hast du den Übertrag beim Rechnen einfach vergessen. Notiere die 1 immer, sonst vergisst du sie beim Weiterrechnen.
 - **Fehlerquelle Übertrag, Null im Ergebnis, falsche Rechenrichtung**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn in einer Spalte die obere Ziffer kleiner als die untere Ziffer ist, dann musst du die obere Ziffer um 10 erweitern und eine 1 in die nächste Spalte übertragen.
 - **Übertrag zur oberen Ziffer addiert**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn du eine 1 in die nächste Spalte übertragen hast, musst du diese 1 zur unteren Ziffer dieser Spalte addieren und von diesem Ergebnis dann zur oberen Ziffer dieser Spalte ergänzen.
 - **mit Übertrag in leere Stelle**

- **Übertrag nicht berücksichtigt**

Klasse, jetzt stimmt dein Ergebnis! Wahrscheinlich hast du den Übertrag beim Rechnen einfach vergessen. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von dieser 1 zur oberen Ziffer.

- **Fehlerquelle Übertrag, Null im Ergebnis, falsche Rechenrichtung**

Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn in einer Spalte die obere Ziffer kleiner als die untere Ziffer ist, dann musst du die obere Ziffer um 10 erweitern und eine 1 in die nächste Spalte übertragen.

- **Übertrag zur oberen Ziffer addiert**

Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von der 1 zur oberen Ziffer.

▪ **mit Ergänzen zur 0**

- **Übertrag nicht berücksichtigt**

Klasse, jetzt stimmt dein Ergebnis! Wahrscheinlich hast du den Übertrag beim Rechnen einfach vergessen. Notiere die 1 immer, sonst vergisst du sie beim weiterrechnen.

- **Fehlerquelle Übertrag, Null im Ergebnis, falsche Rechenrichtung**

Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. In dieser Aufgabe ist in der einen Spalte die obere Ziffer – nämlich die 0 – kleiner als die untere Ziffer. Deswegen musst du die 0 um 10 erweitern. Ergänze dann von der unteren Ziffer zur 10. Und vergiss nicht, eine 1 in die nächste Spalte zu übertragen.

- **Übertrag zur oberen Ziffer addiert**

Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn du eine 1 in die nächste Spalte übertragen hast, musst du diese 1 zur unteren Ziffer dieser Spalte addieren und von diesem Ergebnis dann zur oberen Ziffer dieser Spalte ergänzen.

⇒ **nach unsystematischem Fehler: KR Feedback und Hinweis auf metakognitive Strategie**

▪ **addiert statt subtrahiert**

Klasse, jetzt stimmt es. Kann es sein, dass du einfach addiert hast, anstatt zu subtrahieren. Mein Tipp, kontrolliere dein Ergebnis immer sofort, ob es auch kleiner als die oberste Zahl ist. Denke daran, bei der Subtraktion muss das Ergebnis immer kleiner sein als die oberste Zahl.

▪ **unsystematischer Fehler**

Toll, diesmal war es richtig. Einfach verrechnet, oder? Es tut ganz gut, sein Ergebnis immer noch einmal zu kontrollieren. Mach einfach die Umkehroperation, indem du dein Ergebnis zu der unteren Zahl addierst. Wenn deine Lösung stimmt, muss die oberste Zahl rauskommen. Probier es doch das nächste Mal. Dann brauchst du mich gar nicht mehr, sondern findest deinen Fehler allein.

➤ **Aufgaben falsch gelöst: Feedbacklevel 2: KR-KM Feedback**

⇒ **nach systematischem Fehler bzw. {nach kombiniertem systematischen/unsystematischen Fehler}**

▪ **mit der Null**

Oh, oh dein Ergebnis ist leider nicht ganz richtig. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir wenigstens da zu helfen, wo ich früher auch so meine Probleme hatte, nämlich beim Subtrahieren mit der 0.} Schau dir mal die Spalte mit der 0 an. Überlege, kann im Ergebnis eine [x] stehen, wenn du von der 0 zur [y] ergänzt? Angenommen, ich habe 0 Bonbons in meiner Bonbonschachtel. Wie viele Bonbons muss ich dazulegen, um am Schluss [y] Bonbons in meiner Schachtel zu haben? Versuchs noch einmal.

▪ **mit der Leerstelle**

Upps, hier ist noch was falsch. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann.} Schau dir mal das Fragezeichen an und sag mal ehrlich, ist dein Ergebnis nicht ein wenig klein? Mein Tipp bei dieser Aufgabe, denke dir für die leere Stelle eine 0 und ergänze von der 0 zur oberen Ziffer. Versuchs noch einmal.

▪ **mit gleichen Ziffern**

- **Übertrag bei gleichen Ziffern**

Upps, hier ist noch was falsch. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann.} Schau dir mal die Spalte mit den gleichen Ziffern an. Kann es sein, dass du eine 1 in die nächste Spalte übertragen hast? Ein Übertrag ist aber gar nicht notwendig, weil

beide Ziffern gleich groß sind. Wenn du mit zwei gleichen Zahlen subtrahierst, kommt immer 0 heraus. Versuchs noch einmal.

- **Null erste Ziffer im Ergebnis**

Upps, hier ist noch was falsch. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann.} Denke daran, dass wenn in der ganz linken Spalte des Ergebnisses eine 0 rauskommen würde, diese 0 überhaupt nicht geschrieben wird. Du würdest ja auch nicht schreiben, du bist 08 Jahre alt, sondern du schreibst, du bist 8 Jahre alt. Versuchs noch einmal.

▪ **mit Übertrag**

- **Übertrag nicht berücksichtigt**

Oh, oh dein Ergebnis ist leider falsch. {obwohl mir nicht ganz klar ist, wie du eigentlich gerechnet hast. Ich werde versuchen, dir da zu helfen, wo ich früher auch meine Probleme hatte, nämlich beim Übertrag.} Schau dir mal die Spalte mit der roten Ziffer an. Du scheinst die Merkfziffer beim weiteren Rechnen vergessen zu haben. Solange du die schriftliche Subtraktion mit Übertrag noch nicht wie im Schlaf beherrschst, ist es besser, du schreibst die 1 immer noch in die nächste Spalte. Addiere dann die 1 zur [y] und ergänze zur [x]. Versuchs noch einmal.

- **Fehlerquelle, Null im Ergebnis, falsche Rechenrichtung**

Oh, schade, ich entdecke noch einen Fehler. {obwohl mir nicht ganz klar ist, wie du eigentlich gerechnet hast. Ich werde versuchen, dir da zu helfen, wo ich früher auch meine Probleme hatte, nämlich beim Übertrag.} Schau dir doch einmal die Spalten mit den roten Ziffern an. Du hast sicherlich gemerkt, dass die [y] größer als die [x] ist und dass du deshalb nicht einfach von der [y] zur [x] ergänzen kannst. Deswegen musst du in einem nächsten Schritt die [x] um 10 erweitern, indem du zur [x] 10 dazuzählst. Ergänze dann von der [y] zur [x + 10] und vergiss nicht eine 1 in die nächste Spalte zu übertragen. Addiere dann die 1 zur [y] und ergänze von diesem Ergebnis dann zur [x]. Versuchs noch einmal.

- **Übertrag zur oberen Ziffer addiert**

Oh, schade, ich entdecke noch einen Fehler. {obwohl mir nicht ganz klar ist, wie du eigentlich gerechnet hast. Ich werde versuchen, dir da zu helfen, wo ich früher auch meine Probleme hatte, nämlich beim Übertrag.} Schau dir mal die Spalte mit der roten Ziffer an. Wenn du eine 1 in diese Spalte überträgst, musst du die 1 zur [y] addieren und von diesem Ergebnis dann zur [x] ergänzen. Versuchs noch einmal.

▪ **mit Übertrag in leere Stelle**

- **Übertrag nicht berücksichtigt**

Oh, oh dein Ergebnis ist leider falsch. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann.} Schau dir mal die Spalte mit der roten Ziffer an. Du scheinst die Übertragziffer beim weiteren Rechnen vergessen zu haben. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von dieser 1 zur oberen Ziffer. Versuchs noch einmal.

- **Fehlerquelle, Null im Ergebnis, falsche Rechenrichtung**

Oh, schade, ich entdecke noch einen Fehler. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann.} Schau dir doch einmal die Spalten mit den roten Ziffern an. Wenn du die [x] um 10 erweitern musst, weil die [x] kleiner als die [y] ist, dann musst du eine 1 in die nächste Spalte übertragen. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von dieser 1 zur oberen Ziffer. Versuchs noch einmal.

- **Übertrag zur oberen Ziffer addiert**

Oh, schade, ich entdecke noch einen Fehler. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann.} Schau dir mal die Spalte mit der roten Ziffer an. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von dieser 1 zur oberen Ziffer. Versuchs noch einmal.

▪ **mit Ergänzen zur 0**

- **Übertrag nicht berücksichtigt**

Oh, oh dein Ergebnis ist leider falsch. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann.} Schau dir mal die Spalte mit der roten Ziffer an. Du scheinst die Merkfziffer beim weiteren Rechnen vergessen zu haben. Da die 0 kleiner als die [y] ist, musst du die 0 um 10 erweitern. $0 + 10 = 10$. Ergänze dann von der [y] zur 10. Und vergiss nicht, eine 1 in die nächste Spalte zu übertragen. Addiere dann die 1 zur [y] und ergänze von diesem Ergebnis dann zur [x]. Versuchs noch einmal.

- **Fehlerquelle, Null im Ergebnis, falsche Rechenrichtung**

Oh, schade, ich entdecke noch einen Fehler. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann. } Schau dir doch einmal die Spalten mit den roten Ziffern an. Da die 0 kleiner als die [y] ist, musst du die 0 um 10 erweitern. $0 + 10 = 10$. Ergänze dann von der [y] zur 10. Und vergiss nicht, eine 1 in die nächste Spalte zu übertragen. Versuchs noch einmal.

- **Übertrag zur oberen Ziffer addiert**

Oh, schade, ich entdecke noch einen Fehler. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann. } Schau dir mal die Spalte mit der roten Ziffer an. Wenn du eine 1 in diese Spalte überträgst, musst du die 1 zur [y] addieren und von diesem Ergebnis dann zur [x] ergänzen. Versuchs noch einmal.

⇒ **nach unsystematischem Fehler: KR Feedback und Lösungsbeispiel**

▪ **addiert statt subtrahiert**

Oh, leider ist dein Ergebnis nicht ganz richtig. Kann es sein, dass du addiert hast, anstatt zu subtrahieren. Schau immer auf das Rechenzeichen, wenn ein minus steht, dann musst du subtrahieren. Kontrolliere dein Ergebnis auch, indem du schaust, ob dein Ergebnis kleiner als die oberste Zahl ist. Wenn du Probleme mit der Subtraktion hast, lass mich dir doch helfen. Klicke auf den Rettungsring.

▪ **nach unsystematischem Fehler**

Dein Ergebnis ist leider nicht ganz richtig. Allerdings weiß ich auch nicht so genau, wie du gerechnet hast. Vielleicht kann ich dir ja trotzdem helfen?! Wenn du denkst, klicke auf den Rettungsring.

3. Lösungsversuch

➤ **Aufgabe richtig gelöst: KR Feedback und Hinweis auf richtige Lösungsstrategie (siehe 2. Lösungsversuch)**

➤ **Aufgaben falsch gelöst: Feedbacklevel 3: KR-KH Feedback**

⇒ **nach systematischem und/oder unsystematischem Fehler**

▪ **mit der Null**

- **zweistellig**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. Das Ergebnis lautet [c-d a-b].

- **dreistellig**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. Von der [f] bis zur [e] sind [e-f]. Schreibe [e-f] Das Ergebnis lautet [f-e d-c b-a].

▪ **mit der Leerstelle**

- **zweistellig**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. In der nächsten Spalte befindet sich eine leere Stelle. Denke dir für diese leere Stelle eine 0 und rechne. Von der 0 bis zur [c] sind [c]. Schreibe [c]. Das Ergebnis lautet [c b-a].

- **dreistellig**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. In der nächsten Spalte befindet sich eine leere Stelle. Denke dir für diese leere Stelle eine 0 und rechne. Von der 0 bis zur [e] sind [e]. Schreibe [e] Das Ergebnis lautet [e d-c b-a].

▪ **mit gleichen Ziffern**

- **zweistellig**

gleiche Ziffer in Einer-Spalte

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der

[b] bis zur [b] sind 0. Schreibe 0. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. Das Ergebnis lautet [c-d 0].

gleiche Ziffern in Zehner-Spalte

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur (14-18) [d] sind 0. Da die Null aber am Anfang des Ergebnisses stehen würde, wird die 0 nicht geschrieben. [0 a-b] SÄhe doch auch komisch aus, oder? Das Ergebnis lautet einfach [a-b].

- dreistellig

gleiche Ziffern in Einer- und Zehner-Spalte

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. Und von der [f] bis zur [e] sind [e-f]. Schreibe [e-f]. Das Ergebnis lautet [f-e d-c b-a].

gleiche Ziffern in Hunderter-Stelle

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. Und von der [e] bis zur [e] sind 0. Da die Null aber am Anfang des Ergebnisses stehen würde, wird sie nicht geschrieben. [0 c-d a-b] SÄhe doch auch komisch aus, oder? Das Ergebnis lautet einfach [c-d a-b].

▪ **mit Übertrag (Ergänzen zur 0)**

- zweistellig

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a]. Das geht ja gar nicht, denn ich kann nicht von einer größeren Ziffer zu einer kleineren ergänzen. Also muss ich die [a] um 10 erweitern und eine 1 in die nächste Spalte übertragen. $[a] + 10 = [a+10]$. Jetzt kann ich weiterrechnen. Von der [b] bis zur $a+10$ sind $[10+a-b]$. Schreibe $[10+a-b]$. Jetzt addiere ich die übertragene 1 zur [d]. $1 + [d]$ sind $[1+d]$ und rechne weiter. Von der $[d+1]$ bis zur [c] sind $[c-d+1]$. Schreibe $[c-d+1]$. Das Ergebnis lautet $[c-d+1 b-a]$.

- dreistellig

Übertrag an Einerstelle

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a]. Das geht ja gar nicht, denn ich kann nicht von einer größeren Ziffer zu einer kleineren ergänzen. Also muss ich die [a] um 10 erweitern und eine 1 in die nächste Spalte übertragen. $[a] + 10 = [a+10]$. Jetzt kann ich weiterrechnen. Von der [b] bis zur $a+10$ sind $[a+10-b]$. Schreibe $[a+10-b]$. Jetzt addiere ich die übertragene 1 zu [d]. $1 + [d]$ sind $[1+d]$ und rechne weiter. Von der $[d+1]$ zur [c] sind $[c-d+1]$. Schreibe $[c-d+1]$. Von der [f] bis zur [e] sind [e-f]. Schreibe [e-f]. Das Ergebnis lautet $[e-f c-d+1 a+10-b]$.

Übertrag an Zehnerstelle

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Ich schreibe [a-b] und rechne in der nächsten Spalte weiter. Von der [d] bis zur [c]. Das geht ja gar nicht, denn ich kann nicht von einer größeren Ziffer zu einer kleineren ergänzen. Also muss ich die [c] um 10 erweitern und eine 1 in die nächste Spalte übertragen. $[c] + 10 = [c+10]$. Jetzt kann ich weiterrechnen. Von der [d] bis zur $c+10$ sind $[c+10-d]$. Schreibe $[c+10-d]$. Jetzt addiere ich die übertragene 1 zu [f]. $1 + [f]$ sind $[1+f]$ und rechne weiter. Von der $[1+f]$ bis zur [e] sind $[e-f+1]$. Schreibe $[e-f+1]$. Das Ergebnis lautet $[e-f+1 c+10-d a-b]$.

▪ **mit Übertrag in leere Stelle**

- zweistellig

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a]. Das geht ja gar nicht, denn ich kann nicht von einer größeren Ziffer zu einer kleineren ergänzen. Also muss ich die [a] um 10 erweitern und eine 1 in die nächste Spalte übertragen. $[a] + 10 = [a+10]$. Jetzt kann ich weiterrechnen. Von der [b] bis zur $a+10$ sind $[a+10-b]$. Schreibe $[a+10-b]$. Jetzt addiere ich die übertragene 1 zu der unteren Ziffer. Hoppla, da steht ja gar nichts. Aber ich weiß ja, dass ich eine leere Stelle wie eine Null behandeln muss und rechne weiter $1 + 0$ ist 1. Von der 1 bis zur (15-18) [c] sind [c-1]. Schreibe [c-1]. Das Ergebnis lautet $[c-1 a+10-b]$.

- dreistellig

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Ich schreibe [a-b] und rechne in der nächsten Spalte weiter. Von der [d] bis zur [c]. Das geht ja gar nicht, denn ich kann nicht von einer größeren Ziffer zu einer kleineren Ziffer ergänzen. Also muss ich die [c] um 10 erweitern und eine 1 in die nächste Spalte übertragen. $[c] + 10 = [c+10]$. Jetzt kann ich weiterrechnen. Von der [d] bis zur [c+10] sind. $[c+10-d]$. Schreibe $[c+10-d]$. Jetzt addiere ich die übertragene 1 zu der unteren Ziffer. Hoppla, da steht ja gar nichts. Aber ich weiß ja, dass ich eine leere Stelle wie eine Null behandeln muss und rechne weiter $1 + 0$ ist 1. Von der [1] bis zur [e] sind [e-1]. Schreibe [e-1]. Das Ergebnis lautet [e-1 c+10-d a-b].

KR-KCR Feedback

0. Lösungsversuch - auf Fertig-Button gehen, ohne vorher gerechnet zu haben

- Hey, hey erst rechnen, bevor Du diesen Knopf drückst. Wenn du nicht genau weißt wie, lass es dir doch von mir zeigen. Drücke den Rettungsring.

1. Lösungsversuch

- **Aufgabe richtig gelöst**
 - Spitze, richtig gerechnet!
 - Klasse, Dein Ergebnis ist richtig.
 - Sehr schön, Du hast richtig subtrahiert.
 - Wie Du das wieder gemacht hast. Deine Lösung stimmt.
- **Aufgabe falsch gelöst: Feedbacklevel 1: Knowledge of Response (KR)**
 - Oh, oh dein Ergebnis ist leider falsch. Versuchs noch einmal.
 - Oh, Dein Ergebnis ist leider nicht ganz richtig. Versuchs noch einmal.
 - Upps, hier ist was falsch. Versuchs noch einmal.
 - Oh, da hat sich ein Fehler in Dein Ergebnis eingeschlichen. Versuchs noch einmal.
 - Hoppla, in Deiner Lösung steckt noch ein Fehler. Versuchs noch einmal.
 - Oh, ich habe einen Fehler in deiner Lösung entdeckt. Versuchs noch einmal.

2. Lösungsversuch

- **Aufgabe richtig gelöst**
 - Spitze, richtig gerechnet!
 - Klasse, Dein Ergebnis ist richtig.
 - Sehr schön, Du hast richtig subtrahiert.
 - Wie Du das wieder gemacht hast. Deine Lösung stimmt.
- **Aufgabe falsch gelöst: Feedbacklevel 2: KR-KCR Feedback**
 - Oh, oh dein Ergebnis ist leider falsch. Das Ergebnis lautet [z].
 - Oh, Dein Ergebnis ist leider nicht ganz richtig. Das Ergebnis lautet [z].
 - Upps, hier ist was falsch. Das Ergebnis lautet [z].
 - Oh, da hat sich ein Fehler in Dein Ergebnis eingeschlichen. Das Ergebnis lautet [z].
 - Hoppla, in Deiner Lösung steckt noch ein Fehler. Das Ergebnis lautet [z].
 - Oh, ich habe einen Fehler in deiner Lösung entdeckt. Das Ergebnis lautet [z].

KR-KM Feedback

0. Lösungsversuch - auf Fertig-Button gehen, ohne vorher gerechnet zu haben

- Hey, hey erst rechnen, bevor Du diesen Knopf drückst. Wenn du nicht genau weißt wie, lass es dir doch von mir zeigen. Drücke den Rettungsring.

1. Lösungsversuch

- **Aufgabe richtig gelöst**
 - Spitze, richtig gerechnet!
 - Klasse, Dein Ergebnis ist richtig.
 - Sehr schön, Du hast richtig subtrahiert.
 - Wie Du das wieder gemacht hast. Deine Lösung stimmt.
- **Aufgabe falsch gelöst: Feedbacklevel 1: Knowledge of Response (KR)**
 - Oh, oh dein Ergebnis ist leider falsch. Versuchs noch einmal.
 - Oh, Dein Ergebnis ist leider nicht ganz richtig. Versuchs noch einmal.
 - Upps, hier ist was falsch. Versuchs noch einmal.
 - Oh, da hat sich ein Fehler in Dein Ergebnis eingeschlichen. Versuchs noch einmal.
 - Hoppla, in Deiner Lösung steckt noch ein Fehler. Versuchs noch einmal.
 - Oh, ich habe einen Fehler in deiner Lösung entdeckt. Versuchs noch einmal.

2. Lösungsversuch

- **Aufgabe richtig gelöst: KR Feedback und Hinweis auf richtige Lösungsstrategie**
 - ⇒ **nach systematischem Fehler und nach kombiniertem systematischen/unsystematischen Fehler**
 - **mit der Null**
Klasse, jetzt stimmt dein Ergebnis! Merk dir, wenn du von der 0 zur oberen Ziffer ergänzt, kommt immer die obere Ziffer heraus. Wenn ich meine leere Bonbonschachtel mit [x] Bonbons auffülle, dann habe ich natürlich am Ende [x] Bonbons in meiner Schachtel.
 - **mit der Leerstelle**
Klasse, jetzt stimmt dein Ergebnis! Merk dir für die nächste Aufgabe: Denke dir für die leere Stelle eine 0 und ergänze von dieser 0 zur oberen Ziffer.
 - **mit gleichen Ziffern**
 - **Übertrag bei gleichen Ziffern**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch eine kleine Erinnerung. Wenn zwei gleiche Ziffern übereinanderstehen, ist ein Übertrag nicht nötig.
 - **Null erste Ziffer vom Ergebnis**
Klasse, jetzt stimmt dein Ergebnis! Denke daran, dass wenn in der ganz linken Spalte des Ergebnisses eine 0 rauskommen würde, diese 0 überhaupt nicht geschrieben wird. Du würdest ja auch nicht schreiben, du bist 08 Jahre alt, sondern du schreibst, du bist 8 Jahre alt.
 - **mit Übertrag**
 - **Übertrag nicht berücksichtigt**
Klasse, jetzt stimmt dein Ergebnis! Wahrscheinlich hast du den Übertrag beim Rechnen einfach vergessen. Notiere die 1 immer, sonst vergisst du sie beim Weiterrechnen.
 - **Fehlerquelle Übertrag, Null im Ergebnis, falsche Rechenrichtung**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn in einer Spalte die obere Ziffer kleiner als die untere Ziffer ist, dann musst du die obere Ziffer um 10 erweitern und eine 1 in die nächste Spalte übertragen.
 - **Übertrag zur oberen Ziffer addiert**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn du eine 1 in die nächste Spalte übertragen hast, musst du diese 1 zur unteren Ziffer dieser Spalte addieren und von diesem Ergebnis dann zur oberen Ziffer dieser Spalte ergänzen.
 - **mit Übertrag in leere Stelle**
 - **Übertrag nicht berücksichtigt**
Klasse, jetzt stimmt dein Ergebnis! Wahrscheinlich hast du den Übertrag beim Rechnen einfach vergessen. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von dieser 1 zur oberen Ziffer.

- **Fehlerquelle Übertrag, Null im Ergebnis, falsche Rechenrichtung**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn in einer Spalte die obere Ziffer kleiner als die untere Ziffer ist, dann musst du die obere Ziffer um 10 erweitern und eine 1 in die nächste Spalte übertragen.
- **Übertrag zur oberen Ziffer addiert**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von der 1 zur oberen Ziffer.
- **mit Ergänzen zur 0**
 - **Übertrag nicht berücksichtigt**
Klasse, jetzt stimmt dein Ergebnis! Wahrscheinlich hast du den Übertrag beim Rechnen einfach vergessen. Notiere die 1 immer, sonst vergisst du sie beim weiterrechnen.
 - **Fehlerquelle Übertrag, Null im Ergebnis, falsche Rechenrichtung**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. In dieser Aufgabe ist in der einen Spalte die obere Ziffer – nämlich die 0 - kleiner als die untere Ziffer. Deswegen musst du die 0 um 10 erweitern. Ergänze dann von der unteren Ziffer zur 10. Und vergiss nicht, eine 1 in die nächste Spalte zu übertragen.
 - **Übertrag zur oberen Ziffer addiert**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn du eine 1 in die nächste Spalte übertragen hast, musst du diese 1 zur unteren Ziffer dieser Spalte addieren und von diesem Ergebnis dann zur oberen Ziffer dieser Spalte ergänzen.
- ⇒ **nach unsystematischem Fehler: KR Feedback und Hinweis auf metakognitive Strategie**
 - **addiert statt subtrahiert**
Klasse, jetzt stimmt es. Kann es sein, dass du einfach addiert hast, anstatt zu subtrahieren. Mein Tipp, kontrolliere dein Ergebnis immer sofort, ob es auch kleiner als die oberste Zahl ist. Denke daran, bei der Subtraktion muss das Ergebnis immer kleiner sein als die oberste Zahl.
 - **unsystematischer Fehler**
Toll, diesmal war es richtig. Einfach verrechnet, oder? Es tut ganz gut, sein Ergebnis immer noch einmal zu kontrollieren. Mach einfach die Umkehroperation, indem du dein Ergebnis zu der unteren Zahl addierst. Wenn deine Lösung stimmt, muss die oberste Zahl rauskommen. Probier es doch das nächste Mal. Dann brauchst du mich gar nicht mehr, sondern findest deinen Fehler allein.
- **Aufgaben falsch gelöst: Feedbacklevel 2: KR-KM Feedback**
 - ⇒ **nach systematischem Fehler bzw. {nach kombiniertem systematischen/unsystematischen Fehler}**
 - **mit der Null**
Oh, oh dein Ergebnis ist leider nicht ganz richtig. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir wenigstens da zu helfen, wo ich früher auch so meine Probleme hatte, nämlich beim Subtrahieren mit der 0.} Schau dir mal die Spalte mit der 0 an. Überlege, kann im Ergebnis eine [x] stehen, wenn du von der 0 zur [y] ergänzt? Angenommen, ich habe 0 Bonbons in meiner Bonbonschachtel. Wie viele Bonbons muss ich dazulegen, um am Schluss [y] Bonbons in meiner Schachtel zu haben? Das Ergebnis lautet [z].
 - **mit der Leerstelle**
Upps, hier ist noch was falsch. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann.} Schau dir mal das Fragezeichen an und sag mal ehrlich, ist dein Ergebnis nicht ein wenig klein? Mein Tipp bei dieser Aufgabe, denke dir für die leere Stelle eine 0 und ergänze von der 0 zur oberen Ziffer. Das Ergebnis lautet [z].
 - **mit gleichen Ziffern**
 - **Übertrag bei gleichen Ziffern**
Upps, hier ist noch was falsch. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann.} Schau dir mal die Spalte mit den gleichen Ziffern an. Kann es sein, dass du eine 1 in die nächste Spalte übertragen hast? Ein Übertrag ist aber gar nicht notwendig, weil beide Ziffern gleich groß sind. Wenn du mit zwei gleichen Zahlen subtrahierst, kommt immer 0 heraus. Das Ergebnis lautet [z].
 - **Null erste Ziffer im Ergebnis**
Upps, hier ist noch was falsch. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann.} Denke daran, dass wenn in der ganz linken Spalte des Ergebnisses

eine 0 rauskommen würde, diese 0 überhaupt nicht geschrieben wird. Du würdest ja auch nicht schreiben, du bist 08 Jahre alt, sondern du schreibst, du bist 8 Jahre alt. Das Ergebnis lautet [z].

- **mit Übertrag**

- **Übertrag nicht berücksichtigt**

Oh, oh dein Ergebnis ist leider falsch. {obwohl mir nicht ganz klar ist, wie du eigentlich gerechnet hast. Ich werde versuchen, dir da zu helfen, wo ich früher auch meine Probleme hatte, nämlich beim Übertrag. } Schau dir mal die Spalte mit der roten Ziffer an. Du scheinst die Merzkiffer beim weiteren Rechnen vergessen zu haben. Solange du die schriftliche Subtraktion mit Übertrag noch nicht wie im Schlaf beherrschst, ist es besser, du schreibst die 1 immer noch in die nächste Spalte. Addiere dann die 1 zur [y] und ergänze zur [x]. Das Ergebnis lautet [z].

- **Fehlerquelle, Null im Ergebnis, falsche Rechenrichtung**

Oh, schade, ich entdecke noch einen Fehler. {obwohl mir nicht ganz klar ist, wie du eigentlich gerechnet hast. Ich werde versuchen, dir da zu helfen, wo ich früher auch meine Probleme hatte, nämlich beim Übertrag. } Schau dir doch einmal die Spalten mit den roten Ziffern an. Du hast sicherlich gemerkt, dass die [y] größer als die [x] ist und dass du deshalb nicht einfach von der [y] zur [x] ergänzen kannst. Deswegen musst du in einem nächsten Schritt die [x] um 10 erweitern, indem du zur [x] 10 dazuzählst. Ergänze dann von der [y] zur [x + 10] und vergiss nicht eine 1 in die nächste Spalte zu übertragen. Addiere dann die 1 zur [y] und ergänze von diesem Ergebnis dann zur [x]. Das Ergebnis lautet [z].

- **Übertrag zur oberen Ziffer addiert**

Oh, schade, ich entdecke noch einen Fehler. {obwohl mir nicht ganz klar ist, wie du eigentlich gerechnet hast. Ich werde versuchen, dir da zu helfen, wo ich früher auch meine Probleme hatte, nämlich beim Übertrag. } Schau dir mal die Spalte mit der roten Ziffer an. Wenn du eine 1 in diese Spalte überträgst, musst du die 1 zur [y] addieren und von diesem Ergebnis dann zur [x] ergänzen. Das Ergebnis lautet [z].

- **mit Übertrag in leere Stelle**

- **Übertrag nicht berücksichtigt**

Oh, oh dein Ergebnis ist leider falsch. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann. } Schau dir mal die Spalte mit der roten Ziffer an. Du scheinst die Übertragziffer beim weiteren Rechnen vergessen zu haben. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von dieser 1 zur oberen Ziffer. Das Ergebnis lautet [z].

- **Fehlerquelle, Null im Ergebnis, falsche Rechenrichtung**

Oh, schade, ich entdecke noch einen Fehler. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann. } Schau dir doch einmal die Spalten mit den roten Ziffern an. Wenn du die [x] um 10 erweitern musst, weil die [x] kleiner als die [y] ist, dann musst du eine 1 in die nächste Spalte übertragen. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von dieser 1 zur oberen Ziffer. Das Ergebnis lautet [z].

- **Übertrag zur oberen Ziffer addiert**

Oh, schade, ich entdecke noch einen Fehler. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann. } Schau dir mal die Spalte mit der roten Ziffer an. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von dieser 1 zur oberen Ziffer. Das Ergebnis lautet [z].

- **mit Ergänzen zur 0**

- **Übertrag nicht berücksichtigt**

Oh, oh dein Ergebnis ist leider falsch. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann. } Schau dir mal die Spalte mit der roten Ziffer an. Du scheinst die Merzkiffer beim weiteren Rechnen vergessen zu haben. Da die 0 kleiner als die [y] ist, musst du die 0 um 10 erweitern. $0 + 10 = 10$. Ergänze dann von der [y] zur 10. Und vergiss nicht, eine 1 in die nächste Spalte zu übertragen. Addiere dann die 1 zur [y] und ergänze von diesem Ergebnis dann zur [x]. Das Ergebnis lautet [z].

- **Fehlerquelle, Null im Ergebnis, falsche Rechenrichtung**

Oh, schade, ich entdecke noch einen Fehler. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann. } Schau dir doch einmal die Spalten mit den roten Ziffern an. Da die 0 kleiner als die [y] ist, musst du die 0 um 10 erweitern. $0 + 10 = 10$. Ergänze dann von der [y] zur 10. Und vergiss nicht, eine 1 in die nächste Spalte zu übertragen. Das Ergebnis lautet [z].

- Übertrag zur oberen Ziffer addiert

Oh, schade, ich entdecke noch einen Fehler. {Du machst ja wirklich sehr geheimnisvolle Fehler. Ich versuch dir mal zu helfen, so gut ich kann. } Schau dir mal die Spalte mit der roten Ziffer an. Wenn du eine 1 in diese Spalte überträgst, musst du die 1 zur [y] addieren und von diesem Ergebnis dann zur [x] ergänzen. Das Ergebnis lautet [z].

⇒ nach unsystematischem Fehler: KR Feedback und Lösungsbeispiel**▪ addiert statt subtrahiert**

Oh, leider ist dein Ergebnis nicht ganz richtig. Kann es sein, dass du addiert hast, anstatt zu subtrahieren. Schau immer auf das Rechenzeichen, wenn ein minus steht, dann musst du subtrahieren. Kontrolliere dein Ergebnis auch, indem du schaust, ob dein Ergebnis kleiner als die oberste Zahl ist. Das Ergebnis lautet [z].

▪ nach unsystematischem Fehler

Dein Ergebnis ist leider nicht ganz richtig. Allerdings weiß ich auch nicht so genau, wie du gerechnet hast. Das Ergebnis lautet [z].

KR-KH Feedback

0. Lösungsversuch - auf Fertig-Button gehen, ohne vorher gerechnet zu haben

- Hey, hey erst rechnen, bevor Du diesen Knopf drückst. Wenn du nicht genau weißt wie, lass es dir doch von mir zeigen. Drücke den Rettungsring.

1. Lösungsversuch

- **Aufgabe richtig gelöst**
 - Spitze, richtig gerechnet!
 - Klasse, Dein Ergebnis ist richtig.
 - Sehr schön, Du hast richtig subtrahiert.
 - Wie Du das wieder gemacht hast. Deine Lösung stimmt.
- **Aufgabe falsch gelöst: Feedbacklevel 1: Knowledge of Response (KR)**
 - Oh, oh dein Ergebnis ist leider falsch. Versuchs noch einmal.
 - Oh, Dein Ergebnis ist leider nicht ganz richtig. Versuchs noch einmal.
 - Upps, hier ist was falsch. Versuchs noch einmal.
 - Oh, da hat sich ein Fehler in Dein Ergebnis eingeschlichen. Versuchs noch einmal.
 - Hoppla, in Deiner Lösung steckt noch ein Fehler. Versuchs noch einmal.
 - Oh, ich habe einen Fehler in deiner Lösung entdeckt. Versuchs noch einmal.

2. Lösungsversuch

- **Aufgabe richtig gelöst: KR Feedback und Hinweis auf richtige Lösungsstrategie**
 - ⇒ **nach systematischem Fehler und nach kombiniertem systematischen/unsystematischen Fehler**
 - **mit der Null**
Klasse, jetzt stimmt dein Ergebnis! Merk dir, wenn du von der 0 zur oberen Ziffer ergänzt, kommt immer die obere Ziffer heraus. Wenn ich meine leere Bonbonschachtel mit [x] Bonbons auffülle, dann habe ich natürlich am Ende [x] Bonbons in meiner Schachtel.
 - **mit der Leerstelle**
Klasse, jetzt stimmt dein Ergebnis! Merk dir für die nächste Aufgabe: Denke dir für die leere Stelle eine 0 und ergänze von dieser 0 zur oberen Ziffer.
 - **mit gleichen Ziffern**
 - **Übertrag bei gleichen Ziffern**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch eine kleine Erinnerung. Wenn zwei gleiche Ziffern übereinanderstehen, ist ein Übertrag nicht nötig.
 - **Null erste Ziffer vom Ergebnis**
Klasse, jetzt stimmt dein Ergebnis! Denke daran, dass wenn in der ganz linken Spalte des Ergebnisses eine 0 rauskommen würde, diese 0 überhaupt nicht geschrieben wird. Du würdest ja auch nicht schreiben, du bist 08 Jahre alt, sondern du schreibst, du bist 8 Jahre alt.
 - **mit Übertrag**
 - **Übertrag nicht berücksichtigt**
Klasse, jetzt stimmt dein Ergebnis! Wahrscheinlich hast du den Übertrag beim Rechnen einfach vergessen. Notiere die 1 immer, sonst vergisst du sie beim Weiterrechnen.
 - **Fehlerquelle Übertrag, Null im Ergebnis, falsche Rechenrichtung**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn in einer Spalte die obere Ziffer kleiner als die untere Ziffer ist, dann musst du die obere Ziffer um 10 erweitern und eine 1 in die nächste Spalte übertragen.
 - **Übertrag zur oberen Ziffer addiert**
Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn du eine 1 in die nächste Spalte übertragen hast, musst du diese 1 zur unteren Ziffer dieser Spalte addieren und von diesem Ergebnis dann zur oberen Ziffer dieser Spalte ergänzen.
 - **mit Übertrag in leere Stelle**
 - **Übertrag nicht berücksichtigt**
Klasse, jetzt stimmt dein Ergebnis! Wahrscheinlich hast du den Übertrag beim Rechnen einfach vergessen. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von dieser 1 zur oberen Ziffer.

- **Fehlerquelle Übertrag, Null im Ergebnis, falsche Rechenrichtung**

Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn in einer Spalte die obere Ziffer kleiner als die untere Ziffer ist, dann musst du die obere Ziffer um 10 erweitern und eine 1 in die nächste Spalte übertragen.

- **Übertrag zur oberen Ziffer addiert**

Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn du eine 1 in eine Spalte überträgst, an der die untere Ziffernstelle leer ist, denke dir für diese leere Stelle eine 0 und rechne: $1 + 0 = 1$. Ergänze dann von der 1 zur oberen Ziffer.

▪ **mit Ergänzen zur 0**

- **Übertrag nicht berücksichtigt**

Klasse, jetzt stimmt dein Ergebnis! Wahrscheinlich hast du den Übertrag beim Rechnen einfach vergessen. Notiere die 1 immer, sonst vergisst du sie beim weiterrechnen.

- **Fehlerquelle Übertrag, Null im Ergebnis, falsche Rechenrichtung**

Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. In dieser Aufgabe ist in der einen Spalte die obere Ziffer – nämlich die 0 - kleiner als die untere Ziffer. Deswegen musst du die 0 um 10 erweitern. Ergänze dann von der unteren Ziffer zur 10. Und vergiss nicht, eine 1 in die nächste Spalte zu übertragen.

- **Übertrag zur oberen Ziffer addiert**

Spitze, jetzt stimmt dein Ergebnis! Fürs nächste Mal hier noch einmal eine kleine Erinnerung. Wenn du eine 1 in die nächste Spalte übertragen hast, musst du diese 1 zur unteren Ziffer dieser Spalte addieren und von diesem Ergebnis dann zur oberen Ziffer dieser Spalte ergänzen.

⇒ **nach unsystematischem Fehler: KR Feedback und Hinweis auf metakognitive Strategie**

▪ **addiert statt subtrahiert**

Klasse, jetzt stimmt es. Kann es sein, dass du einfach addiert hast, anstatt zu subtrahieren. Mein Tipp, kontrolliere dein Ergebnis immer sofort, ob es auch kleiner als die oberste Zahl ist. Denke daran, bei der Subtraktion muss das Ergebnis immer kleiner sein als die oberste Zahl.

▪ **unsystematischer Fehler**

Toll, diesmal war es richtig. Einfach verrechnet, oder? Es tut ganz gut, sein Ergebnis immer noch einmal zu kontrollieren. Mach einfach die Umkehroperation, indem du dein Ergebnis zu der unteren Zahl addierst. Wenn deine Lösung stimmt, muss die oberste Zahl rauskommen. Probier es doch das nächste Mal. Dann brauchst du mich gar nicht mehr, sondern findest deinen Fehler allein.

➤ **Aufgaben falsch gelöst: Feedbacklevel 2: KR-KH Feedback**

⇒ **nach systematischem und/oder unsystematischem Fehler**

▪ **mit der Null**

- **zweistellig**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. Das Ergebnis lautet [c-d a-b].

- **dreistellig**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. Von der [f] bis zur [e] sind [e-f]. Schreibe [e-f] Das Ergebnis lautet [f-e d-c b-a].

▪ **mit der Leerstelle**

- **zweistellig**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. In der nächsten Spalte befindet sich eine leere Stelle. Denke dir für diese leere Stelle eine 0 und rechne. Von der 0 bis zur [c] sind [c]. Schreibe [c]. Das Ergebnis lautet [c b-a].

- **dreistellig**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. In der

nächsten Spalte befindet sich eine leere Stelle. Denke dir für diese leere Stelle eine 0 und rechne. Von der 0 bis zur [e] sind [e]. Schreibe [e] Das Ergebnis lautet [e d-c b-a].

- **mit gleichen Ziffern**

- **zweistellig**

- gleiche Ziffer in Einer-Spalte**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [b] sind 0. Schreibe 0. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. Das Ergebnis lautet [c-d 0].

- gleiche Ziffern in Zehner-Spalte**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur (14-18) [d] sind 0. Da die Null aber am Anfang des Ergebnisses stehen würde, wird die 0 nicht geschrieben. [0 a-b] Sähe doch auch komisch aus, oder? Das Ergebnis lautet einfach [a-b].

- **dreistellig**

- gleiche Ziffern in Einer- und Zehner-Spalte**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. Und von der [f] bis zur [e] sind [e-f]. Schreibe [e-f]. Das Ergebnis lautet [f-e d-c b-a].

- gleiche Ziffern in Hunderter-Stelle**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Schreibe [a-b]. Von der [d] bis zur [c] sind [c-d]. Schreibe [c-d]. Und von der [e] bis zur [e] sind 0. Da die Null aber am Anfang des Ergebnisses stehen würde, wird sie nicht geschrieben. [0 c-d a-b] Sähe doch auch komisch aus, oder? Das Ergebnis lautet einfach [c-d a-b].

- **mit Übertrag (Ergänzen zur 0)**

- **zweistellig**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a]. Das geht ja gar nicht, denn ich kann nicht von einer größeren Ziffer zu einer kleineren ergänzen. Also muss ich die [a] um 10 erweitern und eine 1 in die nächste Spalte übertragen. [a] + 10 = [a+10]. Jetzt kann ich weiterrechnen. Von der [b] bis zur a+10] sind [10+a-b]. Schreibe [10+a-b]. Jetzt addiere ich die übertragene 1 zur [d]. 1 + [d] sind [1+d] und rechne weiter. Von der [d+1] bis zur [c] sind [c-d+1]. Schreibe [c-d+1]. Das Ergebnis lautet [c-d+1 b-a].

- **dreistellig**

- Übertrag an Einerstelle**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a]. Das geht ja gar nicht, denn ich kann nicht von einer größeren Ziffer zu einer kleineren ergänzen. Also muss ich die [a] um 10 erweitern und eine 1 in die nächste Spalte übertragen. [a] + 10 = [a+10]. Jetzt kann ich weiterrechnen. Von der [b] bis zur [a+10] sind. [a+10-b]. Schreibe [a+10-b]. Jetzt addiere ich die übertragene 1 zu [d]. 1 + [d] sind [1+d] und rechne weiter. Von der [d+1] zur [c] sind [c-d+1]. Schreibe [c-d+1]. Von der [f] bis zur [e] sind [e-f]. Schreibe [e-f]. Das Ergebnis lautet [e-f c-d+1 a+10-b].

- Übertrag an Zehnerstelle**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind [a-b]. Ich schreibe [a-b] und rechne in der nächsten Spalte weiter. Von der [d] bis zur [c]. Das geht ja gar nicht, denn ich kann nicht von einer größeren Ziffer zu einer kleineren ergänzen. Also muss ich die [c] um 10 erweitern und eine 1 in die nächste Spalte übertragen. [c] + 10 = [c+10]. Jetzt kann ich weiterrechnen. Von der [d] bis zur [c+10] sind [c+10-d]. Schreibe [c+10-d]. Jetzt addiere ich die übertragene 1 zu [f]. 1 + [f] sind [1+f] und rechne weiter. Von der [1+f] bis zur [e] sind [e-f+1]. Schreibe [e-f+1]. Das Ergebnis lautet [e-f+1 c+10-d a-b].

- **mit Übertrag in leere Stelle**

- **zweistellig**

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a]. Das geht ja gar nicht, denn ich kann nicht von einer größeren Ziffer zu einer kleineren ergänzen. Also muss ich die [a] um 10 erweitern und eine 1 in die nächste Spalte übertragen. $[a] + 10 = [a+10]$. Jetzt kann ich weiterrechnen. Von der [b] bis zur [a+10] sind $[a+10-b]$. Schreibe $[a+10-b]$. Jetzt addiere ich die übertragene 1 zu der unteren Ziffer. Hoppla, da steht ja gar nichts. Aber ich weiß ja, dass ich eine leere Stelle wie eine Null behandeln muss und rechne weiter $1 + 0$ ist 1. Von der 1 bis zur (15-18) [c] sind $[c-1]$. Schreibe $[c-1]$. Das Ergebnis lautet $[c-1 \ a+10-b]$.

- dreistellig

Oh, schade, doch noch ein Fehler. Meine Erklärung war wohl nicht ganz so hilfreich. Ok, ich rechne dir mal vor, wie diese Aufgabe gelöst wird. Pass gut auf. Ich beginne in der rechten Spalte. Von der [b] bis zur [a] sind $[a-b]$. Ich schreibe $[a-b]$ und rechne in der nächsten Spalte weiter. Von der [d] bis zur [c]. Das geht ja gar nicht, denn ich kann nicht von einer größeren Ziffer zu einer kleineren Ziffer ergänzen. Also muss ich die [c] um 10 erweitern und eine 1 in die nächste Spalte übertragen. $[c] + 10 = [c+10]$. Jetzt kann ich weiterrechnen. Von der [d] bis zur [c+10] sind $[c+10-d]$. Schreibe $[c+10-d]$. Jetzt addiere ich die übertragene 1 zu der unteren Ziffer. Hoppla, da steht ja gar nichts. Aber ich weiß ja, dass ich eine leere Stelle wie eine Null behandeln muss und rechne weiter $1 + 0$ ist 1. Von der [1] bis zur [e] sind $[e-1]$. Schreibe $[e-1]$. Das Ergebnis lautet $[e-1 \ c+10-d \ a-b]$.