

Einflüsse der Verarbeitungsreihenfolge auf den Wissenserwerb mit Texten und Bildern

**Dissertationsschrift zur Erlangung des akademischen Grades
Doktor der Philosophie (Dr.phil.)**

vorgelegt von
Mark Ullrich

Vom Promotionsausschuss des Fachbereichs Psychologie der Universität
Koblenz-Landau zur Verleihung des akademischen Grades Doktor der
Philosophie (Dr. phil.) genehmigten Dissertation

Datum der wissenschaftlichen Aussprache: 05.05.2011

Vorsitzender des Promotionsausschusses: Prof. Dr. M. Schmitt
Erster Berichterstatter: Prof. Dr. W. Schnotz
Zweiter Berichterstatter: Prof. Dr. H. Horz

Danksagung

An dieser Stelle möchte ich mich bei den Menschen bedanken, die zur Entstehung der vorliegenden Arbeit beigetragen haben. Dabei gilt mein besonderer Dank Prof. Dr. Schnotz, der mich als Betreuer theoretisch und methodisch während meiner ganzen Promotionszeit begleitet hat. Seine Anregungen und Rückmeldungen haben mir bei der Planung, Durchführung und Umsetzung meines Forschungsvorhabens außerordentlich geholfen. Seine Geduld und sein Zutrauen werde ich immer in Erinnerung behalten.

Ebenso danken möchte ich natürlich meinem zweiten Betreuer, Prof. Dr. Holger Horz, der mich zur rechten Zeit motivieren konnte, die Arbeit endlich zu Papier zu bringen. Auch er war gerade in der Schlussphase äußerst langmütig mit mir, wenn ich den Abgabetermin ein ums andere Mal verschoben habe.

Bedanken muss ich mich auch bei all meinen Kolleginnen und Kollegen der Universität Koblenz-Landau, die mir mit Tipps und Ratschlägen stets zur Seite standen. Besonders hervorzuheben sind dabei Dr. Thorsten Rasch, der großen Anteil am Gelingen der Blickbewegungsstudie hatte, Dr. Alwine Lenzner, deren moralische Unterstützung unbezahlbar war, und Carolin Hörnicke, die mir als wissenschaftliche Hilfskraft viel Arbeit bei der Durchführung der Studien abnehmen konnte.

Zum Schluss möchte ich mich bei meinen Eltern, Luzia und Christian Ullrich, bedanken, die immer Geduld und Verständnis für mich hatten. Mein größter Dank gebührt jedoch meiner Frau Karina, die mir vor nur sechs Tagen mit Patrycja und Ella zwei wundervolle Töchter geboren hat.

Frankfurt, 04.01.2011

Mark Ullrich

Inhaltsverzeichnis

1	Einleitung	3
I	Theoretische Vorüberlegungen	7
2	Lernen - Verstehen - Erinnern	9
2.1	Lernen als Wissenserwerb	9
2.2	Verstehen als mentale Repräsentation	10
2.2.1	Deklaratives und Prozedurales Wissen	12
2.2.2	Repräsentation von Wissen	13
2.3	Das Gedächtnis als kognitive Architektur	23
2.3.1	Das Mehrspeichermodell	23
2.3.2	Levels of Processing	25
2.3.3	Das Arbeitsgedächtnis	26
2.3.4	Das Langzeitgedächtnis	30
2.3.5	Die Cognitive Load Theory	37
3	Verarbeitung von Texten und Bildern	43
3.1	Textverstehen	44
3.1.1	Textbasis, zyklische Verarbeitung und semantische Strukturen	46
3.1.2	Modellkonstruktion beim Textverstehen	49
3.2	Bildverstehen	52
3.2.1	Arten und Funktionen von Bildern	53
3.2.2	Kognitive Prozesse bei der Verarbeitung von bildlichen Informationen	57
3.3	Integrierte Verarbeitung von Text- und Bildinformationen	61
3.3.1	Die Theorie der Dualen Codierung	62
3.3.2	Modell des Textlernens unter Verwendung geographischer Karten	66
3.3.3	Kognitive Theorie des multimedialen Lernens	70
3.3.4	Das integrierte Modell des Text- und Bildverstehens	77
4	Fragestellung	87
4.1	Der picture-text-sequencing Effekt	87
4.1.1	Die Kapazitäts-Hypothese	88
4.1.2	Die Interferenz-Hypothese	90
4.2	Überlegungen zur Versuchsplanung	92
4.2.1	Merkmale des Lehrmaterials	92
4.2.2	Das Informationsverhältnis und die Kapazitätshypothese	93

4.2.3	Das Informationsverhältnis und die Interferenzhypothese	94
4.3	Schlussfolgerungen für die Gestaltung des Untersuchungsmaterials	94
II	Empirie	97
5	Untersuchungsmaterial und Instrumente	99
5.1	Die Theorie der Plattentektonik in Texten und Bildern	99
5.1.1	Auswahl der Domäne	99
5.1.2	Prinzipien der Textgestaltung	100
5.1.3	Prinzipien der Bildgestaltung	101
5.1.4	Kongruente und komplementäre Bilder und Texte	103
5.1.5	Evaluation des Lehrmaterials	106
5.2	Instrumente	110
5.2.1	Lern- und Wissenstest (post)	110
5.2.2	Mentale Belastung (post)	120
5.2.3	Vorwissen (prae)	121
5.2.4	FAM Fragebogen zur aktuellen Motivation (post)	122
5.2.5	Verbale und räumliche Intelligenz (prae)	124
5.2.6	Lernstil (prae)	126
5.2.7	Fragen zu Lerngewohnheiten (prae) & Fragen zum Lernen während der Präsentationsphase (post)	128
5.2.8	Soziodemographische Daten (prae)	128
6	Studie I: Interferenzen versus Kapazität	131
6.1	Fragestellung und Hypothesen	131
6.2	Versuchsplanung	134
6.2.1	Design	134
6.2.2	Stichprobe	136
6.2.3	Ablauf und Durchführung der Untersuchung	137
6.2.4	Statistische Auswertung	139
6.3	Ergebnisse	140
6.3.1	Auswahl der der Kontrollvariablen	140
6.3.2	Forschungsfrage 1: Behalten & Verstehen	143
6.3.3	Forschungsfrage 2: Mentale Belastung (<i>cognitive load</i>)	147
6.3.4	Forschungsfrage 3: Lernzeiten	148
6.3.5	Weitere Analysen	153
6.4	Diskussion	156
7	Studie II: Interferenzen versus Konstruktion	163
7.1	Fragestellung und Hypothesen	163
7.2	Versuchsplanung	166
7.2.1	Design	166

7.2.2	Stichprobe	167
7.2.3	Ablauf und Durchführung der Untersuchung	168
7.2.4	Statistische Auswertung	168
7.3	Ergebnisse	169
7.3.1	Text versus Bild	169
7.3.2	Interferenz-Hypothese	172
7.3.3	Konstruktions-Hypothese	177
7.4	Diskussion	181
8	Studie III: Interferenzen & Blickbewegungen	185
8.1	Fragestellung	185
8.2	Versuchsplanung	187
8.2.1	Design	187
8.2.2	Stichprobe	188
8.2.3	Durchführung	188
8.2.4	Auswertung	190
8.3	Ergebnisse	190
8.3.1	Wissenstest, mentale Belastung und Lernzeit	190
8.3.2	Blickbewegungen	193
8.4	Diskussion	198
III	Gesamtdiskussion	203
9	Fazit	205
9.1	Integrierte Deutung der Resultate	205
9.2	Methodische Kritik	211
9.3	Praktische Implikationen	212
9.4	Ausblick	214
IV	Anhang	233
A	Lernmaterial	235
A.1	Kongruente Bilder & Texte	235
A.2	Komplementäre Bilder & Texte	244
B	Instrumente	253
B.1	Wissenstest	253
B.2	Weitere Instrumente	263

Tabellenverzeichnis

5.1	Textlänge der vollständigen und komplementären Versionen	105
5.2	Interne Konsistenz und durchschnittliche Itemschwierigkeit der ersten Version des Lern- und Wissenstests	113
5.3	Auswertungsobjektivität aller Items mit offenem Format	114
5.4	Überblick über die Skalen des Wissenstests	115
5.5	Korrelationen der Skalen des Wissenstests untereinander	117
5.6	Korrelationskoeffizienten zur Prüfung der konvergenten Validität des Lern- und Wissenstest	119
5.7	Eigenschaften der Testskala Vorwissen	122
5.8	Eigenschaften der Testskala aktuelle Motivation	123
5.9	Stichprobenmerkmale der IST 2000 R Intelligenzsskalen	126
5.10	Merkmale der VVQ-Skalen nach Kirby et al. (1988)	127
5.11	Allgemeine Stichprobenmerkmale	128
6.1	Kolmogorow-Smirnow- und Levene-Tests für die Auswahl der Kontrollvariablen	141
6.2	Mittelwerte und Standardabweichung der Kontrollvariablen pro Versuchsbedingung	143
6.3	Kolmogorow-Smirnow- und Levene-Tests für die Skalen des Lerntests	144
6.4	Adjustierte Mittelwerte und deren Standardfehler für die Skalen des Lerntests	145
6.5	Kolmogorow-Smirnow- und Levene-Tests für empfundene Schwierigkeit und mentale Belastung	147
6.6	Kolmogorow-Smirnow- und Levene-Tests für die Lesezeiten der Texte	149
6.7	Adjustierte Mittelwerte und deren Standardfehler für die Lesezeiten der Texte	149
6.8	Kolmogorow-Smirnow- und Levene-Tests für die formatspezifischen Lernzeiten	150
7.1	Kolmogorow-Smirnow- und Levene-Tests für die Auswahl der Kontrollvariablen der zweiten Untersuchung	170
7.2	Kolmogorow-Smirnow- und Levene-Tests für die Auswahl der Kontrollvariablen zur Testung der Interferenz-Hypothese	173
7.3	Kolmogorow-Smirnow- und Levene-Tests für die Auswahl der Kontrollvariablen zur Testung der Konstruktions-Hypothese	178
8.1	Kolmogorow-Smirnow- und Levene-Tests für die Skalen des Lerntests in der dritten Untersuchung	191

Tabellenverzeichnis

8.2	Kolmogorow-Smirnow- und Levene-Tests für die Lernzeiten in der dritten Untersuchung	192
8.3	Kolmogorow-Smirnow- und Levene-Tests für die Anzahl valider Fixationen .	194
8.4	Kolmogorow-Smirnow- und Levene-Tests für die durchschnittliche Fixationsdauer	195
8.5	Mittelwerte und Standardabweichung der Fixationsdauer in Millisekunden .	195
8.6	Kolmogorow-Smirnow- und Levene-Tests für die mittlere Sakkadenamplitude	196
8.7	Mittelwerte und Standardabweichung der Sakkadenamplituden	197
8.8	Korrelationen der Blickbewegungsparameter mit Lernleistung, mentaler Belastung und Lernzeit	197

Abbildungsverzeichnis

2.1	Mehrspeichermodell nach Atkinson & Shiffrin (1968)	24
2.2	Revidierte Fassung des Arbeitsgedächtnisses nach Baddeley (2000, 2001)	28
2.3	Kognitive Komponenten der ACT*-Theorie nach Anderson (1983)	32
3.1	Repräsentationsebenen beim Bildverstehen nach Rasch (2006)	59
3.2	Repräsentationsebenen beim Bildverstehen nach Schnotz & Dutke (2004) .	60
3.3	Modell des verbalen und nonverbalen Systems der Theorie der dualen Kodierung nach Paivio (1986)	64
3.4	Modell der Enkodierung und des Abrufs von Text-Karten-Informationen Kulhavy, Stock & Kealy (1993)	68
3.5	Kognitive Theorie des multimedialen Lernens (nach Mayer, 2005)	73
3.6	Modell des integrierten Text und Bildverstehens (nach Schnotz und Banert, 2003)	78
3.7	Das integrierte Modell des Text und Bildverstehens nach Schnotz (2005) .	83
5.1	Bildseite <i>Plattegrenzen</i> als vollständige Version	103
5.2	Bildseite <i>Plattegrenzen</i> als komplementäre Version	104
5.3	Beispielitem aus dem Beurteilungsbogen zur Einschätzung des Informationsverhältnisses zwischen Bildern und Texten des Stimulusmaterials	108
5.4	Zusammenfassung der Expertenurteile	109
5.5	MC-Item zur Erfassung der Wiedererkennensleistung	111
5.6	Item zur Erfassung des freien Erinnerns	111
5.7	MC-Item zur Erfassung der Fähigkeit Inferenzen zu bilden	112
5.8	Item mit offenem Format zur Erfassung der Fähigkeit Inferenzen zu bilden .	113
5.9	Item zur Erfassung der mentalen Belastung während der Lernphase	120
5.10	MC-Item zur Erfassung des domänenspezifischen Vorwissens	121
5.11	MC-Item zur Erfassung der aktuellen Motivation	123
5.12	Item der Skala Satzergänzung (SE) des I-S-T 2000 R	124
5.13	Item der Skala Würfelaufgaben (WÜ) des I-S-T 2000 R	125
5.14	Item des Verbalizer-Visualizer-Questionnaires	127
6.1	Untersuchungsdesign der ersten Studie	135
6.2	Mittlere adjustierte Gesamtleistung im Wissenstest	146
6.3	Mittlere adjustierte Gesamtlesezeit	151
6.4	Darstellung der Interaktion Zeit*Reihenfolge für die kongruente Bedingung .	152
6.5	Darstellung der Interaktion Zeit*Reihenfolge für die komplementäre Bedingung	153

Abbildungsverzeichnis

6.6	Fragen zum Lernen 1: Details oder Zusammenhänge?	154
6.7	Fragen zum Lernen 3: Texte oder Bilder?	156
7.1	Untersuchungsdesign der zweiten Studie	167
7.2	Mittlere adjustierte Gesamtleistung in der zweiten Studie	172
7.3	Untersuchungsdesign zur Prüfung der Interferenz-Hypothese	172
7.4	Interferenz-Hypothese: mentale Belastung	175
7.5	Interferenz-Hypothese: Gesamtleistung im Wissenstest	176
7.6	Untersuchungsdesign zur Prüfung der Konstruktions-Hypothese	177
7.7	Konstruktions-Hypothese: Gesamtleistung im Wissenstest	180
8.1	Untersuchungsdesign der dritten Studie	188
8.2	Interaktion zwischen formatspezifischer Lesezeit und Reihenfolge unter kongruenter Bedingung	193

Zusammenfassung

Zahlreiche Untersuchungen weisen darauf hin, dass die Kombination von Texten und Bildern beim Wissenserwerb zu besseren Lern- und Behaltensleistungen führt, als wenn nur mit Texten gearbeitet wird. Dieser Multimediaeffekt ist häufig auch dann zu beobachten, wenn beispielsweise Schüler oder Studierende Bild und Text sequenziert lernen. Einige Arbeiten haben dabei gezeigt, dass die Reihenfolge, in der die beiden Formate verarbeitet werden, großen Einfluss auf den Wissenserwerb hat. Erfolgt die Präsentation des Bildes vor dem Text, gelingt es den Lernenden in der Regel mehr Fakten zu erinnern. Das Phänomen des *picture-text-sequencing*-Effekts wurde von zwei Forschern aufgegriffen und anhand alternativer Theoriemodelle auf unterschiedliche Weise erklärt. Während einer der Autoren kapazitätsspezifische Merkmale der mentalen Repräsentation des Bildes als Ursache ansieht, macht der andere Autor Interferenzen für den Einfluss der Verarbeitungsreihenfolge verantwortlich. Das Ziel der Arbeit war es daher, zu untersuchen, welcher der beiden Ansätze der Wahrheit am ehesten entspricht.

Zu diesem Zweck wurde eine Studie durchgeführt, bei der Verarbeitungsreihenfolge und Informationsverhältnis beim Lernen mit multiplen Repräsentationen experimentell variierte. Das Stimulusmaterial bestand aus Texten und Bildern zur Geodynamik, die in einer komplementären und einer kongruenten Version vorlagen. Anders als erwartet, zeigte sich jedoch kein eindeutiges Ergebnis, obwohl die Analyse der Lernzeiten letztlich gegen die Annahme eines besonderen Kapazitätsvorteils sprach. Eine Anschlussuntersuchung erbrachte die Möglichkeit, Personen, die Texte und Bilder erhalten hatten, mit Teilnehmern zu vergleichen, die nur auf ein Format zurückgreifen konnten. Die Auswertungen ergab Unterschiede hinsichtlich der mentalen Belastung bei äquivalenten Leistungen im Wissenstest. Interpretiert wurden diese Resultate als Hinweis für das Auftreten von Interferenzen, die immer dann entstanden, wenn die Versuchspersonen das Bild nach dem Text lernten. In diesem Zusammenhang hatte die dritte und letzte Untersuchung den Zweck, weitere Belege für die Interferenzannahme zu finden. Anhand von Blickbewegungsparametern, die während des sequenzierten Lernens mit den Texten und Bildern erhoben worden waren, konnten jedoch keine Unterschiede festgestellt werden. Insgesamt ließ sich die Schlussfolgerung ziehen, dass hinsichtlich des sequenzierten Lernens mit Texten und Bildern weiterhin ein großer Spielraum für zusätzliche Forschungsarbeiten besteht.

1 Einleitung

Die gemeinsame Verwendung von Bildern und Texten zur Wissensvermittlung ist seit langem gebräuchlich und lässt sich bis in die Epoche der frühen Neuzeit zurückverfolgen. Schon damals waren Gelehrte der Ansicht, dass die Kombination verbaler und visueller Repräsentationsformate den Lernprozess erleichtert und unterstützt. Welche Vorstellungen die Wissenschaftler dieser Zeit von dem menschlichen Geist auch hatten, ihre Einschätzung wurde mittlerweile anhand einer Vielzahl von empirischen Untersuchungen eindrucksvoll belegt. Demnach gilt es innerhalb der Pädagogischen Psychologie als weitgehend gesichert, dass Lernende mehr Wissen erwerben, wenn sie neben einem Text zusätzlich auf ein Bild zurückgreifen können, als wenn sie nur mit einem Text lernen. Dieser Zusammenhang wurde in den 90er Jahren als Multimedia-Effekt bezeichnet und regte im bildungswissenschaftlichen Kontext zahlreiche Studien an. Das Forschungsfeld des Lernens mit Texten und Bildern gewann darüber hinaus zusätzlich an Bedeutung, da durch die Verbreitung von Computern und der Digitalisierung der Datenverarbeitung die Verfügbarkeit von Illustrationen, Grafiken und Abbildungen stark angestiegen ist. War es vor zwei Jahrzehnten noch mit großem Aufwand verbunden, beispielsweise Fotografien oder auch Zeichnungen zu bearbeiten und zu drucken, können Bilder heutzutage mit dem Mobiltelefon erstellt und versendet werden.

Gleichzeitig bietet das Internet einen gewaltigen Fundus an Bildmaterial, den Lernende und Lehrende unter Berücksichtigung urheberrechtlicher Bestimmungen zum Erwerb oder der Vermittlung von Wissen nutzen können. Allerdings ist es irreführend, anzunehmen, dass die kombinierte Präsentation von Bild- und Textmaterial zwangsläufig zu größerem Lernerfolg führt oder den Wissenserwerb weniger anstrengend macht. Wie so oft sind die Zusammenhänge sehr viel komplexer und neben den Inhalten, der Textsorte und der Bildart bestimmen die Merkmale des Individuums über Ausmaß und Richtung der lernförderlichen Wirkung. Nicht wenige Forscher achten daher bei der Entwicklung ihrer theoretischen Konzeptionen auf die Anwendbarkeit ihrer Befunde, indem sie Gestaltungsprinzipien für multimediale Lernumgebungen formulieren. Einer dieser Grundsätze ist das Prinzip der temporären Kontiguität (*temporal contiguity*). Danach sollten Bilder und Texte möglichst gleichzeitig präsentiert werden, da eine sequenzierte Darbietung zu einer unnötigen Belastung des kognitiven Systems und damit zu geringerem Lernerfolg führt.

Manchmal ist es jedoch einfach nicht möglich, Bild und Text gemeinsam darzustellen. So kommt es bei der Gestaltung von multimedialen Dokumenten häufig vor, dass die Abbildung, die man einfügen will, nicht an die Stelle passt, wo der Text Bezug darauf nimmt. Das Bild zu verkleinern ist nicht immer eine sinnvolle Option, da die Leser des Dokumentes die Inhalte noch erkennen sollen. In diesen Fällen müssen die beiden Formate nacheinander dargeboten werden, wobei sich die Frage stellt, ob zuerst das Bild oder zuerst

1 Einleitung

der Text gelesen bzw. betrachtet werden sollte. Obwohl nach einigen Theorien die Verarbeitungsreihenfolge keine Rolle spielen dürfte, haben vereinzelte Studien gezeigt, dass die Sequenz den Lernerfolg sehr wohl beeinflussen kann. Demnach schnitten Schüler, die mit einer geografischen Karte und einem Text lernten deutlich besser ab, wenn sie die Karte vor dem Lesen des Textes erhalten hatten. Dieser Effekt der Verarbeitungsreihenfolge wurde als *picture-text-sequencing*-Effekt bezeichnet, und legt die Schlussfolgerung nahe, Bildinformationen stets vor verbalen Inhalten zu präsentieren.

Allerdings sind Hinweise für die Gestaltung von Lernumgebungen und multimedialen Materialien nur ein Aspekt, zumal andere Bedingungen oder Zielgruppen zu abweichenden Befunden führen können. Für die vorliegende Arbeit war der PTS-Effekt vor allem in theoretischer Hinsicht interessant, da von dem Zusammenhang zwischen Lernerfolg und Verarbeitungsgreihenfolge Rückschlüsse auf die kognitive Architektur erhofft wurden. Dies gilt umso mehr, da zwei unterschiedliche Modelle existieren, die den Einfluss der Sequenz beim Wissenserwerb mit Texten und Bildern zu erklären versuchen. Die Ansätze von Kulhavy, Stock und Caterino (1994) und Schnotz (2002, 2005) widersprechen sich zwar nicht direkt, doch machen die Autoren voneinander abweichende Annahmen zur Beschaffenheit des kognitiven Systems und den darin ablaufenden Prozessen. Durch eine Gegenüberstellung der zugrundeliegenden Konstrukte soll im Rahmen dieser Arbeit geklärt werden, welches der theoretischen Modelle der (empirischen) Wahrheit am ehesten entspricht.

Bevor jedoch auf die Untersuchungen näher eingegangen wird, folgen dieser Einleitung (*Kapitel 1*) zunächst zwei Abschnitte zu den theoretischen Hintergründen. *Kapitel 2* behandelt grundsätzliche Überlegungen zu zentralen Begriffen der pädagogischen Psychologie, wie Lernen und Verstehen. Darüber hinaus wird auf die Bedeutung des Konzepts der Repräsentation eingegangen und mit Propositionen, mentalen Modellen und Schemata Formen beschrieben, wie Wissen repräsentiert werden könnte. Da der Aufbau von Wissenstrukturen innerhalb des kognitiven Systems stattfindet, folgt dieser Darstellung eine Besprechung verschiedener Gedächtnismodelle und -architekturen. Von Mehrpeichermodellen über die Arbeitsgedächtnistheorie bis hin zur *cognitive load theory* werden dabei die wichtigsten Ansätze vorgestellt.

Kapitel 3 widmet sich dann speziell dem Verstehen von Texten und Bildern, wobei im Fall der verbalen Informationen vor allem die visuelle Verarbeitung im Vordergrund steht. So werden zunächst die Besonderheiten beleuchtet, die mit der Verarbeitung der beiden Formate verbunden sind. Die verschiedenen Repräsentationsprinzipien bilden auch die Grundlage für die dann folgenden Abschnitte, in denen es um theoretische Modelle geht, die sich auf den gemeinsamen Wissenserwerb mit Texten und Bildern beziehen. Zwei der diskutierten Ansätze werden dann in *Kapitel 4* aufgegriffen, um anhand der jeweiligen Modellannahmen überprüfbare Vorhersagen abzuleiten. Der Abschnitt beinhaltet zudem die Überlegungen, die der Gestaltung der Untersuchungsmaterials und der Erhebungsinstrumente zugrunde lagen.

Deren Entwicklung und Merkmale werden dann in *Kapitel 5* ausführlich besprochen.

Den Anfang macht die Gestaltung der in den Untersuchungen verwendeten Texte und Bilder. So wird die Auswahl der inhaltlichen Domäne erläutert, Gestaltungsprinzipien diskutiert und schließlich Evaluationsergebnisse vorgestellt. Der zweite Abschnitt des Kapitels ist dann den Erhebungsinstrumenten gewidmet, mit denen psychologische Zustände und Merkmale vor und nach der Intervention erfasst werden sollten. Das betrifft unter anderem ein Fragebogen zur Kontrolle des Lernerfolgs, ein Vorwissenstest oder verbale und räumliche Untertests einer Intelligenzbatterie. Den eigens für die Untersuchungen entwickelten Instrumenten wurde dabei besonders viel Platz eingeräumt, aber auch die standardisierten Verfahren werden kurz besprochen.

Der eigentliche Kern dieser Arbeit sind jedoch die Ergebnisse von drei empirischen Studien. Davon wurde die Erste mit dem Ziel durchgeführt, zwei theoretische Modelle hinsichtlich des PTS-Effekts gegenüber zu stellen, wie in *Kapitel 6* nachzulesen ist. Zu diesem Zweck sollten Versuchspersonen Texte und Bilder zum Wissenserwerb nutzen, wobei die Präsentationsreihenfolge und das Informationsverhältnis der beiden Formate variiert wurde. Die Analysen ergaben jedoch mehrdeutige Resultate, so dass es sinnvoll erschien eine Anschlussuntersuchung durchzuführen (*Kapitel 7*). Im Rahmen dieser zweiten Studie erhielten die Teilnehmer aber jeweils nur ein Repräsentationsformat für den Wissenserwerb. Durch den Vergleich mit den Ergebnissen der ersten Studie konnte auf diese Weise eingeschätzt werden, wie sehr die Versuchspersonen von der Präsentation eines zweiten Formates profitierten. *Kapitel 8* schildert die Befunde einer explorativen Studie, bei der die Auswirkungen der Verarbeitungsreihenfolge und des Informationsverhältnisses auf die Blickbewegungen beim Wissenserwerb mit Texten und Bildern untersucht wurden. Schließlich wird in *Kapitel 9* versucht, die wichtigsten Ergebnisse theoretisch zu deuten und Anregungen für zukünftige Forschungsbemühungen zu geben.

Teil I

Theoretische Vorüberlegungen

2 Lernen - Verstehen - Erinnern

Allgemein wurde der Versuch unternommen, alle theoretischen Ausführungen so ausführlich wie nötig und so knapp wie möglich zu gestalten. Da es sich bei den hier diskutierten Theorien, um Ansätze handelt, die in der psychologischen Forschung und den Bildungswissenschaften durchaus geläufig sind, werden daher nur Aspekte behandelt, welche für die wissenschaftliche Fragestellung von Bedeutung erschienen. Die Darstellung beginnt mit einer wissenschaftlichen Einordnung des Begriffs *Lernen*.

2.1 Lernen als Wissenserwerb

Lernen ist ein Prozess, der in der Entwicklungsgeschichte des Menschen einen zentralen Platz einnimmt. Dies wird umso deutlicher, wenn wir uns vergegenwärtigen, was wir im Laufe eines Lebens alles lernen können bzw. lernen sollen. Motorische Fertigkeiten, das Sprechen in einer Sprache oder auch angemessene Reaktionen in sozialen Situationen sind Verhaltensweisen, die erlernt werden müssen. Vieles von dem, was wir lernen, kann aber nur schwer isoliert betrachtet werden. Verschiedenste Kenntnisse oder Kompetenzen hängen mit anderen Wissensbereichen zusammen oder bauen aufeinander auf. Beispielsweise lernen Grundschüler Schreiben und Lesen, um sich mittels dieser Fertigkeit selbständig weiteres Wissen anzueignen. Selbst als Erwachsener lernt man mehr oder weniger bewusst ständig neue Dinge, so dass manche Autoren von *lebenslangem Lernen* als einer sozialen Realität ausgehen (Kade & Seitter, 2007). Lernen ist also ein komplexer Vorgang, der viele Gegenstandsbereiche hat und von dem viele Formen und Ausprägungen existieren. Umso schwieriger ist es, eine allgemeine Definition zu finden, die alle wichtigen Aspekte zusammenfasst. Eine Lösung für dieses Problems besteht darin, Lernen unter zwei verschiedenen Gesichtspunkten zu betrachten. So kann Lernen ganz allgemein als ein Prozess aufgefasst werden, der bei einem Individuum aufgrund eigener Aktivitäten zu einer relativ überdauernden Verhaltensänderung führt (Steiner, 2006). Dieser Sichtweise bezieht sich vor allem auf die *klassischen* Lerngesetze, deren Erforschung sich Psychologen verschrieben haben, die der Richtung des *Behaviourismus* nahe standen (Spada, Ernst & Ketterer, 1998).

Der zweite Gesichtspunkt des Begriffs Lernen, lässt sich treffend anhand der sprachgeschichtlichen Wurzeln des Wortes ableiten. Etymologisch ist *lernen* mit *lehren* und *List* verwandt und gehört zur Wortgruppe von *leisten*, das ursprünglich einer Spur nachgehen, nachspüren bedeutete. Das Wort Lernen geht sehr wahrscheinlich auf die gotische Bezeichnung für *ich weiß* (*lais*) und das indogermanische Wort für *gehen* (*lis*) zurück (Wasserzieher, 1974). Genau übersetzt bedeutet *lais ich habe nachgespürt* und *laists* steht für *Spur*. Die Herkunft des Wortes deutet darauf hin, dass Lernen ein Vorgang ist, bei dem

man einen Weg zurücklegt und dabei zu Wissen gelangt (Mielke, 2001). Aus etymologischer Sicht ist Lernen also mit Wissenserwerb gleichzusetzen. Eine Auffassung, der sich vor allem viele Vertreter der kognitionspsychologischen Tradition anschließen würden.

Der Aufbau von Verhaltensweisen und der Erwerb von Wissen sind natürlich in vielen Fällen eng miteinander verknüpft. So werden viele bewusste Handlungen von zuvor erworbenem begrifflichen Wissen gesteuert. Dennoch wird aus Gründen der Einfachheit im Rahmen dieser Arbeit *Lernen* mit dem Erwerb von Wissen gleichgesetzt und damit folgendermaßen definiert:

Lernen im Sinne des Wissenserwerbs ist ein bereichsspezifischer, komplexer Prozess, der den Aufbau und die fortlaufende Modifikation von Wissensrepräsentationen umfasst (Steiner, 2006, S.163).

Ergänzend sei an dieser Stelle erwähnt, dass die Veränderungen der Wissensstrukturen, die durch Erwerb oder Modifikation entstanden sind *zeitlich* von Dauer sein müssen, damit man von einem erfolgreichen Lernprozess sprechen kann. Kann Wissen, das beispielsweise beim Lesen eines Textes *verarbeitet* wurde, schon nach kurzer Zeit nicht mehr reproduziert werden, hat auch kein Lernen stattgefunden (Sweller, 2005). Damit liegt eine brauchbare Arbeitsdefinition des Begriff Lernens vor, die zudem an bestimmte Bedingungen geknüpft ist. Allerdings bleibt noch zu klären, was es heißt Wissensstrukturen aufzubauen, welches dabei die entscheidenden Voraussetzungen sind und in welcher Form einmal erworbenes Wissen repräsentiert, abgerufen und verwendet werden kann.

2.2 Verstehen als mentale Repräsentation

Menschen erlangen durch die Auseinandersetzung mit ihrer Umwelt Wissen, das sie nutzen, um sich zurechtzufinden. Dieses Wissen muss in irgendeiner Form mental repräsentiert werden, bevor es in das Gedächtnis gelangen und bei entsprechendem Anlass wieder abgerufen werden kann. Nach Scriven (1972) ist Verstehen nichts anderes als ein Nebeneffekt eines effektiven Informationsverarbeitungssystems, das Mustererkennungs- und Modellierungsvorrichtungen umfasst. Schnotz (1994a) greift diesen Gedanken auf und definiert den Aufbau einer funktionalen Repräsentationen der Umwelt als *Verstehen*. Demnach ist Verstehen ein kognitiver Prozess, bei dem eine in sich stimmige, kohärente mentale Repräsentation eines externen Sachverhaltes konstruiert wird. Das Attribut der Kohärenz bezieht sich dabei auf das Ausmaß, in dem die vorhandenen Elemente, aus denen sich die Repräsentation konstituiert, in sinnvolle Relationen zueinander treten. Texte, aber auch Situationen oder Handlungen sind oft mehrdeutig und komplex, so dass es zum Aufbau einer kohärenten Repräsentation nicht allein ausreicht, Beziehungen herzustellen. Häufig ist es erforderlich, wichtige Informationselemente auszuwählen und unpassende oder widersprüchliche Elemente zu verwerfen (Kintsch, 1998). Diese Prozesse bergen die Gefahr in sich, dass die falschen Elemente ausgewählt oder unterdrückt werden. So ist es durchaus möglich, dass ein Individuum eine kohärent und in sich

stimmige mentale Repräsentation aufbauen kann, die aber den externen Sachverhalt nur inadäquat wiedergibt und damit falsch ist.

Eine inadäquate Repräsentation liegt zum Beispiel dann vor, wenn ein Satz oder eine Aussage *missverstanden* wird. Die mentale Repräsentation muss also nicht nur kohärent, sondern auch dem Gegenstand angemessen bzw. adäquat sein, damit man tatsächlich von Verstehen sprechen kann. Allerdings können gerade bei der Interpretation von Texten oder sozialen Situationen die Grenzen zwischen Verstehen und Missverstehen fließend sein, da beiden Ergebnissen letztlich die gleichen kognitiven Prozesse zugrunde liegen. Ganz ähnlich verhält es sich mit dem Nicht-Verstehen. Hier schlägt allerdings der Aufbau einer kohärenten und in sich stimmigen mentalen Repräsentation fehl. Dies ist immer dann der Fall, wenn es nicht gelingt, Relationen zwischen den Informationselementen herzustellen, die zum Aufbau einer Mentalen Repräsentation zur Verfügung stehen. Beim Verstehen eines Satzes würde das bedeuten, dass die Worte bzw. ihre Bedeutungen zwar einzeln repräsentiert werden, es aber nicht gelingt, diese zu einer kohärenten Struktur zu verbinden. Schnotz (1994a) betont aber, dass zwischen Verstehen und Nicht-Verstehen oft keine scharfe Grenze gezogen werden kann, da vor allem der Grad der Kohärenz eine Ermessensfrage darstellt, die stark von den Personenmerkmalen eines Lesers oder Beobachters abhängt.

Wenn Lernen, wie es in dem Kapitel zuvor definiert wurde, *den Aufbau und die fortlaufende Modifikation von Wissensrepräsentationen* umfasst, dann sind Verstehensprozesse unmittelbar damit verbunden. Allerdings bedeutet allein der Aufbau einer kohärenten mentalen Repräsentation nicht zwangsläufig, dass neue bzw. modifizierte Wissensstrukturen entstehen. Es ist durchaus vorstellbar, dass eine Person eine Textpassage liest, ohne etwas Neues zu lernen. Sie müsste den selben Text nur immer und immer wieder lesen. Würde dieser Vorgang nach einer gewissen Zeit stets in exakt der gleichen Repräsentation resultieren, könnte man behaupten, dass es Verstehen gibt, ohne dass Etwas gelernt wurde. Aber selbst wenn es theoretisch möglich erscheint, dass Verstehen keine Veränderung der Wissensstruktur zur Folge hat, ist es sehr unwahrscheinlich, dass das Verstehen eines Sachverhaltes nicht mit Lernprozessen einhergeht. Ein Text, wird er wiederholt gelesen, bleibt gleich. Dies gilt aber weder für den Leser, der sich von mal zu mal verändert, noch für die Situation, in der gelesen wird. Geht man davon aus, dass die variablen Personen- und Situationsmerkmale nie zum Aufbau ein und derselben mentalen Repräsentation führen, bringt dies zwangsläufig auch eine Modifikation der am Verstehensprozess beteiligten kognitiven Strukturen mit sich. So nimmt bei wiederholter Auseinandersetzung mit dem Text die Kongruenz zwischen der mentalen Repräsentation und dem externen Sachverhalt immer weiter zu. Demnach findet ein akkomodativer Prozess statt, der nicht nur zu einer Festigung des betreffenden Wissens sondern auch zu einer erhöhten Vertrautheit mit dem Sachverhalt führt, der verstanden werden soll (Spiro, 1980). Beides hat wiederum Konsequenzen für weitere Verarbeitungsprozesse, so dass man letztlich davon ausgehen kann, dass Verstehen immer auch mit Lernen *zusammenhängt*.

Allerdings ist es nicht möglich, Verstehen generell als Voraussetzung für Lernprozesse

se anzusehen, da allgemein angenommen wird, dass neben dem bewussten (expliziten) Lernen, auch implizites, also unbewusstes Lernen stattfindet (Reber, 1967). Dies schließt mit ein, dass Regeln, Verhaltensweisen, etc., die unbewusst gelernt wurden, in irgendwelchen Strukturen des kognitiven Systems verfügbar gehalten werden müssen. In diesem Zusammenhang wird von impliziten Gedächtnisinhalten gesprochen, die im Gegensatz zu explizit verfügbaren Wissensstrukturen, nicht willentlich bzw. nur indirekt aktiviert werden können (vgl. Schacter, 1987). Es liegt daher nahe, die empirischen Befunde zum impliziten Lernen mit der Existenz spezifischer Formen von Wissen zu erklären.

2.2.1 Deklaratives und Prozedurales Wissen

Wie bereits dargelegt wurde, besteht Lernen darin, neues Wissen zu erwerben bzw. vorhandene Wissensstrukturen zu modifizieren. Dieses Wissen kann typischerweise historische Fakten, theoretische Annahmen oder auch Vokabeln einer fremden Sprache umfassen. Allerdings ist es eine Sache, zu wissen, dass beispielsweise das polnische Wort *pamięć* Gedächtnis bedeutet, aber es ist etwas ganz anderes, es korrekt auszusprechen. Denn selbst wenn man weiß, wie die einzelnen Laute geformt werden, stellt die Artikulation einer Kombination unvertrauter Laute eine Herausforderung dar. Lernen beinhaltet also nicht nur den Erwerb von Wissen über einen Sachverhalt, Lernen schließt auch den Erwerb von Fertigkeiten bzw. Fähigkeiten ein. In der Psychologie werden diese beiden Formen von Wissen häufig als *deklaratives* und *prozedurales* Wissen unterschieden (J. R. Anderson, 1976). Diese Unterscheidung ist weniger inhaltlicher Natur, sondern beruht auf der Annahme verschiedener Funktionen der beiden Wissensarten. So dienen die deklarativen Wissensstrukturen als Repräsentation der Wirklichkeit bzw. Umwelt, in der das Individuum agiert, während die prozedurale Wissensstruktur so etwas wie ein Handlungsrepertoire darstellt, um diese Umwelt zu beeinflussen. Es ist also durchaus möglich, dass jemand weiß, was man zur Bewältigung einer bestimmten Situation tun muss und vielleicht auch in der Lage ist, einzelne Teilprozesse der Handlung zu beschreiben bzw. zu benennen. Gleichzeitig kann die selbe Person aber außer Stande sein, diese Handlung auszuführen, weil sie nicht über das entsprechende prozedurale Wissen verfügt. Wissen über Prozesse und Handlungen ist damit auch ein Teil des deklarativen Wissens, so dass die häufig zur Veranschaulichung gebrauchte Wendung, prozedurales Wissen sei im Grunde *Wissen, wie...*, zu kurz greift. Am besten lässt sich das prozedurale Wissen anhand der Unterscheidung zwischen implizitem und explizitem Wissen verdeutlichen. Beispielsweise haben N. J. Cohen und Squire (1980) bei einer Studie mit Amnesiepatienten gezeigt, dass diese kognitive Prozeduren erlernen und über einen Zeitraum von mehreren Wochen behalten können. Andererseits konnten die Probanden, den Inhalt dessen, was sie gelernt hatten, schon nach kurzer Zeit nicht mehr erinnern. Es liegt daher nahe, deklaratives Wissen mit explizitem Wissen gleichzusetzen, dessen man sich bewusst ist und über das man sprechen kann.

Prozedurales Wissen ist dagegen oft implizit und es fällt häufig schwer oder ist gar unmöglich, davon zu berichten. Ein Grund könnte sein, dass ein Großteil des prozedurales

Wissen aus perzeptuell-motorischen Fertigkeiten und/oder hoch-automatisierten kognitiven Prozeduren besteht, die meist nicht bewusstseinspflichtig und manchmal gar nicht bewusstseinsfähig sind (Schacter, 1987). So wenden viele Muttersprachler die grammatischen Regeln ihrer Sprache intuitiv zwar richtig an, können diese Regeln einem Fremden aber nicht erklären. Eine scharfe Differenzierung zwischen den beiden Wissensarten ist meist nur auf der theoretischer Ebene möglich, da das implizite Wissen nur schwer zugänglich ist. Außerdem dürfte weitgehend unbestritten sein, dass im Rahmen von bewussten Handlungsausführungen sowohl auf deklaratives als auch auf prozedurales Wissensinhalte zurückgegriffen werden muss (J. R. Anderson, 1983a). Wenn im folgenden näher auf die Repräsentation von Wissen eingegangen wird, so bezieht sich diese Darstellung vor allem auf explizites und damit eher auf deklarative Wissenstrukturen.

2.2.2 Repräsentation von Wissen

Bisher wurde berichtet, dass in der kognitiven Psychologie *Verstehen* als Aufbau einer kohärenten mentalen Repräsentation angesehen werden kann. Allerdings ist bislang noch ungeklärt, was man sich unter einer *Repräsentation* eigentlich vorstellen muss. Zunächst weist der Begriff *Repräsentation* darauf hin, dass ein Sachverhalt für einen anderen Sachverhalt steht. Diese beiden Sachverhalte, die manchmal auch als *Repräsentant* und *Repräsentandum* bezeichnet werden, sind dabei durch eine Abbildungsrelation miteinander verknüpft (Tergan, 1989). Im Gegensatz zu einer Kopie, zeichnet sich eine Repräsentation dadurch aus, dass der Repräsentant niemals alle Eigenschaften und Merkmale des Repräsentandums in sich vereint. Welche Eigenschaften letztlich repräsentiert werden, legt die Abbildungsrelation fest, der in der Regel eine Funktion, Absicht oder ein Ziel zugrunde liegt. Damit ist es also möglich, dass ein und der selbe Gegenstand gleichzeitig mehrere adäquate Repräsentationen besitzen kann, die unter Umständen sehr unterschiedliche Eigenschaften aufweisen.

Versucht man vor diesem Hintergrund verschiedene Repräsentationsformen zu klassifizieren, so ergibt sich eine besonders prägnante Trennlinie zwischen symbolischen und analogen Repräsentationen. Symbolische Repräsentationen zeichnen sich dadurch aus, dass der Repräsentant ein Zeichen ist, das dem Repräsentandum in der Regel arbiträr zugeordnet wird. Durch diese Zuordnung erhält das beliebige Zeichen eine Bedeutung, was es zu einem Symbol macht. Das bekannteste Beispiel für eine solche arbiträre Zuordnung sind die unterschiedlichen Schriften, die Menschen überall auf der Welt verwenden. Schriften stellen regelrechte Symbolsysteme dar, die aufgrund historisch entstandenen Konvention bestimmten Regeln folgen. Ebenso wie die Anzahl verwendbarer Symbole beschränkt ist, können diese nicht zu beliebigen Ketten verbunden werden, denn auch hier liegen Regeln zugrunde. Trotz dieser Einschränkungen, verfügen viele Symbolsysteme über ausreichend Freiheitsgrade, neben konkreten auch abstrakte Sachverhalte zu repräsentieren (T. Rasch, 2006). Sie stellen damit eine sehr mächtige und umfassende Repräsentationsform dar, mit der sich sehr viel beschreiben lässt.

Während sich Symbole dadurch auszeichnen, dass sie keine Gemeinsamkeiten mit

dem Objekt, für das sie stehen, aufweisen müssen, ist *Ähnlichkeit* das zentrale Kennzeichen der analogen Repräsentationsform. Diese Ähnlichkeit zwischen Repräsentation und dem Original kann dabei stark variieren. Beispielsweise enthält die Photographie eines Hauses neben dessen Form und räumlicher Verortung, Oberflächenmerkmale wie Farben und Texturen. Zeichnet ein Architekt oder Bauingenieur hingegen einen Plan von dem Haus, dann sind nur noch die Grundrisse zu erkennen. Beides sind aber Repräsentationen ein und desselben Hauses, auch wenn sich die wahrnehmbare Ähnlichkeit zwischen Repräsentant und Repräsentandum sehr unterscheidet. Um die variierende Ähnlichkeit zwischen Sachverhalt und Repräsentation konzeptionell zu fassen, differenziert Palmer (1978) in Anlehnung an die Überlegungen von Shepard und Chipman (1970) zwischen physischer und funktionaler Isomorphie. Von einem physisch isomorphen Verhältnis zwischen Repräsentation und Original ist die Rede, wenn einzelne Elemente sich wahrnehmbar entsprechen. Beispielsweise kann man anhand einer Farbfotografie, die Farbe eines Gegenstandes ebenso wahrnehmen, wie anhand des Originals. Dagegen ist funktionale Isomorphie bereits gegeben, wenn die Relation zwischen einzelnen Elementen des Repräsentandums auch im Repräsentanten enthalten ist. So lässt sich ein nicht direkt wahrnehmbares Mengenverhältnis zwischen zwei Gruppen als logisches Bild bzw. Venn-Diagramm darstellen. Die Relation ist damit sowohl im Sachverhalt, als auch in dessen Repräsentation gegeben, auch wenn diese auf den ersten Blick keine Ähnlichkeit aufweisen. Dies macht aber deutlich, dass je schwächer die physische Isomorphie ausfällt, umso mehr zusätzliche Informationen werden benötigt, um die betreffende Repräsentation zu verstehen.

Eine weitere Möglichkeit Repräsentationen zu klassifizieren besteht darin, den Ort, an dem sie realisiert werden, als Differenzierungsmerkmal heranzuziehen. Auf diese Weise lassen sich *externe* von *mentalen* Repräsentationen unterscheiden. Externe Repräsentationen sind Informationsstrukturen, die sich außerhalb des kognitiven Systems befinden und damit Bestandteil der physikalischen Realität sind. Typische Beispiele sind Texte oder instruktionale Bilder, denen ein bestimmtes Medium, wie ein Buch oder ein Computer, zugrunde liegt. Schnotz (2003) unterscheidet bei externen Repräsentationen zwischen *deskriptionalen* und *depiktionalen* Formattypen. Deskriptionen sind Repräsentationen, die auf einzelnen Symbolen oder einem Symbolsystem beruhen. So besteht ein gesprochener oder niedergeschriebener Text aus Worten, denen im Laufe der etymologischen Entwicklung relativ beliebig eine Bedeutung zugeordnet wurde. Seine deskriptive Kraft gewinnt die Sprache letztlich dadurch, dass die einzelnen Elemente (Worte) spezifischen Regeln (Grammatik) folgend, in vielfältigen Kombinationen nahezu jeden Sachverhalt beschreiben können. Depiktionen sind im Gegensatz dazu externe Repräsentationen, die über ein analoges Format verfügen und damit einen *anschaulich* Charakter haben. Die Spannweite solch anschaulicher Repräsentationen reicht von einer Photographie (realistisches Bild), über dreidimensionale Modelle bis hin zu abstrakten Diagrammen (logische Bilder).

Diese Vielfalt macht es allerdings schwer eine scharfe Trennlinie zwischen analogen und symbolischen Repräsentationen zu ziehen, da viele anschauliche Formate deskripti-

ve und analoge Elemente umfassen. Diagramme wären ohne eine *Beschriftung* der Achsen meist nicht zu verstehen und auch die Nutzung von Landkarten erfordert in der Regel die Verarbeitung von symbolischen und analogen Inhalten (vgl. Kirby, 1994). Im Grunde stellen viele Depiktionen eine Mischform dar, in der die Prinzipien der anschaulichen und der symbolischen Repräsentation verwirklicht wurden. Der Repräsentationsbegriff wird noch komplexer, wenn man sich den *mentalen* Repräsentationen zuwendet. Da sich das Wissen in den Köpfen der Menschen einer direkten Beobachtung entzieht, können Annahmen darüber nur anhand von Verhaltensäußerungen abgeleitet werden. So konnte man beispielsweise feststellen, dass der Aufbau mentaler Wissensstrukturen in der Regel bedeutungsbezogen und effizient erfolgt (J. R. Anderson, 2001). Sowohl bei Texten als auch bei Bildern werden vor allem Inhalte und Bedeutungen erinnert, während die Form, also der exakte Wortlaut oder grafische Details, meist schnell wieder vergessen werden. Eine weitere Besonderheit bei mentalen Repräsentationen liegt darin begründet, dass diese keine einfache Spiegelung verschiedener Sachverhalte einer externen Umwelt darstellen. Sie werden aktiv konstruiert, bei dem die externen Informationen mit vorhandenen Wissensstrukturen des Individuums interagieren. Auch wenn der Aufbau einer mentalen Repräsentation, wie angenommen wird, zu einer Veränderung der Wissensstruktur führt, ist diese nur für einen begrenzten Zeitraum bewusst vorhanden, bevor sie wieder zerfällt (Schnotz, 1994a). Um sich diesem ebenso komplexen wie schwer zugänglichen Gegenstand zu nähern, bedient sich die Kognitionswissenschaft verschiedener Metaphern, die mehr oder weniger an die konkreten Eigenschaften externer Repräsentationen erinnern (Waskan, 2003). Es wundert daher kaum, dass nicht nur ein Zugang entwickelt wurde, um die Natur der menschlichen Wissensrepräsentation konzeptionell zu erfassen.

Propositionen

Wenn man davon ausgeht, dass eine *Sprache des Geistes* (Fodor, 1975) existiert, durch die dem menschlichen Denken erst Ausdruck verliehen wird, dann liegt es auch nahe, die mentale Repräsentation von Wissensstrukturen mit sprachlichen Mitteln zu analysieren. Diese Analysen beziehen sich in der Regel aber nur auf externe Repräsentationen, da sich die mentalen Wissensstrukturen, welche beispielsweise der Anfertigung eines Textes oder Bildes zugrunde lagen, einer direkten Beobachtung entziehen. Letztendlich wird damit der Versuch unternommen, durch die Anwendung eines Notationssystems die Bedeutung von Sätzen oder Bildern so darzustellen, dass alle wahrnehmungsbezogenen Details abstrahiert werden können. Die zurückbleibende Struktur besteht dann aus einzelnen, *wahrnehmungsbereinigten* Informationseinheiten, die ein Relationsgefüge bilden. Ein in der kognitiven Psychologie weit verbreitetes Notationssystem um bedeutungshaltige Informationen zu analysieren, ist die *propositionale* Darstellungsweise. Wie der Name vermuten lässt, steht im Zentrum dieses Ansatzes die *Proposition*. Der aus der Logik und Linguistik stammende Begriff steht dabei für eine Relation, die zwischen zwei Sachverhalten besteht. Niedergeschrieben werden Propositionen in einer Prädikat-Argument-Struktur, wie ein Beispiel verdeutlichen soll. Der Satz

Catweazel hat zotteliges Haar.

beinhaltet die Propositionsargumente *Catweazel* und *Haar*. Beide Argumente werden durch das Prädikat *Haben* miteinander verbunden. Fügt man noch das Attribut *zottelig* hinzu, so kann man die in dem Satz enthaltene Information in folgende Propositionen zerlegen.

(P1) (*Agent: Catweazel, Objekt: Haar*).

(P2) (*zottelig (Attribut von Haar)*).

Für den übersichtlichen Beispielsatz werden insgesamt zwei Propositionen benötigt, um alle enthaltenen Informationen vollständig abzubilden. Man kann sich daher leicht vorstellen, dass für die Analyse umfangreicherer Informationen sehr viel mehr Propositionen erforderlich sind. Um zusammenhängende Sachverhalte zu beschreiben, wie sie beispielsweise ein Lehrbuchtext enthält, müssen daher auch Relationen zwischen einzelnen Propositionen dargestellt werden können. Das erfordert nicht nur verschiedene *Arten* von Propositionen, sondern gleichzeitig auch eine Darstellungsweise, welche die hierarchischen Beziehungen zwischen den Propositionen reflektiert.

Es gibt in der Tradition der kognitiven Psychologie eine ganze Reihe propositionaler Verfahren bzw. Methoden, um die in externen Repräsentationen vorhandenen Informationen formell zu analysieren. Diese unterscheiden sich sowohl durch die Zahl bzw. Funktion der verwendeten Propositionen, als auch in der Darstellung ihrer hierarchischen Beziehungen. Beispielsweise verwendet Kintsch (1974) im Rahmen seines Ansatzes prädikative, modifikatorische und konnektive Propositionen, die in einem *hierarchischen Kohärenzgraphen* angeordnet werden. Auch wenn andere Autoren alternative Verfahren mit diversen Propositionsdefinitionen entwickelt haben (Frederiksen, 1975) oder *Netzwerke* nutzen, um hierarchische Beziehungen zu symbolisieren, haben die meisten Ansätze doch eine zentrale Gemeinsamkeit. Die kleinste Wissenseinheit, die eine von anderen Einheiten unabhängige Aussage darstellt, ist immer die Proposition. Damit ist sie auch die kleinste Wissenseinheit, die sich sinnvoll als wahr oder falsch beurteilen lässt. Propositionale Darstellungen dienen aber nicht allein der formalen Analyse von Texten oder Informationsmedien. Nicht selten wird die Annahme zugrunde gelegt, dass auch mentale Repräsentation aus hierarchisch geordneten Propositionen bestehen, die über Relationen miteinander verknüpft sind. Es gibt einige empirische Ergebnisse, die einen Zusammenhang zwischen den hypothetischen Merkmalen einer propositionalen Wissenstruktur und beobachtbarem Verhalten nahe legen.

Ein Experiment von Bransford und Franks (1971) zeigte beispielsweise, dass die Informationen in komplexen Sätzen offenbar in einfache abstrakte Einheiten zerlegt werden, um sie besser zu erinnern. Sprachliche Details, die für das *Verständnis* des Inhalts eher unbedeutend sind, werden dagegen schnell vergessen. Auch die Tatsache dass bei gleichlangen Texten die Anzahl der enthaltenen Propositionen die Lesezeit beeinflusst (Kintsch, 1974), kann als Zeichen für ihre psychologische Realität angesehen werden.

Dennoch gibt es auch einige Kritikpunkte und Einwände gegen die vermutete propositionale Struktur mentaler Wissensrepräsentationen. Beispielsweise sprechen verschiedene empirische Ergebnisse für einen Zusammenhang zwischen der Hierarchiehöhe und der Wahrscheinlichkeit, die Bedeutung einer Proposition zu behalten (Kintsch & Keenan, 1973; Kintsch, Kozminsky, Streby, McKoon & Keenan, 1975). Je höher ein Proposition in der Hierarchie des Kohärenzgraphen steht, umso wichtiger ist sie für das Verständnis des ganzen Gefüges und umso eher wird ihr Inhalt erinnert. Da es aber sehr viele verschiedene Möglichkeiten gibt, hierarchische Propositionsstrukturen zu erstellen, können die vorgefundenen Zusammenhänge als die Fähigkeit des jeweiligen Forschers angesehen werden, zentrale und damit gut erinnerbare Inhalte auszuwählen und sie entsprechend weit oben im Kohärenzgraphen einzuordnen (Schnitz, 1994a). Demnach ist der Hierarchieebenen-Effekt nicht unbedingt ein zwingender Beleg dafür, dass Wissen tatsächlich symbolisch repräsentiert wird.

Allgemein muss man feststellen, dass die propositionale Darstellungsweise mentaler Wissensstrukturen sehr eng mit sprachlichen Informationen verknüpft ist und sich eigentlich auch nur auf diese gut anwenden lässt. Schon allein eine simple Wetterkarte von Deutschland ist auf der Grundlage einer rein symbolischen Repräsentation kaum zu analysieren, da räumliche Relationen mit sprachlichen Mitteln in der Regel nur unständig und unvollständig beschrieben werden können. Richtungsangaben wie *oben*, *unten*, *rechts* oder *links* sind sprachliche Begriffe, die sich vor allem im Zusammenhang mit wahrnehmungsbezogenen Informationen sinnvoll nutzen lassen. Versucht man hingegen diese Begriffe in einem rein bedeutungsbezogenen, propositionalen Kontext zu verwenden, der von allen wahrnehmungsbezogenen Einflüssen bereinigt wurde, so führt dies zwangsläufig zu Problemen. Natürlich ist es möglich, die Bedeutungen von *oben* oder *links* zu modifizieren, doch würde allein die vollständige Beschreibung eines Bildes nahezu unendliche viele Modifikationen und Bedeutungsketten erfordern (vgl. Johnson-Laird, 1980). Da es vor dem Hintergrund eher wahrnehmungsbezogener Informationen unwahrscheinlich ist, dass Wissen ausschließlich propositional repräsentiert wird, gehen andere Theorieansätze davon aus, dass stattdessen bzw. zusätzlich interne Repräsentationen konstruiert werden, die einen analogen Charakter haben. Kennzeichen und Eigenschaften dieser *analogen Wissensstrukturen* werden in dem folgenden Abschnitt näher dargestellt.

Mentale Modelle

Modelle sind Objekte, die anhand eines Originals konstruiert werden, um eine spezifische Aufgabe zu erfüllen. Zu diesem Zweck müssen Original und Modell aber hinsichtlich spezifischer Merkmale die gleichen Eigenschaften aufweisen. Beispielsweise fertigen Architekten vor einem größeren Umbau ein dreidimensionales Modell an, um einen besseren Eindruck von den Auswirkungen einzelner Entscheidungen zu gewinnen. Dieses Vorgehen hätte wenig Sinn, wenn die räumlichen Relationen innerhalb des Modells nicht den Relationen des Originals entsprechen würden. Diese Modellmetapher lässt sich auch auf

die Repräsentation interner Wissensstrukturen übertragen. Nur handelt es sich dann um *mentale Modelle* (Johnson-Laird, 1980, 1983), die eine analoge Repräsentation eines Sachverhaltes darstellen. Das zugrundeliegende Prinzip ist bei gegenständlichen und mentalen Modellen aber identisch (Garnham, 1980). Auch hier liegt der Konstruktion ein Original zugrunde, der Aufbau des mentalen Modells ist zweckgebunden und es muss eine Struktur oder Funktionsanalogie bestehen, damit es auf eine sinnvolle Weise genutzt werden kann.

Man kann sich mentale Modelle auch als hypothetische Quasi-Objekte vorstellen, die manipuliert bzw. modifiziert werden können, um Probleme oder Aufgaben zu lösen. Durch die Manipulation des mentalen Modells verändern sich dessen Merkmale, was letztlich Rückschlüsse auf die tatsächlichen Eigenschaften des Originals erlaubt. Dies ist allerdings nur möglich, wenn die Analogierelationen so beschaffen sind, dass sich Original und Modell auch wirklich hinsichtlich bestimmter Charakteristika entsprechen. Ähnlich wie bei gegenständlichen Modellen, werden niemals sämtliche Eigenschaften des darzustellenden Sachverhaltes abgebildet. Das mentale Modell und sein Original können also durchaus in vielerlei Hinsicht unterscheiden, solange die beiden funktional analog sind. Demnach sind mentale Modelle nicht einfach mit bildhaften Vorstellungen gleichzusetzen, obwohl es sich bei diesen auch um analoge, mentale Repräsentationen handelt, die durchaus Modellcharakter haben können (Johnson-Laird, 1983). Zentral ist lediglich, dass die strukturelle Übereinstimmung zwischen Repräsentandum und Repräsentant alle Funktionseigenschaften beinhaltet, welche zur Lösung eines spezifischen Problems gebraucht werden. Dieser Umstand kann dazu führen, dass von ein und dem selben Gegenstand unterschiedliche mentale Modelle konstruiert werden, je nachdem, welche Perspektive zugrunde gelegt wird. Die unterschiedlichen Perspektiven sind dabei eng mit der Aufgaben- oder Problemstellung verbunden, die letztlich auch den Anlass für die Modellkonstruktion darstellt. Doch nicht nur die Perspektive kann bei mentalen Modellen variieren. Je mehr *konzeptuelle Entitäten* (vgl. Greeno, 1983) in einem mentalen Modell enthalten sind, umso eher wird die Komplexität des Gegenstandes erfasst, der repräsentiert werden soll. Mit dem Detailliertheitsgrad steigt aber auch die Anforderungen an den Konstruktionsprozess, weshalb es ratsam erscheint, sich zielführend auf die wichtigsten Eigenschaften eines Sachverhaltes zu konzentrieren. So berichtet Larkin (1983), dass Lernende, die in einer bestimmten Wissensdomäne noch unerfahren sind, dazu neigen, Oberflächenmerkmale zu repräsentieren, die nicht unmittelbar der Bewältigung einer Aufgabe dienen. Konstruieren hingegen Experten ein mentales Modell, so sind darin alle wesentlichen strukturellen und funktionalen Merkmale enthalten. Die Differenziertheit der mentalen Modelle hängt also neben den Anforderungen, die ein bestimmtes Problem oder eine Aufgabe mit sich bringt, in sehr starkem Maße von den Eigenschaften des Individuums ab, das dieses Modell aufgebaut hat. Bei komplexen Gegenständen und Aufgaben ist es darüber hinaus oft erforderlich, mehrere mentale Modelle zu konstruieren, die verschiedene Blickwinkel auf ein Problem vereinen. Diese können dann durch *mapping*-Prozesse aufeinander bezogen werden, um letztlich den Sachverhalt besser zu verstehen (Schnotz, 1994a). Ein Spezialfall tritt dann ein, wenn es darum geht, dynamische Sach-

verhalte wie z.B. den Ablauf einer Kette kausaler Zusammenhänge zu repräsentieren. Das mentale Modell wird dann als Grundlage für eine Art innere Simulation verwendet (Schnotz & Rasch, 2005), deren Ergebnis, wie bei einem statischen Gegenstand, Rückschlüsse auf den realen Sachverhalt zulässt.

Alle diese Annahmen sind jedoch vor allem theoretischer Natur und ein empirischer Nachweis gestaltet sich ebenso komplex, wie die Frage, ob Wissen propositional repräsentiert wird. Zwar würden sicherlich die meisten Menschen der Behauptung zustimmen, dass man sich Bilder mental *vorstellen* kann, doch solche introspektive Aussagen sind problematisch. Allein die Tatsache, dass man die bildliche Vorstellung eines Gegenstandes wahrnehmungsnah erlebt, bedeutet noch nicht, dass die zugrunde liegende Repräsentation auch wirklich einen analogen Charakter haben muss. Aus diesem Grund sieht Pylyshyn (1973, 1981) subjektive bildhafte Vorstellungen als eine Art mentale Wahrnehmungen an, die aber letztlich auf propositional repräsentierten Wissenstrukturen basieren. Dieser Einwand ist ein Baustein in einer seit langem anhaltenden Diskussion, bei der es letztlich um die Frage geht, wie Wissen repräsentiert wird (Waskan, 2003). Die Anhänger der symbolischen Repräsentation sind daher die Schärfsten Kritiker der Anhänger der analogen Repräsentation und umgekehrt. Ein Umstand, welcher der Entwicklung von Theorien über Struktur und Format mentaler Repräsentationen nur genutzt haben dürfte. In diesem Zusammenhang ist auch die Arbeit von Kosslyn (1981) zu sehen, der beide Ansätze innerhalb einer Theorie zu vereinen versucht. Er geht davon aus, dass mentale visuelle Vorstellungen Ausdruck einer vergänglichen Datenstruktur sind, die in einem räumlichen analogen Medium des kognitiven Systems generiert werden. Die mentalen Bildern sind aber lediglich die *Oberfläche* einer Tiefenstruktur, die sich aus abstrakten Wissensrepräsentationen im Langzeitgedächtnis zusammensetzt. Diese abstrakte Wissensrepräsentation setzt sich aus Propositionslisten zusammen, die sowohl konzeptuelle und als auch perzeptuelle Inhalte umfassen. Gleichwohl hat der theoretische Ansatz von (Kosslyn, 1981) die analog-propositional-Debatte keineswegs beendet. Um dennoch die psychologische Realität analoger Repräsentationen nachzuweisen, versuchen verschiedene Forscher neben deren Informationsgehalt auch den zeitlichen Verlauf von mentalen Operationen zu erfassen. Einer der in der experimentellen Psychologie wohl bekanntesten Befunde stammt aus einer Untersuchung von Shepard und Metzler (1971). Die Autoren konnten belegen, dass bei Aufgaben, bei denen ein Objekt in der Vorstellung rotiert werden musste, die Lösungszeit und der Rotationswinkel in einem linearen Zusammenhang standen. Der gleiche Zusammenhang konnte auch für Aufgaben gefunden werden, bei denen Würfel mental gefaltet werden mussten (Shepard & Feng, 1972). Allein diese Ergebnisse weisen scheinbar eindeutig auf das Vorhandensein analoger mentaler Repräsentationen hin. Dennoch ist es möglich, das robuste Phänomen der *mentalen Rotation* auf der Grundlage symbolischer Wissensrepräsentationen zu erklären. So geht Pylyshyn (2002) davon aus, dass die wahrgenommene Bewegung eines ganzen vorgestellten Objektes erst gar nicht stattfindet. Vielmehr wird auf der Grundlage eines symbolischen Repräsentationssystems immer wieder das gleiche Objekt konstruiert, nur dass der Gegenstand jedes mal eine etwas andere räumliche Orientierung aufweist. Wird dieser Vorgang

häufig genug in sehr kurzen Abständen wiederholt, entsteht, wie bei einem Trickfilm, der subjektive Eindruck, einen dreidimensionalen Gegenstand tatsächlich mental zu rotieren.

Das Beispiel macht deutlich, dass es bislang nicht möglich war, eine endgültige empirische Antwort auf die Frage zu finden, ob Wissen nun propositional oder in Form eines mentalen Modells repräsentiert wird. Die einzelnen Befunde lassen sich, auch wenn es manchmal etwas umständlich erscheint, stets sowohl aus der einen, als auch aus der anderen Perspektive erklären. Aus diesem Grund integrieren verschiedene Theorien, die den Wissenserwerb mit Texten und/oder Bildern thematisieren (vgl. Kapitel 3), häufig Propositionen *und* mentale Modelle, wobei die beiden Repräsentationsformate in der Regel aufeinander aufbauen (Schnotz, 1994a). Propositionale Strukturen und mentale Modelle entstehen allerdings keinesfalls aus dem Nichts. Das konzeptionelle Wissen, das dem Aufbau dieser Repräsentationsformate zugrunde liegt, ist Gegenstand des folgenden Abschnitts.

Kognitive Schemata

Eine wichtige Gemeinsamkeit von mentalen Modellen und propositionalen Wissensstrukturen ist der Umstand, dass sie in der Regel zur Bewältigung von aktuellen Aufgaben konstruiert werden. Nachdem die jeweilige Aufgabe jedoch beendet ist, besteht kein Grund mehr, das mentale Modell oder die Propositionen noch länger aufrecht zu erhalten, so dass sie zerfallen und/oder durch neue Repräsentationen ersetzt werden (Mani & Johnson-Laird, 1982; Dutke, 1999). Damit sind mentale Modelle und propositionale Strukturen stark aufgabengebunden und haben einen eher flüchtigen Charakter. Trotz ihrer unbeständigen Natur können diese Repräsentationen in der Regel aber nur dann erfolgreich aufgebaut werden, wenn bestimmtes Wissen bereits zur Verfügung steht. So ist es denkbar schwer, ein funktionales mentales Modell von einem Gegenstand zu konstruieren, den man weder benennen kann, noch dessen Eigenschaften bekannt sind. Wenn man das Beispiel von Seite 16 betrachtet, werden viele Leser keine Ahnung haben, wer oder was *Catweazel* ist. Das Wissen, dass es sich bei *Catweazel* um eine Gestalt aus einer englischen Kinderserie handelt, die ziemlich zerzaustes Haar hat, ist aber entscheidend, um die Gültigkeit der Proposition einschätzen zu können. Damit setzt der Aufbau einer propositionalen Struktur das Vorhandensein von Informationen über die Entitäten voraus, aus denen sie sich zusammensetzt. Um also einen Sachverhalte angemessen zu repräsentieren, muss auf Wissen zurückgegriffen werden, das Konzeptionen von den darzustellenden Objekte enthält und im Bedarfsfall abgerufen werden kann. Das Vorhandensein eines solchen dauerhaften konzeptuellen Wissens wird theoretisch mit der Annahme sogenannter *Schemata* erfasst.

Dieser Begriff wurde vor allem von Bartlett (1932/1995) in die wissenschaftliche Psychologie eingeführt, wobei seine Verwendung philosophiegeschichtlich sehr viel älter ist (vgl. McVee, Dunsmore & Gavelek, 2005). Seit dem wurden kognitive Schemata auf verschiedene Weise spezifiziert, so dass sich anstelle einer einheitlichen Theorie, vielmehr ein theoretischer Rahmen herausgebildet hat, der einzelne Grundannahmen um-

fasst (Schnotz, 1994a). Den meisten Spezifikationen ist jedoch gemein, dass sie kognitive Schemata als Wissensstrukturen auffassen, die durch die Verallgemeinerung konkreter Erfahrungen aufgebaut und gefestigt wurden. Diese Verallgemeinerungen repräsentieren Sachverhalte und Zusammenhänge, wie sie typischer Weise im Laufe eines Lebens erlebt und erfahren werden. Gegenstand und Komplexität der Sachverhalte können dabei nahezu beliebig variieren. Demnach liegen dem Wissen, wie man eine andere Person begrüßt, ebenso ein oder mehrere Schema/ta zugrunde, wie dem Wissen, das man benötigt, um den Begriff *Nachhaltigkeit* zu erklären. Dabei ist es ganz gleich ob das jeweilige Schema soziale Ereignisse oder dem abstrakte Prinzipien umfasst, die Kennzeichen und Eigenschaften bleiben die selben (J. R. Anderson, 2001).

Nach einer Definition von Rumelhart und Ortony (1977) sind Schemata damit Datenstrukturen, die allgemeine Konzepte im menschlichen Gedächtnis repräsentieren. Sie bilden eine Struktur von Leerstellen, die in Anlehnung an einen Fachterminus aus der Informatik, auch als *slots* bezeichnet werden. Diese *slots* stehen für bestimmte Merkmale bzw. Attribute eines Sachverhaltes und können unterschiedliche Ausprägungen annehmen. Die typischen Ausprägungen eines Attributes werden als *Default-Werte* bezeichnet. So gehört beispielsweise zu dem Schema *Vogel* mit hoher Wahrscheinlichkeit das Merkmal, dass diese Lebewesen in der Regel fliegen können. Gleichzeitig ist es aber auch möglich Pinguine dem Schema *Vogel* zuzuordnen, selbst wenn für diese das Merkmal *Fliegen* nicht zutrifft, da Schemata nahezu beliebig ausdifferenziert sein können und mehr als nur ein Merkmal umfassen. So haben Pinguine, einen Schnabel, ein Gefieder und zählen nicht zu der Gattung der Säugetiere, was diese Tiere, trotz ihrer Fluguntüchtigkeit, zweifelsfrei als Vögel qualifiziert. Eine besondere Bedeutung in jedem Schema hat der Slot, der den Oberbegriff enthält und angibt zu welcher übergeordneten Kategorie das Schema gehört. So zählen Pinguine zur Objektklasse der Vögel, Vögel zählen zur Klasse der Tiere und Tiere kann man unter dem Oberbegriff der Lebewesen einordnen. Es entsteht eine Hierarchie, die von unten nach oben immer allgemeiner wird.

Auf den ersten Blick könnte man den Fehler begehen, Schemata als eine etwas komplexere Anordnung verschachtelter Propositionen aufzufassen. Immerhin lassen sich einige Merkmale ohne Schwierigkeiten in einer hierarchischen Struktur darstellen, wie das Beispiel mit den Vögeln nahe legt. Doch nicht allein die Annahme, dass Schemata auch wahrnehmungsbezogene Attribute umfassen, wie Farbe, Größe oder Form zeigt, dass diese mehr sind, als die bloße Erweiterung einer propositionalen Repräsentationsform. Der Unterschied zur Proposition wird offensichtlich, wenn man sich vor Augen führt, dass Schemata keinen bestimmten Sachverhalte repräsentieren, sondern Merkmale und Eigenschaften, die eine Klasse bzw. Kategorie von Sachverhalten in den meisten Fällen gemeinsam hat. Schemata enthalten damit abstraktes Regel-Wissen, das genutzt werden kann, um zu entscheiden, ob ein Sachverhalt einer bestimmten Kategorie angehört oder nicht. Sie stellen damit die Informationen zur Verfügung, welche für den Aufbau und die Beurteilung des Wahrheitsgehaltes einer Proposition benötigt werden. Da Schemata beliebig *tief* verschachtelt sein können, ist es außerdem möglich, sehr komplexe Sachverhalte zu repräsentieren. Hinzu kommt, dass die *slots* mit Default-Werten besetzt sein kön-

nen und das Nichtvorhandenseins eines Attributes nicht automatisch ein Ausschlusskriterium darstellen muss. Dies ist beispielsweise bei den Pinguinen erforderlich, die zwar Vögel sind, aber nicht fliegen. Damit diese dennoch unter der Klasse Vögel eingeordnet werden können, müssen bestimmte Default-Werte der Oberkategorie überschrieben werden. Auf diese Weise entstehen Unterkategorien, mit denen sich auch komplexe Merkmalsrelationen abbilden lassen. Wenn neue Objektklassen erzeugt bzw. vorhandene Strukturen ergänzt und ausdifferenziert werden, sprechen Rumelhart und Norman (1978) von *Lernen*. Diese Ansicht stimmt mit der Definition in Abschnitt 2.1 überein, da es sich bei Schemata um zeitlich relativ stabile Wissensrepräsentationen handelt, die verändert werden.

Gleichzeitig gewährleisten die Default-Werte bei einfachen Sachverhalten eine schnelle und effektive Verarbeitung, da sie die Ausprägungen eines Attributes darstellen, die den höchsten Erwartungswert haben. Erkennen wir, dass ein bestimmtes Objekt einer Objektklasse angehört, kann daraus gefolgert werden, dass es diejenigen Defaultwerte besitzt, die innerhalb des Schemas typisch für die Objektklasse sind. Stehen in einer Situation nicht ausreichend Informationen zur Verfügung, entweder weil beispielsweise ein Text mißverständlich ist oder man sich nicht mehr genau erinnern kann, lassen sich Schemata auch für Schlussfolgerungen heranziehen. Ein Beispiel für solch ein schlussfolgerndes Verhalten, das sich mit den Annahmen der Schematheorie erklären lässt, ist die Untersuchung von R. C. Anderson, Reynolds, Schallert und Geotz (1977). Die Autoren legten zwei Probandengruppen, die aus verschiedenen Bildungskontexten stammten, einen Text vor, der mehrdeutig war und daher *interpretiert* werden musste. Wie sich zeigte, spielte das Vorwissen und die Erfahrungen, auf welche die Versuchsteilnehmer bei der Interpretation der Texte zwangsläufig zurückgreifen mussten, eine entscheidende Rolle. Während die Gruppe weiblicher Psychologiestudentinnen den Text mehr oder minder einheitlich interpretierten, verstanden ihn die Gruppe, die sich aus männlichen Gewichthebern zusammensetzte, ganz anders. R. C. Anderson et al. (1977) macht für diesen Befund die unterschiedlichen Schemata und abweichende Default-Werte verantwortlich, die beim Lesen des Textes aktiviert wurden. Dass Vorwissen oder Erfahrungen neben dem Leseverständnis auch Gedächtnisleistungen beeinflussen können, zeigt eine Studie von Brewer und Treyens (1981). So war die Wiedergabe von Gegenständen eines Experimentalraumes, in dem eine Gruppe von Probanden für kurze Zeit warten musste, erheblich von deren *Büroschema* geprägt. Dies hatte nicht nur zur Folge, dass Gegenstände besser erinnert wurden, deren Vorhandensein den Default-Werten eines *typischen* Büros entsprachen. Es wurden zudem Objekte erinnert, die zwar in einem Büro für gewöhnlich vorkommen, aber in dem tatsächlichen Warteraum nicht vorhanden waren. Folglich wurde die Erinnerung der Versuchsteilnehmer nach Ansicht der Autoren von den Default-Werten ihrer individuellen Büroschemata beeinflusst. Allerdings gibt es auch Einwände, welche die psychologische Realität von Schemata anzweifeln. Ein Kritikpunkt betrifft die Verwendung von ambigen bzw. experimentell veränderten Texten in den Studien. Diese konnten zwar belegen, dass der persönliche Hintergrund und auch die Situation, in der gelesen wird, wichtig für das Verstehen sind. Dennoch sei es allein auf dieser Grundlage problematisch, auf eine schemageleitete Informationsverarbeitung zu schließen (Sadoski,

1981). Auch Alba und Hasher (1983) würdigen die Bedeutung der Schematheorie für die Berücksichtigung des Vorwissens bei der Erforschung der kognitiven Prozesse beim Lesen, sehen aber das größte Problem in der *Unschärfe* der Theorie. Außerdem reichen nach Ansicht der beiden Autoren Schemata nicht aus, um Gedächtnisprozesse zu beschreiben, die umfangreichen, detaillierten und teilweise auch irrelevanten Erinnerungen zugrunde liegen. Schließlich bieten Sadoski, Paivio und Goetz (1991) mit der *Theorie der dualen Kodierung*, die in Kapitel (3) noch ausführlicher diskutiert werden wird, einen konzeptionellen Ansatz an, der die Phänomene aus den Schemata-Studien sparsamer erklären kann. Trotz dieser Einwände bleibt festzuhalten, dass mit der Schematheorie ein gelungener Versuch unternommen wurde, das allgemeine *Weltwissen* theoretisch zu fassen, das der Konstruktion von Propositionen oder mentalen Modellen zugrunde liegt. Dafür spricht nicht nur die Tatsache, dass eine wissenschaftlich fruchtbare Diskussion in Gang gekommen ist, die eine Vielzahl von neuen Forschungsansätzen zur Folge hatte. Wie McVee et al. (2005) zeigen, ist es ungeachtet der vieler Einwände nicht gelungen, die Schematheorie tatsächlich zu ersetzen.

2.3 Das Gedächtnis als kognitive Architektur

Mentale Repräsentationen, wie sie in den vorangegangenen Kapitel behandelt wurden, entstehen keinesfalls aus dem Nichts. Propositionen, mentale Modelle und Schemata sind Gegenstand und/oder Endprodukte einer ganzen Reihe von kognitiven Prozessen, die im Zuge des Erwerbs neuen Wissens ablaufen. Beim Lernen werden Informationen aufgenommen und gespeichert, während gleichzeitig Operationen zur Verarbeitung und Integration von bereits vorhandenen Gedächtnisinhalte zur Verfügung stehen müssen. Um die Gesamtheit dieser Prozesse zu systematisieren, bedient sich die wissenschaftliche Psychologie sogenannter *Gedächtnismodelle bzw. -theorien*, von denen im folgenden einige prominente Vertreter in chronologischer Reihenfolge vorgestellt werden.

2.3.1 Das Mehrspeichermodell

Das erste, umfassende Modelle des menschlichen Gedächtnisses legten Atkinson und Shiffrin (1968) vor. In enger Anlehnung an die Arbeitsweise von Computern unterscheiden die beiden Autoren Strukturmerkmale und Kontrollprozesse. Demzufolge besteht die permanente Struktur des menschlichen Gedächtnis aus drei separaten Speichern: dem sensorischen Register, dem Kurzzeitspeicher und dem Langzeitspeicher. Werden Informationen aus der Umwelt aufgenommen und verarbeitet, durchlaufen diese alle drei Gedächtniskomponenten in einer festgelegten Reihenfolge, wie Abbildung 2.1 verdeutlicht.

Die Abfolge des Informationsflusses unterliegt dabei keiner Kontrolle und verläuft unabhängig von den Inhalten immer gleich. Dem Modell entsprechend, gelangen Informationen aus der Umwelt über die Sinnesorgane zuerst in modalitätsspezifische sensorische Register, wo sie in *Rohform* für sehr kurze Zeit gespeichert werden. Da diese Informationen noch sehr reiznah kodiert sind, findet hier noch keine bedeutungshaltige Analyse

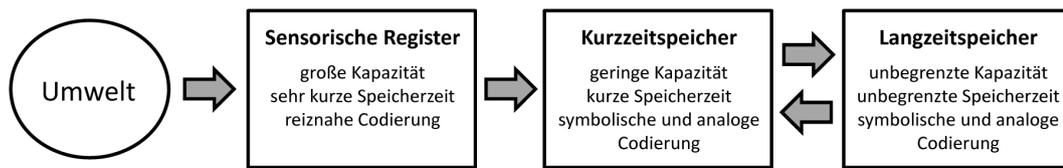


Abbildung 2.1: Mehrspeichermodell nach Atkinson & Shiffrin (1968)

statt. Der eigentlich Zweck der sensorischen Register liegt darin, Reizmuster kurzfristig festzuhalten, damit einzelne Informationen selektiv herausgelesen und in die nachfolgenden Speicher übertragen werden können. Zu diesem Zweck ist die Kapazität der sensorischen Register sehr hoch, wobei die Inhalte innerhalb weniger Millisekunden wieder zerfallen. Da diese kurzfristige Speicherung reiznah erfolgt, postulieren Atkinson und Shiffrin (1968) mehrere modalitätsspezifische Register. So nehmen die Autoren an, dass es eine visuellen oder ikonischen Register (*iconic memory*) und einen auditiven bzw. echoischen Register (*echoic storage*) gibt.

Von den sensorischen Registern gelangen die Informationen zunächst in den Kurzzeitspeicher, in dem aber auch Inhalte aus dem Langzeitspeicher verarbeitet werden, wie die Pfeile in der Abbildung 2.1 andeuten. Letztlich ist er die zentrale Einheit des Modells, da hier die meisten bewusstseinspflichtigen Prozesse ablaufen, die der Weiterverarbeitung dienen. Dennoch ist die Kapazität des Kurzzeitspeichers begrenzt und die Inhalte stehen nur wenige Sekunden zur Verfügung, wenn sich die Aufmerksamkeit auf etwas anderes richtet. Das Format der Informationen kann dabei ganz unterschiedlich sein. So ist es durchaus möglich, den Kurzzeitspeicher als den Ort anzusehen, an dem mentale Modelle oder auch propositionale Repräsentationen konstruiert werden, obwohl die Autoren von einer Dominanz auditiv-verbaler Kodierungsformen ausgehen.

Der letzte strukturelle Bestandteil des Modells von Atkinson und Shiffrin (1968) ist der sogenannte *Langzeitspeicher*. Hier lagert das gesamte Wissen eines Menschen, das er im Lauf seines Lebens erwirbt. Er verfügt nach Ansicht der Autoren über eine unbegrenzte Kapazität und alle Informationen, die einmal dort abgelegt wurden, sind permanent. Die vorherrschende Kodierungsform soll semantischer Natur sein, so dass mit Blick auf den vorangegangenen Abschnitt, die Inhalte des Langzeitspeichers auch in schematischer Form abgespeichert sein könnten. Wäre das Modell gültig, so bestünde letztlich das Ziel aller Lernprozesse darin, die Inhalte des Langzeitspeichers zu verändern, denn nur dann hat *Lernen* auch tatsächlich stattgefunden (vgl. Kapitel 2.1). Ob jedoch eine Übertragung von Inhalten aus dem Kurzzeitsspeicher bzw. den sensorischen Registern in den Langzeitspeicher erfolgt, ist das Ergebnis von bewussten und unbewussten Prozessen. Bei den bewussten Prozessen handelt es sich letztlich um die anfangs erwähnten Kontrollprozesse. Damit sind vor allem Gedächtnisstrategien gemeint, die der bewusst gesteuerten Regelung des Informationsflusses zwischen den drei Speichereinheiten dient. Eine der bekanntesten Strategien ist beispielsweise der Prozess der Wiederholens (*rehear-*

sa/). Diesem Prozess liegt die einfache Überlegung zugrunde, dass durch das Festhalten von Informationen im Kurzzeitspeicher gewährleistet werden soll, dass diese zu einem permanenten Bestandteil des Langzeitspeichers werden.

Auch wenn das Mehrspeichermodell auf den ersten Blick überaus plausibel und übersichtlich erscheint, gilt es mittlerweile als überholt, wobei sich die Kritik vor allem auf dessen strukturelle Merkmale bezieht (Kluwe, 1992). Zum einen bestehen grundsätzliche Zweifel an dem Konzept der sensorischen Register und der Gültigkeit der experimentellen Studien, mit denen diese nachgewiesen wurden (Haber, 1983). So wird die Annahme, dass die ersten Verarbeitungsschritte in einem separaten Gedächtnis geleistet werden, als überflüssig erachtet (Neisser, 1983), da die Wahrnehmung als kontinuierlicher Prozess und nicht als Abfolge von diskreten Schritten angesehen wird. Ein durchaus ähnlicher Kritikpunkt betrifft den Kurzzeitspeicher, dessen Vorhandensein ebenfalls abgelehnt wird. In neueren Modellen tritt an dessen Stelle das sogenannte *Kurzzeitgedächtnis*, das allerdings kein eigenständiges Speichereinheit mehr darstellt. Vielmehr ist das Kurzzeitgedächtnis derjenige Teil des Langzeitspeichers bzw. -gedächtnisses, der gerade aktiviert ist und für bewusste Verarbeitungsschritte zur Verfügung steht. Es handelt sich hier also um einen vorübergehenden Zustand bestimmter Inhalte und Informationselemente des Langzeitgedächtnisses (J. R. Anderson, 1976, 1983a). Auf diese Weise konnten die problematischen Übergänge bzw. der Austausch zwischen den unterschiedlichen Speichereinheiten elegant umgangen werden. Eine Gedächtnistheorie, die anstelle von mehreren separaten Speichern von unterschiedlichen Ebenen der Informationsverarbeitung ausgeht, wird im nachfolgenden Absatz dargestellt.

2.3.2 Levels of Processing

Auch Craik und Lockhart (1972) bezweifeln die Existenz multipler Speicher und schlagen mit den *Levels-of-Processing* einen alternativen theoretischen Ansatz vor, um Gedächtnisphänomene zu erklären. Nach Ansicht der Autoren durchlaufen Informationen, die aus der Umwelt in das kognitive System gelangen, einen kontinuierlichen Prozess, bei dem Wahrnehmungs- und Gedächtnisleistungen ineinander greifen. So werden beispielsweise beim Lesen eines Wortes zunächst nur visuelle Reize wie Linien, Winkel, die Farben Schwarz oder Weiss wahrgenommen. Dieser sensorische Input erhält eine Bedeutung, wenn ein Abgleich mit vorhandenen Gedächtnisinhalten erfolgreich ist und Buchstaben *erkannt* werden. Gelingt es dann der Ansammlung von Buchstaben eine Bedeutung zuzuweisen, können weitere Verarbeitungsschritte erfolgen, indem das Wort bzw. seine Bedeutung *elaboriert* wird. Mit dem Begriff der Elaboration bezeichnen Craik und Lockhart (1972), einen Vorgang, bei dem neue Informationen mit vorhandenen Informationen in Beziehung gesetzt und semantisch analysiert werden. Je mehr Elaborationsprozesse ablaufen, umso mehr Verknüpfungen entstehen zwischen neuen und alten Gedächtnisinhalten und umso tiefer ist die Verarbeitung.

Da einzelne Elaborationsschritte nur mit einer begrenzten Informationsmenge weitergeführt werden können, ist es manchmal notwendig, zusätzliche Inhalte vor einer weiteren

Verarbeitung verfügbar zu halten. Zu diesem Zweck werden Informationen, beispielsweise durch beständiges Wiederholen, *memorierend* verarbeitet. So bleiben die Inhalte zwar auf ein und derselben Ebene, doch führen allein elaborierende Prozesse zum Aufbau dauerhafter Gedächtnisstrukturen. Nach Ansicht der Autoren hinterlässt demnach jeder einzelne Verarbeitungsschritt eine Gedächtnisspur, die dazu genutzt werden kann, um bereits verarbeitete Informationen zu rekonstruieren. Inhalte werden dadurch erinnert, indem man einzelne Ebenen, die eine bestimmte Information durchlaufen hat, nochmals nachvollzieht bzw. wiederverarbeitet. Da bei einer tiefen Verarbeitung mehr Gedächtnisspuren erzeugt werden, die als Ausgangspunkt für eine solche Rekonstruktion verwendet werden können, führt dies im Vergleich zu einer oberflächlichen Verarbeitung zu besseren Behaltensleistungen. Zudem gehen die Autoren davon aus, dass die Gedächtnisspuren der elaborierten Verarbeitung an sich stärker und dauerhafter sind.

Der *Levels-of-Processing*-Ansatz stellte zwar eine brauchbare Alternative zu den vorherrschenden Mehrspeichermodellen dar, doch wurde auch diese Theorie mehrfach kritisiert. Insbesondere die Vorstellung einer linearen Abfolge von Verarbeitungsschritten wurde von verschiedenen Autoren abgelehnt (Schnotz, 1994a). Anstelle eines streng hierarchischen Ablaufs, sprechen verschiedene Befunde dafür, dass sich die Prozesse auf den unterschiedlichen Verarbeitungsebenen wechselseitig beeinflussen. So können Prozesse auf einer niederen Hierarchieebene genau so gut tiefe Elaborationsprozesse initiieren, wie umgekehrt Oberflächenprozesse von tiefen Verarbeitungsebenen gesteuert werden können. Zudem zeigen mehrere Experimenten, dass auch eine rein memorierende Verarbeitung von Informationen zu langfristigem Behalten führen (T. O. Nelson, 1977; Craik & Tulving, 1975). Demzufolge führen nicht allein die Elaborationsprozesse der tiefen Verarbeitungsebenen zu stabilen und überdauernden Gedächtnisspuren. In diesem Zusammenhang weist Eysenck (1978) auf eine weitere Schwäche der ersten Fassung des *Levels-of-Processing*-Ansatzes hin. So ist es nicht das Kriterium der Verarbeitungstiefe, das darüber entscheidet, ob eine Gedächtnisspur zu einer Erinnerung führt oder nicht. Vielmehr hängt das erfolgreiche Erinnern einer bestimmten Information von der Diskriminierbarkeit der Gedächtnisspur und damit von dem Kontext ab, in dem die entsprechenden Inhalte gelernt wurden.

Die vielen Kritikpunkte haben dazu geführt, dass die ursprüngliche Version der *Levels-of-Processing-Theorie* überarbeitet und erweitert wurde (vgl. Craik, 1979). Diese revidierte Fassung integriert kontextabhängige kognitive Prozesse, interagierende Verarbeitungsebenen und die Annahme, dass auch memorierende Prozesse zu dauerhaften Gedächtnisspuren führen können. Während sich die Arbeit von Craik und Lockhart (1972) bzw. Craik (1979) nahezu ausschließlich auf Prozesse konzentriert, stehen im folgenden Abschnitt die Strukturen wieder stärker im Mittelpunkt.

2.3.3 Das Arbeitsgedächtnis

Einer der ältesten und sicherlich auch bedeutensten theoretischen Ansätze zur Modellierung des kognitiven Apparates stammt von Baddeley und Hitch (1974). Die beiden

Autoren griffen das Konzept des Mehrspeichermodells von Atkinson und Shiffrin (1968, siehe auch Kapitel 2.3.1) auf und entwickelten es anhand der Ergebnisse einer ganzen Reihe von Experimenten weiter. Baddeley und Hitch (1974) widmen sich insbesondere dem Kurzzeitspeicher, den sie aufgrund eigener Überlegungen als *working memory* bzw. *Arbeitsgedächtnis* bezeichnen. So ist das Arbeitsgedächtnis jenes System des menschlichen Gehirns, in dem Informationen kurzzeitig gespeichert und in diesem Zeitraum vorwiegend bewusst manipuliert oder verarbeitet werden können. Diese Vorgänge sind notwendig, um komplexe kognitive Anforderungen wie das Verstehen von Sprache, den Erwerb von Wissen oder schlussfolgerndes Denken zu bewältigen.

Im Unterschied zum Kurzzeitspeicher von Atkinson und Shiffrin (1968) besitzt das Arbeitsgedächtnismodell eine eigenständige Substruktur. Demnach beinhaltet die erste Fassung des Modells drei Komponenten: die zentrale Exekutive (*central executive*), den visuell-räumlichen Notizblock (*visuospatial sketchpad*) und die phonologische Schleife (*phonological loop*). Die zentrale Exekutive ist die komplexeste Subkomponente des Modells, die sehr viele Funktionen erfüllt, sich einer genauen Analyse aber weitgehend entzieht. Baddeley (1992) nimmt jedoch an, dass eine ihrer Aufgaben darin besteht, die Informationen zu koordinieren, die in den abhängigen Systemen (den sogenannten *slave systems*) verarbeitet werden. Diese Koordination ist zwingend erforderlich, da Aufmerksamkeit eine *mentale Ressource* ist, die aufgrund der beschränkten Kapazität des Arbeitsgedächtnisses nicht unbegrenzt zur Verfügung steht. Die beiden abhängigen Systeme der ursprünglichen Fassung sind dagegen sehr viel simpler und dienen vereinfacht gesagt der Verarbeitung und kurzzeitigen Aufrechterhaltung modalitätsspezifischer Informationen aus der Umwelt. Während im Subsystem des visuell-räumlichen Notizblocks visuelle Reize und Inhalte verarbeitet und gespeichert werden, dient die phonologische Schleife der Verwertung bzw. Bereitstellung von akustischen Informationen, wie sie beispielsweise mit gesprochener Sprache transportiert werden.

Obwohl die erste Version des theoretischen Ansatzes vor allem auf der Grundlage zahlreicher empirischer Untersuchungen entstanden ist, hat das Modell auch einige Schwächen. Einer der zentralen Kritikpunkte betrifft die nur sehr vage spezifizierten Funktionen der zentralen Exekutive. Baddeley (2001) selbst bezeichnet diese ironischer Weise als *Homunculus* oder theoretischen *ragbag* (=Lumpensack), in den man lästige Fragen hinein tun kann. Hinzu kommt, dass in der ursprünglichen Fassung die Rolle des Langzeitgedächtnisses weitgehend unberücksichtigt bleibt, wodurch einige Befunde sich nicht erklären lassen. Der letzte Kritikpunkt bezieht sich schließlich auf die Tatsache, dass zwei modalitätsgebundene Subsysteme zur Verarbeitung visueller und akustischer Signale existieren, es aber weitgehend unklar bleibt, wie und wo sich Informationen aus beiden Subsystemen integrieren lassen. Vor diesem Hintergrund hat Baddeley (2000, 2001) seinen theoretischen Ansatz überdacht und ein modifiziertes Modell des Arbeitsgedächtnisses entwickelt, dessen Struktur Abbildung 2.2 zeigt.

Wie man auf den ersten Blick erkennen kann, umfasst die aktuelle Fassung des Modells nun weitaus mehr, als nur drei Komponenten. So ist mit dem episodischen Zwischenspeicher (*episodic buffer*) ein zusätzliches Subsystem des Arbeitsgedächtnisses

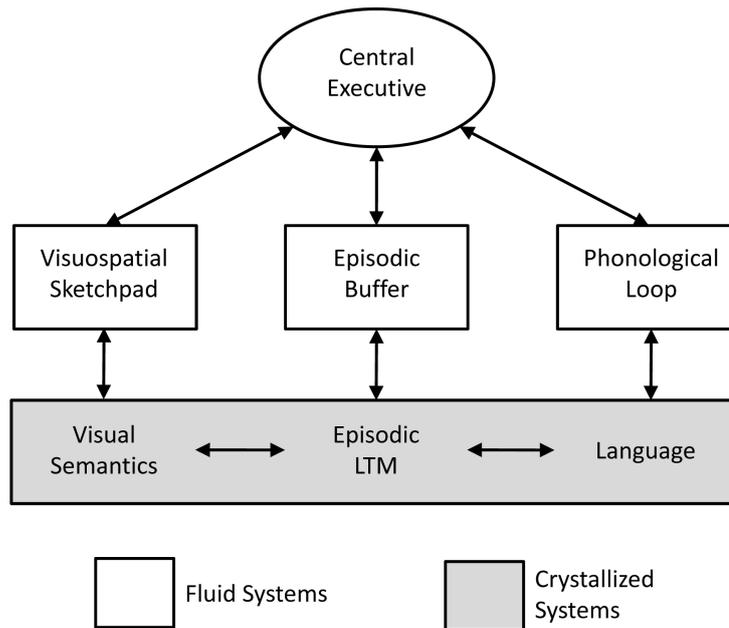


Abbildung 2.2: Revidierte Fassung des Arbeitsgedächtnisses nach Baddeley (2000, 2001)

hinzugekommen, das einen multimodalen Code verwendet und daher die Integration von Informationen aus den anderen beiden abhängigen Systemen zur Aufgabe hat. Der episodische Zwischenspeicher ist damit die Komponente innerhalb der kognitiven Struktur, in der bewusst mentale Modelle bzw. propositionale Strukturen konstruiert werden können. Gleichzeitig fungiert er als eine Art Schnittstelle zwischen Arbeits- und Langzeitgedächtnis, das in dem aktuellen Fassung des Modells eine sehr viel stärkere Berücksichtigung findet. Beispielsweise erfordert das Verstehen bzw. Erinnern komplexer Textpassagen nach Ansicht von Baddeley (2001) Inhalte, die im Langzeitgedächtnis abgelegt sind. Diese Inhalte können über den episodischen Zwischenspeicher aktiviert und damit einer bewussten Verarbeitung zugänglich gemacht werden. Der Ansatz ist den Überlegungen von Ericsson und Kintsch (1995) nicht unähnlich und ermöglicht darüber hinaus die Erklärung von Gedächtnisphänomenen wie etwa dem *chunking* (Miller, 1956).

Doch auch die phonologische Schleife und der visuell-räumliche Notizblock sind in der neuen Version des Modells mit dem Langzeitgedächtnis verbunden. Einerseits werden für das Erkennen von optischen Mustern oder akustischen Reizen Informationen aus dem Langzeitgedächtnis benötigt, andererseits werden neue Inhalte eingespeist. Dieser Austausch, der in Abbildung 2.2 durch bidirektionale Pfeile symbolisiert ist, findet allerdings weitgehend unbewusst statt, da es sich um den Abruf und Aufbau von eher implizitem Wissen handelt (Baddeley, 2001). Mit der Einbindung des Langzeitgedächtnisses und dem episodischen Zwischenspeicher hat sich auch die Rolle der zentralen Exekutiven verändert. Baddeley (2000) sieht diese nicht mehr als eine Art *Hilfsspeicher* oder Schnittstelle zu den Inhalten des Langzeitgedächtnisses an. Nachdem der episodische Zwischenspeicher diese Funktionen übernommen hat, dient die zentrale Exekutive ausschließlich der Steuerung der Aufmerksamkeit bzw. Speicherkapazität, welche auf die

drei Subsysteme verteilt werden muss. Aufmerksamkeit und Speicherkapazität sind dabei als begrenzte Ressourcen zu verstehen, die als *fluid* Teil des kognitiven Systems festlegen, wie viele bewusste Prozesse gleichzeitig stattfinden können. Man kann an den weiß hinterlegten Kästchen erkennen, dass alle Komponenten des Arbeitsgedächtnisses auf diesen fluiden Teil angewiesen sind. Folglich ist dessen Verarbeitungskapazität beschränkt, zumal sich die fluiden Bestandteile nicht durch Lernprozesse verändert lassen. Im Gegensatz dazu gehören die grau unterlegten Strukturen des Langzeitgedächtnisses zum *kristallinen* Teil des kognitiven Systems, das sich durch den Erwerb neuen Wissens sehr wohl verändern lässt. Mit dieser Unterscheidung greift Baddeley (2000, 2001) das Konzept der *fluiden* und *kristallinen* Intelligenz von Cattell (1963) auf und integriert es strukturell in sein Modell.

Bei der abschließenden Betrachtung der überarbeiteten Fassung fällt auf, dass der Ansatz von Baddeley (2001) über ein bloßes Modell des Kurzzeit bzw. Arbeitgedächtnisses weit hinausgeht. Vielmehr handelt es sich um einen umfassenden theoretischen Entwurf, der die kognitive Architektur zum Gegenstand hat. Durch die konzeptionelle Integration des Langzeitgedächtnisses ist das Arbeitsgedächtnismodell eigentlich ein Mehrspeichermodell, auch wenn der Fokus der Überlegungen auf der Struktur und Funktionsweise bewusster kognitiver Prozesse liegt. Doch trotz der umfangreichen Erweiterung des Modells konnten zentrale Kritikpunkte nicht ausgeräumt werden bzw. kommen neu hinzu. So ist die Rolle der zentralen Exekutive immer noch äußerst unscharf definiert und behält weitgehend ihren Status als theoretischer *Homunculus*. Allgemein werden zwar die Funktionsweisen der abhängigen Systeme beschrieben und gedanklich erfasst, gleichzeitig vermeidet es der Autor aber, auf die Prozesse zwischen den einzelnen *Sklavensystemen* einzugehen. Ein weiteres Problem ist die Beschränkung auf die Verarbeitung visueller und akustischer Informationen, so dass beispielsweise haptische und olfaktorische Reize innerhalb des theoretischen Ansatzes unberücksichtigt bleiben. Darüber hinaus ist die Beziehung zwischen der Arbeitsgedächtniskapazität und dem Konzept der fluiden Intelligenz trotz intensiver Forschungsbemühungen immer noch nicht eindeutig geklärt, wie Yuan, Steedle, Shavelson, Alonzo und Oppezzo (2006) in einem Überblicksartikel darstellen. Nach Meinung der Autoren dürfte dies vor allem an den grundlegenden Unterschieden bei der Messung, Terminologie und den eingesetzten statistischen Methoden liegen.

Doch trotz aller Kritik ist es Baddeley und Hitch (1974) bzw. Baddeley (2000, 2001) gelungen, ein Modell vorzulegen, anhand dessen sich viele prüfbare Hypothesen ableiten lassen. Auf diese Weise wurde zu zahlreichen neuen Forschungsansätzen und Theorien angeregt, wie in den folgenden Kapiteln noch zu sehen sein wird. Doch zunächst widmet sich der nächste Abschnitt den Gedächtnismodellen, in deren Mittelpunkt das *Langzeitgedächtnis* steht.

2.3.4 Das Langzeitgedächtnis

Wie bereits angedeutet wurde, führten die theoretischen Probleme mit der Annahme mehrerer Speicher zur Entwicklung einer Reihe theoretischer Ansätze, bei denen nur noch eine *Gedächtnisform* im Mittelpunkt steht. Einer der ersten dieser Ansätze, den Kluwe (1992) als *Ein-Speicher-Modelle* bezeichnet, stammt von Shiffrin und Schneider (1977) und kann als Weiterentwicklung des ursprünglichen Mehrspeicheransatzes (vgl. Atkinson & Shiffrin, 1968) verstanden werden. Dem Modell liegt die Annahme zugrunde, dass die sensorischen Register und der Kurzzeitspeicher keine separaten Systemkomponenten darstellten. Aus diesen Überlegungen resultiert, dass letztlich nur noch eine große Einheit für alle Gedächtnisleistungen verantwortlich ist, das *Langzeitgedächtnis*. Wie im ursprünglichen Modell, umfasst das Langzeitgedächtnis alles Wissen, das ein Mensch im Laufe seines Lebens erwirbt, ganz gleich ob es Sachverhalte, persönliche Erlebnisse oder das Wissen um die eigene Person betrifft. Da es keinen separaten Kurzzeitspeicher mehr gibt, wird angenommen, dass eine bewusste Verarbeitung von Informationen nun durch die Aktivierung einer Teilmenge des Langzeitgedächtnisses stattfindet. Obwohl es sich um einen temporären Zustand handelt, der vorübergehend einzelne Gedächtnisinhalte zugänglich macht, wird diese Teilmenge als *Kurzzeitgedächtnis* bezeichnet.

Grundsätzlich wird angenommen, dass alle Inhalte, die einmal in das Langzeitgedächtnis gelangt sind, dort permanent gespeichert werden. Der ebenso unbestreitbare wie alltägliche Vorgang des *Vergessens* besteht also nicht darin, dass Gedächtnisinhalte zerfallen oder sich in Nichts auflösen (Raaijmakers & Shiffrin, 1981). Gelingt es nicht, sich an etwas zu erinnern, so ist ganz einfach der Versuch fehlgeschlagen, die entsprechenden Inhalte zu aktivieren. Dies wird umso anschaulicher, wenn man sich die (theoretische) Struktur des Langzeitgedächtnis vor Augen führt, wie sie sich Shiffrin und Schneider (1977) vorstellen. Nach Ansicht der Autoren ist das Langzeitgedächtnis ein riesiges Netzwerk, das sich aus einer gewaltigen Anzahl von Knotenpunkten zusammensetzt, die vielfältig miteinander verknüpft sind. Diese Knotenpunkte sind die elementaren Informationseinheiten des Netzwerkes, da in ihnen das eigentliche Wissen enthalten ist. Sie können sowohl eine einzige, einfache Information repräsentieren, als auch aus komplexen, verdichteten Strukturen bestehen, so dass jeder Knotenpunkt unterschiedlich viele Informationen enthält. Lernt man beispielsweise ein neues Wort in einer fremden Sprache, so wird die resultierende Informationseinheit zunächst aus den Buchstaben, den phonetischen Eigenschaften und den wichtigsten Bedeutungen bestehen. Macht der Lernende dann weitere Fortschritte beim Fremdspracherwerb, kommen neue Informationen wie Synonyme oder zusätzliche Verwendungsalternativen hinzu, die zusätzlich in den Knotenpunkt integriert werden. Allerdings machen Shiffrin und Schneider (1977) keine näheren Angaben darüber, welche genauen Merkmale und Funktionen die Verknüpfungen im Netzwerk haben. Immerhin läge es nahe, über die Qualität oder Anzahl der Verknüpfungen zu erklären, wie die Aktivierung von Knotenpunkten funktioniert oder weshalb Menschen etliche Inhalte vergessen.

Etwa zur gleichen Zeit veröffentlichte (J. R. Anderson, 1976) die erste Fassung seiner ACT-Theorie, die später zu ACT* (J. R. Anderson, 1983a) bzw. ACT-R (J. R. Anderson, 1996) weiterentwickelt wurde. Auch dieser Ansatz geht letztlich von einem Langzeitgedächtnis aus, das die Struktur eines Netzwerkes hat und aus miteinander verknüpften Elementen besteht. Ein grundlegender Unterschied zwischen der Gedächtnis-Konzeption von J. R. Anderson (1976, 1983a) und dem Modell von Shiffrin und Schneider (1977) besteht aber vor allem darin, dass die ACT-Theorie von zwei verschiedene Wissensformen im Langzeitgedächtnis ausgeht. So differenziert ACT ausdrücklich zwischen deklarativem und prozeduralem Wissen (siehe auch Kapitel 2.2.1). Dieser Unterscheidung entspricht auch die Annahme, dass den beiden Wissensformen unterschiedliche Repräsentationsformate zugrunde liegen. Während im ACT-Modell das deklarative Wissen die Form eines propositionalen Gefüges bzw. semantischen Netzes annimmt, besteht das prozedurale Wissen aus einer Vielzahl systematisch angeordneter Produktionen. Jede Produktion enthält eine *Anwendungsbedingung* und eine ausführbaren *Aktionen*, die als Wenn-Dann-Anweisung miteinander verknüpft sind, wie das folgende Beispiel verdeutlicht.

Wenn das Lämpchen der Kaffeemaschine rot leuchtet, dann muss Wasser nachgefüllt werden.

Da Produktionen beliebig miteinander verschachtelt sein können, etwa indem zum Erreichen eines übergeordneten Ziels, erst Zwischenschritte geplant und durchgeführt werden müssen, lassen sich auch sehr komplexe kognitive Prozesse mit Hilfe von Produktionssystemen modellieren. Im Gegensatz dazu, setzt sich das deklarative Wissen aus miteinander verbundenen *kognitiven Einheiten* zusammen, die J. R. Anderson (1996) in Anlehnung an Miller (1956) auch als *chunks* bezeichnet. Das englische Wort *chunk* (Klumpen) repräsentiert im Rahmen Gedächtnisforschung eine Wissenseinheit, die aus verdichteten Informationen besteht. Demzufolge kann ein *Knotenpunkt* des deklarativen Netzwerkes nicht nur ein, sondern auch mehrere Informationselemente enthalten, die systematisch oder strukturell zusammenhängen. Trotz dieser konzeptionellen Trennung sind beide Wissensformen eng miteinander verknüpft, denn sowohl die Anwendungsbedingung, als auch die Aktion werden durch die Inhalte der deklarativen Wissensstrukturen benannt und definiert. So kann eine bestimmte Produktion nur dann zur Anwendung kommen, wenn die wahrgenommenen Bedingungen der Definition im deklarativen Wissen entsprechen. Neben dem prozeduralen und deklarativen Langzeitgedächtnis gibt es noch eine weitere wichtige Komponente innerhalb der ACT-Theorie, das *Arbeitsgedächtnis*. Hier findet der eigentliche Austausch mit der Umwelt statt, wie Abbildung 2.3 grafisch verdeutlicht.

Die Überlegungen, die dem Arbeitsgedächtnis von ACT zugrunde liegen, sind durchaus mit den Vorstellungen von Shiffrin und Schneider (1977) vergleichbar. Es stellt keinen separaten Speicher dar, sondern ist ein Teil des Langzeitgedächtnisses, der sich in einem temporären Zustand der Aktivierung befindet. Da immer nur wenige Knotenpunkte bzw. *chunks* gleichzeitig aktiviert sein können, besitzt auch das Arbeitsgedächtnis der ACT-Theorien nur eine begrenzte Kapazität. Ob und welche Knotenpunkte aktiviert und damit einer bewussten Verarbeitung zugänglich werden, hängt allerdings von verschiede-

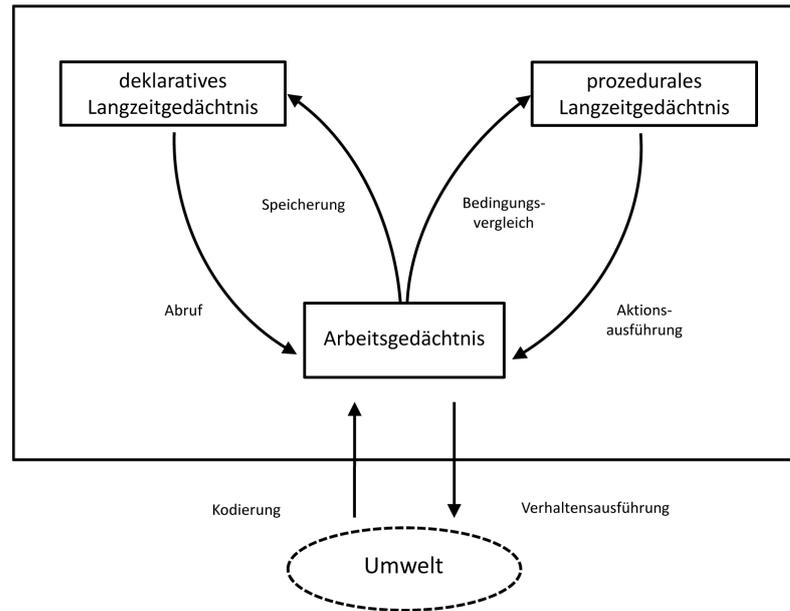


Abbildung 2.3: Kognitive Komponenten der ACT*-Theorie nach Anderson (1983)

nen Bedingungen ab. Die Aktivierung kann die direkte Folge einer Wahrnehmung sein, beispielsweise indem ein bestimmter Reiz *wiedererkannt* und einer entsprechende Bedeutung zugeordnet wird. Kognitive Einheiten können aber auch durch die Anwendung von Produktionen, die bestimmte Such- oder Konstruktionsprozesse zum Inhalt haben, in das Arbeitsgedächtnis gelangen. Befindet sich eine kognitive Einheit erst einmal im Zustand der bewussten Verarbeitung, dann stellt auch sie eine Aktivationsquelle dar.

J. R. Anderson (1983b) geht davon aus, dass sich die Aktivierung, wie in einem neuronalen Netzwerk, von Knotenpunkt zu Knotenpunkt ausbreitet. In diesem Zusammenhang spielt die *Stärke* der *chunks* bzw. Wissensseinheiten eine wichtige Rolle. So nimmt der Autor aufgrund empirischer Ergebnisse an (vgl. J. R. Anderson, 1983b; Ratcliff & McKoon, 1981), dass die Aktivierung immer zu dem anliegenden Knoten *springt*, der die größte Stärke aufweist. Die Stärke eines *chunks* kann in der Logik des ACT-Modells als Maß dafür angesehen werden, wie tief die entsprechende Informationseinheit in den Gedächtnisstrukturen verankert ist. Sie ist umso höher, je häufiger und länger sich eine kognitive Einheit im Arbeitsgedächtnis befindet. Auf diese Weise lässt sich einfach und anschaulich erklären, wie innerhalb von ACT gelernt wird. Reiz- oder Symbolfolgen gelangen über die Schnittstelle Arbeitsgedächtnis in das kognitive System, wo diese enkodiert und mit einer Bedeutung versehen werden. Handelt es sich um bekannte Inhalte, erhöht sich die Stärke der Knoten bzw. *chunks*, in denen diese Informationen ganz oder teilweise enthalten sind. Ist die Information hingegen neu, wird eine neuer Knoten *angelegt*, der allerdings noch sehr schwach in den Wissenstrukturen verankert ist. Doch nicht nur der Erwerb neuen, *deklarativen* Wissens lässt sich mit ACT erklären. Das Modell ermöglicht es zudem, Aussagen über das Erlernen kognitiver Fähigkeiten zu machen. J. R. Anderson

(1976, 1983a) nimmt an, dass Wissen, das zum Erwerb einer kognitiven Fähigkeit erforderlich ist, zunächst weitgehend in deklarativer Form vorliegt. Durch Anwendung und fortschreitende Übung wird dieses deklarative Wissen in prozedurales Wissen überführt bzw. *prozeduralisiert*. Der Vorgang der *Prozeduralisierung* lässt sich dabei durchaus mit dem Erwerb motorischer Fertigkeiten vergleichen. Sind bei den ersten Anwendung noch viele bewusstseinspflichtige Prozesse beteiligt, die vor allem der Steuerung und Überwachung dienen, nehmen diese mit jeder weiteren Ausführung ab. Aus der deklarativen *Handlungsanweisung*, die man aufwendig im Arbeitsgedächtnis aufrecht erhalten und manipulieren muss, werden nach und nach Produktionen. Die weitere Verdichtung und Zusammenfassung der einzelnen Produktionen führt schließlich zur weitgehenden Automatisierung der zuvor noch bewusstseinspflichtigen Prozesse, die damit Teil des prozeduralen Langzeitgedächtnisses werden.

Betrachtet man Abbildung 2.3, dann scheint die ACT-Theorie auf den ersten Blick recht übersichtlich und einfach zu sein, doch dieser Eindruck trügt. Obwohl nur Ausschnitte dargestellt wurden, handelt es sich um den wohl umfassendsten Ansatz in diesem Kapitel. Mehr als jeder andere Theorie, beinhalten die Arbeiten von J. R. Anderson (1976, 1983a) konkrete Aussagen über die Architektur des kognitiven Apparates und den zugrunde liegenden Prozessen. Damit ist ACT weitaus mehr als *nur* ein Gedächtnismodell. Viele Zusammenhänge wurden mathematisch formuliert und soweit präzisiert, dass es möglich ist, die postulierten kognitiven Vorgänge computergestützt zu simulieren. Auf diese Weise lassen sich die Ergebnisse *realen* Verhaltens mit Daten vergleichen, die auf der Grundlage der theoretischen Modells errechnet wurden. Allein die Möglichkeit erscheint zuerst einmal interessant und faszinierend zu sein, doch verbergen sich darin auch Probleme. So formuliert J. R. Anderson (1983a) die grundlegende Annahme, dass es nur ein einziges kognitives System gibt, das die Quelle aller geistigen Leistungen ist. ACT modelliert dieses kognitive System, indem es sparsame, aber sehr präzise Annahme über den Ablauf basaler Informationsverarbeitungsprozesse macht. Da die Theorie ganzheitlich zu verstehen ist, *muss* ihre empirische Prüfung letztlich über eine Simulation und den Vergleich mit tatsächlichem Verhalten erfolgen, womit allerdings eine ganze Reihe von Schwierigkeiten verbunden ist. Zunächst einmal stellt sich die Frage, welche kognitive Leistung man zur Evaluation von ACT heranziehen will und wie sich diese von anderen geistigen Fähigkeiten zumindest theoretisch abgrenzen lässt. Eine noch größere Herausforderung stellt hingegen die Charakterisierung der zugrunde liegenden Prozesse dar, um das konstituierende deklarative und/oder prozedurale Wissen zu bestimmen. Bei einer Analyse von so feiner Granularität, immerhin versucht man sehr basale kognitive Operationen zu erfassen, die sich teilweise nicht einmal bewusstseinspflichtig sind, werden aber zwangsläufig Fehler gemacht. Daher muss ein unpassendes oder abweichendes Simulationsergebnis nicht zwangsläufig bedeuten, dass die Theorie falsch ist. Zu den praktischen Problemen mit der empirischen Prüfbarkeit des Modells, kommen schließlich noch theoretische Einwände, welche die Produktionssysteme an sich betreffen. So nehmen kennektionistische Modelle an (siehe beispielsweise Feldman & Ballard, 1982), dass die menschliche Informationsverarbeitung nicht *sequentiell*, also als Abfolge

von ineinander verschachtelter Produktionen erfolgt. Die Tatsache, dass sich kognitive Prozesse häufig in Form von Wenn-Dann-Verknüpfungen darstellen lassen, ist lediglich das Erscheinungsbild einer subsymbolischen Verarbeitung in *neuronalen Netzwerken*, die auf parallelen Prozessen beruht. Doch dies führt an dieser Stelle zu weit.

Es lässt sich festhalten, dass die ACT-Theorie immer noch ein äußerst faszinierender Ansatz ist, dessen Weiterentwicklung sich im Zuge größerer Rechnerkapazitäten und des technischen Fortschritts beschleunigen dürfte. So ist ihr *Schöpfer*, der weiterhin mit seinem Modell arbeitet, nach einigen Verbesserungen mittlerweile bei Version ACT-R 5.0 angelangt (J. R. Anderson et al., 2004).

Der dritte und letzte Ansatz, bei dem das Langzeitgedächtnis eine zentrale Rolle für die Erklärung von Phänomenen der menschlichen Informationsverarbeitung einnimmt, stammt von Ericsson und Kintsch (1995). Die beiden Autoren versuchen mit der Theorie des *Long-Term-Working-Memory* Befunde zu erklären, bei denen Experten über eine viel höhere Verarbeitungskapazität verfügen, als dies Experimente zur Erfassung Arbeitsgedächtnisspanne nahe legen (vgl. Miller, 1956; Oberauer, Süß, Wilhelm & Wittmann, 2000, 2003). Zwar gehen auch Ericsson und Kintsch (1995) grundsätzlich von einer Unterscheidung zwischen Kurzzeit- und Langzeitgedächtnis aus, doch nehmen sie gleichzeitig an, dass Personen, die spezifische Kenntnisse und Kompetenzen besitzen, das Langzeitgedächtnis als eine Art zusätzlichen Arbeitsspeicher nutzen können. Diese Erweiterung des *Short-Term-Working-Memory* ist allerdings an eine Reihe von Voraussetzungen gebunden. Die erste Voraussetzung betrifft den Umstand, dass nur Experten oder sehr geübte Menschen in der Lage sind, bei einzelnen, domänenspezifischen Aktivitäten ihren Arbeitsspeicher zu erweitern. So setzt die Nutzung des *Long-Term-Working-Memory* umfangreiche und tiefe Wissenstrukturen voraus, da nur diese es erlauben, anfallende Informationen rasch zu integrieren und strukturiert im Langzeitgedächtnis abzulegen. Außerdem gehen Ericsson und Kintsch (1995) davon aus, dass die domänenspezifische Aktivität dem Experten sehr vertraut sein muss, zumal es erforderlich ist, Kontext und Art der zu speichernden Inhalte zu antizipieren. Denn damit sich die Informationen des *Long-Term-Working-Memory* abrufen lassen, werden diese mit Abruf-Reizen, den sogenannten *retrieval cues* versehen.

Bei diesen *retrieval cues* kann es sich um einfache Assoziationen oder kontextbezogene Hinweise handeln, die anzeigen, wann die entsprechenden Inhalte innerhalb des Verarbeitungsprozesses relevant sind. Befinden sich Elemente, die diesen Abruf-Reize ansprechen im *Short-Term-Working-Memory*, dann können die im Langzeitgedächtnis abgelegten Informationen aktiviert und genutzt werden. Im Zusammenhang von komplexen kognitiven Aufgaben ist es in der Regel erforderlich, die *retrieval cues* hierarchisch zu strukturieren. Das Organisationssprinzip dieser hierarchischen *retrieval structures* kann dabei räumlicher, zeitlicher oder auch thematischer Natur sein, je nachdem, um welche Anforderung bzw. Domäne es sich handelt. Damit ist das *Long-Term-Working-Memory* keine zusätzlicher Speicher, der sich bei beliebigen kognitiven Aktivitäten nutzen lässt. Es ist vielmehr ein Mechanismus oder eine Fertigkeit, die erworben wird, wenn effizient

entstrukturierte Wissensrepräsentationen im Langzeitgedächtnis dies ermöglichen. Das bringt allerdings mit sich, dass eine Verwendung des *Long-Term-Working-Memory* für jede Domäne und Aktivität separat erlernt werden muss.

Inhaltlich greift die Theorie auf theoretischer Überlegungen von Ericsson und Chase (1982) zurück, die dann mit Bezug auf das *Construction Integration Model* von Kintsch (1988) zuerst für das Verstehen von Texten entwickelt wurde. Die Autoren gehen aber davon aus, dass sich die Nutzung des *Long-Term-Working-Memory* auch bei anderen kognitive Fertigkeiten beobachten lässt. So begreifen Ericsson und Kintsch (1995) jeden kognitiven Prozess als eine Abfolge nachgeordneter Operationen, bei den Informationen aus der Umwelt und/oder dem Langzeitgedächtnisses verarbeitet werden. Darüber hinaus sind insbesondere komplexe Prozesse auf Inhalte angewiesen, die das Produkt von vorangegangenen Operationen darstellen. Diese Zwischenschritte bzw. Teilergebnisse müssen aber solange verfügbar sein, bis eine Weiterverarbeitung durch nachfolgende Prozesse stattfinden kann. Nach Ansicht von Vertretern der Arbeitsgedächtnistheorie (vgl. Baddeley & Hitch, 1974) werden letztlich alle Informationselemente, welche für die Bewältigung einer Anforderung relevant sind, zeitlich begrenzt im Arbeitsgedächtnis aufrecht erhalten. Dagegen nehmen Ericsson und Kintsch (1995) an, dass bei stark eingeübten Aktivitäten ein großer Teil der nutzbaren Informationen im Langzeitgedächtnis abgelegt wird, der über kontextsensitive Abruf-Reize schnell in das *Short-Term-Working-Memory* gelangen kann.

In diesem Zusammenhang stellt sich allerdings die Frage, wie sich nachweisen lässt, dass eine Versuchsperson beim Lesen eines Textes, tatsächlich auf Inhalte zurückgreift, die eigentlich dem Langzeitgedächtnis entstammen. Eine Unterscheidung zwischen Informationen aus dem *Short-* und dem *Long-Term-Working-Memory* ist nach Ansicht der Autoren dann möglich, wenn eine kognitive Aktivität durch eine Ablenkung an einer bestimmten Stelle des Prozesses unterbrochen wird. Wenn die Unterbrechung lange genug ist, dann sind alle Informationen, die temporär im Arbeitsgedächtnis enthalten waren, unwiederbringlich verloren und die unterbrochene Tätigkeit kann nicht unmittelbar fortgesetzt werden. Konnte hingegen ein großer Teil der aufgabenrelevanten Inhalte im *Long-Term-Working-Memory* gespeichert werden, dann lassen sich diese anhand der *retrieval structure* reaktivieren, so dass eine Fortsetzung der unterbrochenen Aktivität möglich ist.

Um diese Annahmen mit argumentativ zu stützen, beziehen sich Ericsson und Kintsch (1995) vor allem auf empirische Befunde zum Textverstehen. Beispielsweise zeigen verschiedene Studien (vgl. Glanzer, Fischer & Dorfman, 1984; Fischer & Glanzer, 1986; Ericsson, 1988), dass sich die Verstehens- und Gedächtnisleistungen von Versuchspersonen, die während des Lesens unterbrochen werden oder laut denken müssen, dadurch kaum vermindert. Die beiden Autoren sehen in diesen Ergebnissen einen Hinweis dafür, dass umfangreiche Informationen bei der Verarbeitung von Texten im Langzeitgedächtnis abgelegt werden. Den deutlichsten Beleg für die Annahme des *Long-Term-Working-Memory* liefern nach Meinung von Ericsson und Kintsch (1995) aber Forschungsarbeiten, die den Einfluss von verbalen Fertigkeiten und domänenspezifischem Wissen systematisch variiert haben (vgl. Recht & Leslie, 1988; W. Schneider, Körkel & Weinert, 1989).

Diese Studien ergaben, dass die Gedächtnis- und Verstehensleistung allein durch das Vorwissen der Versuchspersonen beeinflusst wurde, nicht aber durch die Lesekompetenz oder allgemeine kognitive Fähigkeiten, wie der Intelligenz. So schnitten beispielsweise Probanden mit großem domänenspezifischem Wissen, aber niedrigem IQ, in den einem nachfolgenden Wissenstest deutlich besser ab, als Versuchsteilnehmer mit hohem IQ, aber niedrigem Vorwissen.

Ericsson und Kintsch (1995) übertragen ihren theoretischen Ansatz auf eine ganze Reihe weiterer Befunde, die sich auf außergewöhnliche kognitive Leistungen in anderen Domänen beziehen. So erklären sie die Überlegenheit von Experten auch beim Kopfrechnen, Schach oder medizinischen Diagnosen mit der erlernten Nutzung von Strukturen des Langzeitgedächtnisses. Allerdings gibt es auch Einwände gegen diese Generalisierung des *Long-Term-Working-Memory* als Grundlage von Expertenleistungen. Beispielsweise merken Vicente und Wang (1998) an, dass man bei der Erforschung von Expertise zwischen zwei unterschiedlichen Domänenarten differenzieren muss. Demnach gibt es Aufgaben, bei denen die Gedächtnisleistung ein inhärenter Bestandteil der zu erbringenden Leistung ist, etwa beim Erinnern von Zahlenreihen. Und es gibt Domänen, bei denen kein *erweitertes Gedächtnis* erforderlich ist, so dass eine Prüfung der Gedächtnisleistung einen künstlichen bzw. forschungstechnisch unökologischen Akt darstellt. Folglich kann man die Fähigkeit eines Schachspielers nur feststellen, indem man ihn tatsächlich Schach spielen lässt und nicht indem man prüft, wie gut er sich an bestimmte Figurenkonstellationen erinnern kann. Wenn außerordentliche Gedächtnisleistungen also kein inhärenter Bestandteil der Expertise sind, muss die Gültigkeit der Theorie des *Long-Term-Working-Memory* bezweifelt werden. Vicente und Wang (1998) führen mehrere Studien an, die nicht mit den Vorhersagen des Ansatzes von Ericsson und Kintsch (1995) übereinstimmen. Sie sollen belegen, dass die Theorie des *Long-Term-Working-Memory* zu kurz greift, da Struktur und Kontext der eigentlichen Expertisentätigkeit nicht ausreichend berücksichtigt werden. Ericsson, Kintsch und Patel (2000) weisen dagegen alle Einwände zurück und interpretieren die Ergebnisse der gegen sie angeführten Studien auf der Grundlage ihres ursprünglichen Ansatzes, ohne diesen verändert zu haben. Aber auch neuere Arbeiten, die sich mit dem Schreiben von Texten (Kellogg, 2001) oder Lesen mit Unterbrechungen (D. W. Schneider & Dixon, 2009) beschäftigen, sprechen für zentrale Annahmen der Theorie.

Zusammenfassend lässt sich festhalten, dass die Annahme eines *Long-Term-Working-Memory* plausible Erklärungsmöglichkeiten bietet, wie große Mengen an Informationen, die sich nicht mehr im Arbeitsgedächtnis befindet, der bewussten Verarbeitung wieder zugänglich gemacht werden können. Der Ansatz geht über ein ressourcenbegrenzt Gedächtnis im traditionellen Sinn hinaus und bietet gleichzeitig die Möglichkeit, den Einfluss von aufgabenrelevanten Vorwissen theoretisch zu erfassen. Eine Theorie, die sich wieder mehr den Kapazitätsbeschränkungen des kognitiven Apparates widmet, wird im Folgenden beschrieben.

2.3.5 Die Cognitive Load Theory

Der letzte Abschnitt dieses Kapitels beschäftigt sich mit einem Ansatz, der innerhalb der Pädagogischen Psychologie wohl zu den bekanntesten und einflussreichsten der letzten anderthalb Jahrzehnte zählt. Dies liegt sicherlich zum großen Teil an der praktischen Ausrichtung der *Cognitive Load Theory*, da sie ausdrücklich auf die optimale Vermittlung von Wissen bzw. Fertigkeiten abzielt. Andererseits lässt sich die Theorie unkompliziert auf *neue* Lernumgebungen und -materialien anwenden, welche im Zusammenhang mit der rasanten technischen Entwicklung entstanden sind.

Das ursprüngliche Konzept stammt von Sweller (1988) bzw. Sweller und Chandler (1991), die sich zunächst intensiv mit Lernprozessen beim Problemlösen beschäftigten. Anhand verschiedener Experimente konnten die beiden Autoren beobachten, dass die selbständige Bearbeitung von Problemen als Instruktionmethode nicht automatisch zu günstigen Lernergebnissen führte. Insbesondere wenn das spezifische Vorwissen fehlte, war der Wissenszuwachs unter den Versuchsteilnehmern vergleichsweise gering. Sweller (1988) bzw. Sweller und Chandler (1991) nehmen an, dass die beschränkte Verarbeitungskapazität des kognitiven Systems für die suboptimalen Lernresultate verantwortlich ist. Demnach ist das Problemlösen für ungeübte Lerner kognitiv so anspruchsvoll, dass sie keine oder nur noch wenige Ressourcen für den aktiven Schemaaufbau aufbringen können. Dies hängt mit der Beschaffenheit des *Arbeitsgedächtnisses* zusammen, in dem immer nur eine begrenzte Anzahl von Elementen gleichzeitig verarbeitet werden kann. Jedes Element beansprucht dabei einen Teil der begrenzten kognitiven Ressourcen und stellt somit eine mentale Belastung dar, die Sweller (1988) in ihrer Summe als *cognitive load* bezeichnet. Neben der Anzahl der Elemente im Arbeitsspeicher sind es die Komplexität der Verarbeitungsprozesse und die Eigenschaften der einzelnen Person, hier insbesondere ihr Vorwissen, die das Ausmaß der kognitiven Beanspruchung mitbestimmen. Damit gehen auch Sweller (1988) bzw. Sweller und Chandler (1991) von einem Arbeitgedächtnis mit beschränkter Kapazität aus. Gleichzeitig nehmen sie aber ein unbegrenztes Langzeitgedächtnis an, dessen Schematastrukturen die bewusste Verarbeitung aber erheblich erleichtern können. Diese Annahmen stellen aus theoretischer Perspektive zunächst keine Neuerung dar, insbesondere wenn man sie den Ansätzen vergleicht, die in den vorangegangenen Abschnitten dieses Kapitels bislang dargestellt wurden. Neu hingegen sind die praktischen Implikationen, die mit der Konzeption des *cognitive loads* verbunden sind. Denn nach Ansicht der beiden Autoren lassen sich Lernprozesse nur dann optimal gestalten, wenn grundlegende Erkenntnisse über die kognitive Architektur des Menschen genutzt werden. In diesem Zusammenhang fordern sie, die Wissensvermittlung so zu gestalten, dass möglichst wenig *cognitive load* durch die *Instruktionmethode* selbst verursacht wird.

Beispielsweise können auch unerfahrene Lerner in der Lage sein, ein spezifisches Problem durch die Anwendung allgemeiner Heuristiken zu lösen. Diese *naive* Strategie ist aber oft umständlich, da sie in Regel aufwendige Mittel-Zweck-Analysen (*means-ends-analysis*) umfasst. Um festzustellen, ob sich durch die Anwendung eines bestimm-

ten Operators der aktuelle Problemzustand dem Zielzustand angenähert hat, muss ein großer Teil des Problemraums mental repräsentiert werden. Die Aufgabe kann so vielleicht erfolgreich bearbeitet werden, doch lässt die mentale Beanspruchung kaum Ressourcen übrig, um effektivere Lösungsansätze zu entwickeln. Verfolgt ein Lehrender aber genau dieses Ziel, so ist allein die Bearbeitung von Problemen keine angemessene Instruktionsmethode.

Eine alternative Vorgehensweise entdeckten Cooper und Sweller (1987), indem sie Personen mit geringen Vorkenntnissen ausgearbeitete Beispielaufgaben vorlegten. Ohne den *cognitive load* der mit der eigenständigen Problembearbeitung verbunden ist, verfügen selbst Anfänger dann über genügend kognitive Ressourcen, um effektive problemspezifische Schemata und Operatoren zu entwickeln. So lernten Novizen durch die Bearbeitung von *worked out examples* mehr, als wenn sie eigenständig Probleme lösten.

Neben dem positiven Auswirkung des Einsatzes von Beispielaufgaben führen Sweller und Chandler (1991) noch weitere Möglichkeiten an, die unnötige Erhöhung des *cognitive loads* beim Wissenserwerb zu vermeiden. So konnten Ward und Sweller (1990) beobachten, dass auch der Gestaltung des Lernmaterials eine große Bedeutung zukommt. Sind Lernende auf Informationen aus mehreren Quellen angewiesen, um bestimmte Inhalte vollständig zu verstehen, dann müssen diese Quellen mental integriert werden. Eine solche Integration ist aber mit *cognitive load* verbunden, da die Aufmerksamkeit geteilt werden muss. Die beiden Autoren nennen das den *split-attention-effect*, dessen Auswirkungen sich aber vermindern lassen, indem die Informationsquellen räumlich in größtmöglicher Nähe dargeboten werden. Beispielsweise sollten die Erläuterungen bzw. die Legende eines Diagrammes direkt in die Abbildung integriert werden und nicht erst auf einer Seite später zu lesen sein.

Im Zusammenhang mit dem *split-attention-effect* konnten Sweller und Chandler (1991) ein weiteres Phänomen identifizieren, das gut in den Rahmen der *Cognitive Load Theory* passt. So führt die Integration zweier separater Informationsquellen, wenn sie inhaltlich redundant sind, tendenziell eher zu einer geringeren Lernleistung. Die Autoren erklären diesen *redundancy-effect* mit der Annahme, dass die Verarbeitung redundanter Informationen mentale Ressourcen bindet, die nicht mehr für den Aufbau von Wissensstrukturen genutzt werden können. Folgerichtig führte eine Beseitigung der Redundanzen zur Verbesserung der Lernergebnisse, wie Chandler und Sweller (1991) in einer experimentellen Studie belegen konnten.

Mit den verschiedenen Effekten, die Sweller und Chandler (1991) im Rahmen ihrer empirischen Arbeiten spezifizierten, entwickelte sich auch ihre Theorie konzeptionell weiter. War zu Beginn noch jede mentale Belastung generell als *cognitive load* bezeichnet worden, nahmen die Autoren mit der Zeit eine Differenzierung vor. Neben dem *extraneous load*, der als Folge des Instruktionsdesigns entsteht identifizieren Sweller und Chandler (1994) noch eine weitere Quelle der mentalen Beanspruchung. Dem liegt die einfache Überlegung zugrunde, dass Problemstellungen, Aufgaben oder auch Lernmaterialien hinsichtlich ihrer Schwierigkeit variieren können. So beansprucht die Auseinandersetzung mit einem schwierigen Text die begrenzte Kapazität des kognitiven Systems weitaus

mehr, als dies ein leichter Text tun würde.

Unabhängig von der Instruktionsform verursacht damit jeder Inhalt, wenn er das Objekt eines Lernprozesses wird, allein aufgrund seiner Beschaffenheit eine mentale Belastung. Sweller und Chandler (1994) bezeichnen diese inhaltsimmanente Form der mentalen Beanspruchung als *intrinsic load*. Seine Höhe wird durch die Interaktivität der Elemente bestimmt (*element interactivity*), die gleichzeitig im Arbeitsgedächtnis gehalten werden müssen. Beispielsweise ist das Lernen von Vokabellisten mit geringem *intrinsic load* verbunden, da die *element interactivity* nahezu Null ist. In der Regel kann die Bedeutung jedes einzelnen Wortes separat erlernt werden, ohne dass Relationen zu vor- oder nachgeordneten Worten in der Liste zu berücksichtigen sind. Liest man hingegen einen vollständigen Satz, dann sind die einzelnen Worte aber keine unabhängigen Elemente mehr. Hier müssen die syntaktischen Relationen (z.B. Subjekt-Prädikat-Objekt) zwischen den Worten berücksichtigt werden. Das macht es wiederum notwendig, mehrere Informationselemente gleichzeitig im Arbeitgedächtnis zu halten, die darüber hinaus noch miteinander interagieren. Der damit verbundene *intrinsic load* ist also vergleichsweise hoch.

Viele Menschen haben beim Lesen aber keineswegs den Eindruck, dass sie gerade eine Tätigkeit ausüben, die mit einer hohen mentalen Belastung verbunden ist. Dies liegt vor allem daran, dass die *element interactivity* nicht allein von den Charakteristika der Aufgabe abhängt. Die Eigenschaften des Lernenden sind mindestens ebenso wichtig, da vorhandene Fertigkeiten und Vorwissen letztendlich (mit-)bestimmen, wie viele interagierende Elemente gleichzeitig repräsentiert werden müssen. So können vier interagierende Elemente im Arbeitgedächtnis eines Anfängers inhaltlich dem einen Element im Arbeitgedächtnis eines Experten entsprechen. Der einzige Unterschied besteht darin, dass das Wissen des Experten einfach besser strukturiert ist. Jede Schätzung des *intrinsic loads* einer bestimmten Aufgabe kann also immer nur unter Berücksichtigung der Lernermerkmale erfolgen. Da eine solche Einschätzung im Einzelfall aufwendige Analysen und Diagnosen nach sich ziehen würde, ist dies oft nur sehr eingeschränkt möglich.

Aus diesem Grund ist es einfacher, grob darauf zu achten, dass die inhaltliche Komplexität den Voraussetzungen des Lernenden zumindest einigermaßen entspricht. Praktisch wurde der *intrinsic load* als mehr oder weniger stabil angesehen, auch wenn dieser über Aufgaben und Personenmerkmale variierte. Die Bemühungen der meisten Forscher konzentrierten sich vor allem auf die Reduktion des *extraneous load*, da hier die besten Möglichkeiten gesehen wurden, um hilfreiche Instruktionsstrategien zu entwickeln (vgl. Mayer & Moreno, 2003). Eine ähnliche Absicht hatten vielleicht auch Paas und van Merriënboer (1994), als sie den Zusammenhang zwischen der Variabilität von Beispielaufgaben und Lernerfolg untersuchten. Dabei zeigte sich, dass Versuchsteilnehmer, die viele verschiedene *worked-out-examples* zum Wissenserwerb nutzen konnten, auch bessere Ergebnisse erzielten. Erstaunlich war aber, dass die Variabilität der Beispielaufgaben den *cognitive load* während der Lernphase zwar erhöhte, aber gleichzeitig der Transferleistung der Probanden zugute kam. Eine mentale Belastung beim Lernen muss demnach nicht zwangsläufig negative Konsequenzen für den Wissenserwerb haben. Um diese Befunde in die Theorie zu integrieren, führten Sweller, van Merriënboer und Paas (1998)

eine positive Form des *cognitive loads* ein, den sogenannten *germane load*.

Als *germane load* kann demnach jede mentale Belastung angesehen werden, die infolge des Aufbaus oder der Modifikation kognitiver Schemata entsteht. Sie tritt immer dann auf, wenn ein Teil der Kapazitäten des Arbeitsgedächtnisses dazu genutzt wird, um im Sinne der Definition aus Kapitel 2.1 zu *lernen*. Als praktische Konsequenz aus der Einführung des *germane load*, ergibt sich, dass bei der Gestaltung von Lernumgebungen und -materialien vor allem zwei Formen der mentalen Beanspruchung berücksichtigt werden müssen. Während der *extraneous load* nach wie vor minimiert werden soll, begünstigt eine Erhöhung des *germane load* die Effektivität der Wissensvermittlung. Dabei gilt grundsätzlich das Prinzip der *Additivität* zwischen den unterschiedlichen Load-Formen (Sweller, 2005). So ergibt die Summe aus *intrinsic*, *extraneous* und *germane load* die mentale Gesamtbelastung, der eine Person beim Lernen ausgesetzt ist. Sind beispielsweise *intrinsic* und *extraneous load* so hoch, dass sie die Kapazität des Arbeitsgedächtnisses übersteigen, bleiben keine Ressourcen für den Aufbau kognitiver Schemata und es wird nichts bzw. kaum etwas gelernt.

Obwohl es unzweifelhaft ist, dass die *Cognitive Load Theory* einen wichtigen Beitrag zur Erforschung von Lern- und Instruktionsprozessen leistet, gibt es auch einige Kritikpunkte. Ein zentrales Problem betrifft die Frage, wie sich die mentale Belastung überhaupt messen und damit nachweisen lässt. Gerade in den frühen Studien wurde der *Cognitive Load* zunächst indirekt durch die Manipulation der Versuchsbedingungen erfasst. Mit dem wachsenden praktischen Anspruch der Theorie entstand aber das Bedürfnis, die mentale Belastung auch außerhalb eines streng kontrollierten experimentellen Kontextes zu bestimmen. Vor diesem Hintergrund entwickelten Paas, van Merriënboer und Adam (1994) eine Skala zur Selbsteinschätzung, auf der Lernende retrospektiv den wahrgenommenen *load* berichten. Dieses Verfahren ist einfach und lässt sich ohne großen Aufwand anwenden, was für die ökologische Validität dieser Methode spricht. Ein solches Vorgehen unterliegt aber grundsätzlich der Problematik aller *Ratingskalen* und den damit einhergehenden Urteilsfehlern subjektiver Einschätzungen (Bortz & Döring, 1995). Daher gibt es zusätzliche Bemühungen, *objektivere* Messverfahren zu entwickeln (vgl. Brünken, Steinbacher, Plass & Leutner, 2002), um die zentrale Konzeption der Theorie auf eine breitere empirische Basis zu stellen. Allerdings existiert bislang noch keine Methode, welche die einzelnen *load*-Formen getrennt erfassen kann (J. van Merriënboer, Schuurman, Crook & Paas, 2002; Schnotz & Kürschner, 2007). Die meisten Ergebnisse, die im Zusammenhang mit *Cognitive Load*-Ansätzen berichtet werden, beziehen sich also auf eine in der Regel *erlebte* Gesamtbelastung.

Ein weiterer Kritikpunkt betrifft die stark vereinfachte Annahme, dass eine Reduktion der mentalen Belastung zwangsläufig zu besserem Lernen führt. Schnotz und Rasch (2005, 2008) beobachteten, dass Versuchsteilnehmer mit geringem Vorwissen von den Hilfestellungen, die eine Animation bot, nicht profitieren konnten. Nach Meinung der Autoren war in diesem Fall jedoch eine zu niedrige, und nicht eine zu hohe mentale Beanspruchung für die schlechten Resultate verantwortlich. Der Wissenserwerb mit einer animierten Simulation verlangte den Lernenden einfach zu wenig ab. Anstatt die dargestellten

Prozesse mental nachzuvollziehen, verfolgten die Lernenden die Geschehnisse auf dem Bildschirm, ohne sich selbst dabei anzustrengen. Der *intrinsic load* der Animationsbedingung war für Versuchsteilnehmer, die über wenig Vorwissen verfügten, offensichtlich zu niedrig.

Das Beispiel zeigt, dass die Annahme eines stabilen *intrinsic loads* für die optimale Gestaltung von Lernumgebungen und -materialien nur wenig zuträglich ist. In diesem Zusammenhang sehen auch J. J. G. van Merriënboer und Sweller (2005) es als überaus wichtig an, Lernanforderungen und Personenmerkmale flexibel aufeinander abzustimmen. Das kann allerdings nur gelingen, wenn zuverlässige Methoden und Verfahren vorhanden sind, die sowohl die Voraussetzungen der Lernenden, als auch den *intrinsic load* erfassen können. Und genau hier sehen die Autoren einen Schwerpunkt zukünftiger Forschungsarbeiten im Rahmen der *Cognitive Load Theory*. Auf der Grundlage zuverlässiger Messungen wäre es dann auch möglich, *adaptive* Lernumgebungen zu entwickeln, die sich an die wachsende Expertise einer Person anpassen.

Ein weiterer Kritikpunkt, der häufig in Bezug auf die *Cognitive Load Theory* geäußert wird, betrifft die Vernachlässigung motivationaler Aspekte beim Lernen. So sehen J. J. G. van Merriënboer und Sweller (2005) auch auf diesem Feld noch einigen Forschungsbedarf, selbst wenn es erste Ansätze gibt, diese Lücke zu füllen (vgl. Paas, Tuovinen & Darabi, 2005). Dennoch betonen Schnotz, Fries und Horz (2009), dass der Zusammenhang zwischen der erlebten Beanspruchung und der Motivation der Lernenden innerhalb der theoretischen Konzeption bislang zu wenig berücksichtigt wurde. Grundsätzlich kann nicht ausgeschlossen werden, dass motivierte Personen über höhere Ressourcen verfügen oder eine niedrigere Belastung empfinden. Eine Anstrengung wird nicht als solche wahrgenommen, wenn man die damit verbundenen Tätigkeit gerne ausübt.

Wie dem auch sei, zusammenfassend lässt sich festhalten, dass die *Cognitive Load Theory* trotz einiger konzeptioneller Schwächen einen überaus wichtigen Beitrag zur Lehr-Lern-Forschung leistet. Neben den Fortschritten bei der Entwicklung fortschrittlicher Lernumgebungen bietet sie einen theoretischen Rahmen, um neue Lernmedien und -angebote, die im Zuge der rasanten technischen Entwicklung entstanden sind, kritisch zu evaluieren. Eine weitere Stärke des Ansatzes besteht zudem darin, dass sie sich ständig weiterentwickelt und neue Gedanken integriert. Dazu zählt auch, dass Sweller und Sweller (2006) die vornehmlich kognitiv-psychologische Theorie in einen größeren Gesamtzusammenhang stellen und die menschlichen Informationsverarbeitung aus einer evolutionsbiologischen Perspektive betrachten.

3 Verarbeitung von Texten und Bildern

Es ist wahrscheinlich nicht übertrieben, wenn man behauptet, dass Menschen seit mehreren tausend Jahren versuchen, Wissen jenseits der gesprochenen Sprache zu kommunizieren. Von der Höhlenmalerei, über ägyptische Hieroglyphen bis hin zur Mathematik, wurden und werden dabei Formate verwendet, die auf den ersten Blick nicht unterschiedlicher sein könnten. Dennoch beruhen letztlich alle Arten der externen Wissensrepräsentation auf zwei basalen Formatprinzipien (Schnotz, 2005), was sich an einer exemplarischen Gegenüberstellung von Texten und Bildern veranschaulichen lässt. Während schriftliche Texte als typische Vertreter der *deskriptiven* Repräsentationsformate anzusehen sind, haben Bilder in der Regel einen eher *depiktionalen* Charakter.

Der entscheidende Unterschied zwischen *Deskription* und *Depiktion* liegt vor allem darin begründet, dass sie auf unterschiedlichen Zeichensystemen beruhen. So setzen sich Texte aus Symbolen zusammen, wohingegen Bilder vornehmlich aus sogenannten *Icons*¹ bestehen (Peirce, 1906; Schnotz, 2002). Die Symbole eines Textes haben dabei eine arbiträre Struktur, die das Ergebnis kulturspezifischer Entwicklungen ist und den damit verbundenen Konventionen unterliegt. Oder mit anderen Worten, die Beziehung zwischen einem Symbol und seiner Bedeutung ist grundsätzlich beliebig, wenn man die Konventionen außer Acht lässt. Besonders deutlich zeigt sich dieser arbiträre Charakter, wenn man Worte aus verschiedenen Sprachen vergleicht. Sowohl das polnische *pies*, als auch das englische *dog* haben im Deutschen die Bedeutung *Hund*. Beiden Worten haben nicht einen Buchstaben gemeinsam und repräsentieren dennoch das selbe Konzept.

Dagegen erschließt sich die Bedeutung eines *Icons* aus den strukturellen Übereinstimmungen mit dem Objekt, das es repräsentiert. *Ähnlichkeit*, wie sie in der Regel zwischen einem Gegenstand und dessen Photographie besteht, ist ein Ausdruck dieser Übereinstimmung. Darüber hinaus gibt es strukturelle Gemeinsamkeiten zwischen einem Objekt und seiner *depiktiven* Repräsentation, die über eine *äußerliche* Ähnlichkeit hinausgehen. Beispielsweise lässt sich der mathematische Ausdruck $x^2 + 2x + 4$ in einem Schaubild als Kurve darstellen, die mit der Formel äußerlich rein gar nichts gemeinsam hat. Dennoch sind beide Repräsentationen durch ihre Übereinstimmung auf abstrakter Ebene eindeutig miteinander verknüpft.

Dieser grundlegende Unterschied zwischen deskriptiven und depiktiven Formaten hat natürlich Konsequenzen, wenn es darum geht, bestimmte Informationen zum Zweck der Wissenskommunikation aufzubereiten. So sind Sprache und Schrift sehr mächtige Repräsentationsinstrumente (Schnotz, 2002), die neben Verallgemeinerungen auch abstrakte, also nicht wahrnehmbare Inhalte darstellen können. Die Vermittlung eines *relativ* einfa-

¹An dieser Stelle wird auf eine Übersetzung aus dem Englischen verzichtet, da der Begriff *Ikone* im Deutschen religiös behaftet ist

chen Begriffes wie Freiheit allein auf der Grundlage von Bildern ist dagegen um einiges umständlicher, wenn auf den Gebrauch konventioneller Symbole (z.B. zerbrochene Ketten) verzichtet werden soll. Allerdings ist es ein besonderes Merkmal graphischer Darstellungen, dass sie spezifische Inhalte in einer Gesamtheit und Detailfülle enthalten können, die kein Text zu beschreiben vermag. Form, Farben und die Orientierung im Raum sind Informationen, die sich selbst bei überichtlichen Bildern, schriftlich nur unzureichend erfassen lassen. Aus diesem Grund ist es in vielen Fällen angebracht, beide Repräsentationsformate zur Wissenvermittlung zu nutzen, da sich deren Vorzüge durchaus ergänzen können. Bevor jedoch von der integrierten Verarbeitung von Text- und Bildinformationen die Rede sein wird, geht es im folgenden Kapitel zunächst um das Verstehen von Texten. Da im Rahmen der empirischen Untersuchungen alle sprachliche Informationen schriftlich dargeboten wurden, beziehen sich die Ausführungen vor allem auf das *Leseverstehen*.

3.1 Textverstehen

Berücksichtigt man die Definition aus Kapitel 2.1, dann wird ein gelesener Text *verstanden*, wenn es gelingt, eine kohärente mentale Repräsentation der darin enthaltenen Informationen aufzubauen. Dieser Prozess ist allerdings in der Regel an eine Aufgabe geknüpft, zu deren Bewältigung man sich mit dem Text erst auseinandersetzt (Schnotz, 2002). Gerade aus dem Blickwinkel der pädagogischen Psychologie besteht diese Aufgabe häufig darin, sich Wissen anzueignen oder ein bestimmtes Problem zu lösen. Die Entnahme und Speicherung von Informationen aus einem Text ist jedoch ein komplexer Vorgang, was an den spezifischen Charakteristika des Repräsentationsformates liegt. Lesen ist zuerst einmal eine kulturspezifische Kompetenz, die im Kindesalter erworben werden muss, um mit vielen Anforderungen des Alltags umgehen zu können. Kenntnis der Sprache, Schrift und Wortbedeutungen sind dabei lediglich Minimalvoraussetzungen für den Aufbau einer kohärenten mentalen Repräsentation. Allgemein sind vorhandene Wissenstrukturen für das Lesen von grundlegender Bedeutung, da im Text enthaltene Informationen mit dem Vor-, Welt und Sprachwissen verbunden werden müssen (Artelt, Stanat, Schneider & Schiefele, 2001). Damit stellt sich die Verarbeitung von Texten als ein zielgerichteter, aktiver Prozess dar, bei denen der Leser und die Inhalte des Textes in Interaktion treten.

Der Grund, weshalb das Vorwissen bei der Verarbeitung schriftlicher Informationen ein so große Bedeutung hat, liegt in dem *unvollständigem* Charakter von Texten (T. Rasch, 2006). Schriftliche Darstellungen können immer nur eine Auswahl von Aspekten umfassen, die einen Gegenstand oder ein Thema kennzeichnen. Versuche, simple Gegenstände (zum Beispiel eine Kaffeetasse) *erschöpfend* zu beschreiben, sind entweder unmöglich oder enden als umfangreiches und schwer lesbares Dokument, dem in der Regel eine pragmatische Rechtfertigung fehlt. Texte bestehen somit aus einer begrenzten Anzahl von Informationen, die mit einem bestimmten Vermittlungsziel ausgesucht, strukturiert und angeordnet wurden. Diese reichen allerdings in vielen Fällen nicht aus, um eine kohärente, mentale Repräsentation aufzubauen, so dass zusätzlich Informationen ergänzt

werden müssen. Das kann auf der Grundlage bestehender Wissenstrukturen bzw. Schemata geschehen oder anhand der zuvor gelesenen Inhalte des Textes. Das bereits repräsentierte Wissen wird demnach genutzt, um *neue* Wissens Elemente abzuleiten, die einen Beitrag zur Kohärenz der mentalen Repräsentation leisten. Einen solchen Prozess, bei denen Informationen generiert werden, bezeichnen Rickheit und Sichelschmidt (1999) als *Inferenz*. Inferenzen müssen dabei keineswegs das Ergebnis angestrebter Überlegungen sein, wie es oft beim schlussfolgernden Denken im Zusammenhang mit logischen Problemstellungen der Fall ist. Beispielsweise gilt es schon als Schlussfolgerung, wenn man das Wort Wachhund liebt und annimmt, dass es sich bei dem betreffenden Tier um einen Dobermann handelt, obwohl der Text keine Informationen zu der Rasse des Hundes enthält (T. N. McNamara, Miller & Bransford, 1996).

Neben dem Vorwissen, hängt es natürlich auch von den Eigenschaften der schriftlichen Darstellung ab, ob diese verstanden wird oder nicht. Hier kommt es vor allem darauf an, wie die Informationen eines Textes von dem Autor aufbereitet und in Beziehung gesetzt wurden. Grundsätzlich fördert es die Verständlichkeit, wenn die Inhaltelemente *verknüpft*, logisch *konsistent* und aus der Perspektive des Lesers *relevant* sind. Da Sprache in der Regel linear verarbeitet werden muss, ist häufig diese Verknüpfung wichtiges Merkmal eines gut lesbaren Textes. Semantische und Grammatische Bezüge zwischen Sätzen erleichtern es dem Leser, die darin enthaltenen Informationen in Beziehung zu setzen, was letztlich den Aufbau einer kohärenten mentalen Repräsentation unterstützt. Gibt es viele bzw. ausreichend solcher Verbindungen spricht man in der Forschung davon, dass ein Text *kohärent* ist. Die *Kohärenz* eines Buchkapitels oder Lehrtextes kann allerdings individuell sehr unterschiedlich wahrgenommen werden. Geübte Leser oder Personen mit ausreichendem Vorwissen, werden einen Text als kohärenter einschätzen, als dies ungeübte Leser oder Laien tun würden. Damit ist die *Textkohärenz* kein verallgemeinerbares Qualitätskriterium, das über alle Individuen hinweg Gültigkeit hat.

Beispielsweise konnten D. S. McNamara, Kintsch, Songer und Kintsch (1996) feststellen, dass viele inhaltlich Verknüpfungen allein Lesern mit niedrigem Vorwissen nützen. Personen mit hohem Vorwissen profitieren dagegen eher von Texten mit geringer *Kohärenz*, da sie stärker dazu angeregt wurden, ihre bestehenden Wissensstrukturen für Inferenzprozesse zu nutzen. *Kohärenz* ist damit nicht wirklich ein Merkmal von Texten, sondern das Ergebnis eines kognitiven Verarbeitungsprozesses, den Schnotz und Dutke (2004) als Kohärenzbildung bezeichnen. Diese mentale Kombination und Verknüpfung von Informationen während des Lesens erfolgt dabei auf zwei hierarchisch nachgeordneten Ebenen. Die Erste ist *lokal*, hier werden die Bedeutungen von zwei unmittelbar aufeinanderfolgenden Sätzen mental repräsentiert. Die Zweite ist eine *globale* Ebene, auf der ein Leser die Inhalte größerer Textabschnitte und -passagen in einen kohärenten Zusammenhang bringen muss. Beim Textverstehen genügt es demnach nicht, aus einzelnen Wörtern Sätze zu bilden, deren Inhalte dann mental miteinander verbunden werden. Die wahre Herausforderung besteht darin, eine übergeordnete Bedeutung zu konstruieren, was aber nicht allen Menschen gleichermaßen leicht fällt (vgl. Hess, Foss & Caroll, 1995). Zwar unterstützen semantische und grammatikalische Verknüpfung ein-

3 Verarbeitung von Texten und Bildern

zelter Sätze die *Kohärenzbildung* auf lokaler Ebene, doch bedarf es mehr, um einen Text tatsächlich zu verstehen. So muss der Leser die logische Konsistenz und Relevanz der Informationen beurteilen können, wozu in der Regel aufgaben- bzw. themenspezifisches Vorwissen benötigt wird. Die Kohärenzbildung auf globaler Ebene hängt damit zu einem großen Teil von bereits vorhandenen kognitiven Schemata ab.

Kohärenzbildung ist letztlich das zentrale Konzept der meisten einflussreichen Theorien zum Textverstehen, auch wenn sich Art und Zahl der involvierten Prozesse oder Operationen unterscheiden. Ebenso verbreitet ist die Annahme, dass eine Verarbeitung von schriftlichen Informationen zu einer mentalen Repräsentationen führt, die mehrere Ebenen umfassen (Graesser, Hoffman & Clark, 1980). Dennoch lassen sich viele Arbeiten finden, die sich lediglich auf eine dieser Ebenen konzentrieren, während die anderen unberücksichtigt bleiben (Perrig & Kintsch, 1985). Im folgenden werden theoretische Ansätze erläutert, die das Textverstehen umfassend, als Aufbau multipler Repräsentationen begreifen. Da die einzelnen Arbeiten aufeinander aufbauen, bietet es sich an, diese in chronologischer Reihenfolge zu diskutieren. Die Darstellung beginnt mit einem Ansatz von Kintsch und van Dijk (1978), die bereits Ende der siebziger Jahre eine grundlegende Theorie zum Textverstehen entworfen haben. Obwohl das Modell im Laufe der Zeit mehrfach überarbeitet und durch konzeptionelle Elemente ergänzt wurde, dienen zentrale Annahmen weiterhin als Basis für aktuelle theoretische Entwicklungen.

3.1.1 Textbasis, zyklische Verarbeitung und semantische Strukturen

Eines der ersten umfassenderen Modelle zum Verstehen und zur Produktion von Texten stammt von Kintsch und van Dijk (1978). Die beiden Autoren verfolgten das Ziel, möglichst ausführlich das System mentaler Operationen zu beschreiben, welches der Verarbeitung sprachlicher Informationen zugrunde liegt. Die Theorie bewegt sich dabei ausschließlich auf einer semantischen Ebene, so dass Dekodierungsprozesse und syntaktische Analysen ausgeklammert wurden. Obwohl sich Kintsch und van Dijk (1978) durchaus bewusst sind, dass Vorwissen bei der Bedeutungskonstruktion eine entscheidende Rolle spielt, macht das Modell zudem keine präzisen Angaben darüber, wie bestehende Wissensstrukturen genutzt werden oder wie Inferenzprozesse ablaufen.

Die Autoren konzentrierten sich auf die semantische Verarbeitung von sprachlichen Informationen, die nach ihrer Auffassung mit dem mentalen Aufbau einer sogenannten *Textbasis* beginnt. Diese besteht aus inhaltlich verschachtelten Propositionenlisten, welche die Bedeutung des gelesenen Textes strukturiert repräsentieren. Die Textbasis stellt die Grundlage aller weiteren Verarbeitungsschritte dar und muss daher schon in ihrer einfachsten Form kohärent sein, damit ihre Inhalte auch auf abstrakteren Verständnisebenen genutzt werden können. Aus diesem Grund wird beim Aufbau einer Textbasis ständig geprüft, ob die Propositionen tatsächlich eine kohärente Struktur bilden. Wenn (semantische) Lücken auftreten, die sich nicht durch Informationen aus dem Text schließen lassen, müssen über Inferenzprozesse Propositionen *generiert* werden, die eine kohärente Repräsentation ermöglichen. Die Prüfung kann jedoch nicht anhand der vollständige Text-

basis in ihrer Gesamtheit erfolgen, da die Kapazitäten des Arbeitsgedächtnisses dazu nicht ausreichen. Stattdessen wird Sprache sequentiell verarbeitet, so dass die einzelnen Informationen erst nach und nach in das kognitive System gelangen. Texte, die in lateinischer Schrift verfasst wurden, liebt man gewöhnlich Zeile für Zeile, von links nach rechts und von oben nach unten. Diese Ordnung spiegelt sich auch in den Propositionslisten der Textbasis wieder, deren Aufbau damit das Ergebnis *zyklischer* Verarbeitungsprozesse ist.

Kintsch und van Dijk (1978) nehmen an, dass diese Verarbeitung auf der Grundlage von *chunks* (vgl. Miller, 1956) erfolgt, die jeweils eine begrenzte Anzahl von Propositionen umfassen. Wie viele Propositionen tatsächlich enthalten sind, wird unter anderem von den Satz- und Phrasengrenzen beeinflusst, was Aaronson und Scarborough (1977) belegen konnten. Neben den Textcharakteristika determinieren das Vorwissen und die kognitive Fähigkeiten den Umfang der *chunks*, die während des Lesens gebildet werden. Die zyklische Verarbeitung der Textbasis erfordert es, einzelne Propositionen semantisch miteinander zu verknüpfen, damit eine kohärente Gesamtrepräsentation der Textinhalte entstehen kann. In diesem Zusammenhang gehen die Autoren davon aus, dass ein Teil des Arbeitsgedächtnisses beim Lesen als *short-term memory buffers* genutzt wird. Bei der Verarbeitung eines *chunks*, werden hier die Propositionen *zwischen gespeichert*, die für eine Verknüpfung mit nachfolgenden Informationen zur Verfügung stehen sollen. Da die Kapazität des *short-term memory buffers* beschränkt ist, kann bei umfangreichen Inhalten immer nur eine begrenzte Auswahl bereit gehalten werden. Gelingt es dem Leser, eine semantische und/oder grammatische Übereinstimmung zwischen den Propositionen im Zwischenspeicher und denen eines aktuell verarbeiteten *chunks* herzustellen, dann entsteht lokale Kohärenz. Scheitert hingegen der Versuch, müssen ressourcenintensive Suchprozesse in Gang gesetzt werden, um alternative Verknüpfungsmöglichkeiten zu finden. Diese Suche kann bereits im Langzeitgedächtnis abgelegte Propositionen betreffen, oder es werden einzelne Textabschnitte wiederholt gelesen. Die fehlende Übereinstimmung kann auch durch Inferenz überwunden werden, so dass eigens generierte Propositionen eine kohärente Verbindung aufeinander folgender Inhalte ermöglicht.

Die Autoren gehen davon aus, dass alle Propositionen, die während eines Zyklus verarbeitet werden, mit einer bestimmten Wahrscheinlichkeit in Langzeitgedächtnis gelangen. Ob diese sich von dort abrufen lassen, wenn es beispielsweise darum geht, die Inhalte eines zuvor gelesenen Textes wiederzugeben, hängt von verschiedenen Faktoren ab. Neben der Art und Umfang der Wiedergabe, ist es vor allem die Verarbeitungshäufigkeit, die eine Speicherung wahrscheinlicher macht. Demnach kann ein und dieselbe Proposition das Arbeitsgedächtnis in mehreren Zyklen durchlaufen. Dies ist dann der Fall, wenn in einem Lehrtext besonders wichtige Inhalte wiederholt oder am Ende eines Kapitels zusammengefasst werden. Zudem ist es möglich, dass Propositionen mit zentraler Bedeutung mehrere Male in den *short-term memory buffer* geladen werden, etwa um Kohärenz auf einer globaleren Ebene zu erzeugen. Welche Propositionen Gegenstand einer mehrfachen Verarbeitung werden, hängt dabei von dem Vorgehen des jeweiligen Lesers ab, wobei Kintsch und van Dijk (1978) anmerken, dass viele der Prozesse automatisch und mehr oder weniger gleichzeitig ablaufen. Erst wenn es nicht gelingt, einzelne Inhalte

3 Verarbeitung von Texten und Bildern

kohärent miteinander zu verbinden, setzen bewusstseinspflichtigen Such- und Inferenzprozesse ein, die sich willentlich gesteuert werden müssen.

Der Aufbau einer kohärenten Textbasis ist nach Ansicht von Kintsch und van Dijk (1978) aber lediglich ein erster Schritt beim Verstehen sprachlicher Medien. Durch die Konstruktion eines solchen Propositionsgefüges auf lokaler Ebene entsteht eine semantische *Mikrostruktur*, welche vor allem den wörtlichen Inhalt repräsentiert. Texte besitzen in der Regel aber auch eine semantische *Makrostruktur*, die beispielsweise das Thema oder die Quintessenz eines Buchkapitels wiedergibt. Grundlage dieser *Makrostruktur* ist eine kohärente Textbasis, deren detaillierte Informationen durch die Anwendung spezifischer Operatoren bzw. Regeln reduziert werden. Eine solche Reduktion auf die wesentliche Bedeutung eines Textes wird erreicht, indem überflüssige Propositionen *gelöscht*, *generalisiert* oder dazu genutzt werden, um neue Propositionen über Inferenzprozesse zu konstruieren. Daraus folgt allerdings nicht, dass die betroffenen *Mikropropositionen* tatsächlich aus dem Gedächtnis entfernt werden müssen. Sie sind nach Auffassung der Autoren nur kein Element der semantischen *Makrostruktur* des Textes, können aber sonst ganz normal gespeichert und erinnert werden. Dies ist beispielsweise dann der Fall, wenn man beim Lesen einer Zeitung einzelne Formulierungen oder Fakten wiedergeben kann, die für die eigentliche Kernaussage des Artikels ohne Bedeutung sind. Gleichzeitig sollte ein geübter Leser aber in der Lage sein, die Inhalte eines Artikels in einem globaleren Zusammenhang zu sehen und kritisch zu reflektieren. Fehlt es an Übung oder dem notwendigen Vorwissen, gelingt dies vielen Menschen jedoch nicht und sie bleiben auf einer lokalen Verständnisebene hängen (Hess, Foss & Carroll, 1995). Die *erfolgreiche* mentale Repräsentation von Textinformationen umfasst den Aufbau von Mikro- und Makropropositionen, wobei diese allerdings verschiedenen Abstraktionsstufen angehören.

Kintsch und van Dijk (1978) gehen davon aus, dass die Transformation der Textbasis in eine *Makrostruktur* durch kognitive Schemata kontrolliert wird. Diese Schemata enthalten neben dem Sprach- und Weltwissen, das auch der lokalen Kohärenzbildung dient, die persönlichen Ziele des Lesers. Mit einer solchen Kontrolle soll letztlich gewährleistet werden, dass eine Anwendung der Operatoren rechtzeitig endet und die Reduktionsprozesse nicht zu Makrostrukturen führen, die so abstrakt sind, dass sie keinen Sinn mehr ergeben. Obwohl diese Annahmen darauf hinweisen, dass den beiden Autoren sehr wohl die Bedeutung vorhandenen Wissens bewusst ist, konzentriert sich der ursprüngliche Ansatz von Kintsch und van Dijk (1978) noch sehr auf die Repräsentation des Textes an sich. So macht das Modell der zyklischen Verarbeitung keine Aussagen darüber, welchen Anteil die Erfahrungen und Gedächtnisinhalte des Lesers beim Aufbau der Mikro- und Makrostrukturen haben. Die Repräsentation der Situation oder Geschehnisse, die ein Text beschreibt, spielen für das tatsächliche Verständnis des Textes nur eine untergeordnete und diffuse Rolle. Wie sich diese theoretische Perspektive innerhalb von nur wenigen Jahren wandelte, wird in dem folgenden Abschnitt erläutert, welcher der Konstruktion *mentaler* Modelle bei der Verarbeitung sprachlicher Informationen gewidmet ist.

3.1.2 Modellkonstruktion beim Textverstehen

Bereits sechs Jahre vor dem Modell der zyklischen Verarbeitung (Kintsch & van Dijk, 1978), konnten Bransford, Barclay und Franks (1972) die Bedeutung *semantischer Situationen* für das Erinnern sprachlicher Informationen belegen. Bei ihrem Experiment, das die drei Forscher mit gesprochenem Stimulusmaterial durchführten, wurde Versuchspersonen zunächst 14 einzelne Sätze präsentiert. Nach einer kurzen Pause mussten sich die Probanden weitere Sätze anhören und beurteilen, ob diese mit denen identisch waren, die sie kurz davor gehört hatten. Bransford et al. (1972) manipulierten dabei entweder die Oberflächenstruktur oder den semantischen Gehalt der zweiten Stimulusreihe, wobei die Unterschiede immer nur einzelne Worte betrafen. So wurde im Fall der Oberflächenstruktur zwar ein Wort durch ein anderes ersetzt, doch blieb die ursprüngliche Bedeutung des Satzes erhalten. Bei der Manipulation des semantischen Gehaltes hingegen veränderte das ausgetauschte Wort den Sinn des Satzes, so dass eine abweichende Situation entstand. Wie sich zeigte, gelang es den Versuchspersonen sehr viel besser, die Sätze korrekt zu klassifizieren, wenn der Unterschied dem semantischen Gehalt betraf. Die Autoren erklärten ihren Befund mit der Annahme, dass Menschen bei der Verarbeitung sprachlicher Informationen *semantische Beschreibungen der Situation* konstruieren. Sätze, die zu abweichenden *Situationskonstrukten* führen, konnten aus diesem Grund sehr viel leichter identifiziert werden. Betrifft der Unterschied aber lediglich die Oberflächenstruktur, so führt die Verarbeitung zu identischen mentalen Situationsbeschreibungen und die Versuchsperson kann diese nicht mehr auseinander halten.

Vor dem Hintergrund dieses Befundes griffen van Dijk und Kintsch (1983) die Vorstellung eines mental konstruierten Modells auf und verbanden sie mit der bereits bestehenden theoretischen Konzeption der zyklischen Verarbeitung. Die beiden Autoren gelangten zu der Einsicht, dass es beim Textverstehen weniger darum geht, den Text selbst, als vielmehr seine Inhalte aufnehmen und wiedergeben zu können. So spielt die Konstruktion kohärenter Mikro- und Makrostrukturen auch weiterhin eine überaus wichtige Rolle, doch wurde die ursprüngliche Konzeption um eine weitere Repräsentationsebene ergänzt. Nach van Dijk und Kintsch (1983) wird beim Verstehen sprachlicher Informationen neben der bereits erwähnten Textbasis ein mentales Modell der beschriebenen Situation konstruiert. Die Grundlage des sogenannten *Situationsmodells* sind bestehende Wissenstrukturen, die zur Konstruktion herangezogen werden können, wenn es beispielsweise einem Leser gelingt, diese mit den Inhalten des Textes zu verknüpfen. Das Individuum wird quasi durch die sprachliche Informationen an vergangene Situationen erinnert, die es so ähnlich möglicherweise auch selbst erlebt hat. Dementsprechend handeln viele Texte von Personen, Objekten, Situation, Orten, Vorgängen usw., die dem Leser durch vergleichbare Erfahrungen bereits vertraut sind. van Dijk und Kintsch (1983) nehmen an, dass es diese Erfahrungen sind, ganz gleich ob sie tatsächlich erlebt oder nur durch Lesen indirekt erworben wurden, die im Gedächtnis *cluster* bilden. Die sich teilweise überlappenden *cluster* strukturieren sich nach der Art und Ähnlichkeit der einzelnen Erfahrungen, wobei es durch Prozesse der Dekontextualisierung auch zu Verallgemeine-

rungen kommen kann.

Nach Ansicht der beiden Autoren greift ein Individuum auf diese *Erfahrungscluster* zurück, wenn es sprachliche Informationen verarbeitet. Die *cluster* bilden dann die Bausteine des Situationsmodells, das parallel zum Aufbau der Textbasis konstruiert wird. Enthält ein Text Inhalte, die dem Leser bereits bekannt sind, reicht in der Regel ein Modell aus. Ist das sprachliche Material dagegen etwas anspruchsvoller, kann es notwendig sein, mehrere nur teilweise relevante Situationsmodelle zu konstruieren, die dann als Bausteine für die eigentliche Modellkonstruktion dienen. Das *Situationsmodell* bzw. das Wissen und die Erfahrung, aus denen es sich zusammensetzt, ist damit die Grundlage des Textverstehens. Es enthält alle Informationen, die in einem Text nur implizit enthalten sind oder auf andere Weise vorausgesetzt werden. Dabei spielt es keine Rolle, ob der Konstruktion des *Situationsmodells* allgemeines Wissen oder spezifische Erfahrungen zugrunde liegen. Beide Quellen können für den Aufbau einer entsprechenden mentalen Repräsentation genutzt werden, auf deren Grundlage sich dann *neue* Inhalte verarbeiten und aufnehmen lassen.

Ein vergleichbarer theoretischer Ansatz, der sich auch auf das Verstehen sprachlicher Informationen übertragen lässt, stammt von Johnson-Laird (1983, 1980). Auch er nimmt an, dass die kohärente mentale Repräsentation der Inhalte eines Textes untrennbar mit der Konstruktion eines *mentalen Modells* verbunden ist. Allerdings gibt es einen grundlegenden Unterschied zwischen dem *Situationsmodell* von van Dijk und Kintsch (1983) und den mentalen Modellen, wie sie sich Johnson-Laird (1983) vorstellt. Bereits in Kapitel 2.1 war erläutert worden, dass *mentale Modelle* Formate sind, die eine analoge Beziehung zwischen der Struktur eines Objekts und seiner Repräsentation herstellen. Dagegen setzen sich die Situationsmodelle in ihrer ursprünglichsten Fassung allein aus Propositionen zusammen, denen ein symbolisches Repräsentationsprinzip zugrunde liegt. Zwar geht auch Johnson-Laird (1983) davon aus, dass sprachliche Informationen zunächst propositional repräsentiert werden, doch ist dies nur ein textnaher Zwischenschritt. Das erstrebenswerte Endprodukt aller Verarbeitungsprozesse bleibt ein, falls erforderlich, multidimensionales mentales Modell das temporäre, räumliche und kausale Zusammenhänge auf analoge Weise abbilden kann.

Dieser anfängliche Unterschied zwischen den beiden Ansätzen verwischte sich jedoch mit der Zeit zunehmend, da die Annahmen hinsichtlich des Formates der Situationsmodelle *glockert* wurde. Ein Grund für diese Entwicklung ist vor allem in experimentellen Befunden zu sehen, die gegen eine reine propositionale Repräsentation von Textinhalten sprechen. Beispielsweise gaben Perrig und Kintsch (1985) Versuchspersonen Texte zu lesen, die Informationen zur räumlichen Anordnung von Objekten enthielten. Die Ergebnisse der Studie legten nahe, dass einige Probanden Repräsentationen des sprachlichen Inhaltes konstruiert hatten, die mentalen Landkarten entsprachen und damit ein analoges Format hatten. Kintsch (1988) greift diesen Befund auf und hält fest, dass Situationsmodelle sich nicht zwangsläufig aus Propositionen zusammensetzen müssen. Vielmehr können sie unter bestimmten Umständen die gleichen Eigenschaften besitzen, wie die mentalen Modelle eines Johnson-Laird (1983). Darüber hinaus führt Kintsch (1998) eini-

ge Varianten ein, die er, je nach Gegenstand, als *räumliche* Situationsmodelle, *images*, Problemmodelle, die mathematisch definiert sind, oder als Systemmodelle bezeichnet. Diese Vielzahl möglicher Modellvarianten trägt nicht unbedingt zur begrifflichen Vereinheitlichung bei, so dass es wenig überrascht, wenn die Bezeichnungen nahezu synonym verwendet werden (vgl. Zwaan & Radvansky, 1998; Glenberg, Meyer & Lindem, 1987).

Einzelne Forscher sehen diese terminologische Unschärfe zwar kritisch (T. Rasch, 2006), doch lassen sich darin auch positive Aspekte finden. So kann allein die Tatsache, dass van Dijk und Kintsch (1983) und Johnson-Laird (1983) mehr oder weniger zeitgleich zu ähnlichen Schlüssen kamen, als ein Hinweis auf die Gültigkeit ihrer grundlegenden Annahmen interpretiert werden. In diesem Zusammenhang verwundert es nicht, dass Mehrebenentheorien den wissenschaftlichen Diskurs weitgehend dominieren. Die meisten Ansätze unterscheiden die Repräsentationsebenen der Textoberfläche, eine propositionale Inhaltstruktur und ein (mentales) Situationsmodell, das Vor- und Kontextwissen in den Verstehensprozess integriert (Dutke, 2003). Obwohl sich mit Sicherheit ganz trefflich darüber streiten lässt, welches Format der *höchsten* Repräsentationsebene nun tatsächlich zugrundeliegt (vgl. Waskan, 2003), ist es möglich, diese Frage auch aus einer eher pragmatischen Perspektive zu betrachten. Repräsentation werden konstruiert, um sich mit den Anforderungen der Umwelt auseinanderzusetzen, so dass man sie auch als zweckgebundene mentale Quasi-Objekte bezeichnen könnte. Die Anforderungen der Umwelt variieren jedoch sehr stark, was umso mehr gilt, wenn man sich die Fülle von Aufgaben vor Augen führt, die ein Individuum tagtäglich bewältigen muss. Es scheint daher durchaus praktikabel zu sein, wenn Menschen bei der mentale Modellkonstruktion auf unterschiedliche Formate zurückgreifen können. Das Modell und damit auch das zugrundeliegende Format muss vor allem auf die strukturellen Anforderungen einer entsprechenden Aufgabe passen, so dass die Repräsentation zielführend verwendet werden kann. Es verwundert daher wenig, wenn sowohl Johnson-Laird (1983) als auch Kintsch (1988) versuchen, den teils sehr unterschiedlichen Kontexten der Modellkonstruktion mit mehr oder weniger genau spezifizierten Modell-Typologien zu begegnen.

Darüber hinaus hat Kintsch (1988) seine Überlegungen nochmals weiterentwickelt und mit dem *Construction Integration Model* einen Ansatz vorgelegt, der sich speziell dem Verhältnis von *top-down* und *bottom-up* Prozessen bei der Sprachverarbeitung widmet. Jedoch sind seine detaillierten Überlegungen in Hinblick auf die Zielsetzung der vorliegenden Arbeit zu feinkörnig, weshalb es nur wenig Sinn macht, das Modell an dieser Stelle ausführlich zu diskutieren. Ganz ähnlich verhält es sich mit eine Reihe neuerer Theorien zum Textverstehen, die im Rahmen der *Embodied Cognition* entstanden sind (vgl. Zwaan, 2004; Glenberg & Kaschak, 2002). Die Annahme, dass ein Individuum bei der Verarbeitung von sprachlichen Informationen weitaus stärker involviert ist, als es der Aufbau mentaler Repräsentationen nahe legt, ist zwar überaus interessant, aber führt letztlich ebenfalls zu weit. Damit endet hier das Kapitel zum Textverstehen. Die nun folgenden Abschnitte thematisieren theoretische Ansätze, die sich speziell mit der kognitiven Verarbeitung von bildlichen Informationen beschäftigen.

3.2 Bildverstehen

Das Sprichwort, *ein Bild ist mehr wert als tausend Worte* ist eine vermeintlich chinesische Weisheit, von der allerdings die Chinesen noch nie etwas gehört haben sollen (Larkin & Simon, 1987). Damit liegt der Verdacht nahe, dass es sich bei der unterstellten Überlegenheit von Bildern um so etwas wie *Folklore* handelt, zumal nicht spezifiziert ist, worin der *Mehrwert* eines Bildes bemessen wird. Aus der Perspektive der pädagogischen Psychologie würde eine bildliche Darstellung (oder auch ein Text) dann wertvoll sein, wenn das Format den Lernenden beim Erwerb neuen Wissens unterstützen würde. Obwohl sich vor diesem Hintergrund untersuchen ließe, ob nun Texte oder Bilder lernförderlicher sind, findet man nur vereinzelt Studien, bei denen der *Wert* bildlicher und sprachlicher Lernmaterialien miteinander verglichen wurde (z.B. Dixon & Saltz, 1977). Dieser Umstand verwundert nicht wirklich, wenn man bedenkt, dass bei einer Gegenüberstellung von deiktionalen und deskriptionalen Informationsträgern letztlich Äpfel mit Birnen verglichen werden. Die zugrunde liegenden Zeichensysteme sind schlichtweg zu verschieden (siehe Seite 43), so dass in der Regel nur eine komplementäre Verwendung von Texten und Bildern didaktisch sinnvoll ist. So existieren mittlerweile zahlreiche Untersuchungen zu den Auswirkungen, die Illustrationen auf den Wissenserwerb mit Texten haben (vgl. J. Levin, Anglin & Carney, 1987; Carney & Levin, 2002), doch wurden Verstehensprozesse, die sich ausschließlich auf die kognitive Verarbeitung von graphischen Darstellungen richteten, lange Zeit vernachlässigt (Weidenmann, 1988).

Während die kognitive Verarbeitung von sprachlicher Informationen bereits eingehend untersucht wurde und die Entwicklung umfassender theoretischer Modelle angeregt hat (siehe Kapitel 3.1), steht die Erforschung des Verstehens von Bildern oder Grafiken vergleichsweise noch am Anfang (Schnotz & Dutke, 2004). Das ist umso erstaunlicher, wenn man bedenkt, dass sich in den letzten Jahrzehnten der Zeichengebrauch der Menschen stark verändert hat, was Mitchell (1995) auch als *pictorial turn* umschreibt. Diese *visuelle Wende* besteht nicht nur darin, dass die Anzahl von Bildern in den täglich konsumierten Medien stark zugenommen hat, was sicherlich auch auf die Digitalisierung und rasante Entwicklungen der Computer- und Kommunikationstechnologien zurückzuführen ist. Der *pictorial turn* geht zusätzlich mit einer Aufwertung grafischer Darstellungen gegenüber der Sprache einher (Ballstaedt, 2005). Bilder sind damit nicht länger nur ein Ersatzcode für kleine Kinder oder Menschen, die (noch) nicht lesen können. Grafische Darstellungen dienen durchaus der Vermittlung komplexer Wissensinhalte (z.B. Schaltpläne) oder werden dazu genutzt, neue Erkenntnisse zu generieren oder abzuleiten. Um Bilder aber auf diese Weise verwenden zu können, muss ein Individuum in der Lage sein, diese zu verstehen.

Dem Verstehen von Bildern lässt sich dabei die gleiche Definition (siehe Seite 10) zugrunde legen, die im Zusammenhang mit der Verarbeitung von Texten diskutiert wurde (siehe Kapitel 3.1). So wird ein Bild dann *verstanden*, wenn es dem Rezipienten gelingt, eine kohärente mentale Repräsentation der darin enthaltenen Informationen aufzubauen. Wie beim Lesen von Texten, ist dieser Prozess in der Regel mit einer Aufgabe oder

Absicht verknüpft, die den Aufbau der Repräsentation zweckorientiert steuert (Schnotz, 2002). Denkt man an die Bilderflut, die im Alltag auf die Menschen einwirkt, so fällt es zunächst schwer, die intentionale Verarbeitung von grafischer Darstellungen nachzuvollziehen. Sehr häufig werden Bilder dazu eingesetzt, Emotionen anzusprechen, selbst wenn diese von Seite des Rezipienten nicht wirklich gewollt sind. Beispielsweise sollen illustrierte Werbeplakate zum Kauf von Produkten anregen, Abbildungen von Raucherbeinen den Konsum von Tabakwaren unterbinden und einige Menschen stellen Familienfotos auf ihre Büroschreibtische, um sich heimischer und damit wohler zu fühlen. In der Pädagogischen Psychologie werden natürlich auch die emotionalen Konsequenzen des Einsatzes von Bildern untersucht (vgl. Lenzner, 2009), doch können grafische Darstellungen ebenso kognitive Funktionen erfüllen. In diesem Zusammenhang dienen Bilder, ganz ähnlich wie Texte, primär der Vermittlung und dem Aufbau von Wissen. Aber selbst wenn man sich auf den Wissenserwerb konzentriert und affektiv-motivationale Aspekte unberücksichtigt lässt, werden im Rahmen von Lehr- und Lernprozessen sehr unterschiedliche Bilder verwendet, die meist recht spezifische Zwecke erfüllen. Aus diesem Grund wird in dem nun folgenden Abschnitt der Versuch unternommen, Bildarten und ihre Funktionen zu systematisieren.

3.2.1 Arten und Funktionen von Bildern

Eine Ursache, weshalb es bislang nur wenige theoretische Arbeiten zu reinen bildspezifischen Verstehensprozessen gibt, mag darin begründet liegen, dass Bilder sehr unterschiedliche Formen annehmen können. Vergleicht man beispielsweise ein Linien-Diagramm mit dem bereits erwähnten Familienfoto, dann besteht die einzige offensichtliche Gemeinsamkeit darin, dass es sich offensichtlich nicht um (reinen) Text handelt. Eine umfassende Theorie zum *kognitiven* Bildverstehen müsste demnach die Verarbeitung von sehr unterschiedlichen grafischen Darstellungen integrieren können, was jedoch ein durchaus ehrgeiziges Unterfangen darstellt. Erschwerend kommt hinzu, dass es keine allgemein akzeptierte Terminologie der verschiedenen Bildarten gibt und viele Autoren ihre eigenen oder abweichende Begrifflichkeiten verwenden. Die auf den ersten Blick einfachste Klassifikation lässt sich vornehmen, indem man *informierende* Bilder von künstlerischen oder unterhaltenden Bildern abgrenzt. Nach einer Definition von Weidenmann (1994b) handelt es sich bei informierenden Bildern um Illustrationen, die primär erstellt wurden, um Wissen zu vermitteln, wohingegen ästhetische und emotionale Aspekte der Rezeption in den Hintergrund treten. Allerdings lassen sich auch künstlerische oder unterhaltende Bilder als historische Quellen verwenden, da sie Informationen über eine bestimmte Zeit, Epoche oder die Absichten hinter dem Werk enthalten. Zwar betont (Weidenmann, 1994b), dass eine solche Rezeptionsweise eher untypisch ist und den Intentionen des Künstlers in der Regel widerspricht, doch zeigt der Einwand, wie schwer eine klare Abgrenzung bereits in diesem einfachen Fall ist.

Dessen ungeachtet stammt einer der ersten Versuche, *informierende* Bilder systematisch zu klassifizieren, von Knowlton (1966), der *realistische Bilder*, *logische Bilder* und

3 Verarbeitung von Texten und Bildern

Analogiebilder unterscheidet. Diese Dreiteilung haben jedoch nicht alle Autoren, die sich ihrerseits an einer terminologischen Abgrenzung versucht haben, übernommen. Dennoch finden sich bei den meisten Wissenschaftlern die konzeptionelle Unterscheidung zwischen realistisch-anschaulichen und logisch-analytischen Bildern, wobei auch diese Bezeichnung nicht unproblematisch ist (vgl. Schnotz, 1994b). Von beiden Bildtypen ist das realistische Bild, das vereinzelt auch als darstellendes Bild (Peeck, 1994) oder Abbild (Ballstaedt, 1997) bezeichnet wird, sicherlich am häufigsten in Schul- und Lehrbüchern anzutreffen. Beispiele realistischer Bilder sind Zeichnungen, naturalistische Gemälde, Fotos, Cartoons, Piktogramme oder auch Landkarten. Sie werden vorwiegend dazu verwendet, konkrete Sachverhalte zu veranschaulichen, so dass realistische Bilder in den Regel den Dingen, die sie repräsentieren, möglichst stark ähneln. Schnotz (2006) weist in diesem Zusammenhang darauf hin, dass es sich auch dann um ein realistische Darstellung handeln kann, wenn der abgebildete Sachverhalt allein der Vorstellungskraft des Zeichners oder Künstlers entspringt und in der realen Welt gar nicht existiert. Bilder von vermeintlichen Außerirdischen sind zwar Produkte der Fantasie, doch besitzen auch sie einen anschaulichen Charakter, wenn ein Betrachter Kopf, Gliedmaßen usw. zuordnen kann.

Die Verwendung *logischer Bilder* hat eine vergleichsweise junge Geschichte, ist aber aus dem Kontext der Wissensvermittlung nicht mehr wegzudenken (Schnotz, 2006). Im Gegensatz zu realistischer Darstellungen, weisen logische Bilder keine offensichtliche Ähnlichkeit mit dem Sachverhalt auf, den sie repräsentieren. Aus diesem Grund eignen sich logische Bilder besonders gut zur Veranschaulichung abstrakter Sachverhalte, bei denen Zusammenhänge oder andere strukturelle Eigenschaften sich in den meisten Fällen einer direkten Wahrnehmung entziehen. Typische Vertreter dieser Bildart sind Struktur-, Kreis-, Säulen-, Linien- und Streudiagramme, die zur Visualisierung qualitativer oder quantitativer Beziehungen eingesetzt werden. Eine etwas andere Typologisierung nehmen Hegarty, Carpenter und Just (1996) vor, die insbesondere logische Bilder etwas differenzierter betrachten. Während in der Terminologie von Hegarty et al. (1996) die *iconic diagrams* weitgehend den realistischen Bilder entsprechen, gehen die die Autoren von zwei zusätzlichen *logischen* Bildtypen aus, die von der Klassifikation, wie sie etwa (Knowlton, 1966) vorgenommen hat, abweichen. So werden grafische Darstellungen wie Schaltpläne, Fluss- oder Baumdiagrammen als *schematic diagrams* bezeichnet. Diese bestehen aus abstrakten Komponenten oder Entitäten, die als Ganzes eine Struktur bilden, die funktionale Beziehungen repräsentiert. Da in der Regel eine physische Ähnlichkeit zwischen dem Konzept und dem Sachverhalt, für den es steht, nicht zu erkennen ist, unterliegt die Interpretation solcher schematischen Darstellungen bestimmten Konventionen, die erlernt werden müssen. Nach Hegarty et al. (1996) werden *schematic diagrams* vor allem dazu verwendet, um strukturelle oder qualitative Beziehungen abstrakter Konzepte zu veranschaulichen. Quantitative Zusammenhänge lassen sich hingegen durch die zweite Kategorie logischer Bilder, die *graphs & charts* darstellen. Bei diesem Typus werden bestimmte Sachverhalte der realen Welt auf die quantifizierbaren Attribute eines Bildes übertragen, etwa durch die Länge von Balken oder die Position in einem Koordina-

tensystem. Typische Vertreter der *graphs & charts*-Kategorie sind Linien-, Balken- oder Streudiagramme.

Eine weitere Bildart, die von Hegarty et al. (1996) und anderen Forschern jedoch nicht als eigenständige Klasse angeführt wird, sind die bereits erwähnten *Analogiebilder* (Knowlton, 1966; Schnotz, 2006). Der Grund, weshalb einige Autoren bildliche Analogien den realistischen Bildern zuordnen, liegt womöglich in der oberflächlichen Ähnlichkeit dieser beiden Darstellungstypen begründet. So thematisieren auch Analogiebilder anschaulichen Gegenstände, die Teil der menschlichen Gedanken- und Umwelt sind. Das besondere an bildlichen Analogien ist jedoch, dass sie eine Bedeutung besitzen, die über den dargestellten Sachverhalt hinausgeht. Demnach soll ein Betrachter die Abbildung nicht dazu nutzen, um mehr über den Gegenstand zu erfahren, der darauf eigentlich zu erkennen ist. Vielmehr ist der Rezipient eines Analogiebildes dazu angehalten, die Bedeutungszusammenhänge, die er aus der Darstellung entnehmen kann, auf einen intendierten Sachverhalt zu übertragen, der so direkt nicht beobachtbar ist, aber vergleichbare Eigenschaften besitzt. Auf diese Weise können neue Themengebiete gedanklich erschlossen werden, indem der Lernende auf vorhandenes Wissen zurückgreift und strukturelle oder funktionale Beziehungen zwischen den Elementen eines Systems auf die Komponenten eines anderen System überträgt. Eine in didaktischen Kontexten häufig verwendete Analogie ist die Veranschaulichung eines Stromkreises anhand eines Wasserkreislaufs (Issing, 1994). So kann der fließende Strom zwar mithilfe entsprechender Instrumente gemessen werden, doch entzieht sich der Ladungstransport einer direkten sinnlichen Wahrnehmung, so dass es vielen Schülerinnen und Schülern schwer fällt, sich die ablaufenden Prozesse vorzustellen. Durch den Hinweis auf die strukturell-funktionale Analogie mit dem sinnlich wahrnehmbaren Wasserkreislauf gewinnt der fließende Strom so etwas wie eine sinnliche Qualität und die betreffenden Vorgänge lassen sich besser einordnen und damit auch verstehen. Der Gebrauch von Analogien zur Erkenntnisgewinnung und Wissensvermittlung ist bereits seit der Antike geläufig und wird meist dazu verwendet, um komplexe Sachverhalte zu veranschaulichen (Issing, 1994). Allerdings gilt es zu berücksichtigen, dass diese Form informierender Bilder hohe Anforderungen sowohl an die didaktische Gestaltung, als auch an die kognitive Verarbeitung stellt. Neben dem Verstehen der konkreten Abbildung muss dem Betrachter ein Analogieschluss auf die eigentlichen Zielinhalte gelingen, was entsprechende Anleitungen und Dekodierhinweise notwendig macht.

Gerade im Zusammenhang mit Analogiebildern wird deutlich, dass eine trennscharfe Kategorisierung kein einfaches Unterfangen darstellt. Zwar handelt es sich äußerlich um realistische Darstellungen, welche die Eigenschaften eines konkreten Sachverhaltes veranschaulichen, doch erst die erklärende Funktion von Analogiebildern machen den Unterschied aus. Damit eröffnet sich eine weitere Möglichkeit, Bilder zu kategorisieren. Während in den bisherigen Ausführungen oftmals die Ähnlichkeit zwischen einem Sachverhalt und dessen grafischen Darstellung Grundlage der Klassifikation war, lassen sich Bilder auch nach ihrer Funktion einteilen. Bleiben sämtliche motivationalen und affektiven Aspekte unberücksichtigt, unterscheiden beispielsweise J. Levin, Anglin und Carney

3 Verarbeitung von Texten und Bildern

(1987) vier kognitive Funktionen, die visuelle Darstellungen beim Wissenserwerb erfüllen können.

Am häufigsten werden nach Meinung der Autoren Bilder eingesetzt, um einzelne Gegenstände, Objekte oder Zusammenhänge zu konkretisieren. J. Levin et al. (1987) sprechen in diesem Zusammenhang von einer *representation function*, da einzelne Inhalte, die etwa in einem Text beschrieben sind, gezielt veranschaulicht werden. Handelt es sich um einen Text beispielsweise von einer bestimmten Vogelart, dann enthalten viele Lehrbücher eine Photographie oder realistische Zeichnung, die das entsprechende Tier zeigen. Darüber hinaus können Bilder dazu dienen, komplexe Sachverhalte so zu organisieren, dass der Lernende einen Überblick erhält und das zu erwerbende Wissen strukturieren kann. Diese *organization function* ist besonders geläufig beim Lernen von Prozeduren, wie man sie oftmals in Kochbüchern oder bei Bauanleitungen findet. Neben einer schriftlichen Darstellung des Vorgehens, wird die Abfolge besonders komplexer Arbeitsschritte anhand einer Reihe von Bildern verdeutlicht, wobei jedem Handgriff eine eigene Darstellung gewidmet ist. Ein weiteres Beispiel für grafische Darstellungen mit organisierender Funktion sind geographische Karten, da auch die Abbildung der räumlichen Relationen zwischen einzelnen Elementen eine Form der Organisation ist.

Bildern, die den Lernenden gezielt beim Verstehen unterstützen sollen, wird von J. Levin et al. (1987) eine *interpretation function* zugeschrieben. Zwar kann auch die adäquate Organisation spezifischer Inhalte das Verständnis fördern, doch liegt dem eine andere *Lernmechanismus* zugrunde. Bei interpretativen Bildern werden Inhalte und Zusammenhänge erinnert, weil sie das Verstehen des Sachverhaltes ermöglichen bzw. erleichtern, während bei der Organisationsfunktion positive Gedächtniseffekte vor allem auf die strukturelle Ordnung des zu erwerbenden Wissens zurückzuführen sind. Typisch für grafische Darstellung mit erklärender Funktion sind die bereits erwähnten Analogiebilder, die vor allem dazu verwendet werden, um abstrakte oder komplexe Themen zu verdeutlichen, die man sich nur schwer vorstellen kann. Ein Beispiel ist der bereits erwähnte Ladungstransport innerhalb eines Stromkreises.

Eine besondere Funktion haben die sogenannten *transformational pictures*, bei deren Gestaltung gezielt mnemotechnische Methoden eingesetzt werden, um das Behalten bestimmter Inhalte zu fördern (J. R. Levin, Shriberg & Berry, 1983). Zu den gedächtnisstützenden Techniken zählen das Umkodieren von Informationen, die Bildung von Relationen und die Verwendung von Hinweisreizen, die einen intentionalen Abruf erleichtern sollen. Natürlich haben auch die bereits erwähnten Bildfunktionen zur Folge, dass Gelerntes besser behalten und erinnert wird, doch sehen J. Levin et al. (1987) dies eher als einen indirekten Effekt an. Bilder mit Transformationsfunktion sind jedoch ziemlich selten in Lehr- oder Schulbüchern zu finden, obwohl deren positive Auswirkung auf die Behaltensleistung früh dokumentiert werden konnte (Bellezza, 1981; Pressley, Levin & Delaney, 1981). Zudem ist es fraglich, ob es sich hierbei tatsächlich um eine Funktion und nicht um einen spezifischen Bildtypus handelt, da die *transformational pictures* eigens konstruierte, grafische *Eselsbrücken* darstellen. Vielleicht unterscheidet Weidenmann (2002a) aus diesem Grund *nur* drei kognitive Funktionen, wobei die von ihm definierte *Zeigefunktion*

weitgehend der *representational function* von (J. Levin et al., 1987) entspricht. Ähnlich verhält es sich mit der *Konstruktionsfunktion*, die dadurch gekennzeichnet ist, dass Bilder den Aufbau mentaler Modelle unterstützen können und somit im Sinne der *interpretation function* das Verstehen fördern. Allerdings erfüllen grafische Darstellungen, die bei den Betrachtern gezielt Vorwissen oder Alltagserfahrungen aktivieren sollen, nach Weidenmann (2002a) eine *situierende* Funktion. Diese dritte und letzte Kategorie unterscheidet sich damit von der *organizational function*, da hier das Bildmaterial keine Organisationshilfe ist, sondern durch die Aktivierung von Kontextinformationen den Wissenserwerb positiv beeinflusst. Fasst man die verschiedenen Definitionen zusammen, die J. Levin et al. (1987) und Weidenmann (2002a) aufzählen, dann ergeben sich damit insgesamt fünf kognitive Bildfunktionen.

Abschließend lässt sich festhalten, dass die verschiedenen Arten und Funktionen von Bildern den Weg zu einer allgemeinen Theorie des Bildverstehens deutlich erschweren. Dennoch gibt es einige kognitive Prozesse, die theoretisch jeder Verarbeitung bildlicher Informationen zugrunde liegen, wie in dem nun folgenden Abschnitt ausführlich dargestellt wird.

3.2.2 Kognitive Prozesse bei der Verarbeitung von bildlichen Informationen

Um sich dem Prozess des Bildverstehens zu nähern, bietet es sich an, erneut die Verarbeitung von sprachlichen Informationen zu betrachten. Obwohl in der vorangegangenen Darstellung zum Textverstehen subsemantische Prozesse ausgeklammert wurden, beruhen die meisten der in Kapitel 3.1 diskutierten Ansätze auf der Annahme, dass vor bzw. zeitgleich mit dem Aufbau semantischer Repräsentationen eine sogenannte *Textoberflächenstruktur* konstruiert wird. Diese Textoberflächenstruktur repräsentiert die wortwörtliche Bedeutung der verbalen Informationen und ist das Ergebnis von Wahrnehmungs- und basalen Gedächtnisprozessen, von denen angenommen wird, dass sie weitgehend unbewusst ablaufen. So müssen Buchstaben und Worte zunächst diskriminiert und identifiziert werden, bevor die Zuweisung einzelner Bedeutungen und das Verstehen ganzer Sätze erfolgen kann. Wahrnehmungsprozesse und der Zugriff auf das mentale Lexikon laufen dabei in so dichter Folge ab, dass ein Leser oder Zuhörer sich dessen kaum bewusst wird. Aufgrund des linearen Charakters sprachlicher Informationen und deren grammatischen Regulierung, ist es noch relativ einfach, diese Prozesse nachzuvollziehen und theoretische Annahmen zu prüfen.

Etwas anders verhält es sich mit der Verarbeitung bildlicher Informationen, da das Betrachten von grafischen Darstellungen sehr viel weniger reglementiert und vorgezeichnet ist. Dennoch kann angenommen werden, dass auch das Verstehen von Bildern mit einer Reihe von präattentiven, subsemantischen Verarbeitungsprozessen beginnt, die vor allem der Diskrimination und der Identifikation von grafischen Komponenten oder Objekten dient (Winn, 1994). Ähnlich wie Buchstaben zu Worten zusammengesetzt werden, lassen sich auch Punkte, Linien und Flächen mental zu bedeutungstragenden piktoralen Einheiten verbinden. Die Regeln, nach denen Menschen in einer Zeichnung mit nur

3 Verarbeitung von Texten und Bildern

wenigen Strichen beispielsweise einen Baum erkennen, werden noch heute mit den Gestaltgesetzen von Wertheimer (1923) erklärt, auf die an dieser Stelle aber nicht näher eingegangen werden soll. Seitdem wurden vereinzelt weitere theoretischer Ansätze zur grundlegenden Bildwahrnehmung und -verarbeitung entwickelt vgl. Pinker (1984), doch auch diese führen zu sehr ins Detail. Festzuhalten bleibt, dass in dem Frühstadium der Verarbeitung piktorialer Informationen schwierig ist, eine klare Trennlinie zwischen Perzeption und Kognition zu ziehen. Immerhin kann man davon ausgehen, dass ein Großteil der frühen Prozesse automatisiert und stimulusgeleitet abläuft, weshalb höhere kognitive Fähigkeiten und vorhandene Wissensstrukturen diese kaum beeinflussen dürften (Ullman, 1984).

Das Ergebnis der präattentiven Verarbeitungsschritte ist schließlich eine perzeptuelle Repräsentation des Bildes oder Diagramms, die sich mit der Textoberflächenstruktur verbaler Informationen vergleichen lässt. Schnotz und Dutke (2004) nehmen an, dass spätestens nach dem Aufbau dieser *Bildoberflächenrepräsentation* bewusstseinspflichtige Prozesse einsetzen, welche die Aufmerksamkeit des Betrachters erfordern und damit die Kapazitäten des Arbeitsgedächtnisses in Anspruch nehmen. Das *visual sketchpad*, in dem piktoriale Informationen wahrnehmungsnah abgebildet werden (Baddeley, 1992), unterliegt dabei Kapazitätsbeschränkungen, weshalb die kognitive Verarbeitung komplexer oder umfangreicher Bildmaterialien mehrere Arbeitsschritte benötigen kann. Während beim Lesen von Texten, die Reihenfolge der Prozesse in Abhängigkeit von Grammatik und Schriftart vorgegeben ist, wird der Betrachter einer grafischen Darstellung sehr viel weniger angeleitet (Winn, 1990). Bei der Betrachtung eines Bildes ist demzufolge nicht immer klar, welche der dargestellten Elemente für das Verständnis wichtig sind und welche nicht. Zwar kann man davon ausgehen, dass bei der Bildverarbeitung mental zuerst eine globale Gesamtstruktur aufgebaut wird (Navon, 1977), doch ist die Sequenz, in der die einzelnen grafischen Komponenten analysiert werden, weitgehend der Entscheidung des Betrachters überlassen. Damit bestimmt letztlich das Vorwissen über Auswahl und Reihenfolge der entnommenen piktorialen Informationen, da die Erfahrungen des Betrachter festlegen, in welchen Kontext die Abbildung einzuordnen ist.

Im Zusammenhang mit realistischen Bildern kann dabei auf kognitive Schemata zurückgegriffen werden, die vor allem das Ergebnis alltäglicher Wahrnehmungen sind (Weidenmann, 1988). Beim Verstehen von logischen Bildern ist ein solcher Rückgriff dagegen in der Regel nicht möglich, da hier keine perzeptuelle Ähnlichkeit mit dem repräsentierten Sachverhalt vorliegt. Beispielsweise werden zur Verarbeitung eines Liniendiagrammes spezifische Schemata benötigt, mit deren Hilfe ein Lernender die Relationen von Achsen und Kurven mental integriert, um eine entsprechende Information ermitteln zu können. Diese müssen aber, ähnlich wie die Fertigkeit zu lesen, erlernt werden. Schnotz (1994b) nimmt an, dass ein Betrachter zum vollständigen Verstehen eines logischen Bildes zwei unterschiedliche Arten von Schemata benötigt. Die *lokalen Schemata* beziehen sich auf Details, wie das Erkennen von Minima und Maxima oder das Ablesen eines bestimmten einzelnen Wertes. Für das Erkennen von Entwicklungstrends, Relationen zweiter Ordnung oder andere komplexer Muster sind dagegen *globale Schemata* erforderlich,

über das vor allem Personen mit entsprechendem inhaltlichen Vorwissen verfügen (Lowe, 1994). Somit ist auch das Verstehen von logischen und realistischen Bildern stark vorwissensabhängig, da zum Aufbau einer kohärenten mentalen Repräsentation in der Regel Informationen verlangt werden, die in der entsprechenden grafischen Darstellungen nicht enthalten sind.

Was die Repräsentationsform anbelangt, so wird gerade von bildlichen Lehr- bzw. Lernmaterialien erwartet, dass sie aufgrund ihres anschaulichen Charakters die Konstruktion mentaler Modelle besonders gut unterstützen (Gyselinck & Tardieu, 1994; Glenberg & Langston, 1992). T. Rasch (2006) weist jedoch darauf hin, dass sich text- und bildbasierte Konstruktionsprozesse unterscheiden, da das Verstehen eines Bildes mehr oder weniger direkt zum Aufbau eines mentalen Modells führt. Nach seiner Ansicht entfällt die propositionale Repräsentationsebene, zumal der Aufbau des mentalen Modells in erster Linie auf der Grundlage der wahrnehmungsnahen Bildoberflächenrepräsentation vollzogen wird. Beide Repräsentationsformen sind ihrer Natur nach analog, weshalb der Übergang von der Wahrnehmung zum Verstehen eines Bildes eigentlich leicht fallen sollte. Allerdings schließt der Autor nicht aus, dass bei der Verarbeitung piktorialer Informationen Propositionen beteiligt sein können, doch werden diese erst anhand des mentalen Modells abgelesen, wie Abbildung 3.1 verdeutlicht. Trotz der strukturellen Ähnlichkeit zwischen

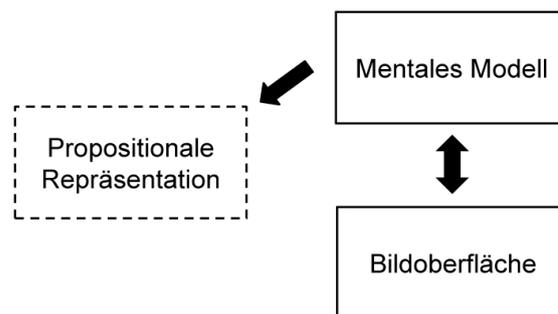


Abbildung 3.1: Repräsentationsebenen beim Bildverstehen nach Rasch (2006)

Oberflächen- und Modellrepräsentation ist anzunehmen, dass die Prozesse, die von einer Ebene zur nächsten führen, vielschichtig und komplex sind. So müssen einzelne graphische Komponenten entdeckt, identifiziert und aufgrund der Kapazitätsbeschränkungen der Arbeitsgedächtnisses sequentiell in den Konstruktionsprozess eingebracht werden. Von einem schlichten Kopiervorgang kann in diesem Zusammenhang keine Rede sein.

Eine davon leicht abweichende Position vertreten Schnotz und Dutke (2004), in deren Vorstellung die Konstruktion des mentalen Modells und eine konzeptgeleitete Analyse mehr oder weniger parallel ablaufen. Grundsätzlich basiert die Verarbeitung piktorialer Informationen zunächst vor allem auf analogen Strukturabbildungsprozessen (vgl. Gentner, 1983), bei denen ein Betrachter anforderungsrelevante Charakteristika einer grafischen Darstellungen mental nachbildet. Die erfolgreiche Identifikation einzelner Bildkomponenten macht es jedoch wahrscheinlich, dass bereits während der Konstruktion des mentalen

3 Verarbeitung von Texten und Bildern

Modells verschiedenen Elementen eine Bedeutung zugewiesen wird. Mit dieser konzeptgeleiteten Analyse ist somit auch eine deskriptionale Enkodierung verbunden, so dass die propositionale Verarbeitung von Beginn an zum Bildverstehen dazu gehört. Piktoriale Inhalte können dabei sowohl anhand der perzeptionsnahen Oberflächenrepräsentation, als auch auf der Grundlage des mentalen Modells abgelesen und propositional enkodiert werden. Demzufolge nimmt ein Betrachter saliente oder bedeutsame Bildkomponenten zunächst wahr und analysiert bzw. interpretiert sie mithilfe kognitiver Schemata, welche er aus dem Vorwissen abrufen. Ist der Interpretationsprozess erfolgreich, führt dies unter anderem zu einer propositionalen Repräsentation der entnommenen piktorialen Informationen, mit denen sich bereits vorhandene Propositionsgefüge ergänzt werden können. Die beteiligten Prozesse sind dabei miteinander verschachtelt und operieren seriell, so dass sich eine klare Abgrenzung der einzelnen Verarbeitungsschritte selbst theoretisch kaum verwirklichen lässt. Abbildung 3.2 fasst die Annahmen von Schnotz und Dutke (2004) noch einmal grafisch zusammen.

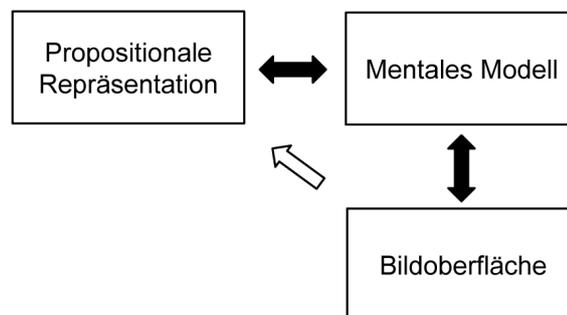


Abbildung 3.2: Repräsentationsebenen beim Bildverstehen nach Schnotz & Dutke (2004)

Der weiße Pfeil zwischen der Bildoberfläche und der propositionalen Repräsentation in der Abbildung soll verdeutlichen, dass diese Verbindung sehr viel schwächer ausgeprägt ist, als es ein schwarzer Pfeil anzeigen würde. So ist anzunehmen, dass Propositionen, die auf der Grundlage der Bildoberfläche generiert werden, in erster Linie der Identifikation bzw. Benennung einzelner Bildelemente dienen. Die eigentliche *Propositionalisierung* piktorialer Zusammenhänge erfolgt dagegen auf der semantischen Ebene des mentalen Modells, an dem sich Informationen ablesen bzw. inferieren lassen. Schnotz und Dutke (2004) nehmen an, dass diese *neuen* Informationen als Propositionen enkodiert und zu der bereits bestehenden Strukturen hinzugefügt werden. Ist die propositionale Repräsentation elaboriert genug bzw. kohärent, kann sie ihrerseits als Ausgangspunkt fungieren, um das mentale Modell zu inspizieren und weiter zu entwickeln. Es findet also eine Interaktion zwischen den beiden Formaten auf der semantischen Repräsentationsebene statt. Unabhängig von dem zugrunde liegenden Zeichensystem ist eine propositionale Repräsentation ebenso ein Bestandteil der kognitiven Bildverarbeitung, wie die Konstruktion eines mentalen Modells beim Textverstehen.

Allerdings unterschätzen Lernende häufig den Informationsgehalt grafischer Darstel-

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

lungen, so dass die zum Verständnis notwendige konzeptionelle Analyse oft zu oberflächlich ausfällt. Ähnlich wie beim Lesen von Texten, kann damit auch beim Betrachten von Bildern eine Art *illusion of knowing* entstehen (vgl. Glenberg, Wilkinson & Epstein, 1982), ohne dass die dargestellten Sachverhalte wirklich begriffen werden. Nach Schnotz und Dutke (2004) ist dieses oberflächliche Verständnis ein Zeichen dafür, dass ein Lernender zu wenige oder nur unzureichende Ableseprozesse an seinem mentalen Modell vorgenommen hat. Die direkte Folge ist eine propositionale Repräsentation, die nicht genügend elaboriert ist, um für eine weitere Modellkonstruktion und -inspektion verwendet zu werden. Der Aufbau des mentalen Modells bleibt damit unvollendet und die entsprechenden Informationen werden nur scheinbar kohärent repräsentiert. Allgemein besteht beim Lernen mit Bildern die Gefahr, dass ein Betrachter glaubt, alle relevanten Informationen mit wenigen Blicken aufnehmen und verarbeiten zu können (Mokros & Tinker, 1987). Gleichzeitig verfügen viele Lernende nicht über die erforderlichen metakognitiven Monitoringstrategien oder wenden sie bei grafischen Darstellungen nicht an, so dass die Kontrolle, ob der abgebildete Sachverhalt tatsächlich verstanden wurde, weitgehend entfällt (Pettersson, 1988). Aus diesem Grund ist es oft notwendig, die tiefgehende Verarbeitung eines Bildes zu fördern, indem man diese ausdrücklich instruiert.

Obwohl einige Ansätze bestehen, zeigen die bisherigen Ausführungen, dass eine umfassende und detaillierte Theorie, wie sie beispielsweise für das Verstehen von Texten existiert, bislang fehlt. Ein Grund für diese bestehende Lücke ist sicherlich mit dem Umstand verbunden, dass es sehr verschiedene Arten von Bildern gibt, die konsequenter Weise eine spezifische theoretische Betrachtung erfordern. Beispielsweise wird angenommen, dass realistische Bilder aufgrund von Schemata verstanden werden, die in der alltäglichen Auseinandersetzung mit der Umwelt zuvor erworben wurden. Dagegen muss ein Individuum den korrekten Umgang mit logischen Bildern, vergleichbar mit der Lesekompetenz, erst erlernen. Hinzu kommt, dass grafische Darstellungen sehr unterschiedliche kognitive Funktionen wahrnehmen können, die darüber hinaus nicht immer mit Absichten desjenigen übereinstimmen, der sie gestaltet hat (Pettersson, 1988). Bilder sind oftmals mehrdeutig und werden daher gerade im Zusammenhang mit dem Wissenserwerb in der Regel nicht isoliert eingesetzt. So beziehen sich Abbildungen in Lehr- oder Schulbüchern meistens auf Texte oder enthalten selbst symbolische Zeichen wie Beschriftungen oder Legenden, die den Betrachter anleiten sollen. Aus diesem Grund ist es nicht erstaunlich, dass einige ausführliche theoretische Modelle existieren, die sich mit der integrativen Verarbeitung von Text- und Bildinformationen beschäftigen. Diesen Modellen und dem Wissenserwerb mit multiplen Informationsquellen ist das nachfolgende Kapitel gewidmet.

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

Die Verwendung von Bildern in Texten besitzt eine lange Tradition und lässt sich bis in die frühe Neuzeit zurückverfolgen (Schnotz, 2005). Schon damals waren einzelne Gelehrte der Ansicht, dass die Verwendung multipler Repräsentationsformate den Wissenserwerb erleichtern bzw. unterstützen kann. Diese Einschätzung hat sich mittlerweile auch empirisch

3 Verarbeitung von Texten und Bildern

risch als zutreffend erwiesen, da wissenschaftliche Überblickdarstellungen aus neuerer Zeit belegen, dass häufig bessere Wiedergabe- und Verständnisleistungen erzielt werden, wenn Lernende beim Wissenserwerb Texte und Bilder und nicht nur Text heranziehen (Levie & Lentz, 1982; J. Levin, Anglin & Carney, 1987; Carney & Levin, 2002). Das Phänomen der Überlegenheit einer Kombination multipler Repräsentationsformate wurde von Mayer (1997) als *Multimedia-Effekt* bzw. als *Multimedia-Prinzip* bezeichnet und hat seit dem zahlreiche theoretische und empirische Arbeiten innerhalb der pädagogischen Psychologie angeregt. Der Begriff *Multimedia* wird allerdings auch außerhalb des wissenschaftlichen Kontextes und dort meist im Zusammenhang mit technischen Neuerungen oder der Unterhaltungselektronik verwendet, was gerade unter Laien zu Verwirrungen führen kann. Daher ist es weniger missverständlich von *multimedialem Lernen* zu sprechen, was sich schlicht auf die Tatsache bezieht, dass Schüler oder Studenten während des Wissenserwerbs auf Repräsentationsformate zurückgreifen, denen unterschiedliche Zeichensysteme zugrunde liegen. Dazu braucht es aber nicht unbedingt einen Computer oder andere technische Gerätschaften. In seiner einfachsten Form handelt es sich schon um multimediales Lernen, wenn ein Text gelesen wird, der mindestens ein Bild enthält, dem eine kognitive Funktion zugeschrieben wird.

Obwohl der Multimedia-Effekt als empirisch gut abgesichert gilt, ist es irreführend anzunehmen, dass sich mit der Kombination von Texten und Bildern der Lernerfolg zwangsläufig steigern lässt. Wie so oft sind die Zusammenhänge sehr viel komplizierter und neben den Charakteristika des Lehr-/Lernmaterials bestimmen die Merkmale des Lernenden über Richtung und Ausmaß des Einflusses. Im Laufe der Zeit sind daher einige theoretische Arbeiten entstanden, um die teilweise widersprüchlichen experimentellen Ergebnisse zu erklären und in ein stimmiges Gesamtkonzept zu integrieren. Von diesen werden im folgenden vier Modelle näher vorgestellt, die sich ganz speziell der gemeinsamen Verarbeitung von Text- und Bildinformationen widmen.

3.3.1 Die Theorie der Dualen Codierung

Einer der ersten Ansätze, der explizit die kognitive Verarbeitung verbaler und bildlicher Informationen thematisiert, ist die Theorie der dualen Kodierung oder DCT (*dual coding theory*) von Paivio (1986). Sie wurde auf der Grundlage von Beobachtungen entwickelt, die der Autor bei einer Reihe von Gedächtnisexperimenten machte (vgl. Paivio, 1991). So konnte Paivio, Rogers und Smythe (1968) unter anderem nachweisen, dass Bilder besser erinnert werden, als Worte, was für eine mnemonische Überlegenheit piktorialer Informationen spricht. Dieser *Bildüberlegenheitseffekt* führte Paivio (1971) zu der Annahme, dass die Entstehung und Verarbeitung mentaler Repräsentationen auf zwei, theoretisch voneinander abgrenzbaren Systemen beruht. In einem ersten, verbalen System werden demnach alle sprachlichen Informationen verarbeitet, ganz gleich ob sie gelesen, gehört oder, wie im Fall der Brailleschrift, ertastet werden. Man könnte in diesem Zusammenhang auch von einem symbolischen System sprechen, da die Kodierung der eingehenden oder produzierten Informationen einen arbiträren Charakter besitzen und

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

ihre Bedeutung durch kulturelle Konventionen bestimmt ist.

Alle nicht linguistische Reize wie Bilder, Klänge, haptische und olfaktorische Reize werden dagegen in einem non-verbalen System verarbeitet, das sich von der Funktionsweise des verbalen Systems unterscheidet. Während sprachliche Stimuli sequentiell verarbeitet werden müssen, erlauben die non-verbalen Repräsentationen eine parallele oder ganzheitliche Kodierung der entsprechenden Informationen (Clark & Paivio, 1991). Außerdem zeichnen sich die Formate des non-verbalen Systems durch Ähnlichkeit und Wahrnehmungsnähe und nicht durch einen symbolhaften Charakter aus. Gerade das Repräsentationsformat von Bildern besitzt dabei Eigenschaften, die stark an die mentalen Modelle von Johnson-Laird (1983) erinnern, auch wenn die Betonung struktureller Analogien fehlt. Beispielsweise kann ein Schüler in seiner Vorstellung einen zuvor wahrgenommenen Würfel rotieren lassen oder durch mentale Simulation das Ergebnis einer Carambolage beim Billard vorhersagen. Solche räumlichen Transformationen sind dagegen mit dem verbalen Repräsentationsformat nicht möglich. Etwas unklar bleibt hingegen die Verarbeitung von Gerüchen oder nicht sprachlichen Lauten, die zwar auch als olfaktorische oder auditive Bilder (*images*) bezeichnet werden, allerdings keine eigenständige kognitive Funktion besitzen.

Die den beiden Systemen zugrunde liegenden Kodierungsformate bezeichnete Paivio (1971) ursprünglich als verbale und imaginale Repräsentationen, benannte sie aber etwas später in *logogene* und *imagene* um. Obwohl das verbale und nonverbale Subsystem aufgrund der zugrunde liegenden Formate grundsätzlich verschieden sind, nimmt Paivio (1986) an, dass die Verarbeitung eines bildlichen Stimulus (*imagen*) auch immer zur Aktivierung entsprechender *logogene* führt. Dies geschieht durch *referentielle Verknüpfungen*, die immer dann möglich sind, wenn der Betrachter eines Bildes den dargestellten Sachverhalt mit einem oder mehreren Worten benennen oder umschreiben kann. Erblickt man beispielsweise bei einem Spaziergang einen Hund, werden parallel zur visuellen Wahrnehmung entsprechende sprachliche Repräsentationsmuster aktiviert. Diese Aktivierung der *logogene* kann Rasse, Größe und Aussehen des Tiers betreffen, so dass die eigentlich piktorialen Informationen in beiden Systemen verarbeitet und gespeichert werden.

Die duale Kodierung von Informationen ist aber auch dann möglich, wenn die Grundlage ein Text ist. Hier aktivieren einzelne Worte entsprechende *imagene* und bilden referentielle Verknüpfungen zwischen den verschiedenen Repräsentationsformaten. Bei den aktivierten nonverbalen Gedächtnisinhalten kann es sich um Erinnerungen an Gerüche, Laute oder visuelle Vorstellungen bzw. mentale Bilder handeln, je nachdem, welche perzeptiven Eindrücke mit dem Begriff assoziiert werden können. Abbildung 3.3 verdeutlicht diesen Zusammenhang und hebt hervor, dass eine duale Kodierung von Informationen auf der Grundlage beider Systeme möglich ist.

Allerdings regt nicht jedes Wort, das Bestandteil des menschlichen Wortschatzes ist, gleichermaßen zu einer referentiellen Verknüpfung zwischen den beiden Systemen an. Gerade bei Begriffen, die einen abstrakten Sachverhalt repräsentieren, ist es weniger wahrscheinlich, dass perzeptuelle Eindrücke vorhanden sind, die für eine nonverbale Ko-

3 Verarbeitung von Texten und Bildern

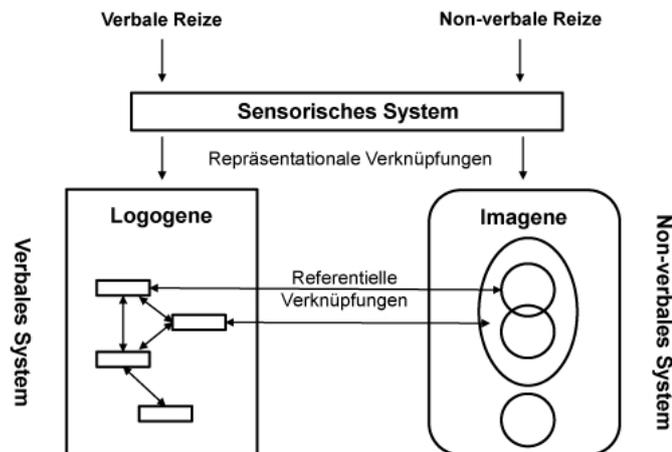


Abbildung 3.3: Modell des verbalen und nonverbalen Systems der Theorie der dualen Kodierung nach Paivio (1986)

dierung genutzt werden können. Damit ist es letztlich die Anschaulichkeit (*concreteness*) der verbalen Inhalte, die über das Ausmaß und die Intensität der referentiellen Verknüpfungen entscheidet. Enthält ein Zeitungsartikel beispielsweise die Geschichte eines Hundes, in der Rasse, Aussehen und Verhalten ausführlich beschrieben sind, werden sich viele Leser das Tier bildlich vorstellen können. Fehlen hingegen die entsprechenden sprachlichen Details, dann ist aufgrund der vielen Freiheitsgrade lediglich eine verschwommene Vorstellung möglich, selbst wenn der betreffende Zeitungsleser viel über Hunde weiß. Je abstrakter, je unspezifischer oder wahrnehmungsferner eine sprachliche Umschreibung ist, umso schwerer fällt die Verarbeitung in beiden Systemen. Das kann in einigen Fällen zu Konsequenz haben, dass sich einzelne Begriffe oder Zusammenhänge ausschließlich verbal und damit einfach kodieren lassen.

Nach Paivio (1991) ist es jedoch gerade die gemeinsame Nutzung des verbalen und des non-verbalen Systems bzw. der darin kodierten Repräsentationseinheiten, die einen positiven Einfluss auf das Erinnerungsvermögen hat. Stark vereinfacht geht die DCT davon aus, dass nach einer Kodierung in beiden Systemen zwei Gedächtnisspuren existieren, die quasi additiv für Abrufprozesse genutzt werden können. Ursprünglich hat Paivio (1967) diese *Additivitätshypothese* formuliert, um zu erklären, weshalb konkrete Wörter besser erinnert werden, als etwa abstrakte, doch weitete er seine Überlegungen nur wenig später auf den *Bildüberlegenheitseffekt* aus (Paivio, Rogers & Smythe, 1968). Demnach ist eine doppelte Kodierung am wahrscheinlichsten, wenn dem Lernenden eine Bild vorgelegt wird, da zur Identifikation und Kategorisierung piktorialer Informationen das verbale System fast zwangsläufig benötigt wird. Hinzu kommt, dass Bilder in der Regel sehr viel konkreter sind als Texte. Beispielsweise zeigt die Photographie eines Hauses genau ein bestimmtes Haus und keine anderes, während selbst eine ausführliche sprachliche Umschreibung bei jedem Leser eine individuell einzigartige Vorstellung erzeugen wird.

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

So nimmt auch bei Worten, die konkrete Sachverhalte bezeichnen, die Wahrscheinlichkeit für eine Nutzung beider Kodierungsformen ab. Allerdings ist es möglich duale Kodierungsprozesse der entsprechenden sprachlichen Informationen durch Imaginationsinstruktionen (*Stellen Sie sich die bezeichneten Gegenstände bitte bildlich vor...*) gezielt anzuregen. Diese Option gibt es im Zusammenhang mit abstrakten sprachlichen Inhalten hingegen nicht, da hier kongruente perzeptuelle Wahrnehmungen fehlen, was die Nutzung beider Kodierungsformate eher unwahrscheinlich macht.

Obwohl die DCT vor allem durch ihre Anschaulichkeit besticht und der Bildüberlegenheitseffekt empirisch als gut abgesichert gilt (Engelkamp, 1991; D. L. Nelson, Reed & Walling, 1976), provoziert der Ansatz auch Kritik. So sieht Engelkamp (1994) besonders die variablen Wahrscheinlichkeiten, mit der Informationen dual enkodiert werden, als problematisch an, da sie seiner Meinung nach erst *post hoc* in das theoretische Konzept integriert wurden. Ein weiterer Kritikpunkt betrifft den Effekt von Imaginationsinstruktionen beim Lernen konkreter Worte auf die Gedächtnisleistung. Eigentlich sollte durch die Erzeugung bildlicher Vorstellung der Bildüberlegenheitseffekt verschwinden, was nach Ansicht von Engelkamp (1994) und Weidenmann (2002b) jedoch nicht der Fall ist. Dieser Einwand steht jedoch im Widerspruch zu experimentellen Befunden von Paivio und Csapo (1973), die schon früh belegen, dass durch Imaginationsanweisungen der Bildüberlegenheitseffekt tatsächlich aufgehoben werden kann.

Wie dem auch sei, ein weitere problematischer Aspekt ist die Tatsache, dass Paivio (1986) die spezifischen Eigenschaften der einzelnen Sinnesmodalitäten des nonverbalen Systems weitgehend unberücksichtigt lässt. Jeder nicht-sprachliche Sinneseindruck mehr oder weniger ein *imagen*, selbst wenn der entsprechende Reiz beispielsweise akustischer Natur ist. Laut Engelkamp (1994) wäre dieser Umstand nicht weiter problematisch, wenn das Klingeln eines Telefon die gleichen repräsentativen Prozesse zur Folge hätte, wie dessen Anblick. Da er diese vereinfachende Annahme aber bezweifelt, müssten modalitätsspezifische Reize ihr mentale Repräsentation innerhalb der DCT weitaus mehr Beachtung finden, was jedoch nicht der Fall ist. Ein anderer wichtiger Einwand gegen die Theorie von Paivio (1971, 1986) betrifft ihre Brauchbarkeit zur Erklärung komplexer Lernprozesse. Auch wenn gelegentlich versucht wird, die DCT auf *reale* Lehr-/Lernkontexte anzuwenden (vgl. Clark & Paivio, 1991), so ist ihre praktische Relevanz ziemlich gering. Letztlich beziehen sich alle Annahmen auf Experimente, bei den einfache Wortlisten frei oder assoziiert wiedergegeben werden mussten. Sicherlich ist es ein überaus beachtenswerter Befund, wenn der Bildüberlegenheitseffekt durch gezielte kognitive Prozesse überwunden werden kann, doch lassen sich daraus nur wenige Konsequenzen ableiten, die beispielsweise für die Unterrichtsgestaltung relevant wären. Verstehen und bedeutungsvolles Lernen sind Konzepte, die nachträglich und nur am Rande innerhalb der DCT aufgegriffen werden (Paivio, 1991).

Doch unabhängig von all diesen Kritikpunkten bleibt Paivio's (19971, 1986) Arbeit wegweisend, da er als einer der ersten ein theoretisches Modell zu Diskussion gestellt hat, das eine getrennte Verarbeitung sprachlicher und bildlicher Informationen annimmt. Diese zentrale Annahme, wurde zumindest teilweise von neueren Ansätzen übernommen,

wie die nun folgenden Abschnitte zeigen werden.

3.3.2 Modell des Textlernens unter Verwendung geographischer Karten

Die Überschrift zu diesem Unterkapitel ist zugegebenermaßen ein klein wenig sperrig, was zu einem großen Teil an der englischen Originalbezeichnung liegt, die *Model of text learning using geographic maps* lautet (Verdi & Kulhavy, 2002). Auch mag das Modell bei weitem nicht so einflussreich sein, wie etwa die Arbeiten von (Paivio, 1986) oder den noch folgenden Autoren. Dennoch umfasst es interessante theoretische Annahmen hinsichtlich der kognitiven Verarbeitung von Text- und Bildinformationen, weshalb diesem Ansatz ein eigenes Kapitel gewidmet ist. Seine Wurzeln hat das Modell des *Textlernens unter Verwendung geographischer Karten* in der DCT von Paivio (1971) und in der *conjoint retention theory* (CRT) von Kulhavy, Lee und Caterino (1985), deren Bezeichnung sich etwas umständlich mit *Theorie der verbundenen Aufrechterhaltung* übersetzen lässt. Da der Ansatz von Paivio (1971, 1986) im vorangegangenen Kapitel ausführlich behandelt wurde, konzentriert sich die nun folgende Darstellung in erster Linie auf die CRT.

Bereits Davis und Hunkins (1968) konnten im Rahmen eines Lernexperimentes feststellen, dass geographische Karten den Wissenserwerb mit Texten positiv beeinflussen. Damit dies aber geschehen kann, müssen sich beide Medien inhaltlich aufeinander beziehen und zu einem gewissen Grad redundant² sein. Der Effekt erweist sich auch dann als stabil, wenn Lernende die Karten selbst zeichnen oder den Text erst nach dem Betrachten der Karte lesen (Dean & Kulhavy, 1981). Eine naheliegende Erklärung für diesen Befunde ist die Annahme, dass die abgebildeten Elemente in der Abbildung die Aufmerksamkeit auf relevante Inhalte des Textes lenken, was zu einer besseren Behaltensleistung führt. Allerdings widersprechen experimentelle Ergebnisse von Schwartz und Kulhavy (1981) einem solchen Erklärungsversuch ebenso, wie der Vermutung, dass die räumliche Anordnung geographischer Karten als eine Art *advanced organizer* fungiert. Kulhavy et al. (1985) sehen den eigentlichen Grund für den Effekt in einer Interaktion der Kartenstruktur mit den semantischen Elementen der darauf bezogenen Textinhalte.

Ihre *conjoint retention* Hypothese greift dabei zentrale Annahmen von Paivio (1971) auf und überträgt sie auf räumlich organisierte Abbildungen bzw. geographische Karten. So gehen die Autoren davon aus, dass Bilder als perzeptionsnahe und räumliche Repräsentationen enkodiert werden, die zumindest quasi-piktoriale Merkmale des ursprünglichen Stimulus besitzen (Kosslyn & Pomerantz, 1977). Später präzisieren Kulhavy, Stock, Verdi, Rittschof und Savenye (1993) diese Annahme und sprechen von *images*, die vor allem räumliche Informationen der zugrundeliegenden Karte bewahren können. Dagegen erfolgt die Verarbeitung von sprachlichen Reizen in einem separaten Speicher, wobei das zugrunde liegende Format Propositionen sind, deren Enkodierung sequentiell abläuft. Die verbalen und bildlichen Repräsentationen erfüllen jedoch unterschiedliche Funktionen, so dass die mit ihnen verbunden kognitiven Prozesse in der Regel unabhängig stattfinden.

²Zur Problematik des Begriffs Redundanz, wenn es um das Informationsverhältnis von Bild- und Textinformationen geht, siehe Kapitel 7.1.2 auf Seite 77)

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

Dennoch ist es nach Ansicht von Kulhavy, Stock, Verdi et al. (1993) möglich, zwischen den beiden Repräsentationsformaten assoziative Verknüpfungen zu bilden. Dies gilt umso mehr, wenn Lernende dazu angehalten werden, die Inhalte von Text und Karte inhaltlich aufeinander zu beziehen. Bei einer erfolgreichen Verknüpfung reicht es aus, dass eine der assoziierten Repräsentationen in das Arbeitsgedächtnis gelangt, um die andere im Langzeitspeicher aktivieren und abzurufen zu können. Wird beispielsweise das *image* einer Karte bewusst erinnert, dient es als Aktivationsquelle für Textinformationen, die mit den verschiedenen Bildelementen verknüpft wurden.

Das Besondere an der *conjoint retention theory*, die sich bis zu diesem Punkt kaum von der DCT unterscheidet, ist der spezielle Status, den die Enkodierung und Repräsentation visueller Informationen besitzen. Demnach setzten sich geographische Karten nach Ansicht von Kulhavy et al. (1985) aus einzelnen Bildelementen (*features*) zusammen, die räumlich angeordnet sind. Diese räumliche Anordnung wird von den Autoren als kartenimmanente Struktur angesehen, die für den Lernenden eine zusätzliche Informationsquelle darstellt. Während es sich bei den *features* in der Regel um separate Bilder, Zeichnungen und Symbole handelt, die eine mehr oder weniger anschauliche Bedeutung besitzen, gibt die Struktur Auskunft über die Relationen zwischen den verschiedenen Bildelementen. Erst durch Abstände, Richtungen, und Abgrenzungen entstehen so Beziehungen, die vor dem inneren Auge des Betrachters ein Gesamtbild erzeugen. Auf diese Weise werden auf der Grundlage von strukturellen Informationen aus vielen kleinen Bildern ein großes Bild geformt, das sehr viel mehr ist, als eine willkürliche Ansammlung oder Liste visueller Einzelobjekte. Die Encodierung einer solchen Struktur hat natürlich auch Konsequenzen für die kognitiven Prozesse, die den Wissenserwerb betreffen. So nehmen die Autoren an, dass selbst die mentale Repräsentation komplexer Karten aus nur einem einzigen *image* bestehen und gleichzeitig sämtliche *features* umfassen kann. Dies ist allerdings nur dann gegeben, wenn alle Strukturinformationen vollständig und korrekt abgebildet werden. Resultiert der Lernprozess jedoch in einem *intakten image* der Karte (vgl. Kulhavy, Stock, Verdi et al., 1993), dann beansprucht es nicht mehr Arbeitsgedächtniskapazität, als ein einziger *chunk* (Miller, 1956)).

Trotz dieser Informationskomprimierung bleibt es nach Meinung von (Kulhavy et al., 1985) möglich, auf sämtliche Inhalte der piktorialen Repräsentation zurückzugreifen. Dies geschieht, indem die Aufmerksamkeit von einem Bildelement auf das nächste gerichtet wird, wobei die mitcodierten Strukturinformationen den Abrufprozess leiten und die einzelnen Positionen vorgeben. Somit besitzen räumlich organisierte Abbilder gegenüber sprachlichen Informationen einen deutlichen Verarbeitungsvorteil. Denn obwohl deren mentalen Repräsentationen sehr viel mehr Informationen enthalten können, als dies etwa bei einer Proposition der Fall ist, nehmen sie etwa die gleichen mentalen Ressourcen in Anspruch. Besonders bedeutsam werden diese Eigenschaften, wenn neben einer Karte auch verbales Material für den Wissenserwerb zur Verfügung steht (Kulhavy, Stock & Kealy, 1993). Dies liegt zum einem Teil an den bereits erwähnten referentiellen Verknüpfungen zwischen Bild- und Textelementen und der damit verbundenen Nutzung von zwei Speichersystemen. Wirklich entscheidend bleibt jedoch die Encodierung der Kartenstruk-

3 Verarbeitung von Texten und Bildern

tur, da letztlich sie bestimmt, ob das *image* der Karte intakt ist oder nicht. Nur wenn alle Strukturinformationen vollständig repräsentiert werden, entsteht in der Vorstellung des Betrachter ein Bild, dessen Aufrechterhaltung nur wenig Kapazität beansprucht und das gleichzeitig für assoziative Prozesse genutzt werden kann. Ist die Encodierung der Kartenstruktur hingegen unvollständig, lassen sich die Bildelemente bzw. features der Karte mental nicht mehr sinnvoll zusammensetzen. Anstelle eines großen Bildes müssen nun mehrere kleine Bilder bewusst aufrechterhalten werden, was die begrenzte Kapazität des Arbeitsgedächtnisses sehr viel stärker belastet.

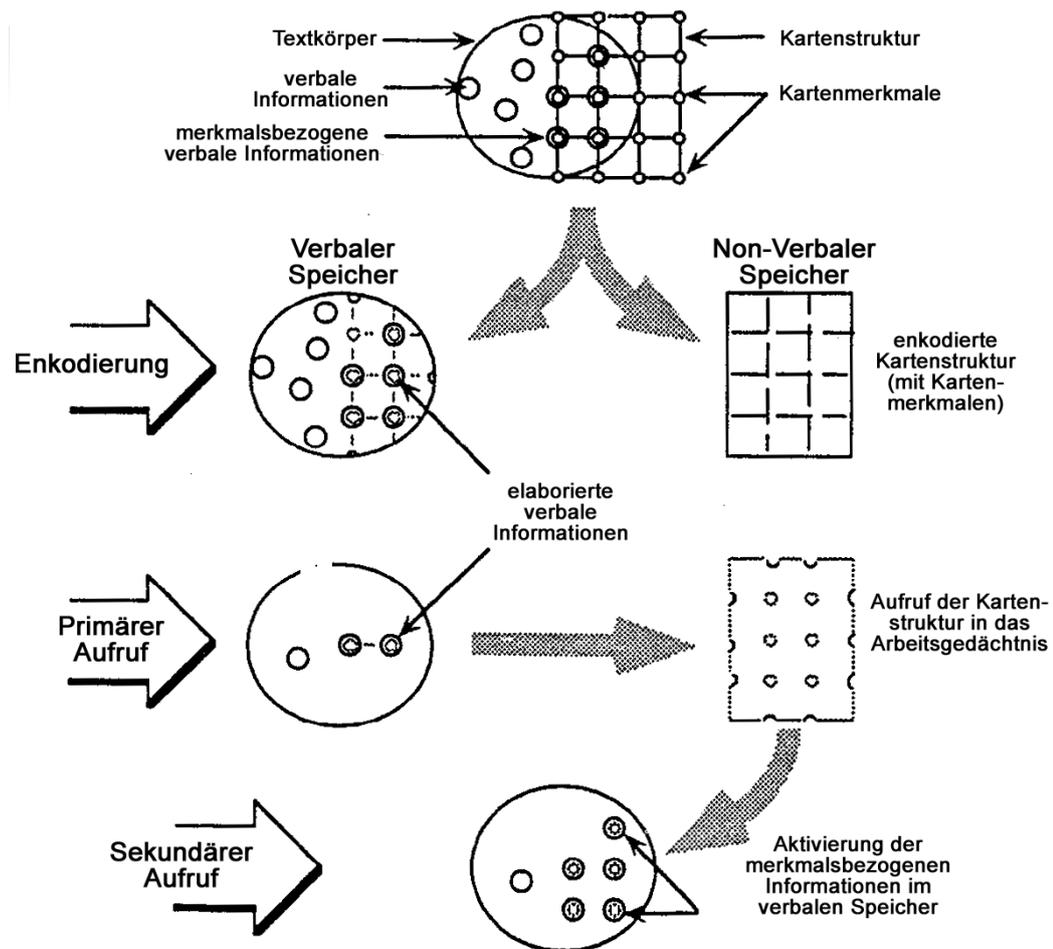


Abbildung 3.4: Modell der Enkodierung und des Abrufs von Text-Karten-Informationen Kulhavy, Stock & Kealy (1993)

Abbildung 3.4 fasst die bisherigen Ausführungen zusammen und setzt sie in den Kontext einer gemeinsamen Verarbeitung von Text- und Karteninformationen. Die erste Zeichnung, welche sich am oberen Ende der Abbildung befindet, verdeutlicht zunächst die inhaltliche Beziehung der beiden Formate. Es ist zu erkennen, dass sich die Informationen in Text und Karte überlappen, was referentielle Verbindungen zwischen verbalen und bildlichen Inhalten grundsätzlich möglich macht. Der eigentliche Verarbeitungsprozess beginnt dann mit der formatspezifischen Enkodierung, die in zwei separaten Speichern

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

des Arbeitsgedächtnisses erfolgt. So werden die Inhalte des Textes als Propositionen re-präsentiert, während die Verarbeitung des Bildes zu einem mentalen *image* führt, das neben den einzelnen Bildelementen auch sämtliche strukturellen Informationen umfasst. Obwohl der Vorgang in Abbildung 3.4 nicht dargestellt wird, entstehen in dieser Phase die referentiellen Verknüpfungen zwischen den verbalen Informationen und den einzelnen Bildelementen. Alle sprachlichen Inhalte, die mit einem *feature* der Karte assoziiert wurden, sind daher durch einen doppelten Kringel gekennzeichnet. Kulhavy, Stock und Kealy (1993) sprechen in diesem Zusammenhang auch von *elaborierten* verbalen Informationen, was sich jedoch allein auf die Verknüpfung mit einem entsprechenden Bildelement bezieht. Nachdem der Enkodierungsprozess abgeschlossen ist, werden die beiden Repräsentationen getrennt in das Langzeitgedächtnis überführt, von wo sie im Bedarfsfall abgerufen werden können.

Wie die Abbildung zeigt, unterscheiden Kulhavy, Stock und Kealy (1993) beim Abruf eine primäre von einer sekundären Phase. Während des primären Abrufs gelangen die enkodierten Inhalte erneut in das Arbeitsgedächtnis, wo in separaten Speichern versucht wird, die gesuchten Informationen zu extrahieren. Da ein intaktes *image* der Karte die begrenzte Kapazität vergleichsweise wenig belastet, können bildliche und verbale Repräsentationen gleichzeitig aufrechterhalten werden. In dieser Phase erfolgt die Extraktion der Zielinformationen noch weitgehend ohne bewussten Zugriff auf die zuvor gebildeten referentiellen Verknüpfungen. Beispielsweise kann eine Person zunächst versuchen, allein die assoziativen Verkettungen zwischen den Propositionen für den Abruf zu nutzen. Ist die Repräsentation der verbalen Inhalte jedoch unvollständig, misslingt der Zugriff und die benötigte Information wird nicht erinnert. In diesem Fall ist es möglich, das *image* der Karte im Rahmen eines sekundären Abrufs zu nutzen. Kulhavy et al. (1985) nehmen an, dass dabei die räumliche Repräsentation mental nach Hinweisen (*cues*) abgesehen wird, über die eine Aktivierung der entsprechenden verbalen Informationen im Langzeitspeicher schließlich doch noch gelingen kann. Der sekundäre Abruf ist damit eine direkte Konsequenz der referentiellen Verknüpfungen, die während der Enkodierungsphase zwischen Text und Bildinformationen etabliert wurden.

Kulhavy, Stock und Kealy (1993) selbst bezeichnen ihr Modell als eine *Version* der DCT, auch wenn einzelne Annahmen über die klassische Theorie hinausgehen. Gleichzeitig lässt sich jedoch kaum einer der Kritikpunkte, die den ursprünglichen Ansatz von (Paivio, 1986) betreffen, auf das theoretische Konzept zur Verarbeitung von geographischen Karten übertragen (siehe 65). Dies liegt in erster Linie daran, dass die Autoren für ihr Modell nur einen engen Geltungsbereich beanspruchen. Schwartz und Kulhavy (1981) experimentieren zwar mit Texten, die akustisch dargeboten wurden, doch konzentrieren sich die meisten der nachfolgenden Untersuchungen in der Regel auf die visuelle Sinnesmodalität. Ähnliches gilt für die pikorialen Stimuli, auf deren Grundlage der Ansatz entstanden ist. In ihrer weitesten Definition werden diese als organisierte räumliche Abbilder (*organized spatial displays*) bezeichnet. Darunter verstehen Verdi und Kulhavy (2002) neben den geographischen Karten, Diagramme, Advanced Organizer und andere bildlichen Formate, die eine räumliche Struktur aufweisen. Die Frage, ob sich das theoretische Kon-

zept, auch auf andere *Bildarten* anwenden lässt, bleibt dabei weitgehend unbeantwortet. Allerdings dürfte dies durchaus der Fall sein, da Illustrationen ohne räumliche Organisation nur schwer vorstellbar sind, selbst wenn die Struktur nicht auf den ersten Blick zu erkennen ist.

Da zahlreiche empirischen Befunde den positiven Einfluss von Karten auf den Wissenserwerb belegen, liegt es natürlich nahe, konkrete Anwendungsmöglichkeiten für den schulischen Unterricht zu entwickeln. Dennoch reichen die praktischen Implikationen, die mit der *conjoint retention* Theorie verbunden sind, nicht sonderlich weit. Zu einen lassen sich Karten nur in wenigen Fächern regelmäßig einsetzen, was gerade die Verwendungsmöglichkeiten im Sprachunterricht, wo besonders viel Text gelesen wird, stark einschränkt. Hinzu kommt, dass die Präsentation einer Karten nur dann wirklich Sinn macht, wenn die Schülerinnen und Schüler deren Bedeutung tatsächlich erfassen. Das Lesen einer Karte ist jedoch eine Kulturtechnik, die erlernt werden muss und dabei von individuellen Merkmalen, wie etwa dem Vorwissen, der allgemeinen Intelligenz oder der Motivation abhängt. Das größte Problem ist jedoch, dass die theoretische Konzeption von Verdi und Kulhavy (2002) ein sehr eingeschränktes Konzept des Wissenserwerbs verfolgt. In ihrem Modell geht es ausschließlich um die Aufnahme und Wiedergabe (*retention*) von Informationen, während komplexe kognitive Prozesse, die ein wirkliches *Verstehen* der Inhalte erfordern, keine Rolle spielen. Aus einer pragmatischer Perspektive ist es jedoch überaus interessant, zu welchen mentalen Operationen das *image* einer Karte genutzt werden kann, die über den Aufbau referentieller Verknüpfungen hinausgehen. Kulhavy, Stock, Verdi et al. (1993) konzentrieren sich jedoch in erster Linie auf die invarianten (räumlichen) Eigenschaften der Kartenrepräsentation und vernachlässigen dabei Annahmen zu ihrer Funktionalität. Gerade der theoretische Vergleich zwischen dem *image* einer Karte und dem mentalen Modell von Johnson-Laird (1983) wäre überaus spannend gewesen.

Wie dem auch sei, die Autoren haben viele Annahmen des Modells empirisch gut belegt und gleichzeitig versucht, ihre Theorie in einen praktischen Kontext zu stellen. Außerdem spricht nur wenig dagegen, den Geltungsbereich etwas allgemeiner zu fassen und auf andere Bildtypen oder -arten zu übertragen. Trotz einiger Schwächen ist die CRT damit eine solide Theorie, auf deren Grundlage sich verschiedene Phänomene der Text-Bild-Verarbeitung plausibel erklären lassen. Dennoch ist das Modell im Vergleich zu den anderen theoretischen Arbeiten, die in diesem Kapitel angeführt werden, relativ unbekannt geblieben. Ganz anders verhält es sich mit der Theorie, der das folgende Kapitel gewidmet ist.

3.3.3 Kognitive Theorie des multimedialen Lernens

Das wahrscheinlich bekannteste Modell, das sich speziell auf den kombinierten Wissenserwerb mit Texten und Bildern konzentriert, ist die *kognitive Theorie des multimedialen Lernens* von Mayer (2005). Wie der Titel bereits vermuten lässt, ist die CTML (*cognitive theory of multimedia learning*) sehr pragmatisch ausgerichtet. So orientierte sich ihre Entwicklung nicht allein an den Kriterien der Überprüfbarkeit und empirischen oder theo-

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

retischen Plausibilität. Nach Ansicht des Autors sollte die Theorie auch anwendbar sein und einen direkten Beitrag zur Verbesserung bei der Gestaltung von Lernumgebungen leisten. Ein Grund für diese Anliegen ist die eingeschränkte Gültigkeit des *Multimedia-Prinzips*, das Mayer (1997) noch vor dem ersten vollständigen Entwurf seiner Konzeption eingeführt hat. Wie bereits einleitend erwähnt wurde, zieht die Kombination sprachlicher und grafischer Medien nicht zwangsläufig ein besseres Lernergebnis nach sich. Obwohl der positive Einfluss, den Bilder auf den Wissenserwerb haben können, gut dokumentiert ist, gibt es auch experimentelle Befunde, bei denen ganz andere Effekte beobachtet wurden (vgl. Kalyuga, Ayres, Chandler & Sweller, 2003).

Letztlich besitzt jeder, der Lernmaterialien gestalten oder entwerfen will, zumindest implizite Vorstellungen über die Funktionsweise des kognitiven Systems. Solche Annahmen sind jedoch oft naiv oder unzutreffend, was häufig dazu führt, dass bei der Kombination von Texten und Bildern grundlegende Fehler gemacht werden, die den Wissenserwerb eher behindern als fördern. Nach Ansicht von Mayer (2005) führt bei der Gestaltung effektiver Lernumgebungen daher kein Weg daran vorbei, kognitionspsychologische Forschungsergebnisse zu berücksichtigen. Dementsprechend beinhaltet auch die CTML drei grundlegende theoretische Annahmen über das kognitive System. Die erste dieser Annahmen besteht zunächst aus der einfachen Feststellung, dass Informationen in zwei unterschiedlichen Kanälen verarbeitet werden. In Anlehnung an die Theorien von Paivio (1986) bzw. Baddeley (1992) unterscheidet Mayer (2005) dabei zwischen einem *verbal-auditiven* und einem *visuell-piktoralen* Kanal. Die Unterschiede zwischen den beiden Kanälen betreffen sowohl die zugrunde liegende Sinnesmodalität (*Hören versus Sehen*), als auch das Format (*Text versus Bild*) der Informationen, wobei eine scharfe Trennung nicht immer möglich ist. Wird beispielsweise ein Text akustisch dargeboten, erfolgt dessen Verarbeitung in dem dafür vorgesehenen *verbal-auditiven* Kanal. Dagegen ist es beim Lesen erforderlich, dass die betreffenden Buchstaben und Worte zunächst visuell aufgenommen werden. Da im *visuell-piktoralen* Kanal eine Weiterverarbeitung sprachlicher Inhalte aber nicht vorgesehen ist, müssen die Informationen in das verbale System gelangen. Daher wird angenommen, dass sich die Ergebnisse der visuellen Enkodierungsprozesse in ein *verbal-auditives* Format konvertieren bzw. umwandeln lassen. So könnten Schüler beim Lesen eines Lehrbuchtextes dessen Sätze *innerlich* artikulieren und auf diese Weise die Informationen in den *verbal-auditiven* Kanal einspeisen.

Die zweite grundlegende Annahme der CTML betrifft den Umstand, dass die Menge an Informationen, die gleichzeitig in beiden Kanälen verarbeitet werden kann, grundsätzlich begrenzt ist. Beispielsweise kann selbst der aufmerksamste Zuhörer einer Vorlesung immer nur wenige Worte gleichzeitig im Arbeitsgedächtnis aufrecht erhalten. Auch wenn von dort (hoffentlich) viele Informationen in den Langzeitspeicher gelangen, so müssen die Inhalte dennoch portionsweise verarbeitet werden, was eine wortgetreue Repräsentation und Wiedergabe erschwert. Einen ähnlichen Zusammenhang nimmt Mayer (2005) für den *visuell-piktoralen* Kanal an. So geht er davon aus, dass nicht die vollständige Kopie eines Bildes, sondern immer nur wenige Bestandteile gleichzeitig verarbeitet werden können. Diese Ansicht steht in Widerspruch zur theoretischen Konzeption von (Kulhavy,

3 Verarbeitung von Texten und Bildern

Stock & Kealy, 1993), die im vorangegangenen Kapitel erläutert wurde. Demnach werden Bilder nicht als intakte *images* kodiert und müssen ähnlich wie Sätze oder Worte zunächst zerlegt werden. Als Konsequenz besitzen geographische Karten innerhalb der CTML keine Kapazitätsvorteile und belasten das Arbeitsgedächtnis in gleicher Weise, wie dies Worte bzw. Propositionen tun würden. Die Vorstellung von einer begrenzten Verarbeitungskapazität des kognitiven Systems besitzt in der psychologischen Forschung eine lange Tradition, wobei sich Mayer (2005) vor allem auf die Arbeiten von Baddeley (1992) und Chandler und Sweller (1991); Sweller et al. (1998) bezieht. Zudem nimmt er an, dass sich die Anzahl der Informationselemente, die gleichzeitig im Arbeitgedächtnis gehalten werden kann, zwischen 5 und 7 *chunks* bewegt (vgl. Miller, 1956; Simon, 1974). Zwar lassen sich diese *chunks* durch Übung inhaltlich anreichern, doch ändert das nichts an der grundsätzlichen Begrenzung der Arbeitsgedächtnisspanne. Diese Einschränkungen zwingen Lernende dazu, ihr kognitiven Ressourcen gezielt und sparsam einzusetzen. Damit ist unter anderem auch die Entscheidung verbunden, welchen Informationen Aufmerksamkeit geschenkt werden soll und welchen eher nicht.

Die dritte und letzte Annahme der CTML knüpft an diese Auffassung an und betont die Bedeutung des Individuums innerhalb des theoretischen Modells. So ist der Wissenserwerb nach Ansicht von Mayer (2005) in erster Linie ein aktiver Prozess, der den Aufbau einer kohärenten mentalen Repräsentation zum Ziel hat. Lernende sind demnach keine Gefäße, in die Wissen unbegrenzt *hineingeschüttet* und konserviert werden kann. Bedeutungen und Zusammenhänge müssen vielmehr durch eine bewusste Auseinandersetzung des Individuums mit den entsprechenden Sachverhalten konstruiert werden. Dieser Konstruktionsprozess lässt sich auch als Modellbildung ansehen, da die resultierenden Repräsentationen häufig strukturelle Eigenschaften umfassen, die einen funktionalen Charakter besitzen. Beispielsweise wird ein Sachverhalt nur dann wirklich verstanden, wenn es dem Lernenden gelingt, kausale Zusammenhänge und Prozessabläufe mental zu modellieren. Dabei ist es überaus hilfreich, wenn schon die präsentierten Lehrmaterialien eine kohärente Struktur aufweisen und den Aufbau einer entsprechenden mentalen Repräsentation anleiten.

Sicherlich erleichtern adäquat gestaltete Materialien das Lernen neuer Inhalte, doch die eigentliche Arbeit kann dem Individuum damit nicht abgenommen werden. Daher ist der aktive Wissenserwerb durch Prozesse der Informationsselektion, -organisation und -integration gekennzeichnet (Mayer, 1996). Meistens besteht die erste Aufgabe des Lernenden darin, relevante Informationen zu selektieren und einer bewussten Verarbeitung zugänglich zu machen. Welche Inhalte relevant sind und welche nicht, hängt von den individuellen Lernzielen ab und erfordert mehr oder weniger bewusste Entscheidungen. Sind die betreffenden Informationen einmal ausgewählt, können sie innerhalb des Arbeitsgedächtnisses verschiedenen Organisationsprozessen unterzogen werden. Dazu zählen neben der Knüpfung von strukturellen Relationen, vor allem Prozesse der Kategorisierung und Generalisierung. Um den Lernprozess abzuschließen, müssen die Informationen zu guter Letzt in bestehende Wissenstrukturen integriert werden. Dieser Vorgang umfasst die Aktivierung und den Transfer von relevanten Wissensseinheiten aus dem Langzeitspei-

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

cher in das Arbeitgedächtnis, wo der Lernende neue und alte Informationen miteinander verknüpfen kann.

Die Annahme von zwei Sinneskanälen, einer begrenzten Verarbeitungskapazität und einer aktiven Rolle der Lernenden bilden das Grundgerüst der CTML, wobei (Mayer, 2005) noch weitere Aspekte in seine Modell integriert, wie Abbildung 3.5 zeigt. Die vier-

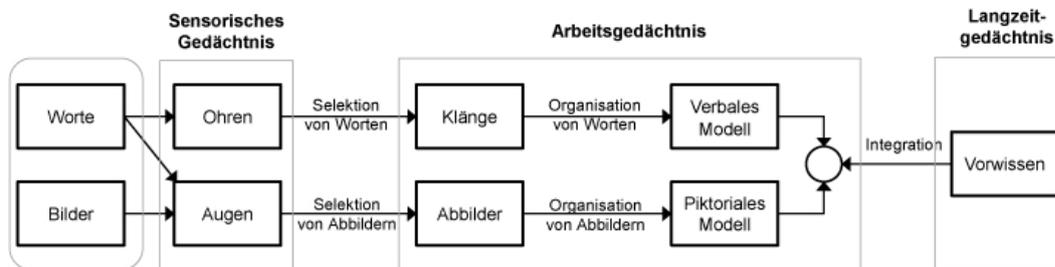


Abbildung 3.5: Kognitive Theorie des multimedialen Lernens (nach Mayer, 2005)

eckigen Kästchen stehen für die verschiedenen Gedächtniskomponenten des kognitiven Systems, während der Kasten mit den abgerundeten Ecken das Lehrmaterial repräsentiert. So werden Bilder und Texte zunächst wahrgenommen und gelangen in den sensorischen Speicher, wo eine exakte, perceptionsnahe Kopie für sehr kurze Zeit aufrecht erhalten wird (siehe auch Atkinson & Shiffrin, 1968). Die beiden Pfeile, die von dem Wort-Kästchen ausgehen, verdeutlichen dabei, dass neben einer akustischen Wahrnehmung verbale Informationen auch gelesen werden können. Sprachliche Inhalte, die in schriftlicher Form vorliegen, gelangen dann über den visuellen Kanal des sensorischen Speichers in das Arbeitsgedächtnis. Welche Informationen dort weiterverarbeitet werden, ist das Ergebnis der bereits erwähnten Selektionsprozesse, die weitgehend bewusstseinspflichtig sind. Der Lernende richtet dabei seine Aufmerksamkeit auf für ihn relevante Worte oder Bildbestandteile, und transferiert diese in das Arbeitsgedächtnis. Dort werden die noch wahrnehmungsnahen Repräsentationsformate durch Organisationsprozesse in kohärente Wissenstrukturen umgewandelt, die Mayer (2005) als *verbales* bzw. als *piktoriales Modell* bezeichnet.

Der sicherlich bedeutendste Vorgang beim multimedialen Lernen besteht jedoch in der Integration der sprach- und bildbasierten Repräsentationen. Dabei werden aus dem verbalen und dem piktorialen Modell eine integrierte mentale Repräsentation geformt, die aus referentielle Verknüpfungen zwischen korrespondierenden Elementen und Relationen besteht. In Abbildung 3.5 ist zu erkennen, dass bei diesem Prozess auch Vorwissen aus dem Langzeitgedächtnis herangezogen wird, um ein möglichst kohärentes Modell aufzubauen. Die Verknüpfungen zwischen beiden Repräsentationsformaten und den Wissenstrukturen des Langzeitspeichers findet im verbalen und visuellen Teil des Arbeitgedächtnisses statt, was eine aufwendige Koordination erfordert. Aus diesem Grund kann der Integrationsprozess nur dann gelingen, wenn die die knappen kognitiven Ressourcen effizient und zielführend eingesetzt werden. Nach Ansicht von Mayer (2005) spielt dabei das Vorwissen eine besondere Rolle, da vorhandene Wissenstrukturen den Aufbau einer

3 Verarbeitung von Texten und Bildern

integrierten Repräsentation begünstigen.

Hinsichtlich der Repräsentationsformate, die den Prozessen des multimedialen Lernens zugrunde liegen, macht die CTML keine genauen Angaben. Zwar wird davon ausgegangen, dass die Inhalte des sensorischen Speichers mehr oder weniger exakte Kopien des Textes bzw. des Bildes darstellen, doch welche Eigenschaften diese besitzen, bleibt weitgehend unklar. Gleiches gilt für die Repräsentationsformate im Arbeitsgedächtnis, deren Charakteristika ebenfalls nicht näher spezifiziert werden. Eine Ausnahme sind in diesem Zusammenhang die Wissenstrukturen des Langzeitspeichers, von denen Mayer (2005) in Anlehnung an Sweller, van Merriënboer und Paas (1998) annimmt, dass sie aus *Schemata* bestehen. Auch grafisch unterstreicht Abbildung 3.5, dass Bilder und Texte parallel verarbeitet werden, wobei keines der beiden Formate einen besondere Eigenheit aufweist. Der einzige Unterschied besteht lediglich darin, dass die entsprechenden Prozesse in modalitätsspezifischen Kanälen bzw. Systemen des kognitiven Apparates stattfinden.

Viel wichtiger als die Frage nach dem Format der mentalen Bild- und Textrepräsentationen, scheinen dem Autor die praktischen Konsequenzen zu sein, die sich anhand des Modells ableiten lassen. Auf der Grundlage der CTML und zahlreichen experimentellen Befunden hat Mayer (2001) daher insgesamt sieben Prinzipien formuliert, die bei der Gestaltung von multimedialen Lernumgebungen und -materialien zu beachten sind. Nicht alle dieser Grundsätze sind für die Fragestellung der vorliegenden Arbeit von der gleichen Bedeutung, weshalb an dieser Stelle nur ein Teil näher dargestellt wird. Bei dem ersten Gestaltungsprinzipien handelt es sich um das bereits erwähnte *multimedia principle*, welches die Verwendung von Text-Bild-Kombinationen gegenüber Texten ohne Illustration favorisiert. Den Vorteil einer gemeinsamen Verwendung von Texten und Bildern erklärt die CTML in erster Linie durch die Nutzung von zwei Kanälen bzw. Systemen des Arbeitsgedächtnisses. Allerdings geht die oftmals beobachteten besseren Lernleistung nicht allein auf doppelt kodierte Informationen zurück, wie es etwa die DCT erklären würde (vgl. Pavió, 1986). Das entscheidende beim multimedialen Lernen ist, dass bildliche und verbale Inhalte herangezogen werden, um eine integrierte, kohärente Repräsentation aufzubauen. Auf diese Weise resultiert ein Modell, das im Idealfall sehr viel *reichhaltiger* ist, als wenn es allein auf der Basis eines Textes entstanden wäre. Gelingt es dem Lernenden das kohärente Modell eines Sachverhaltes zu konstruieren und es in bestehende Wissensstrukturen zu integrieren, dann hat er den entsprechenden Sachverhalt auch wirklich verstanden. Verstehen, wie es Mayer und Moreno (2003) definieren, besteht jedoch nicht darin, gelernte Informationen oder Inhalte einfach wiederzugeben. Im Sinne der CTML kann man von bedeutungsvollem Lernen oder tiefem Verstehen (*deep understanding*) sprechen, wenn der Lernende sein erworbenes Wissen auf neue Zusammenhänge und Problemstellungen transferieren kann.

Die Herausforderung besteht also darin, kohärente Repräsentation aufzubauen, obwohl die Integrationsprozesse sehr aufwendig und die mentalen Ressourcen begrenzt sind. Damit dies gelingen kann, sollte das multimediale Lehr/Lernmaterial aber bestimmte Eigenschaften erfüllen. Beispielsweise müssen nach dem Prinzip der *spatial contiguity*

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

korrespondierende Bild- und Textinformationen räumlich möglichst nah präsentiert werden. Auf diese Weise entfallen Ressourcen zehrende Such- und Scanningprozesse, die bei einer getrennten Präsentation anfallen würden. Somit ist der Lernende nicht dazu gezwungen seine Aufmerksamkeit zwischen unterschiedlichen Medien zu teilen, und es kommt zu keinem *split attention* Effekt (Mayer & Moreno, 1998). Ähnliches gilt für das Prinzip der *temporal contiguity*, wobei hier die gleichzeitige Präsentation von Text und Bild angemahnt wird. Dagegen bedeutet es eine zusätzliche Belastung der Kapazität, wenn Lernende dazu gezwungen sind, die beiden Formate sequentiell zu lernen. Um bei einer sukzessiven Präsentation verbale und bildliche Informationen zu integrieren, muss die Repräsentation des ersten Formates während der Verarbeitung des zweiten Formats im Arbeitsgedächtnis aufrecht erhalten werden. Dabei können die vorhandenen Ressourcen schnell aufgebraucht sein, so dass erforderliche Selektions-, Organisations- und Integrationsprozesse nicht mehr möglich sind.

Damit die vorhandenen Kapazitäten generell geschont werden, empfiehlt das *modality* Prinzip eine kombinierte Nutzung des verbal-auditiven und des visuell-piktorialen Kanals. Sprachliche und grafische Informationen lassen sich demnach am besten verarbeiten, wenn die Präsentation statischer oder bewegter Bilder von akustisch dargebotenem Text begleitet wird. Der Lernende muss sich dann nicht mehr entscheiden, ob er lieber den Text lesen oder die Illustration betrachten will, er kann beides mehr oder weniger gleichzeitig tun. Die Verarbeitung der verbalen und piktorialen Informationen erfolgt dabei weitgehend parallel, ohne dass es Konflikte zwischen den beteiligten Systeme geben sollte. Durch die simultane Inanspruchnahme beider Sinneskanäle wird die kognitive Belastung optimal auf die verbalen und piktorialen Bereiche des Arbeitsgedächtnisses verteilt. Auf diese Weise bleiben im Idealfall genügend mentale Ressourcen übrig, die für Prozesse der Modellkonstruktion und -integration genutzt werden können.

Ob die bereits beschriebene Prinzipien tatsächlich greifen, hängt jedoch nicht allein von der Gestaltung des Lehr/Lernmaterials ab. Eine ebenso wichtige Größe, die den Erfolg des multimedialer Instruktionen stark beeinflussen kann, ist der Lernende selbst, bzw. seine Eigenschaften und Fähigkeiten. Mayer (2001) spricht in diesem Zusammenhang von dem *individual differences* Prinzip, das ähnlich wie die anderen Grundsätze berücksichtigt werden muss. Die wahrscheinlich bedeutsamste Lernervariabel ist das Vorwissen, da es den Ausgangspunkt für die meisten Koordinations- und Integrationsprozesse darstellt (Kalyuga, 2005). Verfügt ein Individuum über keine oder zu wenige relevante Wissenstrukturen, wird der Aufbau einer kohärenten mentalen Repräsentation mit hoher Wahrscheinlichkeit misslingen. In der Regel ist es also notwendig, dass Lernende bei der Modellkonstruktion zumindest auf elementare Kenntnisse zurückgreifen können, da sonst grundlegende Prozess nicht umgesetzt werden können.

Damit ist das Vorwissen sehr wahrscheinlich die wichtigste, jedoch nicht die einzige Lernervariable, welche den Wissenserwerb mit multiplen Repräsentationen beeinflusst. Beispielsweise haben Mayer und Sims (1994) in einer Untersuchung feststellen können, dass bei Individuen mit ausgeprägten räumlichen Fähigkeiten ein starker *temporal contiguity* Effekt zu beobachten ist. Für Personen, deren Fähigkeiten eher schwach ausge-

3 Verarbeitung von Texten und Bildern

prägt sind, ist es dagegen fast unerheblich, ob Bild und Text gleichzeitig oder sukzessive gelernt werden. Nach Meinung von Mayer und Moreno (2003) hängen viele grundlegenden Prozesse, die mit der Verarbeitung von visuellen Informationen zusammenhängen, von dem räumlichen Denkvermögen ab. So muss das Individuum generell in der Lage sein, die Repräsentation eines Bildes lange genug im Arbeitsgedächtnis zu halten, um Verknüpfungen mit sprachlichen Informationen oder den Inhalten des Langzeitspeichers zu etablieren. Je stärker die räumlichen Fähigkeiten ausgeprägt sind, umso weniger kognitive Ressourcen nehmen diese grundlegenden Vorgänge dabei in Anspruch. Die verbleibenden Kapazitäten können dann für den Aufbau eines integrierten Modells oder andere Prozesse des bedeutungsvollen Lernens verwendet werden.

Auch wenn sich die meisten Studien auf das Vorwissen oder kognitive Fähigkeiten beschränken, ist es durchaus wahrscheinlich, dass noch weitere Lernermerkmale den Wissenserwerb mit multiplen Repräsentationsformaten beeinflussen. Denkbare Variablen, die Effekte verstärken oder abschwächen können, sind die Ausgangsmotivation, Modalitätspräferenzen oder spezifische Einstellungen gegenüber bestimmten Themen. Daher lässt sich recht allgemein aus dem Prinzip der *individual differences* die Konsequenz ableiten, dass die Lernermerkmale auch im Kontext des multimedialen Lernens berücksichtigt werden sollten. Im Grund geht es um eine möglichst gute Passung zwischen den Voraussetzungen des Individuums und den Anforderungen, die mit dem Lehrmaterial verbunden sind. Diese Vorstellung weist Ähnlichkeiten mit der *Zone der nächsten Entwicklung* auf, da auch bei Wygotski (1964) die Abstimmung von Anforderung und Entwicklungsstand ein wichtiger Bestandteil pädagogischen Handelns ist.

Abschließen betrachtet, ist die CTML von Mayer (2005) ein theoretischer Ansatz, der sowohl durch Einfachheit, als auch durch seine pragmatische Orientierung besticht. Aufgrund des ausgeprägten Bezugs auf Konzepte der kognitiven Psychologie und einer starken empirischen Fundierung besitzt das Modell zudem eine beachtliche Plausibilität. Allerdings ist keine Theorie perfekt und auch die CTML gibt genügend Anlass zu kritischen Betrachtungen. Insbesondere die Parallelität der Text und Bildverarbeitung wird von Schnotz und Bannert (1999) als problematisch angesehen, da sprachliche und visuelle Informationen auf unterschiedlichen Zeichensystemen und Repräsentationsprinzipien beruhen. Diese Kritik bezieht sich zwar auf eine frühere Version (vgl. Mayer, 1997), doch auch die aktuelle Fassung der CTML hält an der Annahme einer parallelen Verarbeitung von sprachlichen und visuellen Informationen fest. Bis auf die Tatsache, dass die entsprechenden Prozesse in unterschiedlichen Kanälen bzw. kognitiven Systemen des Arbeitsgedächtnisses ablaufen, sind sprachliche und visuelle Informationen damit weitgehend austauschbar. Diese Vereinfachung ist jedoch nicht wirklich überzeugend, da sprachliche und visuelle Medien letztlich zu verschieden sind.

Ein anderes Problem betrifft die praktische Relevanz des theoretischen Modells und der damit verbundenen Prinzipien des multimedialen Lernens. Alle Experimente, die zur Konzeption der CTML beigetragen haben, wurden mit sehr sorgfältig entwickelten Lehr-Lernmaterialien durchgeführt. Das ist alles andere als ein Fehler, doch können experimentelle Manipulationen die Generalisierbarkeit der Ergebnisse einschränken. Beispiels-

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

weise verwendete Mayer, Heiser und Lonn (2001) eine Animation über die Entstehung von Blitzen, die sich aus bewegten Bildern und verbalen Kommentaren zusammensetzte. Durch die gekonnte Abstimmung und Segmentierung, wurde ein optimales Gleichgewicht zwischen den visuellen und sprachlichen Informationen hergestellt. Da pro Bild bzw. Animationsabschnitt in der Regel nicht mehr als ein Satz präsentiert wurde, dürfte sich auch die kognitive Belastung in Grenzen gehalten haben. Aus der Perspektive eines wissenschaftlichen Psychologen ist es zwar überaus beeindruckend, dass Mayer (2001) mit seiner Animation einen Modalitätseffekt nachweisen konnte, doch sind die praktischen Implikationen dieses Befundes vergleichsweise gering. Allein der Aufwand für die Entwicklung solch kleinteiliger Lernumgebungen ist immens, so dass sich automatisch die Frage nach einer Kosten-Nutzen-Abwägung stellt. Bei umfangreichen und/oder komplexen Sachverhalten ist es zudem gar nicht möglich, für alle Textabschnitte ein äquivalente Visualisierung zu finden oder zu generieren. Ebenso erschweren abstrakte Inhalte und Themen eine weitreichende Umsetzung des *multimedia* bzw. des *modality* Prinzips.

Ein weiterer, etwas unklarer Aspekt betrifft die Frage, inwieweit einzelne Grundsätze oder Effekte an bestimmte inhaltliche Domänen gebunden sind. Beim Vorwissens ist ein solcher Zusammenhang offensichtlich, doch scheint auch die Ausprägung des räumlichen Denkvermögens einen starken Aufgabenbezug zu besitzen, wie Sims und Mayer (2002) feststellen konnten. Damit ist es jedoch nicht länger möglich, den *temporal contiguity* Effekt allein mit allgemeinen räumlichen Fähigkeiten zu erklären, auch wenn dies bei der Verarbeitung von visuellen Informationen nahe liegt. Dieser Einwand ändert allerdings nichts an der grundsätzlichen Bedeutung der räumlichen und zeitlichen Kontiguität für die Gestaltung von illustrierten Texten. So lässt sich kaum bezweifeln, dass Text-Bild-Scheren oder eine sequentielle Präsentation integrative Verarbeitungsprozesse erschweren, was zu schlechteren Leistungen führen kann. Ebenso wenig verliert die Berücksichtigung der Zielgruppemerkmale an Relevanz, falls man eine Über- oder Unterforderung der Lernenden verhindern will.

Es ist daher sicherlich nicht übertrieben, wenn man die Kognitive Theorie des multimedialen Lernens von Mayer (2005) als einen Meilenstein in der Pädagogischen Psychologie bezeichnet. Einzelne Aspekte mögen zwar umstritten oder zu einfach ausgefallen sein, was eine Übersetzung des kognitionspsychologischen Modells in konkrete Handlungsanweisungen schwieriger macht, als es Mayer's Prinzipien auf den ersten Blick suggerieren. Dennoch war und ist die CTML Grundlage für unzählige Untersuchungen und Forschungsarbeiten, die auf eine Förderung des bedeutungsvollen Lernens abzielen. Das Verstehen von Sachverhalten steht auch im Mittelpunkt des letzten theoretischen Ansatzes, der im folgenden Abschnitt vorgestellt wird und das Kapitel beschließt.

3.3.4 Das integrierte Modell des Text- und Bildverstehens

In vielen früheren Untersuchungen war der Wissenserwerb mit Texten und Bildern vor allem dann erfolgreich, wenn es den Versuchspersonen gelang, möglichst viele der zuvor gelernten Inhalte fehlerfrei wieder zu geben (vgl. Paivio, 1986; Kulhavy, Stock & Kealy,

3 Verarbeitung von Texten und Bildern

1993). Lernen wurde damit in erster Linie als ein rezeptiver Vorgang aufgefasst, was sich auch in den entsprechenden theoretischen Modellen ausdrückt, die primär die Speicherung und den Abruf von Informationen thematisieren. Verstehen, im Sinne des Aufbaus kohärenter mentaler Repräsentationen, wurde dagegen lange Zeit kaum berücksichtigt. Das änderte sich im Zuge neuerer lernpsychologischer Ansätze, die den Wissenserwerb als einen konstruktiven Prozess ansehen. Demnach ist das lernende Individuum nicht länger ein passives Gefäß, in das Informationen hinein gefüllt werden. Beim Wissenserwerb müssen Menschen die benötigten Informationen aktiv suchen, diese evaluieren und selektieren, damit unter Rückgriff auf das Vorwissen Bedeutungen konstruiert werden können (Wittrock, 1989).

Einer der ersten, der diese konstruktivistische Sichtweise auf den Wissenserwerb mit Texten und Bildern übertrug, war Mayer (1997), dessen kognitive Theorie des multimedialen Lernens im Kapitel zuvor diskutiert wurde. Obwohl der Ansatz überaus populär ist, kritisierten Schnotz und Bannert (1999) schon früh die unspezifische Parallelität der Text- und Bildbearbeitung (siehe Seite 76). Die beiden Autoren beließen es jedoch nicht bei dieser Kritik, sondern entwickelten ein alternatives theoretisches Konzept, das die unterschiedlichen Repräsentationsprinzipien sprachlicher und visueller Informationen sehr viel stärker berücksichtigt. Wie die Bezeichnung bereits verrät, geht es auch beim *integrierten Modell des Text- und Bildverstehens* nicht um die Erklärung einfacher, mnemonischer Phänomene, ganz im Gegenteil. Im Zentrum steht der Aufbau kohärenter mentaler Repräsentationen, wobei die Verarbeitung sprachlicher und bildlicher Informationen sehr differenziert betrachtet wird. Abbildung 3.6 gibt das Modell grafisch wieder.

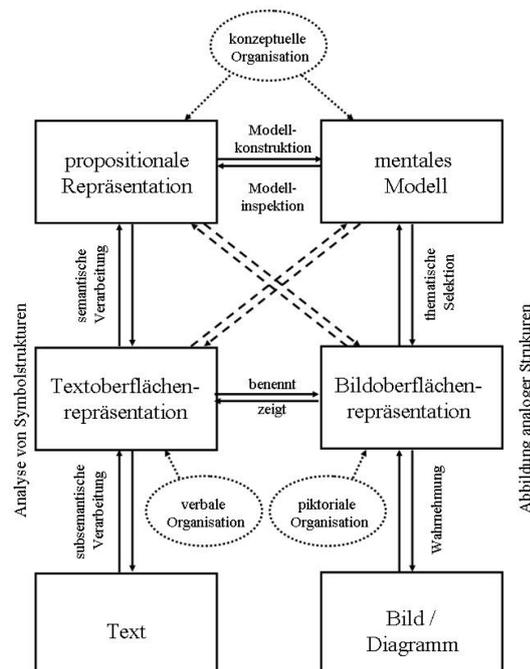


Abbildung 3.6: Modell des integrierten Text und Bildverstehens (nach Schnotz und Bannert, 2003)

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

Auf den ersten Blick scheint auch das Modell von Schnotz und Bannert (1999) zwei modalitätsspezifische Kanäle zu besitzen, doch dieser Eindruck täuscht. Anders als bei Mayer (2005) handelt es sich bei dieser frühen Version um *Repräsentationszweige*, die nicht ausdrücklich an eine spezifische Sinnesmodalität gebunden sind. Der linke, deskriptionale Zweig umfasst den Text, die mentale Repräsentation der Textoberflächenstruktur und die ebenfalls mentale propositionale Repräsentation, die den semantischen Inhalt bzw. die Bedeutung des Textes wiedergibt. Alle Prozesse, die von dem geschriebenen Wort bis hin zur Proposition führen, beruhen dabei auf der Analyse und Verarbeitung von Symbolstrukturen. Dagegen beinhaltet der depiktionale Zweig das externe Bild bzw. Diagramm, dessen erste wahrnehmungsnahere Repräsentation und schließlich ein mentales Modell, das strukturelle Eigenschaften des visuellen Sachverhaltes widerspiegelt. Der Aufbau der Repräsentationen des depiktionalen Zweiges erfolgt anhand von Prozessen der analogen Strukturabbildung, wie sie etwa Gentner (1983) annimmt.

Dabei ist jedoch zu beachten, dass die propositionale Repräsentation des Textes und das mentale Modell in der Konzeption von Schnotz und Bannert (1999) keineswegs das Endprodukt ihres jeweiligen Repräsentationszweiges darstellen. Wie in Abbildung 3.6 zu erkennen ist, führt die Verarbeitung sprachlicher Informationen zunächst zu einer Repräsentation der Textoberflächenstruktur, die das Ergebnis subsemantischer Analysen und verbaler Organisationsprozesse ist. Auf dieser Ebene werden nur syntaktische und morphologische Charakteristika des Textes mental abgebildet, aber noch keine Bedeutungen zugewiesen oder generiert. Eine Repräsentation des semantischen Gehalts entsteht erst durch konzeptuelle Organisationsprozesse, welche die Textoberfläche zu einem propositionalen Gefüge weiterverarbeiten. Dieses propositionale Gefüge bildet dann die *Textbasis*, auf deren Grundlage es dann möglich ist, ein mentales Modell der beschriebenen Sachverhalte zu konstruieren (van Dijk & Kintsch, 1983; Schnotz, 1994a). Obwohl das Vorwissen in Abbildung 3.6 nicht gesondert angeführt wird, erfolgen sämtliche Konstruktionsprozesse vor dem Hintergrund auf- und absteigender Schemaaktivierungen.

Ähnlich verhält es sich mit der Verarbeitung von visuellen Inhalten. So selektiert der Betrachter eines Bildes zunächst anhand absteigender Schemaktivierungen einzelne Informationen, die durch automatisierte visuelle Routinen enkodiert werden (Ullman, 1984). Zentrale Bestandteile dieser Enkodierungsvorgänge sind Prozesse, bei denen einzelne grafische Elemente diskriminiert und identifiziert werden, um sie anschließend in übergeordneten Strukturen zu organisieren (Wertheimer, 1923; Winn, 1994). Auf diese Weise entsteht eine mentale Repräsentation, die Schnotz und Bannert (1999) als *visuelle Wahrnehmung* oder als *Vorstellung* bezeichnen. Mit dem zweiten Begriff berücksichtigen die beiden Autoren den Umstand, dass Lernende in der Lage sind, eine perzeptionsnahe mentale Repräsentation allein aufgrund vorhandener Wissenstrukturen zu erzeugen. Da angenommen wird, dass visuellen Wahrnehmungen und Vorstellungen vergleichbare kognitive Mechanismen zugrunde liegen, dürften die beiden Formate keine gravierenden Unterschiede aufweisen (Kosslyn, 1994).

Auf der Ebene der visuellen Wahrnehmung bzw. Vorstellungen werden vor allem perzeptive Strukturen, aber noch keine wirklichen Bedeutungen repräsentiert. Daher erfor-

3 Verarbeitung von Texten und Bildern

dert echtes Verstehen semantische Verarbeitungsprozesse, welche die wahrnehmungsnahe Repräsentation in ein mentales Modell überführen. Dies geschieht, indem Elemente und räumliche Relationen der visuellen Wahrnehmung auf mentale Elemente und semantische Relationen abgebildet werden. Nach Ansicht von Schnotz und Bannert (1999) ist Bildverstehen damit ein Prozess der schemageleiteten analogen Strukturabbildung, bei dem ein System visuell-räumlicher Relationen auf ein System von semantischen Relationen übertragen wird. Im Grunde wird eine oberflächliche Wahrnehmung eines Bildes bzw. ein Ausschnitt daraus abstrahiert. Die entstehenden mentalen Modelle besitzen deshalb nur noch strukturelle oder funktionale Ähnlichkeiten mit den Sachverhalten, die sie repräsentieren. Welche Strukturen genau abgebildet werden, hängt von den Anforderungen ab, zu deren Bewältigung das jeweilige Modell konstruiert wurde.

Auch hinsichtlich des Informationsgehaltes unterscheiden sich mentale Modelle von den visuellen Wahrnehmungen und Vorstellungen. Neben der funktionalen Abstraktion, bei der die meisten perceptiven Details verloren gehen dürften, erfordert der Konstruktionsprozess einen Rückgriff auf bereits vorhandene Wissenstrukturen. Anhand von Schemata aus dem Langzeitgedächtnis wird das mentale Modell elaboriert, so dass es letztlich mehr oder zumindest andere Informationen umfassen kann, als die ursprüngliche Wahrnehmung. Hinzu kommt, dass durch Prozesse der Modellsinspektion neue Informationen generiert bzw. abgelesen werden können. Dies geschieht beispielsweise dann, wenn ein Lernender Schlussfolgerungen zieht oder eine Abfolge von kausalen Zusammenhängen vor seinem inneren Auge simuliert, um ein Ergebnis vorherzusagen. Schnotz und Bannert (1999) nehmen an, dass die Ergebnisse der Modellinspektion als Propositionen enkodiert werden. Als Bestandteil der propositionalen Repräsentation, lassen sich die neu generierten Informationen schließlich dazu verwenden, das mentale Modell weiter zu verbessern. Auf diese Weise entsteht eine Art Kreislauf von Prozessen der Modellkonstruktion und -inspektion, der vor allem der Kohärenzentwicklung dient.

Bilder und Texte führen damit sowohl zu einer deskriptionalen als auch zu einer de-piktionalen internen Repräsentation, wobei allein die Ausgangspunkte verschieden sind. Dennoch steht das integrierte Modell des Text- und Bildverstehens nicht in der Tradition der dualen Kodierung, auf die sich etwa Mayer (2005) bezieht. Innerhalb der Konzeption von Schnotz und Bannert (1999) resultiert die Verarbeitung sprachlicher Informationen auch dann in einer doppelten Enkodierung, wenn zuvor keine explizite Aufforderung erfolgte, sich die verbalen Inhalte bildlich vorzustellen. Darüber hinaus ist ein mentales Modell aber sehr viel mehr als eine zusätzliche (visuelle) Repräsentation, die für Verknüpfungen mit verbalen Inhalten genutzt werden kann. Der Unterschied zu den *imaginen* von Paivio (1986) ist dabei vor allem qualitativer Natur. So spiegelt ein mentales Modell die strukturellen und funktionalen Charakteristika eines Sachverhaltes wieder, die das Ergebnis anforderungsabhängiger Selektionsprozesse sind. Auf diese Weise wird es möglich, Wissen, das im Zusammenhang mit einer bestimmten Aufgabe erworben wurde, auf andere Problemstellungen zu übertragen, die eine ähnlich Anforderungsstruktur besitzen. Die Fähigkeit zu einem solchen Transfer wird häufig auch als Zeichen für *bedeutungsvolles Lernen* oder *Verstehen* angesehen. Da der Wissenstransfer nur stattfinden kann,

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

wenn zuvor eine entsprechend abstrakte Repräsentation konstruiert wurde, lässt sich Verstehen mit der Konstruktion mentaler Modelle gleichsetzen.

Lernende können jedoch durchaus in der Lage sein, Inhalte oder kognitive Operationen in einem engen Rahmen zu wiederholen bzw. anzuwenden, ohne die entsprechenden Sachverhalte auch wirklich verstanden zu haben. In diesem Fall ist der Aufbau eines kohärenten mentalen Modells gescheitert und das erworbene Wissen liegt nicht in der Form vor, die einen Transfer ermöglicht. Es ist vorstellbar, dass der Konstruktionsprozess misslingt, da erforderliche Vorkenntnisse oder bestimmte kognitive Fertigkeiten fehlen. Als Folge kommt der Lernenden trotz guten Willens über eine propositionale Repräsentation der Inhalte nicht hinaus. Auf dieser Grundlage ist dann zwar die Wiedergabe bestimmter Propositionen möglich, doch verhindert ein unvollendeter Modellaufbau alle Inspektionsprozesse, so dass keine Elaboration der Informationen stattfindet. Manche Autoren sprechen in diesem Zusammenhang auch von *trägem Wissen* (*inert knowledge*) (Renkl, Mandl & Gruber, 1996), dessen Entstehung sich mit dem integrierten Modell des Text-Bildverstehens sehr gut erklären lässt. Allerdings soll an dieser Stelle nicht der Eindruck entstehen, dass die propositionale Repräsentation eine Stufe der kognitiven Informationsverarbeitung ist, die mit bedeutungsvollem Lernen nichts zu tun hat. Auch wenn das Verstehen eines Sachverhaltes in erster Linie mit dem Aufbau des mentalen Modells assoziiert werden kann, sehen Schnotz und Bannert (1999) beide Formate als notwendig an, da sie sich gegenseitig ergänzen.

Diese theoretischen Überlegungen haben zur Folge, dass Lernende beim gemeinsamen Wissenserwerb mit Texten und Bildern zwei Informationsquellen komplementär nutzen können. Da es jedoch möglich ist, das mentale Modell sowohl anhand der verbalen als auch anhand der piktorialen Informationen zu konstruieren, müssen sie sich entscheiden, auf welches Format wann zurückgegriffen wird. Schnotz und Bannert (1999) gehen davon aus, dass der antizipierte kognitive Aufwand eine wichtige Rolle bei dieser Entscheidung spielt. Nach Ansicht der beiden Autoren ist die Belastung, die mit dem Aufbau eines mentalen Modells auf der Grundlage eines Textes verbunden ist, größer, als wenn die Konstruktion anhand eines informationsäquivalenten Bildes erfolgt. Textverstehen beinhaltet demnach die Umwandlung einer propositionalen Struktur, die auf symbolischen Repräsentationsprinzipien beruht, in das analoge Format eines mentalen Modells. Eine solche Umwandlung entfällt jedoch, wenn die Konstruktion anhand eines Bildes stattfindet, da hier ein analoges Format in ein anderes analoges Format überführt wird. Auf den ersten Blick scheint es daher sinnvoll zu sein, immer erst das Bild für den Aufbau des mentalen Modells zu nutzen. Anschließend verarbeitet verbale Informationen lassen sich dann dazu verwenden, die Repräsentation weiter auszubauen, zu spezifizieren oder zu korrigieren. Eine so nahe liegende und einfache Schlussfolgerung ist jedoch nicht möglich, da der Konstruktionsaufwand interindividuell sehr unterschiedlich sein kann. Wie viel Mühe die Konstruktion eines mentalen Modells einem Lernenden abverlangt, ergibt sich letztlich aus der Interaktion zwischen dem Lehrmaterial (Komplexität, Qualität etc.) und dem Individuum (Fähigkeiten, Vorwissen, etc.), weshalb keine allgemeingültigen Empfehlungen gegeben werden können..

3 Verarbeitung von Texten und Bildern

Ähnlich wie andere theoretische Konzeptionen hat sich auch der Ansatz von Schnotz und Bannert (1999) in den letzten Jahren weiterentwickelt. Abbildung 3.6 zeigt daher eine etwas überarbeitete Fassung, die jedoch nur leicht von dem ursprünglichen Modell abweicht. Das betrifft vor allem eine Reihe von Kreuz- und Querverbindungen, die (Schnotz & Bannert, 2003) nachträglich eingefügt haben und die durch Pfeile zwischen den verschiedenen Repräsentationsformaten angedeutet wird. Inhaltlich stellen diese Verbindungen keine grundlegende Änderung dar, vielmehr unterstreichen die Autoren nochmals den integrativen Charakters ihres Modells. Während die ursprüngliche Fassung vor allem die Interaktion zwischen der propositionalen Repräsentation und dem mentalen Modell hervorhob, berücksichtigt die neuere Version auch Wechselbeziehungen mit den Formaten der vorgelagerten Ebene. So ist es durchaus plausibel, dass parallel zum Aufbau einer visuellen Wahrnehmung eine Reihe von Propositionen gebildet wird, die den visuellen Inhalte eine explizite Bedeutung zuordnen und möglicherweise die Aufmerksamkeit lenken. Ebenso können mentale Modelle den Aufbau von Textoberflächenstrukturen anleiten, indem sie beispielsweise helfen, syntaktische Mehrdeutigkeiten auf der Grundlage des semantischen Gesamtzusammenhangs zu lösen. Die Pfeile, die auf solche Querverbindungen hinweisen, sind jedoch nur gestrichelt. Damit betonen Schnotz und Bannert (2003), dass die Modellkonstruktion und -inspektion auch in der neuen Fassung auf in erster Linie auf interaktiven Prozessen zwischen dem mentalen Modell und der propositionalen Repräsentation beruhen.

Sehr viel mehr Veränderungen brachten Modifikationen mit sich, die den ursprünglichen Ansatz zu dem Modell der *Integrated Comprehension of Text and Pictures* (kurz ICTP) erweiterten (Schnotz, 2005). Gegenüber den ersten beiden Fassungen wurden eine Reihe zusätzlicher Annahmen gemacht, die vor allem das Vorhandensein mehrerer Gedächtnissysteme (Atkinson & Shiffrin, 1968), ein Arbeitsgedächtnis mit begrenzter Kapazität (Baddeley, 1992) und die duale Kodierung von Informationen betrafen (Paivio, 1986). Abbildung 3.7 zeigt die aktuelle Version des Modells.

Wie man erkennen kann, unterscheidet die neue Fassung zunächst eine perzeptuelle von einer kognitiven Verarbeitungsebene. Der sensorische Speicher fungiert dabei als Schnittstelle, die Wahrnehmungen aus der Umwelt übersetzt und in das Arbeitsgedächtnis weiterleitet. Die verschiedenen Informationen können dabei über den auditiven oder über den visuellen Kanal in das kognitive System gelangen, wo dann die ersten bewussteinspflichtigen Prozesse ablaufen. Liest eine Lernender beispielsweise einen Text, so werden dessen Inhalte über den visuellen Kanal des sensorischen Speichers in das visuelle Arbeitsgedächtnis transferiert. Schnotz (2005) geht davon aus, dass die daraus resultierenden Informationsmuster einer Repräsentation der Textoberflächenstruktur entsprechen. Durch Filter, die in Abbildung 3.7 als kleine schwarze Dreiecke dargestellt sind, werden dem visuellen Arbeitsgedächtnis verbale Informationen entnommen und in den propositionalen Teil des Arbeitsgedächtnisses weitergeleitet. Vor dem Hintergrund bestehender Wissenstrukturen aus dem Langzeitspeicher erfolgt dort eine Analyse der Inhalte, was letztlich den Aufbau einer propositionalen Struktur zur Folge haben kann, die ihrerseits die Konstruktion eines mentalen Modells anregt.

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

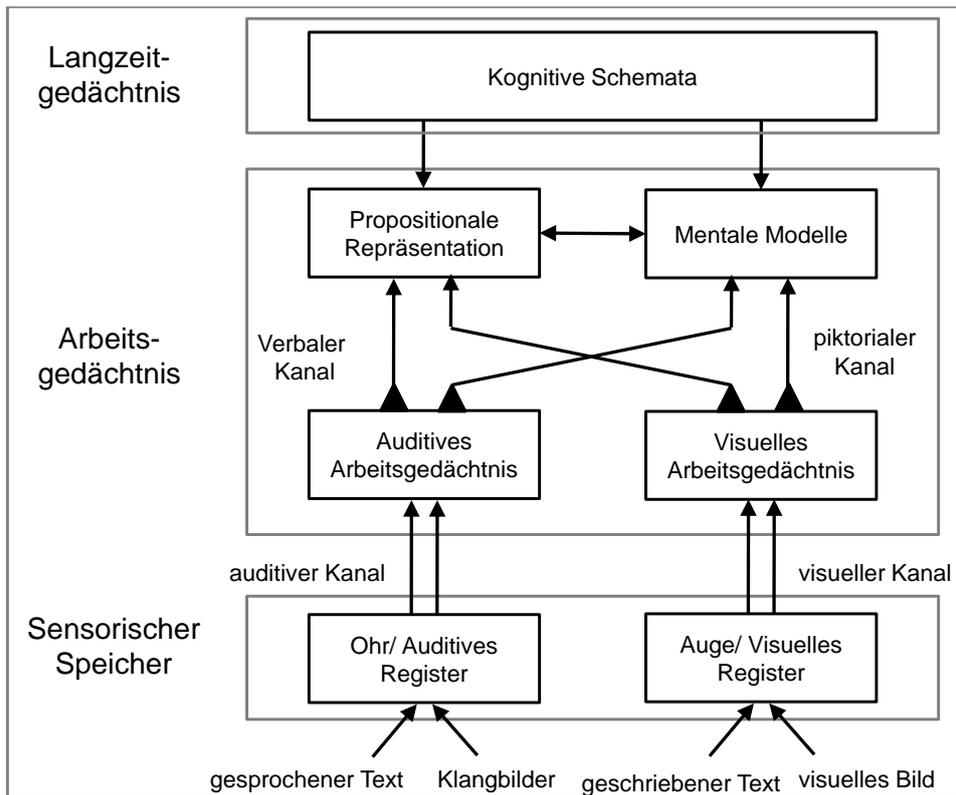


Abbildung 3.7: Das integrierte Modell des Text und Bildverstehens nach Schnotz (2005)

Mit der Verarbeitung von gesprochenem Worten verhält es sich recht ähnlich, nur gelangen die verbalen Informationen in diesem Fall über den auditiven Kanal in das kognitive System. Zudem befindet sich die Textoberflächenstruktur, auf deren Grundlage letztlich die propositionale Repräsentation entsteht, im auditiven Teil des Arbeitsgedächtnisses. Betrachtet ein Lernender hingegen ein Bild, dann werden die piktorialen Informationen durch den visuellen Kanal aufgenommen und in den visuellen Teil des Arbeitsgedächtnisses weitergeleitet. Das resultierende Repräsentationsformat ist dann die bereits erwähnte *visuelle Wahrnehmung*, deren Informationen gefiltert und selektiert werden müssen, bevor sie sich für den Aufbau eines mentalen Modells verwenden lassen. Ist es dem Individuum schließlich gelungen, ein kohärentes mentales Modell eines Sachverhaltes zu konstruieren, dann können im Zuge von Inspektionsprozessen zusätzlich Propositionen gebildet werden.

Besonders interessant an dem ICTP-Modell ist der Umstand, dass auch sogenannte *Klangbilder* (sound images) zur Konstruktion eines mentalen Modells führen. Demnach gelangen non-verbale Laute zunächst in das auditive Arbeitsgedächtnis, wo eine perzeptionsnahe Repräsentation für weitere Verarbeitungsschritte kurzfristig verfügbar gehalten wird. Wiederum setzen Filter- und Selektionsprozesse ein, deren Ziel es ist, anforderungsrelevante Informationen für die Modellkonstruktion bereit zu stellen. Gelingt der Aufbau, lässt sich auch dieses Ergebnis zumindest teilweise in eine propositionale Repräsentation überführen. Allerdings fällt die Vorstellung nicht leicht, dass aus nonverbalen

3 Verarbeitung von Texten und Bildern

Geräuschen eine komplexe mentale Repräsentation konstruiert werden kann. Ein gutes Beispiel gibt Schnotz (2005), der die Modellkonstruktion im Zusammenhang mit Klangbildern anhand von Vogellauten veranschaulicht. Kann eine Lernende die Schreie von zwei Vögeln aufgrund seines Vorwissens identifizieren und handelt es sich bei den beiden Tieren um einen Raubvogel und seine mögliche Beute, so könnte das mentale Modell aus einem Jagdszenario bestehen.

Mit der Erweiterung des ursprünglichen Modells hat sich Schnotz (2005) der CTML von Mayer (2005) zwar angenähert, doch bleiben zentrale Unterschiede weiterhin bestehen. So sind Textinformationen nicht zwangsläufig mit dem auditiven Kanal und Bilder nicht zwingend mit dem visuellen Kanal assoziiert, wie es etwa bei der CTML der Fall ist. Der wichtigste Unterschied betrifft Aussagen, die das ICTP-Modells zu den einzelnen Repräsentationsformaten macht. Während Mayer (2005) annimmt, dass die Verarbeitung von Texten und Bildern zu zwei formatspezifischen Modellen führt, geht Schnotz (2005) von nur einem mentalen Modell aus. Dieses kann sowohl auf der Grundlage verbaler, als auch auf der Grundlage visueller Informationen konstruiert werden. Zwar nimmt der Autor auch an, dass Lernende Inhalte zusätzlich propositional repräsentieren, was überaus wichtig für bereits erwähnten Konstruktions- und Inspektionsprozesse ist. Dennoch bleibt die eigentliche *Zielrepräsentation* das mentale Modell, an dessen Güte sich zeigt, ob der entsprechende Sachverhalt tatsächlich verstanden wurde.

Natürlich besitzen die Ansätze von Schnotz (2005) und Mayer (2005) auch viele Gemeinsamkeiten, so dass die zwei Autoren bei der Interpretation empirischer Befunde in zahlreichen Fällen zu den gleichen Ergebnissen gelangen. Doch gerade die differenzierte Betrachtungsweise des ICTP-Modells ermöglicht die Formulierung weiterer Prinzipien, die über den theoretischen Rahmen der CTML hinausgehen. Beispielsweise konnten Schnotz und Bannert (2003) in einem Experiment zeigen, dass Bilder nur dann ein tiefes Verständnis fördern, wenn deren Visualisierung den zu erwartenden Anforderungen entspricht. Werden jedoch keine anforderungsrelevanten Strukturen visualisiert, fällt es dem Lernenden äußerst schwer, ein aufgabenadäquates mentales Modell zu konstruieren. Dies gilt selbst dann, wenn alle dazu notwendigen Informationen in dem Bild enthalten sind. Da die Verarbeitung von Bildern zu großen Teilen auf Prozessen der Strukturabbildung beruht, beeinflusst die Form der Visualisierung auch die Struktur des mentalen Modells. Schnotz (2005) bezeichnet diesen Befund als *structure mapping* Effekt, der besonders von den Entwicklern illustrierter Lehr-Lernmaterialien beachtet werden sollte.

Generell lassen sich anhand des ICTP-Modells auch negative Befunde erklären, die im Rahmen des multimedialen Lernens auftreten und grundlegenden Prinzipien scheinbar widersprechen. So berichten Kalyuga, Ayres, Chandler und Sweller (2003) von verschiedenen Untersuchungen, bei denen die Verwendung von Bildern und Texten keine oder sogar negative Effekt auf den Wissenserwerb hatten. Interessanter Weise betraf dieses Problem vor allem Personen, die umfangreiche Vorkenntnisse besaßen. Nach Auffassung von Schnotz (2005, 2010) ist dieser *expertise reversal effect* eine Folge unnötiger Redundanzen, die der Einsatz multipler Repräsentationen erzeugt. Demnach benötigen domänenspezifischen Expertise oftmals nur ein Format, um ein adäquates Modell der

3.3 Integrierte Verarbeitung von Text- und Bildinformationen

beschriebenen bzw. abgebildeten Sachverhalte aufzubauen. In der Regel reichen die sprachlichen oder bildlichen Informationen aus, so dass die zweite Repräsentation keinen Mehrwert für den Lernenden besitzen. Werden trotzdem beide Formate verarbeitet, zehren die überflüssigen Integrationsprozesse an den mentalen Ressourcen, was jedoch nur indirekt mit einer zusätzlichen Beanspruchung des Arbeitsgedächtnisses einhergeht (vgl. Schnotz, 2010). Innerhalb des ICTP-Modells wird dieser Zusammenhang auch als *general redundancy* Effekt bezeichnet, der jedoch nicht mit dem *redundancy* Prinzip der CTML verwechselt werden sollte. Die Entdeckung von Mayer (2005), dass die gleichzeitige Darbietung des Textes in schriftlicher und gesprochener Form beim multimedialen Lernen zu schlechteren Resultaten führt, wird daher von Schnotz (2005) als *specific redundancy* Effekt bezeichnet.

Eine weitere, abweichende Annahme des ICTP-Modells betrifft die allgemeine Gültigkeit des *modality* Prinzips. Demnach verliert der *split attention* Effekt an Bedeutung, wenn anstelle von Animationen statische Bilder verwendet werden und den Lernenden ausreichend Zeit zur Verfügung steht, um den Text zu lesen. Dabei ist vor allem entscheidend, dass mit dem Lesen eine weitreichende Kontrolle des kognitiven Verarbeitungsprozesses einhergeht, die bei einer akustischen Präsentation nicht gegeben ist. Unter der Voraussetzung, dass die Lernzeit nicht beschränkt ist, können daher schwierige Passagen so lange wiederholt werden, bis sie endlich verstanden sind. Diesen Vorteil von geschriebenen gegenüber gesprochenen Texten bezeichnet Schnotz (2005) als *control of processing* Prinzip. Von einer akustischen Präsentation eines Textes ist folglich immer dann abzuraten, wenn es sich um eine Kombination von statischen Bildern mit einem sehr schwer verständlichen Text handelt. Auch im Fall einer unbegrenzten Lernzeit ist es häufig angebracht, verbale Informationen schriftlich zu vermitteln, da der Lernende auf diese Weise sein Tempo beim Wissenserwerb selbst bestimmen kann.

Das ICTP-Modell von Schnotz (2005) ist nicht nur der jüngste, sondern auch der detaillierteste der hier diskutierten theoretischen Ansätze zum Text-Bildverstehen. Doch trotz seiner relativen Komplexität, enthält es nach Ansicht des Autors immer noch zu viele vereinfachende Annahmen, die einer Überarbeitung bedürfen. So ist es durchaus möglich, dass grundsätzlich mehrere Repräsentationsebenen existieren. Im Zusammenhang mit der propositionalen Repräsentation wäre es unter Umständen angebracht, zwischen Mikro- und Makropropositionen zu unterscheiden (van Dijk & Kintsch, 1983). Auch das Vorhandensein von mehreren mentalen Modellen, die in einer hierarchischen Beziehung zueinander stehen, ist nicht auszuschließen. Gerade bei der Verarbeitung komplexer Sachverhalten könnte eine adäquate Repräsentation aus mehreren mentalen Modellen zusammensetzen, die verschiedene Aspekte abdecken oder unterschiedliche Detailebenen wiedergeben.

Ein weiterer Kritikpunkt, den Schnotz (2005) selbst nennt, betrifft die Interaktionsprozesse zwischen dem verbalen und dem piktorialen Kanal. In Abbildung 3.7 ist zu erkennen, dass diese Interaktion ausschließlich zwischen der propositionalen Repräsentation und dem mentalen Modell stattfindet. Dagegen räumt der Autor ein, dass es in manchen Fällen möglich sei, mentale Modelle auch direkt anhand der Textoberflächenstruktur zu

3 Verarbeitung von Texten und Bildern

konstruieren. Ebenso könnten visuelle Wahrnehmungen, als eine vergleichbare percepti-
onsnahe Repräsentation visueller Informationen, direkt in eine propositionale Repräsen-
tation überführt werden. *Umwege* über ein mentales Modell oder eine propositionale Reprä-
sentation, wie es das ICTP-Modell vorsieht, würden daher entfallen. Interessanter Weise
enthält die überarbeitete Fassung des ersten Modells (Schnotz & Bannert, 2003), die in
Abbildung 3.6 abgebildet ist, entsprechende Kreuz- und Querverbindungen.

Trotz dieser Kritik bietet das *integrierte Modell des Text- und Bildverstehens* einen theo-
retischen Rahmen, mit dem sich viele Phänomene des multimedialen Lernens erklären
und in einen größeren Gesamtzusammenhang einordnen lassen. Es ist sehr wahrschein-
lich der fortschrittlichste Ansatz, der bislang entwickelt wurde, auch wenn Schnotz (2005)
selbst noch Potential für eine weitere Überarbeitung sieht. Vor allem die ausführliche Be-
rücksichtigung von Befunden aus der Textforschung und grundsätzliche Überlegungen zu
den Repräsentationsprinzipien visueller und verbaler Informationen machen den *Mehr-
wert* der Theorie aus. Die zusätzlichen Annahmen des ICTP-Modells erlauben einen dif-
ferenzierteren Blick auf bereits bekannte Phänomene und ermöglichen gleichzeitig die
plausible Interpretation von Befunden, die dem Multimediaprinzip scheinbar widerspre-
chen. Als Konsequenz kann der Autor eine Reihe weiterer Prinzipien und Effekte identi-
fizieren, die für das Lernen mit multiplen Repräsentationen von Bedeutung sind.

4 Fragestellung

Mit jeder wissenschaftlichen Arbeit ist ein Ziel und eine entsprechende Fragestellung verbunden, die einen bestimmten Forschungsgegenstand betreffen. Ganz allgemein betrachtet, ist der Forschungsgegenstand dieser Arbeit der Wissenserwerb mit Texten und Bildern. Dabei geht es allerdings nicht allein darum, Hinweise für eine optimale Gestaltung von Lernumgebungen und -materialien zu gewinnen. Von mindestens ebenso großem Interesse sind mögliche Rückschlüsse auf die kognitive Architektur der Lernenden und die damit verbundenen theoretischen Implikationen. Das *multimediale Lernen* ist jedoch mittlerweile zu einem sehr weiten Forschungsfeld angewachsen, weshalb es durchaus hilfreich ist, die eigene Fragestellung etwas näher einzugrenzen. Gerade im Zusammenhang mit der kognitiven Verarbeitung von Texten und Bildern bietet es sich daher an, eines der zahlreichen Prinzipien herauszugreifen und näher zu untersuchen. Daher orientiert sich auch die Fragestellung der vorliegenden Arbeit an einem Phänomen, das beim multimedialen Lernen beobachtet werden konnte und von Schnotz (2005) als *picture-text-sequencing* Effekt bezeichnet wurde. Was es damit auf sich hat, wird in dem nun folgenden Kapitel ausführlich dargestellt.

4.1 Der picture-text-sequencing Effekt

Ein wichtiger Grundsatz, den Mayer (2005) für die Gestaltung multimedialer Lernumgebungen formuliert hat, ist das Prinzip der *temporal contiguity*. Demzufolge sollten Texte und Bilder möglichst gleichzeitig präsentiert werden, da eine sequenzierte Darbietung das kognitive System unnötiger Weise belastet. Liest ein Lernender beispielsweise einen Text, ist das Resultat eine mentale Repräsentation der darin enthaltenen Informationen, die im Terminus der CTML als verbales Modell bezeichnet wird. Soll nachträglich noch ein Bild verarbeitet werden, dessen Inhalte den ergänzen, stellt das eine zusätzliche Belastung der begrenzten Kapazitäten dar. Um die verbale und piktoriale Informationen zu verknüpfen bzw. zu integrieren, muss der Lernende das verbale Modell während der Bildverarbeitung im Arbeitgedächtnis aufrecht erhalten. In Abhängigkeit vom Vorwissen oder den kognitiven Fähigkeiten übersteigt diese Anforderung aber die vorhandenen mentalen Ressourcen und eine Integration scheitert. Werden das verbale und piktoriale Modell nicht integriert, dann ist die Repräsentation des Sachverhaltes unvollständig, was sich vor allem in schlechten Transferleistungen ausdrückt. Solch ein Zusammenhang konnte mehrfach empirisch belegt werden (vgl. Mayer & Anderson, 1992; Mayer & Sims, 1994; Mayer, Moreno, Boire & Vagge, 1999), so dass es nahe lag ein entsprechendes Prinzip zu formulieren, das die Bedeutung der zeitliche Kontiguität hervorhebt.

Wie bereits an anderer Stelle erwähnt wurde, geht Mayer (2005) von einer weitge-

4 Fragestellung

hend parallelen Verarbeitung von Text- und Bildinformationen aus. Innerhalb der CTML werden damit visuelle und verbale Repräsentationsformate als mehr oder weniger gleichwertig angesehen, wobei zugrunde liegende Repräsentationsprinzipien unberücksichtigt bleiben. Folglich ist es daher irrelevant, ob zuerst ein Bild oder ein Text verarbeitet wird. Eine Verletzung des Prinzips der *temporal contiguity* führt in beiden Fällen zu einer vergleichbar schlechten Lernleistung. Dieser Ansicht widerspricht jedoch Kulhavy, Stock und Caterino (1994), der sich dabei auf zwei eigene Studien bezieht. Zwar leugnet der Autor nicht, dass die simultane Präsentation von Texten und Bildern den Wissenserwerb weit mehr fördert, als wenn sequenziert gelernt wird. Doch gibt es seiner Ansicht nach sehr wohl empirische Hinweise darauf, dass die Reihenfolge, in der die Verarbeitung visueller und verbaler Informationen erfolgt, die Lernleistung maßgeblich beeinflusst.

Bei seinen beiden Studien, die leider unveröffentlicht blieben, verwendete Kulhavy, Stock und Caterino (1994) geographische Karten und Texte, die sich inhaltlich aufeinander bezogen. Alle beteiligten Versuchspersonen hatten dann die Aufgabe, so viele visuelle und verbale Informationen wie möglich zu lernen, wobei lediglich die Reihenfolge der Stimuli variiert wurde. Die Gruppen, die zuerst die Karte gesehen hatten, erinnerten bis zu 30% mehr Fakten, als die Gruppen, die zuerst den Text gelesen hatten. Nach Ansicht des Autors zeigte sich damit auf recht eindeutige Weise, dass die sequenzierte Verarbeitung von visuellen und verbalen Informationen leichter fällt, wenn zuerst das Bild gelernt wird. Schnotz (2005) greift diese Befunde im Rahmen seines ICTP-Modells auf und bezeichnet den Zusammenhang zwischen Stimulusreihenfolge und Lernergebnis als *picture-text sequencing* Effekt (PTS-Effekt). Das aus wissenschaftlicher Perspektive interessante am PTS-Effekt ist, dass sowohl Schnotz (2005), als auch Kulhavy, Stock und Caterino (1994) alternative Erklärungen anbieten. Diese widersprechen sich zwar nicht direkt, doch beruhen beide Erklärungsansätze auf unterschiedlichen theoretischen Modellvorstellungen. Wie die Autoren den PTS-Effekt genau ableiten und welche verschiedenen Annahmen sie dabei zugrunde legen, wird in den folgenden Abschnitten dargestellt.

4.1.1 Die Kapazitäts-Hypothese

In Kapitel 3.3.2 wurde das Modell des *Textlernens unter Verwendung geographischer Karten* erläutert, mit dem Kulhavy, Stock, Verdi et al. (1993) detailliert ausführen, weshalb räumlich organisierte Bilder eine positive Wirkung auf den Wissenserwerb besitzen. Kern ihrer theoretischen Konzeption ist die Annahme, dass verbale und bildliche Informationen in verschiedenen kognitiven Systemen verarbeitet werden. Dennoch ist es möglich, Elemente der verbalen und piktorialen Repräsentationen miteinander zu verknüpfen, so dass bestimmte Inhalte *doppelt kodiert* werden, was deren Abrufwahrscheinlichkeit deutlich erhöht. Außerdem nehmen die Autoren an, dass sprachliche und bildliche Informationen zu unterschiedlichen Repräsentationsformaten führen. Während Textinhalte als Propositionen enkodiert werden, resultiert die Verarbeitung von Bildern in sogenannten *images*.

In der Vorstellung von Kulhavy, Stock, Verdi et al. (1993) lassen sich diese *images* men-

talen Kopie der entsprechenden visuellen Reize vergleichen. Sie umfassen daher neben den Elementen (Linien, Icons, etc.), aus denen sich beispielsweise eine geographische Karte zusammensetzt, auch alle strukturellen Informationen. Diese beinhalten räumlichen und metrische Distanzen, Verortungen oder auch auf die Grenzen der Abbildung. Werden alle Elemente und strukturellen Informationen vollständig enkodiert, dann ist die resultierende Repräsentation nach Auffassung der Autoren *intakt*. Das bedeutet, dass ein *image* mit allen darin enthaltenen Informationen als Einheit verarbeitet werden kann, Dabei beansprucht es nicht mehr von der begrenzten Kapazität des Arbeitsgedächtnisses, als dies eine einzige Proposition tun würde. Gleichzeitig besteht jedoch die Möglichkeit auf sämtliche Informationen, die in dem *image* enthaltenen sind, zurückzugreifen. Angeleitet durch die mitkodierte Struktur, kann ein Lernender seine mentale Kopie der Karte regelrecht *absuchen*, ohne darauf sämtliche mentale Ressourcen zu verwenden.

Diese Annahmen haben auch Konsequenzen für die Reihenfolge, in der verbale und visuelle Informationen nach dem Modell verarbeitet werden sollten. Demnach führt die Präsentation des Bildes vor dem Text zunächst zum Aufbau einer intakten Repräsentation (*image*), die im Langzeitspeicher abgelegt wird. Liest ein Lernender dann einen Text, der sich auf die Inhalte des Bildes bezieht, wird das *image* aktiviert und in das Arbeitsgedächtnis gerufen. Dort lässt es sich, ohne große Ressourcen zu beanspruchen, auch während des Lesens aufrecht erhalten. Die kognitive Verarbeitung der Textinhalte erfolgt linear, so dass die Propositionen auf ihrem Weg in den Langzeitspeicher das Arbeitsgedächtnis seriell durchlaufen. Da die mentale Repräsentation des Bildes während des ganzen Leseprozesses zur Verfügung steht, lassen sich referentielle Verbindungen zwischen den Text- und Bildinformationen knüpfen. Informationen, die auf diese Weise *doppelt kodiert* wurden, besitzen damit zwei Abrufreize (*cues*), was die Wahrscheinlichkeit, dass die entsprechenden Inhalte erinnert werden, erhöht.

Ganz anders verhält es sich dagegen, wenn der Lernende den Text vor dem Bild liest. Die sprachlichen Informationen werden zunächst seriell als Propositionen enkodiert und im Langzeitspeicher abgelegt. Betrachtet der Lernende im Anschluss dann ein Bild, wird er versuchen möglichst viele verbale Inhalte, die das Verstehen der grafischen Darstellung unterstützen sollen, aufzurufen. Allerdings ist es nach Ansicht von Kulhavy et al. (1994) mit einem beträchtlichem kognitiven Aufwand verbunden, mehrere Propositionen gleichzeitig im Arbeitsgedächtnis zu halten. In dieser Situation muss der Lernende sich daher entscheiden, ob er seine Aufmerksamkeit auf die Propositionen oder das Bild richtet. Will er die Informationen des Bildes lernen, gehen die Propositionen verloren, da deren Aufrechterhaltung eine aktive Verarbeitung erfordert (van Dijk & Kintsch, 1983). Entschieden sich der Lernende hingegen für die propositionale Repräsentation, dann reicht die begrenzte Kapazität des Arbeitsgedächtnisses nicht aus, um ein intaktes *image* zu konstruieren. In beiden Fällen wird es demnach nicht gelingen, referentielle Verknüpfungen zwischen Bild und Textinformationen zu knüpfen. Als direkte Folge verfügt der Lernende über deutlich weniger *cues* bzw. Abrufreize, was dazu führt, dass weniger Inhalte erinnert werden können.

Der Erklärungsansatz von Kulhavy, Stock und Caterino (1994) beruht damit vor allem

4 Fragestellung

auf der Vermutung, dass die (intakte) mentalen Repräsentationen visueller Inhalte einen besonderen Verarbeitungsvorteil besitzen. Da der positive Effekt eines Bildes unmittelbar mit der Kapazität verbunden ist, die ein *image* im Arbeitsgedächtnis beansprucht, soll dieser Zusammenhang im folgenden als *Kapazitätshypothese* bezeichnet werden.

4.1.2 Die Interferenz-Hypothese

Eine alternative und auch etwas einfachere Erklärung des PTS-Effekts bietet Schnotz (2002, 2005) auf der Grundlage seines Modells des integrierten Text-Bildverstehens an. Im Gegensatz zu Kulhavy, Stock und Caterino (1994) geht der Autor nicht davon aus, dass die Verarbeitung verbaler und visueller Informationen grundsätzlich zu verschiedenen Repräsentationsformaten führt. So ist die Zielrepräsentation in beiden Fällen ein kohärentes mentales Modell, unabhängig davon, ob es anhand eines Textes oder anhand eines Bildes konstruiert wird. Allerdings unterscheidet sich die Ausgangspunkte bzw. die Wege, die dem Aufbau eines solchen Modells vorausgehen. So entsteht beim Lesen zunächst eine Repräsentation der Textoberflächenstruktur, die durch konzeptuelle Organisationsprozesse in eine propositionales Gefüge umgewandelt wird. In der Regel ist diese propositionale Repräsentation dann die Grundlage für den Aufbau des mentalen Modells, was den Übergang von einem symbolischen in ein analoges Format erfordert. Bei einem Bild kann das mentale Modell direkt anhand der visuellen Wahrnehmung konstruiert werden, ohne dass es irgendwelcher Zwischenschritte oder Umwandlungen bedarf. Obwohl Schnotz und Bannert (1999) deshalb annehmen, dass der Konstruktionsprozess mit piktorialen Informationen leichter fallen dürfte, gibt es damit grundsätzlich zwei Möglichkeiten ein mentales Modell zu konstruieren.

Daher muss sich ein Lernender, dem für den Wissenserwerb Bilder und Texte zur Verfügung stehen, die mehr oder weniger informationsäquivalent sind, entscheiden, mit welchem Medium er beginnen will. Eine solche Entscheidung entfällt jedoch, wenn die Reihenfolge in der verbale und piktorialen Informationen verarbeitet werden, nicht frei wählbar ist. In diesem Fall ist dann das Format, das als erstes präsentiert wird, die Grundlage für das zu konstruierende mentale Modell, während im Anschluss verarbeitete Informationen der Ergänzung dienen. Da Bilder und Texte aber zu einer vergleichbaren Zielrepräsentation führen, könnte man vermuten, dass die Reihenfolge innerhalb des theoretischen Konzepts von Schnotz (2005) letztlich keine Rolle spielt. Das tut sie aber, denn die unterschiedlichen Repräsentationsprinzipien, die verbalen und piktorialen Informationen zugrunde liegen, beeinflussen das Zusammenspiel von Konstruktions- und Integrationsprozesse beim Aufbau des mentalen Modells.

Verbale Beschreibungen eines Sachverhaltes sind stets durch einen gewissen Grad an Unbestimmtheit gekennzeichnet, die nicht selten zu abweichenden Interpretationen führen. Das deskriptive Repräsentationsprinzip eines Textes lässt es schlichtweg nicht zu, Objekte so detailliert und anschaulich zu erfassen, wie das mit einem Bild möglich wäre. Als Konsequenz ist die Verarbeitung verbaler Inhalte sehr stark auf bestehende Wissenstrukturen angewiesen, damit trotz aller prinzipiellen Unschärfen ein kohärentes

mentales Modell aufgebaut werden kann. Informationslücken, die ein Text zwangsläufig enthält, werden dabei durch die Vorkenntnisse des Lernenden ergänzt. Da vorhandene Wissensstrukturen und persönlichen Erfahrungen von Individuum zu Individuum variieren, ist jedes mentale Modell, das selbst auf der Grundlage ein und des selben Textes konstruiert wird, so etwas wie ein Unikat. Dieser Zusammenhang hat unmittelbare Folgen, wenn piktoriale und verbale Informationen, die sich auf den selben Sachverhalt beziehen, nacheinander gelernt werden.

Verarbeiten ein Lernender zuerst den Text, führt das zum Aufbau eines mentalen Modells, das mit sehr hoher Wahrscheinlichkeit von dem Bild abweicht, das im Anschluss präsentiert wird. Diese Abweichungen dürften vor allem strukturelle Elemente betreffen, die für den funktionalen Charakter des Modells jedoch nicht unbedingt wichtig sein müssen. Da der Lernenden aber bestrebt ist, eine kohärente Repräsentation des dargestellten Sachverhaltes aufzubauen, wird er versuchen, das interne mentale Modell und das externe Bild zu integrieren. Zwar ist vorstellbar, dass durch analoge Strukturabbildungsprozesse fehlenden Modellelemente ergänzt werden können, doch führen selbst nichtige Abweichungen oder Widersprüche zu kognitiven Interferenzen. Diese Interferenzen behindern den Konstruktionsprozess, so dass es dem Lernenden nicht gelingt, eine kohärente mentale Repräsentation aufzubauen. Als Folge wird der Sachverhalt nicht richtig verstanden, was sich in einer schlechteren Lernleistung ausdrückt.

Ganz anders verhält es sich hingegen, wenn zuerst die piktorialen Informationen präsentiert werden. In diesem Fall erfolgt der Aufbau des mentale Modells mehr oder weniger direkt anhand der visuellen Wahrnehmung, wobei ein perzeptionsnahes analoges Format in ein abstraktes analoges Format überführt wird. Das dürfte vielen Lernenden etwas leichter fallen, als wenn sie das Modell anhand eines Textes konstruieren müssen (Schnotz & Bannert, 1999). Dennoch ist der entscheidende Vorteil bei dieser Reihenfolge ein anderer. Wird das Bild zuerst gelernt, dann ist das resultierende Modell in der Regel zunächst unvollständig, da es den Sachverhalt ohne die verbalen Informationen nicht lückenlos abdecken kann. Für eine kohärente und vollständige Repräsentation müssen daher zusätzlich die entsprechenden Textinhalte verarbeitet und integriert werden. Das mentale Modell, das auf der Grundlage des Bildes entstanden ist, fungiert dabei als eine Art Blaupause bzw. als Entwurf oder Gerüst für den Aufbau der eigentlichen Zielrepräsentation. Es gibt zentrale Strukturen vor, während einzelne Modellkomponenten durch die serielle Verarbeitung der verbalen Informationen nach und nach ergänzt oder spezifiziert werden. Auf diese Weise ist der Text Grundlage für Prozesse der Modellkonstruktion und -inspektion, ohne dass es zu Abweichungen zwischen den externen und internen Formaten kommt.

Schnotz (2002, 2005) erklärt den PTS-Effekt in erster Linie durch die Entstehung kognitiver Interferenzen, die immer dann auftreten, wenn der Text vor dem Bild präsentiert wird. Damit liegt es nahe, den zweiten Erklärungsansatz unter der Bezeichnung *Interferenzhypothese* zusammen zu fassen.

Abschließend lässt sich festhalten, dass beide Herleitungen des PTS-Effekts logisch konsistent und plausibel sind, so dass sich konsequenter Weise die Frage stellt, wel-

4 Fragestellung

cher der Autoren nun Recht hat. Diese Frage steht im Mittelpunkt der vorliegenden Arbeit, mit der letztlich eine Antwort gefunden werden soll. Da es jedoch nicht möglich ist, allein anhand theoretischer Annahmen eine Entscheidung für oder wider einen der beiden Erklärungsansätze zu treffen, muss eine empirische Untersuchung Abhilfe schaffen. Zu diesem Zweck bietet es sich an, experimentelle Bedingungen zu schaffen, in denen die Kapazitäts- und die Interferenzhypothese zu jeweils verschiedenen Vorhersagen führen, die sich dann überprüfen lassen. Die allgemeinen Überlegungen, die dem Untersuchungsablauf und -design zugrunde liegen, werden in dem folgenden Kapitel näher ausgeführt.

4.2 Überlegungen zur Versuchsplanung

Die Planung einer experimentellen Studie, die Aussagen und Aspekte kognitiver Theorien untersucht, ist keine triviale Angelegenheit. Das Hauptproblem besteht sicherlich darin, dass man die postulierten Prozesse nicht direkt beobachten kann, was die Entwicklung eines Versuchsplanes erfordert, der Rückschlüsse auf die entsprechenden psychologischen Vorgänge zulässt. Grundlage für diese Schlussfolgerungen sind in der Regel Beobachtungen von konkreten Verhaltensweisen oder die Messung physiologischer Parameter, von denen angenommen werden kann, dass sie mit den kognitiven Prozessen zusammenhängen. Im Fall des PTS-Effekts bietet es sich daher an, den Lernerfolg zu erfassen, da die Reihenfolge der Präsentation von Text und Bild nachweislich Auswirkungen auf den Wissenserwerb hat. Doch allein anhand eines Instruments, mit dem bilanziert werden kann, wie viel die Versuchspersonen gelernt haben, lässt sich noch nicht feststellen, welcher der konkurrierenden Erklärungsansätze nun tatsächlich zutrifft. Zu diesem Zweck müssen die Texte und Bilder, die in der Untersuchung verwendet werden sollen, bestimmte Merkmale aufweisen. Um welche Merkmale es sich dabei genau handelt, ist Thema des folgenden Abschnitts.

4.2.1 Merkmale des Lehrmaterials

Es liegt nahe, dass eine Studie zum PTS-Effekt verbales und pictoriales Stimulusmaterial beinhaltet, das den Versuchspersonen in einer bestimmten Reihenfolge präsentiert wird. Doch allein anhand der Feststellung, dass Lernende tatsächlich bessere Leistungen zeigen, wenn sie ein Bild vor einem Text betrachten, lässt sich die zentrale Fragestellung dieser Arbeit nicht beantworten. Will man herausfinden, ob nun die Kapazitäts- oder die Interferenzhypothese den treffenderen Erklärungsansatz darstellt, bedarf es daher einer gezielten Manipulation der experimentellen Bedingungen. Da jedoch die Reihenfolgen die Ursache des Effekts ist, an der nichts verändert werden sollte, bietet es sich an, das Lehr-Lernmaterial zu modifizieren. Grundlage solcher Eingriffe müssen natürlich theoretische Überlegungen sein, die nach ihrer Umsetzung experimentell valide Schlussfolgerungen zulassen.

Eine Größe, an der die Modifikation des Stimulusmaterials ansetzen kann, ist das Informationsverhältnis von Text und Bild. So ist es gerade für Kulhavy, Stock, Verdi et al. (1993) nicht ohne Bedeutung, wie viele gemeinsame Inhalte die beiden Formate enthalten, da zentrale Annahmen seiner Modellvorstellungen auf der DCT beruhen (Paivio, 1986). Demnach müssen sich Text und Bild auf den selben Sachverhalt beziehen, damit referentielle Verbindungen zwischen verbalen und piktorialen Informationen überhaupt geknüpft werden können. Ist dies nicht der Fall, dann verliert die zusätzliche Verwendung des Bildes teilweise seine positive Wirkung auf den Wissenserwerb. Die visuellen Inhalte werden dann zwar immer noch *doppelt kodiert*, doch zusätzliche Abrufreize textitcues für Texinformationen entfallen. Vorausgesetzt die Annahmen treffen zu, dann müsste eine geringe Informationsäquivalenz zwischen Text und Bild zu weniger referentiellen Verknüpfungen führen, was sich negativ auf die Wiedergabeleistung auswirkt.

Als Konsequenz werden Versuchspersonen, die auf multimediale Lernmaterialien mit einer hohen inhaltliche Überlappung zurückgreifen können, mehr lernen, als Versuchspersonen, die für den Wissenserwerb komplementäre Formate nutzen. Diese Vorhersage steht im Einklang mit der DCT und der CTML, doch ist es anhand der beiden Ansätze nicht möglich, das Merkmal der Informationsäquivalenz auf den PTS-Effekt zu übertragen. Dazu müssten bestimmte Annahmen hinsichtlich der verschiedenen Repräsentationsformate gemacht werden, die aus der Verarbeitung von Bildern und Texten resultieren. Solche Annahmen finden sich jedoch nur in den theoretischen Konzepten von Kulhavy, Stock und Caterino (1994) oder Schnotz (2002, 2005), wo diese Frage ausführlich behandelt wird. Da die beiden Autoren allerdings nicht explizit darauf eingehen, welche Auswirkungen es für den PTS-Effekt hat, wenn Text und Bild kaum oder gar keine inhaltliche Überlappung aufweisen, muss dieser Zusammenhang anhand der jeweiligen Hypothese bzw. anhand des dahinter stehenden Modells hergeleitet werden. Zu welchen Ergebnissen diese Herleitungen führen wird in den folgenden beiden Kapiteln dargestellt.

4.2.2 Das Informationsverhältnis und die Kapazitätshypothese

Wie bereits mehrfach angedeutet wurde, nehmen auch Kulhavy, Stock, Verdi et al. (1993) an, dass die Vorteile die ein Bild beim Wissenserwerb haben kann, eng mit dem Aufbau referentieller Verknüpfungen zwischen piktorialen und verbalen Informationen zusammenhängen. Allerdings geht das theoretische Modell der Autoren über die klassische DCT hinaus, indem sie den mentalen Repräsentationen der Bilder (*images*) besondere Eigenschaften zuschreiben. Demnach können die *images* ohne große Kapazitätsbelastung im Arbeitsgedächtnis gehalten werden, was Kulhavy, Stock und Caterino (1994) letztlich als Ursache für den PTS-Effekt ansieht. Dieser Verarbeitungsvorteil müsste auch dann vorhanden sein, wenn sich Bild und Text inhaltlich nur wenig überlappen. Solange sich die beiden Formate auf den selben thematischen Kontext beziehen, lassen sich zumindest indirekte Verbindungen herstellen, bei denen die Lernenden auftretende Lücken mithilfe ihres Vorwissens schließen. Dies ist jedoch nur möglich, wenn das *image* im Arbeitsgedächtnis gehalten wird, während die serielle Verarbeitung der Textinformationen

4 Fragestellung

erfolgt. Zwar ist davon auszugehen, dass die Anzahl der direkten referentiellen Verknüpfungen abnehmen wird, was nach Ansicht von Kulhavy et al. (1994) eine verminderte Wiedergabeleistung zur Folge hat. Doch sollten Versuchspersonen, die das Bild zuerst sehen, immer noch besser abschneiden, da sie verbale und piktoriale Informationen inhaltlich besser aufeinander beziehen können. Treffen die Annahmen der Kapazitätshypothese zu, ist der PTS-Effekt auch bei einem komplementären Informationsverhältnis zwischen Text und Bild zu beobachten.

4.2.3 Das Informationsverhältnis und die Interferenzhypothese

Im Gegensatz zu Kulhavy et al. (1994) geht Schnotz (2002, 2005) nicht davon aus, dass die Verarbeitung von Bild- und Textinformationen zu grundsätzlich verschiedenen Repräsentationsformaten führt. Dementsprechend orientiert sich sein Ansatz stärker an der Relation zwischen verbalen und piktorialen Informationen, was letztlich in einem engen Zusammenhang mit seiner Erklärung des PTS-Effekts steht. Wie bereits in Kapitel 4.1.2 ausführlich dargestellt wurde, sind nach Ansicht des Autors Interferenzen dafür verantwortlich, dass Lernende schlechter abschneiden, wenn sie den Text vor dem Bild lesen. Damit diese Interferenzen aber überhaupt auftreten, müssen sich verbale und piktoriale Informationen nicht nur auf den selben Sachverhalt beziehen, sondern möglichst informationsäquivalent sein. Nur wenn die Verarbeitung des Textes zu einem mentalen Modell führt, das inhaltlich mit dem nachfolgenden Bild wirklich konkurriert, kann es zu kognitiven Konflikten kommen. Weisen Text und Bild jedoch keine oder nur geringe inhaltliche Überlappungen auf, dürften bei einer sequenzierten Präsentation des Lernmaterials keine Interferenzen entstehen. Treffen die Annahmen der Interferenzhypothese zu, dann ist der PTS-Effekt bei einem komplementären Informationsverhältnis zwischen Text und Bild nicht zu beobachten.

4.3 Schlussfolgerungen für die Gestaltung des Untersuchungsmaterials

Die Konsequenzen, die sich aus den bisherigen Überlegungen ziehen lassen, sind denkbar einfach. Will man letztlich entscheiden, ob nun Schnotz (2005) oder Kulhavy et al. (1994) Recht behält, muss man ein Versuchsdesign umsetzen, bei dem zwei unterschiedliche Ausführungen von Lehrmaterialien verwendet werden. Eine erste Version sollte aus Texten und Bildern bestehen, die sich inhaltlich stark überlappen bzw. einen hohen Anteil äquivalenter Informationen enthalten. Die zweite Version müsste dagegen Texte und Bilder umfassen, die sich zwar auf das selbe Themen beziehen, aber unterschiedliche Aspekte des betreffenden Sachverhaltes abdecken, womit ein komplementäres Informationsverhältnis erzeugt wird. Dabei ist jedoch noch nicht ganz klar, welche Arten von Bildern sich bei Untersuchung sinnvoller Weise verwenden lassen.

Um die Herleitung zu vereinfachen, war in den vorangegangenen Abschnitten ganz allgemein von Bildern die Rede, was für das ICTP-Modell von Schnotz (2005) grundsätzlich kein Problem darstellt. Fraglicher wird diese Verallgemeinerung jedoch bei der Kapazi-

4.3 Schlussfolgerungen für die Gestaltung des Untersuchungsmaterials

tätshypothese, da sich Kulhavy et al. (1994) bei ihren Annahmen ausdrücklich auf geographische Karten bzw. räumlich strukturierte Abbildungen beziehen. Um dennoch eine *faire* Gegenüberstellung der beiden Erklärungsansätze zu gewährleisten, müssten daher alle verwendeten Bilder dieser Kategorie angehören. Das favorisiert natürlich den Einsatz geographischer Karten, doch genügt auch eine Darstellung, die den Querschnitt durch die Erdkruste zeigt, dem Kriterium der räumlichen Strukturierung. So sind beispielsweise Dicke oder Lage der einzelnen Schichten Informationen, die sich in erster Linie aus der Anordnung und den Relationen zwischen verschiedenen Bildelementen erschließen lassen. Auch schematische Zeichnungen oder realistische Bilder besitzen damit eine räumliche Struktur, weshalb man die Kapazitätshypothese ebenso auf diesen Abbildungstypus anwenden kann.

Eine weitere Unschärfe zwischen den beiden Erklärungsansätzen betrifft die Resultate der Lernprozesse. Wenn beispielsweise Kulhavy et al. (1994) den Lernerfolg erfassen, dann beziehen sich die Autoren in erster Linie auf die Wiedergabe (*retention*) von Fakten. In der Vorstellung von Schnotz (2005) zielt der Wissenserwerb hingegen auf die Konstruktion einer kohärenten mentalen Repräsentation, was eher mit Verstehen und bedeutungsvollem Lernen gleichzusetzen ist. Natürlich spielt dabei auch die Verfügbarkeit von einzelnen Informationen eine nicht zu unterschätzende Rolle, doch liegt der Schwerpunkt eindeutig auf dem tiefen Verständnis eines Sachverhaltes. Dennoch kann man davon ausgehen, dass der Aufbau eines reichhaltigen mentalen Modells mit dem Erwerb entsprechender Wissensstrukturen verbunden ist, so dass sich auch anhand der Wiedergabe von einzelnen Inhalten bemessen lässt, ob die Inferenzhypothese zutrifft oder nicht.

Im Fall der Kapazitätshypothese und der Leistung bei Transfer- oder Inferenzaufgaben kann ein ähnlicher Zusammenhang vermutet werden, selbst wenn Kulhavy et al. (1994) das Konzept Verstehen gar nicht aufgreifen. So müssen bestimmte Fakten bekannt sein bzw. erinnert werden, damit eine valide Schlussfolgerung gelingen kann. Ob der Lernende dann wirklich in der Lage ist, bestimmte Problemstellungen durch Inferenzen zu lösen, hängt natürlich auch von dem prozeduralem Wissen ab, über das ein Individuum verfügt. Dennoch lässt sich auf der Grundlage der Kapazitätshypothese mutmaßen, dass der PTS-Effekt sich auch bei Aufgaben zeigen wird, die über eine reine Wiedergabe von einfachen Fakten hinausgehen.

An dieser Stelle endet das Kapitel zur Fragestellung und den grundlegenden theoretischen Gedanken hinter der Untersuchung. Wie die verschiedenen Überlegungen im Detail experimentell umgesetzt wurden, kann in den Abschnitten nachgelesen werden, die den einzelnen Studien gewidmet sind. Im folgenden Kapitel, mit dem der empirische Teil der Arbeit beginnt, wird jedoch zunächst die Entwicklung des Untersuchungsmaterials und der Instrumente zur Erfassung der Personenmerkmale dargestellt.

Teil II

Empirie

5 Untersuchungsmaterial und Instrumente

Neben der theoretischen Fundierung einer empirischen Arbeit bestimmen die verwendeten Materialien und Methoden maßgeblich deren Qualität. Aus diesem Grund werden die im Rahmen der Untersuchung eingesetzten Texte und Bilder sowie alle Instrumente, mit denen psychologische Zustände und Merkmale erfasst wurden, im Folgenden ausführlich beschrieben. Da zuerst die Lernmaterialien entstanden, beginnt dieses Kapitel mit deren Charakterisierung.

5.1 Die Theorie der Plattentektonik in Texten und Bildern

Bei Studien, die zum Ziel haben, Bedingungen zu identifizieren, die den Wissenserwerb beeinflussen können, kommt dem verwendeten Stimulusmaterial besondere Bedeutung zu. Deshalb wird in den folgenden Abschnitten detailliert dargestellt, welche Überlegungen der Gestaltung der Texte und Bilder zugrunde lagen, die den Versuchsteilnehmern im Rahmen der Untersuchung präsentiert wurden. Die finalen Versionen der Texte und Bilder sind im Anhang A enthalten.

5.1.1 Auswahl der Domäne

Während es bei Studien im schulischen Kontext häufig sinnvoll ist, auf curriculare Inhalte zurückzugreifen, um unter anderem das Vorwissen der Schüler zu kontrollieren (vgl. Lenzner, 2009), haben Forscher bei der Auswahl einer Wissensdomäne für Lernexperimente mit Erwachsenen sehr viel mehr Freiheitsgrade. In der Regel wird aber aus ökonomischen Gründen versucht, auf existierendes Material zurückzugreifen, das sich bereits bewährt hat und nicht neu entwickelt werden muss. Beispielsweise verwendet vgl. Butcher (2006) bei ihrer Untersuchung einen expositorischen Text über das Herzkreislaufsystem, der bereits von Wolfe et al. (1998) für einen ähnlichen Kontext verfasst worden war. Dennoch ist es oft unumgänglich, vorhandenes Lehrmaterial zu modifizieren oder auch vollkommen neu zu entwerfen. In der Regel muss das Stimulusmaterial ganz bestimmte Eigenschaften aufweisen, um in der betreffenden Studie sinnvoll eingesetzt werden zu können.

Aufgrund der Verfügbarkeit bot es sich im Rahmen der vorliegenden Arbeit an, Material zum Thema Vulkanismus zu nutzen, das bereits von Hauck (2004) verwendet worden war. Allerdings wurde bei der näheren Sichtung des Materials schnell deutlich, dass die vorhandenen Diagramme und Texte den Anforderungen der eigenen Fragestellung nicht genügten. Die einzelnen Textbausteine waren sehr umfangreich, während die vorhandenen grafischen Darstellungen dagegen nur einen Bruchteil des schriftlichen Materials abdeckten und ausschließlich eine ergänzende Funktion hatten. Außerdem enthielten die

5 Untersuchungsmaterial und Instrumente

Texte zu wenige Schilderungen kausaler Zusammenhänge und Prozesse, so dass es schwierig geworden wäre, eine angemessene Zahl von Fragen und Aufgaben zu formulieren, die ein tieferes Verständnis erfassten. Dennoch wurde durch die Auseinandersetzung mit dem Material zum Vulkanismus eine inhaltliche Domäne entdeckt, die aufgrund ihres anschaulichen Charakters dazu geeignet war, möglichst viele Sachverhalte grafisch darzustellen. Darüber hinaus bot es sich an, bei der Gestaltung des Lehrmaterials weitere Themenbereiche der Geodynamik zu berücksichtigen. So kam es, dass sich der inhaltliche Fokus von der Entstehung der Vulkane mehr und mehr wegbewegte und gegen Ende die Theorie der Plattentektonik thematisch im Zentrum stand. Das Themengebiet der Geodynamik umfasst von phänomenologischen Beschreibungen, über geophysische Forschungsmethoden bis hin zu mathematischen Kausalmodellen eine Vielzahl von Inhalten, die sich *relativ* einfach in einen kohärenten Gesamtzusammenhang stellen lassen. Vor diesem Hintergrund konnte allein durch die Auswahl entsprechender Themen die Komplexität des Stimulusmaterials variiert werden. Die endgültige Entscheidung, welche inhaltlichen Aspekte schließlich in den Texten und Bildern umgesetzt werden sollten, wurde während der Gestaltung des Lehrmaterials getroffen, die in den nachfolgenden Abschnitten erläutert wird.

5.1.2 Prinzipien der Textgestaltung

Der erste Schritt bei der selbständigen Gestaltung des Lehrmaterials bestand schließlich darin, einen expositorischen Text (Chafe, 1994) zu verfassen, der verschiedene Phänomene und Theorien der Plattentektonik beschrieb bzw. erklärte. Inhaltlich orientierte sich dieser Text an zwei Lehrbüchern zur Geodynamik bzw. Geophysik, die im Zusammenhang mit der Ausbildung von Erdkunde- und Geographielehrern an der Universität Koblenz-Landau verwendet werden. Als Grundlage für den Text über den Schalenbau der Erde diente ein Aufsatz von Powell (1995), während die Darstellungen der übrigen Abschnitte auf der Arbeit von Stüwe (2000) basierte. Da in beiden Lehrbüchern die Geophysik eine zentrale Rolle einnahm, wurden viele der Zusammenhänge darin mathematisch beschrieben bzw. als mathematische Modelle dargestellt.

So bot es sich an, einzelne Formeln und numerische Ausdrücke in das Stimulusmaterial mit aufzunehmen, doch wurde aus verschiedenen Gründen darauf verzichtet. Entscheidend war letztendlich die Überlegung, dass mit der Integration mathematischer Ausdrücke eine *weitere Repräsentationsform* zur Wissensvermittlung verwendet worden wäre. Im Zentrum der Fragestellung stand aber die Integration von Bild und *verbalen* Informationen. Sicherlich wäre es möglich gewesen, eine weitere Indikatorvariable zu erheben, um die mathematischen Fähigkeiten der Versuchsteilnehmer statistisch einigermaßen zu kontrollieren. Dies hätte aber nicht nur den Aufwand der Untersuchung erhöht, sondern auch die Interpretation der Ergebnisse erschwert, so dass alle Zusammenhänge schließlich allein verbal umschrieben wurden.

Mit Blick auf die Population von Psychologiestudenten, aus der die Stichprobe gezogen werden sollte, schien es zudem angebracht zu sein, von einem eher geringen Vorwissen

der potentiellen Versuchsteilnehmer auszugehen. Da die Verständlichkeit eines Textes keine invariante Eigenschaft ist, sondern stets in Bezug zu den Fähigkeiten und Kenntnissen des einzelnen Lesers beurteilt werden muss (Schnotz, 2006), sollte der Lehrtext deshalb so gestaltet werden, dass er gerade für Personen mit geringem Vorwissen möglichst verständlich war. Aus diesem Grund orientierte sich die schriftliche Umsetzung des Lehrtextes an den Dimensionen *Gliederung und Ordnung*, *sprachliche Einfachheit* sowie *Kürze/Prägnanz* des Hamburger Verständlichkeitskonzepts (Langer, Schulz von Thun & Tausch, 2002). Für die Umsetzung der Dimension *Gliederung/Ordnung* bot es sich inhaltlich an, den Text in vier Abschnitte zu unterteilen, die jeweils mit aussagekräftigen Überschriften versehen wurden. Damit es möglichst leicht war, den *roten Faden* auch innerhalb der einzelnen Textbausteine aufzunehmen (und beizubehalten), war es das erklärte Ziel, jeden Abschnitt thematisch kontinuierlich aufzubauen (Schnotz, 1994a). Zu diesem Zweck wurden die einzelnen Sätze so formuliert, dass sie sowohl lokal als auch global in einem kohärenten Zusammenhang standen.

Da es letztlich inhaltlich zwingend war, Fachtermini aus der Geophysik und damit eine ganze Reihe von Fremdwörtern zu verwenden, musste die Dimension *sprachliche Einfachheit* vor allem durch den Verzicht auf komplexe Schachtelsätze und den Gebrauch eines möglichst unkomplizierten Satzbaus umgesetzt werden. Zur Verwirklichung der Dimension *Kürze/Prägnanz* wurde darauf geachtet, dass die Länge der schriftlichen Darstellung in einer angemessenen Relation zu ihrer Funktion stand. Allein aus Rücksicht auf die Fragestellung der Untersuchung durfte der Text nicht zu umfangreich sein, da möglichst viele der darin enthaltenen Informationen in einem instruktionalen Bild dargestellt werden sollten. Außerdem orientierte sich die Länge des schriftlichen Lehrmaterials nicht zuletzt daran, dass sowohl die Bereitschaft, als auch die Fähigkeit der studentischen Versuchsteilnehmer sich konzentriert Wissen anzueignen, ihre Grenzen haben.

Um schließlich einen kompakten Sachtext zu erhalten, der den Studenten der Form nach auch in ihrem Alltag begegnen könnte, wurde von der Berücksichtigung der vierten Dimension *zusätzliche Stimulanz* des Hamburger Verständlichkeitskonzepts abgesehen. Durch diesen Verzicht auf gezielte Maßnahmen zur Erzeugung von Interesse und Lesemotivation, wie sie Langer et al. (2002) anführen, sollte neben der angestrebten Prägnanz vor allem der Lehrbuchcharakter der Texte betont werden. Am Ende umfasste das schriftliche Material insgesamt 1027 Worte, die sich auf vier inhaltlich kohärente Einheiten verteilten. Diese inhaltlichen Einheiten dienten auch als Grundlage zur Gestaltung der instruktionalen Bilder, wie im folgenden Abschnitt dargestellt wird.

5.1.3 Prinzipien der Bildgestaltung

Bei der Gestaltung des Bildmaterials wurde das übergeordnete Ziel verfolgt, möglichst viele der in den Texten enthaltenen Informationen in einem instruktionalen Bildes zu veranschaulichen. Mit Rücksicht auf die bestehende Gliederung, musste daher für jeden der vier Texte eine grafische Darstellung entworfen werden, welche die Inhalte der einzelnen Abschnitte so vollständig wie möglich wiedergab. So war es unumgänglich, die vier

5 Untersuchungsmaterial und Instrumente

Abbildungen aus mehreren kleineren Bildern und grafischen Elementen *zusammenzusetzen*, um den Informationsgehalt der Textvorlage angemessen reflektieren zu können. Gleichzeitig musste aber darauf geachtet werden, dass die *Bildkompositionen* nicht zu umfangreich gerieten. Die Versuchsteilnehmer sollten die instruktionalen Grafiken auf einem Computerbildschirm betrachten können, ohne scrollen oder blättern zu müssen.

Als Vorlagen für die Gestaltung des Lehrmaterials dienten Abbildungen und Grafiken aus den Geographielehrbüchern von Powell (1995) und Stüwe (2000), die eingescannt und mit der Bildbearbeitungssoftware Adobe Photoshop CS2 aufbereitet wurden. Allerdings konnte das auf diese Weise gewonnene grafische Material so gut wie nie unverändert übernommen werden, so dass in den meisten Fällen mehr oder weniger aufwendige Arbeitsschritte erforderlich waren, um die *Rohbilder* der schriftlichen Vorlage anzupassen. Diese Arbeitsschritte reichten vom Entfernen von Beschriftungen, über das Ein- und Umfärben einzelner Bildelemente, bis hin zur Anfertigung eigenständiger Zeichnungen. Besonders häufig wurden farbige Pfeile als *directing codes* eingefügt (Weidenmann, 1994a), um Bewegungsrichtungen und dynamische Prozesse zu signalisieren.

Für eine Angleichung des Informationsgehaltes zwischen den Bildern und dem schriftlichen Lehrmaterial war es zudem unumgänglich, gezielt Textfragmente in die grafischen Darstellungen zu integrieren. Neben thematischen Überschriften, die der Orientierung und der Zuordnung der Bilder zu den entsprechenden Textabschnitten dienen sollten, gaben beispielsweise Zahlen auf einer Weltkarte an, mit welcher Geschwindigkeit sich die gewaltigen Gesteinsplatten bewegten. Darüber hinaus wurden Legenden verwendet, um bedeutsame Bildelemente zu spezifizieren, die für das Verständnis der Sachverhalte notwendig waren. Gerade bei der bildlichen Darstellung von Prozessen und kausalen Zusammenhängen konnte auf eine ausführlichere *direkte Beschriftung* (Schnotz, 2006) nicht verzichtet werden, da sich deren Bedeutung ohne großes Vorwissen sonst nur schwer hätte erschließen lassen.

Grundsätzlich wurde darauf geachtet, direkte Beschriftungen so einzufügen, dass sie sich in unmittelbarer Nähe zu den grafischen Objekten befanden, die sie beschrieben. Auf diese Weise sollte das Prinzip der *räumlichen Kontiguität* (vgl. Mayer & Moreno, 2003) realisiert werden, um den Lernenden die unnötige Suche nach Informationen zu ersparen. Letztendlich bestand das depiktionale Lehrmaterial aus vier HTML-Seiten, die geografische Karten, schematische Zeichnungen und Textfragmente umfassten. Entsprechend der Definitionen von Knowlton (1966) handelte sich bei den verwendeten grafischen Elementen ausschließlich um realistische Bilder. Auch wenn in der Regel das schriftliche Material als Vorlage für das Bild diente, so wurde mit dem Ziel einer möglichst hohen Informationsüberlappung der Text in einigen Fällen nachträglich an die Gegebenheiten der Bilder angepasst. Auf die gezielte grafische Darstellung abstrakter Textelemente oder -aussagen wurde mit Blick auf den damit verbundenen konzeptionellen und planerischen Aufwand verzichtet, so dass die Abbildungen allein anschauliche Sachverhalte zeigen.

5.1.4 Kongruente und komplementäre Bilder und Texte

Abbildung 5.1 zeigt die vollständige Version einer Bildseite, die den Zusammenhang zwischen den tektonischen Plattengrenzen und Phänomenen wie Erdbeben oder Vulkanismus veranschaulicht. Links, links oben und oben in der Mitte sind drei unterschiedliche *Grenztypen* abgebildet, deren Eigenschaften mit einigen Worten näher charakterisiert werden.

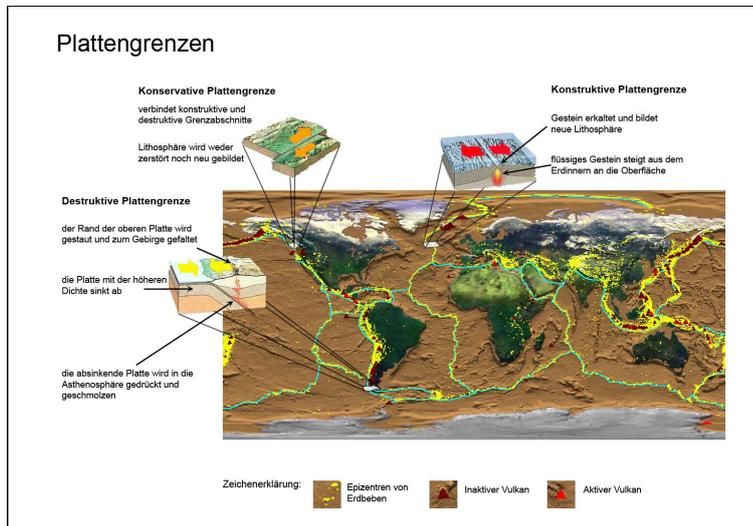


Abbildung 5.1: Bildseite *Plattengrenzen* als vollständige Version

Im Hintergrund ist eine Weltkarte zu erkennen, in der Erdbeben als gelbe Punkte und Vulkane als rote Dreiecke eingezeichnet sind. Die Bedeutungen der Symbole auf der Weltkarte können in der Legende unter dem Bild abgelesen werden. In der Textvorlage, auf deren Grundlage die Abbildung *Plattengrenzen* entstanden ist, wird dementsprechend der Zusammenhang zwischen den Plattengrenzen und dem Auftreten von Erdbeben und Vulkanen ausführlich erläutert. Darüber hinaus enthält der Text eine Beschreibung der verschiedenen Plattengrenzen und ihrer Bedeutung für die Entstehung einzelner Phänomene der Geodynamik. Prinzipiell ist es also möglich, bestimmte Informationen sowohl dem Bild zu entnehmen, als auch im Text nachzulesen. Aller codespezifischen Eigenheiten der beiden Formate zum Trotz, werden damit die selben Sachverhalte gleichzeitig deskriptiv und depiktional repräsentiert.

Wenn ein Leser bzw. Betrachter beide Repräsentationen zum Wissenserwerb heranzieht, überschneiden sich die Informationen im Bild mit den Informationen des Textes. In Anlehnung an Mayer et al. (2001) könnte man dieses Informationsverhältnis zwischen Bild und Text als *redundant* bezeichnen, allerdings bezieht sich der *redundancy effect* vor allem auf die gleichzeitige Darbietung desselben Textes in schriftlicher und gesprochener Sprache. Da eine solch starke Äquivalenz zwischen Bild und Textinformationen aber nicht angenommen werden kann (Hamrodi & Lieber, 2008) und dem Begriff Redundanz etwas Negatives anhaftet, erschien er zur Charakterisierung des Informationsverhältnisses zwischen Bild und Text ungeeignet. Eine passendere Umschreibung verwendet

5 Untersuchungsmaterial und Instrumente

Ballstaedt (2005), der eine Bild-Text-Relation dann als *kongruent* bezeichnet, wenn die kognitive Verarbeitung der Bild und Textinformationen zu inhaltlich vergleichbaren mentalen Repräsentationen führt. Dagegen sieht er inhaltliche *Komplementarität* als gegeben an, wenn Bild und Text zwar unterschiedliche mentale Repräsentationen aktivieren, diese aber durch einen gemeinsamen Kontext miteinander verknüpft sind.

Mit Blick auf die Fragestellung bestand nun der dritte und letzte Schritt bei der Entwicklung des Lehrmaterials darin, das bereits erzeugte kongruente Informationsverhältnis zwischen Bildern und Texten in ein inhaltlich komplementäres Verhältnis umzuwandeln. Ein komplementäres Verhältnis war dann gegeben, wenn sich Bild und Text auf den selben Themenabschnitt bezogen (z.B. auf die Plattengrenzen), zentrale Informationen aber immer nur einem Repräsentationsformat entnommen werden konnten. Unter der Annahme dass beide Informationsquellen beim Wissenserwerb zur Verfügung standen, musste der Lernende gleichzeitig aber in der Lage sein, durch die integrative Verarbeitung der sich ergänzenden Bild- und Textinformationen eine kohärente mentale Repräsentation der dargestellten Sachverhalte aufzubauen. Um dies sicherzustellen, wurden das bereits entwickelte *kongruente* Lehrmaterial als Grundlage für die Erstellung der komplementären Bilder und Texte verwendet.

Da es sehr viel leichter war, nachträglich einen Text umzuschreiben bzw. anzupassen, als ein Bild anhand einer neuen Textvorlage zu verändern, wurde letztlich vor allem auf der Grundlage des Bildes bestimmt, welche Inhalte deskriptiv und welche Inhalte depiktorial dargeboten wurden. Daher mussten zuerst Abschnitte bzw. Elemente innerhalb der Bildseiten identifiziert werden, die für sich genommen, abgeschlossene Informationseinheiten darstellten. Anschließend wurde dann entschieden, welche Teile entfernt werden konnten, ohne die Bedeutung oder den Informationsgehalt der verbleibenden Bildelemente nachhaltig zu beeinflussen. Abbildung 5.2 zeigt die komplementäre Version der Bildseite *Plattengrenzen*.

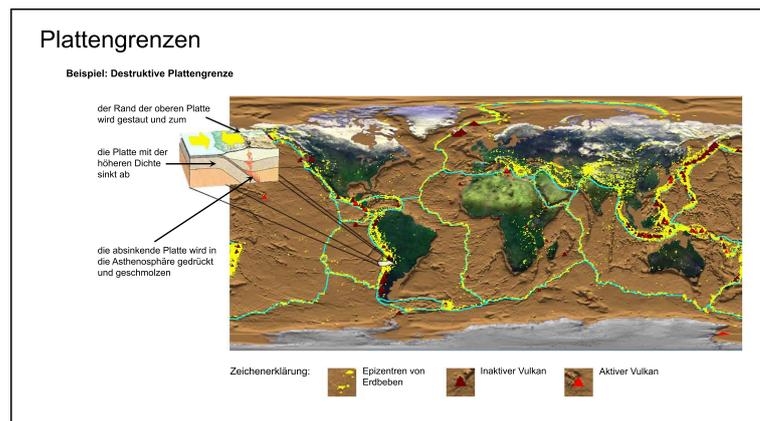


Abbildung 5.2: Bildseite *Plattengrenzen* als komplementäre Version

Vergleicht man diese mit der vollständigen Abbildung 5.1, so fällt auf, dass hier neben dem Hintergrundbild nur noch ein plattentektonischer Grenztyp zu erkennen ist. Die anderen beiden Grenztypen wurden samt Beschriftung entfernt, da sie für sich genommen

relativ unabhängige Komponenten darstellten. Nachdem alle Bildseiten auf ähnliche Weise *bearbeitet* worden waren, bestand nun der zweite Schritt darin, die Texte *anzupassen*. Zu diesem Zweck wurden alle Abschnitte, welche die im Bild verbliebenen Komponenten beschrieben, entfernt. Beibehalten wurden nur die Sätze, welche sich auf Bildelemente bezogen, die zuvor aus der entsprechenden grafischen Darstellung gelöscht worden waren. Dabei wurde darauf geachtet, exakt die Formulierungen zu verwenden, die auch der vollständige Text enthielt. So sollte gesichert werden, dass sich der Unterschied zwischen den beiden Textversionen allein auf die Anzahl der darin enthaltenen Sätze bzw. Informationen bezog. Da es schwierig war, den Informationsgehalt der Bilder zu quantifizieren, diente die Anzahl der verbliebenen Worte in den Texten als grobe Orientierungshilfe, um das Ausmaß der Kürzungen einzuschätzen. Bei der Gestaltung des komplementären Materials wurde angestrebt, die eine Hälfte der Informationen im Text und die andere Hälfte der Informationen im Bild darzustellen. Dies würde bei einer idealen Umsetzung dazu geführt haben, dass die komplementären Textabschnitte etwa die halbe Anzahl der Worte enthalten hätten, wie die vollständigen Versionen. Dieses Ziel konnte allerdings nicht realisiert werden, wie die folgende Tabelle 5.1 zeigt.

Tabelle 5.1: Textlänge der vollständigen und komplementären Versionen

Thema/Inhalt	Anzahl der Worte vollständige Version	Anzahl der Worte komplementäre Version
<i>I Die Theorie der Plattentektonik</i>	255	200
<i>II Bedeutung der Plattengrenzen</i>	264	137
<i>III Messungen im Inneren der Erde</i>	255	178
<i>IV Mechanismen der Plattenbewegung</i>	253	143

Es fällt auf, dass die vollständigen Texte im Durchschnitt etwas mehr als 250 Worte umfassten, die Textlänge der komplementären Version jedoch sehr viel stärker variierte. Diese Schwankungen der Textlänge hatte mehrere Ursachen. Entweder ließ der Aufbau des Bildes eine simple Halbierungen der Informationen nicht zu, oder aber die Struktur des Textes schränkte die Freiheitsgrade bei den Kürzungen ein. Immerhin sollten die komplementären Texte noch gut lesbar sein und nicht unvollständig oder wie ein Lückentext wirken. Hinzu kommt, dass die einzelnen Texte in unterschiedlichem Maße anschauliche und abstrakte Sachverhalte thematisierten. Während darauf verzichtet worden war, abstrakte Inhalte grafisch darzustellen, kamen für die Gestaltung der komplementären Versionen allein Textabschnitte in Frage, die sich auf anschauliche Inhalte bezogen. Sätze, die abstrakte Sachverhalte beschrieben, wurden dagegen unverändert übernommen.

Beispielsweise enthielt der Text *Die Theorie der Plattentektonik* einen Abschnitt mit einleitenden Worten, der die theoretische Bedeutung der Plattentektonik für die Geologie und Geophysik hervorhob. Diese Inhalte waren bei der Bildgestaltung nicht umgesetzt worden und konnten deshalb unverändert in der komplementären Version verwendet werden. Stattdessen bot es sich an, einen Textabschnitt zu entfernen, der viele Informationen beschrieb, die auch das Bild enthielt. Der Abschnitt war jedoch inhaltlich so eng mit

der *abstrakten* Einleitung verknüpft, dass er nicht herausgestrichen werden konnte, ohne große Teile des restlichen Textes neu zu schreiben. Eine *Umformulierung* kam allerdings nicht in Frage, da die komplementäre Version exakt die gleichen Sätze enthalten sollte, wie die vollständige Version. Also wurde eine andere Textpassage entfernt, die sehr viel kürzer war, obwohl die Struktur des Bildes prinzipiell eine gleichmäßigere Aufteilung der Informationen ermöglicht hätte. In diesem Zusammenhang wurde die Lesbarkeit der Bilder und Texte als wichtiger angesehen, als deren Länge oder relativer Informationsgehalt. Auch wenn es hinsichtlich der Fragestellung in erster Linie darum ging, die Komplementarität zwischen den deskriptiven und depiktionalen Repräsentationsformaten zu maximieren, wurde dennoch immer auf die Verständlichkeit des Lehrmaterials geachtet.

5.1.5 Evaluation des Lehrmaterials

Nachdem die Gestaltung des Lehrmaterials aus Entwicklersicht abgeschlossen war, sollten die Texte und Bilder vor einer Verwendung evaluiert werden. Ein zentrales Anliegen dieser Evaluation bestand darin, vorhandene inhaltliche Schwächen, orthographische Fehler und mögliche Verständnisprobleme zu identifizieren und zu beheben. Für diese Aufgabe wurden vier unabhängige *Experten* verpflichtet, von denen die ersten beiden ein Hochschulstudium der Germanistik (Hauptfach), die anderen beiden ein Studium der Geographie (Hauptfach) als *magister artium* bzw. mit dem ersten Staatsexamen abgeschlossen hatten. Während von den Germanisten eine kompetente Beurteilung der Qualität der Textoberfläche (Grammatik, Orthographie und Verständlichkeit) erwartet wurde, sollten die beiden Geographen vor allem die fachliche Richtigkeit der Inhalte und die Güte der Karten bzw. schematischen Darstellungen evaluieren.

Allen Experten wurde das vollständige Lehrmaterial in kongruenter und komplementärer Ausfertigung als HTML-Seiten an einem Rechner präsentiert. Zuvor hatten sie die Instruktion erhalten, das Lehrmaterial gründlich auf inhaltliche und formale Fehler zu prüfen. Jedem Experten lagen die Texte und Bilder zusätzlich als Ausdrucke vor, so dass alle Anmerkungen und Korrekturvorschläge mit einem Bleistift direkt in den Text oder die entsprechenden Bilder eingetragen werden konnten. Von kleineren Anmerkungen abgesehen, die ausschließlich vereinzelte orthographische Mängel betrafen, hatten aber weder die Geographen, noch die beiden Germanisten größere inhaltliche oder strukturelle Beanstandungen zu machen. Die Überarbeitung der fehlerhaften Stellen nahm daher wenig Zeit in Anspruch und konnte recht zügig abgeschlossen werden.

Wie in den Abschnitten zuvor bereits erläutert wurde, war die Entwicklung des Lehrmaterials weitgehend regelgeleitet erfolgt. Dennoch konnte allein anhand von Gestaltungsprinzipien und theoretischen Überlegungen nicht festgestellt werden, ob es tatsächlich gelungen war, komplementäre bzw. kongruente Text-Bild-Paare zu konstruieren. Da das Informationsverhältnis aber im Zentrum der wissenschaftliche Fragestellung stand, erschien es notwendig zu sein, diese Annahme im Vorfeld der eigentlichen Studie zu prüfen. Dem Anliegen lag die Überlegung zugrunde, dass alle entworfenen Texte-Bild-Paare, die sich auf das selbe inhaltliche Thema bezogen, zu einem gewissen Grad komplementär

tär sein mussten. Das ergab sich allein aus der Tatsache, dass nur anschauliche Inhalte grafisch umgesetzt worden waren, obwohl die Texte auch abstrakte Sachverhalte beinhalteten. So stellte sich letztlich nicht die Frage, ob das Informationsverhältnis zwischen den Texten und Bildern komplementär war, sondern welche Größenordnung diese komplementäre Beziehung hatte.

Der offensichtlichste Weg bestand darin, alle Informationen, die sowohl dem Bild als auch dem Text entnommen werden konnten, einfach zu zählen. Wäre dann die Anzahl der gemeinsamen Informationen durch die Anzahl aller Informationen, die beide Formate zusätzlich enthielten, geteilt worden, so hätte dieser Wert das Informationsverhältnis zwischen Text und Bild zum Ausdruck gebracht. Allerdings scheiterte der Versuch, sämtliche Informationen des Lehrmaterials zu erfassen, bereits an den Charakteristika der Bildseiten. Ähnlich wie bei einem Liniendiagramm, aus dem sich theoretisch beliebig viele Einzeldaten ablesen lassen, konnten allein der Weltkarte, die im Hintergrund von Abbildung 5.1 dargestellt ist, eine schier unendliche Menge von Detailinformationen entnommen werden. Mit Blick auf den thematischen Kontext und das Verständnis der dargestellten Sachverhalte spielen diese Detailinformationen aber im besten Fall eine mittelbare Rolle. Hinzu kommt, dass aufgrund der Mehrdeutigkeit von Bildern (Ballstaedt, 2005), eine vollständige Aufzählung von mehreren unabhängigen Personen hätte durchgeführt werden müssten, so dass der Aufwand in keinem Verhältnis zum den Ergebnissen gestanden hätte.

Aus diesem Grund wurde ein sehr viel einfacheres Vorgehen gewählt, um die Informationsrelation der Text-Bild-Paare in einen numerischen Wert zu überführen. Über ein Urteilsverfahren sollte demnach *geschätzt* werden, wie groß der Anteil der Inhalte war, die beiden Formaten entnommen werden konnten. Damit diese Einschätzung möglichst strukturiert erfolgte, wurde für jede Version und jeden Themenabschnitt auf der Grundlage der Texte acht Beurteilungsbögen erstellt, die zwischen 9 und 16 Items umfassten. Jedes Item bestand aus einem Satz (dem *Anker*) und einer Ratingskala, auf der eingeschätzt werden musste, wie viele der Informationen des Satzes in dem entsprechenden Bild abgelesen werden konnten. Als Anker wurden die Originalsätze der verschiedenen Textversionen verwendet, wobei darauf verzichtet worden war, die Sätze umzuformulieren oder in kleinere *Informationseinheiten* zu zerlegen. Auf diese Weise sollten Bedeutungsänderungen vermieden werden, was die Reliabilität der Urteile eingeschränkt hätte. Außerdem wurde die Reihenfolge der Sätze, die sie in den Textversionen hatten, in den Beurteilungsbögen beibehalten. Abbildung 5.3 zeigt eines der Items.

Die Beurteilung erfolgte in Einzelsitzungen und wurde von den selben vier *Experten* durchgeführt, die zuvor das Lehrmaterial auf formale und inhaltliche Fehler geprüft hatten. Jeder Experte erhielt kurz vor der Sitzung die Instruktion, sich bei der Einschätzung auf die Inhalte zu konzentrieren, die seiner Meinung nach für das Verständnis der Sachverhalte wesentlich waren. Zudem sollten sie versuchen, Informationen, die sich nicht unmittelbar ablesen ließen, auf der Grundlage des Bildes zu erschließen. Nachdem alle noch offenen Fragen geklärt worden waren, wurden die Bilder als HTML-Seiten an einem Computer präsentiert. Während dieser Präsentation trugen die Experten ihre Einschät-

Wie viele Informationen, die in Satz (10) enthalten sind, können aus **Bild_I_k** entnommen werden?

(10) *Die hohe Temperatur und das geschmolzene Gestein machen die Asthenosphäre flexibel und verformbar.*

keine <input type="checkbox"/>	wenige <input type="checkbox"/>	etwa die Hälfte <input type="checkbox"/>	viele <input type="checkbox"/>	alle <input type="checkbox"/>
-----------------------------------	------------------------------------	---	-----------------------------------	----------------------------------

Abbildung 5.3: Beispielitem aus dem Beurteilungsbogen zur Einschätzung des Informationsverhältnisses zwischen Bildern und Texten des Stimulusmaterials

zungen mit einem Bleistift direkt in die Beurteilungsbögen ein, die ihnen in Papierform vorlagen. Da für jede der acht Bildseiten ein eigener Beurteilungsbogen mit insgesamt 90 Sätzen bzw. Items ausgefüllt werden musste, dauerten die Sitzungen bis zu zwei Stunden, wobei immer wieder kleine Erholungspausen eingelegt wurden.

Als die Einschätzungen schließlich abgeschlossen waren, sollten für die Text-Bild-Paare die mittlere Skalenwerte berechnet werden. Durchschnittliche Einschätzungen mehrerer Urteiler sind in der Regel zwar reliabler und valider als Einzelurteile (vgl. Horowitz, Inouye & Siegelman, 1979; Strahan, 1980), doch setzt die Berechnung des Mittelwertes die hinreichende Übereinstimmung der Individualurteile voraus (Bortz, Lienert & Boehnke, 1990). Aus diesem Grund bot es sich, die *Urteilskonkordanz* (vgl. Bortz & Döring, 1995) als Indikator für die Übereinstimmung der Einschätzungen zu ermitteln. Da die Messwerte Ordinalskalenniveau hatten und es galt, die Reliabilität der Ratings über mehrere Urteiler hinweg zu bewerten, wurde Kendall's W (Kendall & Smith, 1939) als Konkordanzkoeffizient bestimmt.

Kendall's W kann Werte zwischen 0 (keine Übereinstimmung) und 1 (vollkommene Übereinstimmung) annehmen und ist ähnlich zu interpretieren, wie zum Beispiel das Reliabilitätsmaß α von Cronbach (1951). Zusätzlich kann für den Koeffizienten W eine χ^2 -verteilte Prüfgröße berechnet werden, welche die Konkordanzannahme gegen den Zufall absichert. Für die Urteile der vier Experten wurde ein $W = 0.87$ ermittelt, ein Wert, der nach Bortz und Döring (1995) einer hohen Reliabilität entspricht. Folglich war auch die Prüfgröße mit einem $\chi^2 = 311.9$ bei 89 Freiheitsgraden signifikant ($p < 0.001$), was eine rein zufällige Übereinstimmung der Urteile äußerst unwahrscheinlich machte. Um die kongruenten und komplementären Versionen zu vergleichen, wurde nach der Zusammenfassung der Urteile schließlich für jedes Text-Bild-Paar ein Durchschnittswert berechnet. Mit Blick auf die Hauptuntersuchung war es von entscheidender Bedeutung, dass die Experten bei den kongruenten Versionen möglichst viele Informationen, die in den Sätzen enthalten waren, auch in den Bildern wieder finden würden. Dagegen sollte bei den komplementären Bild-Text-Paaren, wenn überhaupt, nur vereinzelte Informationen sowohl im

Text, als auch im Bild enthalten sein.

Da Ballstaedt (2005) keine Angaben dazu macht, wann eine Informationsrelation zwischen deskriptiven und depiktionalen Formaten als kongruent und wann sie als komplementär angesehen werden muss, war es notwendig, eine eigenständige Festlegung vorzunehmen. Vor dem Hintergrund der verwendeten Ratingskalen (vgl. Abbildung 5.3) konnten die Mittelwerte, die das Informationsverhältnis eines Text-Bild-Paares wieder spiegelten, einen Wert zwischen 0 und 4 annehmen. Während ein Wert von 0 vollständiger Komplementarität entsprach, hätte ein Wert von 4 bedeutet, dass alle der in den Sätzen enthaltenen Informationen in dem Bild abgelesen werden konnten. Aufgrund des Vorgehens bei der Gestaltung des Lehrmaterials war ein Wert nahe Null für die komplementären Versionen durchaus möglich und mit Blick auf die Untersuchung auch erstrebenswert. Hingegen musste man ein vollständig kongruentes Verhältnis aufgrund der abstrakten Inhalte und des Umfangs der dargestellten Sachverhalte als unrealistisch einstufen. So erschien bereits ein durchschnittlicher Wert von ≥ 2 ausreichend zu sein, um von einem kongruenten Verhältnis zwischen Text und Bild auszugehen. Immerhin bedeutete dieser Wert, dass mehr als die Hälfte der Informationen eines Textes auch in den Bildern enthalten waren, was die Relation aufgrund des Mengenverhältnisses mehrheitlich kongruent gemacht hätte. Wie das Informationsverhältnis der Text-Bild-Paare von den vier Experten tatsächlich beurteilt wurde, zeigt Abbildung 5.4.

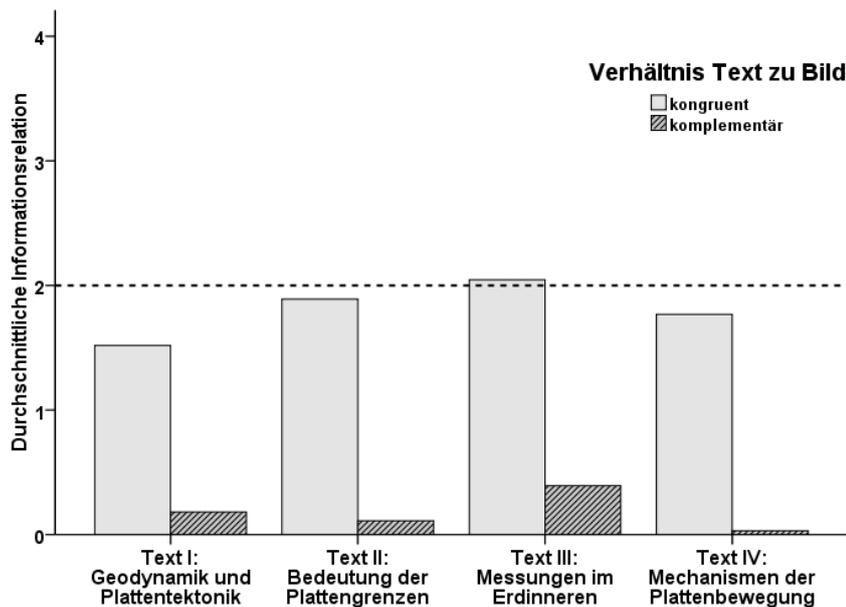


Abbildung 5.4: Zusammenfassung der Expertenurteile

Geordnet nach den vier thematischen Abschnitten wurden hier die kongruenten und komplementären Versionen nebeneinander dargestellt. Es fällt sofort auf, dass nur eine der insgesamt vier kongruenten Themenabschnitte das Kriterium eines durchschnittlichen Informationsverhältnisses von ≥ 2 , in Abbildung 5.4 durch eine gestrichelte Linie symbo-

lisiert, erfüllte. Ähnliches gilt für die kompletären Versionen, die alle mehr oder weniger deutlich über dem Wert 0 lagen. Dennoch waren die beiden Ausführenden hinsichtlich der Informationsrelationen zwischen Text und Bild augenscheinlich so verschieden, dass es gerechtfertigt war, von zwei Versionen auszugehen, die man als kongruent und komplementär bezeichnen konnte. Auf der Grundlage dieser Ergebnisse wurde schließlich darauf verzichtet, das Stimulusmaterial noch einmal zu überarbeiten.

5.2 Instrumente

Nachdem Entwurf und Evaluation des Lehr- bzw. Stimulusmaterials abgeschlossen waren, wurden die Instrumente zusammengestellt oder entwickelt, mit denen im Hinblick auf die Fragestellung relevante psychologische Zustände und Merkmale erfasst werden sollten. Alle Tests, die speziell für die Untersuchung konstruiert wurden, sind vollständig im Anhang B enthalten. Standardisierte Verfahren, sofern sie nicht speziell für die Untersuchung adaptiert oder verändert wurden, werden dagegen nicht im Anhang angeführt. Damit zu erkennen ist, ob die Instrumente vor oder nach der Präsentation des Stimulusmaterials bearbeitet werden mussten, sind die Überschriften mit einem in Klammern stehenden *prae* (= vor der Präsentation der Texte und Bilder) oder *post* (= nach der Präsentation der Texte und Bilder) versehen.

5.2.1 Lern- und Wissenstest (post)

Im Hinblick auf die wissenschaftliche Fragestellung war der Fragebogen zur Erfassung des Lernerfolgs das zentrale Instrument der Untersuchung, weshalb die Darstellung auch mit diesem beginnt. Da für das spezifische Lehrmaterial nicht auf ein bewährtes Instrument zurückgegriffen werden konnte, musste ein *Lerntest* erst entwickelt werden. Die Konstruktion des Instruments zur Erfassung von Lernergebnissen basierte weitgehend auf den theoretischen Überlegungen, die in Kapitel 2.1 abgehandelt wurden. Demnach lässt sich Lernen als Prozess des Wissenserwerbs definieren, der die kognitiven Leistungen Behalten und Verstehen umfasst und sich über diese auch messen lässt. Bei der Konzeption der Aufgaben, deren erfolgreiche Bearbeitung das Behalten von Fakten und Details erforderte, wurden zwei verschiedene Formate entwickelt, die unterschiedliche Teilleistungen abbilden sollten. Die erste Teilleistung war das *Wiedererkennen* von Sachverhalten, das möglichst ohne Schlussfolgern, also durch reines Erinnern, zu einer richtigen Lösung führte. Die Bezeichnung Wiedererkennen schien gerechtfertigt zu sein, weil bei diesem Aufgabentyp mehrere Antwortalternativen in einem Multiple Choice-Format vorgegeben wurden und die richtige Aussage bzw. Antwort wiedererkannt werden konnte. Abbildung 5.5 zeigt eine Beispielaufgabe.

Um die Frage richtig zu beantworten, musste die Versuchsperson sich an einen Ausschnitt aus dem ersten Text bzw. Bild erinnern, die das Ausmaß der Bewegung der Lithosphärenplatten thematisierten. Wie leicht zu erkennen ist, wurden bei dieser Aufgabe drei mögliche Antworten vorgegeben. Die Anzahl Antwortalternativen variierte zwischen

Um wie viele cm bewegen sich die Gesteinsplatten jährlich?

- 1,7 bis 19,4 cm
 - 4,1 bis 17,2 cm
 - 1,7 bis 17,2 cm
-

Abbildung 5.5: MC-Item zur Erfassung der Wiedererkennensleistung

zwei und vier, je nachdem wie viele sinnvolle Aussagen auf der Basis des Lernmaterials gebildet werden konnten. Um einer Verzerrung der Ergebnisse durch ein zufälliges Antwortverhalten vorzubeugen, enthielt die Instruktion zur Bearbeitung des Lerntests unter anderem die Bitte, nicht zu raten und nur dann ein Kreuzchen zu machen, wenn die Versuchsteilnehmer davon überzeugt waren, die richtige Lösung zu wissen.

Diese Problematik war bei dem zweiten Aufgabentyp, der das freie Erinnern erfassen sollte, nicht gegeben. Hier wurde von den Versuchsteilnehmern verlangt, Fragen zu den Texten und Bildern zu beantworten, ohne dass Vorgaben den Erinnerungsprozess erleichtern konnten. Abbildung 5.6 zeigt ein Item mit offenem Format, zu dessen Lösung Inhalte aus dem Gedächtnis abgerufen werden mussten, die im dritten thematischen Abschnitt enthalten waren.

Wie bezeichnet man die Methode, mit der Messungen im Erdinneren gemacht werden können?

Abbildung 5.6: Item zur Erfassung des freien Erinnerns

Auch wenn beide Antwortformate in erster Linie auf den *einfachen* Abruf von Gedächtnisinhalten abzielten, wurden diese auf unterschiedlichen Skalen angeordnet. So enthielt die Skala *Wiedererkennen* (recognition) alle MC-Items, deren Lösung kein schlussfolgerndes Denken erforderte, während alle Gedächtnis-Aufgaben mit offenem Format, der Skala *freies* Erinnern (recall) zugeordnet wurden. Für dieses Vorgehen sprachen empirische Arbeiten, die auf qualitative Unterschiede zwischen Wiedererkennen und freiem Erinnern hindeuten. So encodierten Lernende Text-Bild-Stimuli unterschiedlich, je nachdem ob sie erwarteten einen Test zum Wiedererkennen oder freiem Erinnern bearbeiten zu müssen (Tversky, 1973). Darüber hinaus legen neuro-psychologische Studien nahe, dass beim freien Erinnern und beim Wiedererkennen unterschiedliche Hirnareale aktiviert werden (Staresina & Davachi, 2006), was für eine Beteiligung unterschiedlicher kognitiver Prozesse spricht.

Natürlich ist eine scharfe Abgrenzung der beiden Gedächtnisleistungen durch die beschriebenen Aufgabenformate nicht vollständig zu realisieren. Beispielsweise ist denkbar, dass ein Versuchsteilnehmer das Item in 5.5 beantwortet, indem er/sie die Frage

5 Untersuchungsmaterial und Instrumente

liest und ohne auf die Distraktoren zu achten, die Lösung *frei* generiert. Zudem legen J. R. Anderson und Brown (1972) dar, dass Wiedererkennen ein Teilprozess des freien Erinnerns ist. So müssen Worte und Konzepte zuerst einmal wiedererkannt bzw. verstanden werden, um das Item aus 5.6 zu lösen zu können. Zudem sind Aufgaben, die freies Erinnern erfordern in der Regel einfach schwerer, weil Lerninhalte ohne konkreten Hinweisreiz abgerufen werden müssen (Haist, Shimamura & Squire, 1992). Letztlich wurde der Konzeption der beiden Gedächtnis-Skalen die Arbeitsdefinition zugrunde gelegt, dass die Bearbeitung der MC-Items in erster Linie Prozesse des Wiedererkennens erfordert, während Items mit offenem Format dem Versuchsteilnehmer zusätzliche Gedächtnisleistungen abverlangen.

Eine ähnliche Problematik bestand bei der Konstruktion der Items die das tiefere Verständnis erfassen sollte. In Abgrenzung zu den (reinen) Gedächtnisaufgaben reichte es bei *Verständnisitems* nicht mehr aus, allein Text- oder Bildinhalte zu erinnern bzw. wiederzugeben. Bei diesem Aufgabentyp war es ein zentrales Merkmal, dass eine richtige Antwort Schlussfolgerungen bzw. Inferenzen voraussetzte. Abbildung 5.7 zeigt ein Verständnis-Item mit Multiple-Choice-Format.

Eine Erdbebenwelle bewegt sich von der äußeren Kruste exakt auf den Mittelpunkt der Erde zu...

- ...sie wird in Abhängigkeit der Konvektionsströme tendenziell langsamer
- ...sie bewegt sich mit konstanter Geschwindigkeit auf den Erdkern zu
- ...sie wird in Abhängigkeit der Konvektionsströme tendenziell schneller

Abbildung 5.7: MC-Item zur Erfassung der Fähigkeit Inferenzen zu bilden

Um die richtige Antwortalternative auszuwählen, mussten Sachverhalte erinnert werden, die im dritten Bild und/oder Text erläutert und dargestellt wurden. Diesen konnte entnommen werden, dass Erdbebenwellen, wenn sie durch heißes Gestein wandern, an Geschwindigkeit verlieren. Zusätzlich musste der Versuchsteilnehmer wissen bzw. gelernt haben, dass das innere der Erde (wahrscheinlich (vg. Powell, 1995)) aus verschiedenen Schalen besteht und die Temperatur stetig zunimmt je näher man dem Erdmittelpunkt kommt. Durch die Kombination beider Informationen ließ sich dann schließen, dass Erdbebenwellen, wenn diese sich genau auf den Mittelpunkt der Erde zubewegen, tendenziell langsamer werden.

Während MC-Items, wie in Abbildung 5.7 eine einfache Auswertung versprachen und den Vorteil hatten, dass sie den Versuchsteilnehmer in seinem Antwortverhalten lenkten, bestand bei den offenen Items die Gefahr, dass die oft kurz gehaltenen Fragen missverstanden wurde und die Antwort nicht auf den Gegenstand abzielte, den sie eigentlich adressierte. Andererseits war anzunehmen, dass MC-Items deutlich leichter waren als das offene Format, selbst wenn die Teilnehmer den Anweisungen des Versuchsleiters nachkamen und die Frage nur dann beantworteten, wenn sie glaubten die Antwort zu

wissen. Aus diesem Grund wurden zusätzlich Inferenzaufgaben mit offenem Format in den Lerntest aufgenommen. Abbildung 5.8 zeigt eines dieser Items mit offenem Format.

Warum sind in der Geodynamik die Plattengrenzen besonders wichtig?

Abbildung 5.8: Item mit offenem Format zur Erfassung der Fähigkeit Inferenzen zu bilden

Die Frage bezog sich auf die Inhalte des zweiten Textes bzw. Bildes. Als Antwort sollten die Versuchsteilnehmer kurz ausführen, dass die meisten Phänomene der Geodynamik nur an den Plattengrenzen beobachtet werden können und Rückschlüsse über den inneren Aufbau der Erde allein über diese Phänomene möglich ist. Allerdings gibt es auf diese Frage auch andere *richtige* Antworten, die auf elaborierte Wissensstrukturen schließen lassen. Beispielsweise könnte in etwas abgewandelter Form argumentiert werden, dass die meisten Erdbeben an Plattengrenzen auftreten und nur das Vorhandensein von Erdenbenenwellen Messungen im Erdinneren erst möglich macht.

Insgesamt wurden 40 Items bzw. Fragen konzipiert, welche die Versuchsteilnehmer im Anschluss an eine Lernphase mit dem bereits beschriebenen Lehr-Material beantworten sollten. Vor einer Verwendung im Rahmen der Hauptuntersuchung, wurde der Test zuerst 20 Studenten vorgelegt, nachdem diesen das finale Lehrmaterial in kongruenter und komplementärer Form an einem Rechner präsentiert worden war. Anschließend wurden die Teilnehmer dazu aufgefordert, Fehler und Probleme bei der Bearbeitung zu nennen, was sich allerdings als wenig ergiebig erwies, da die Studenten den Test für machbar und nicht zu schwer hielten. Um einen groben Eindruck von der Reliabilität der Arbeitsversion des Wissenstests zu erhalten, wurden die Itemschwierigkeiten, die Trennschärfen mit *part-whole*-Korrektur und die interne Konsistenz der Skalen berechnet (Cronbach, 1951). Tabelle 5.2 gibt einen Teil der Ergebnisse dieser Skalen-Analyse wieder.

Tabelle 5.2: Interne Konsistenz und durchschnittliche Itemschwierigkeit der ersten Version des Lern- und Wissenstests

Skala	Cronbachs α	$\bar{p} = \frac{\sum_{k=1}^I x_k}{I}$
Wiedererkennen	.385	.42
freies Erinnern	.593	.35
Inferenz MC-Format	.354	.25
Inferenz offenes Format	.294	.31

Bei der Auswertung fiel auf, dass bis auf die Skala zum freien Erinnern, alle übrigen sehr niedrige interne Konsistenzen aufwiesen und damit schlichtweg unreliabel waren. Dies galt umso mehr, wenn man berücksichtigte, dass beiden Verstehensskalen erst dann einen positiven α -Wert annahmen, nachdem jeweils das Item mit der negativsten

5 Untersuchungsmaterial und Instrumente

Trennschärfe aus der Berechnung ausgeschlossen worden war. Dieses Ergebnis machte deutlich, dass einige Aufgaben überarbeitet bzw. ersetzt werden mussten. Dabei richtet sich die Aufmerksamkeit vor allem auf Items mit negativer Trennschärfe oder einer Trennschärfe von $p \leq .10$. Insgesamt wurden 17 der 40 Items verbessert bzw. neu entwickelt, wobei nur in wenigen Fällen offensichtliche Fehler für die mangelhafte Güte identifiziert werden konnten. Auf eine erneute Evaluation des Testes wurde in Anbetracht der hohen Anzahl von Items und der damit verbundenen Möglichkeit, einzelne fehlerhafte Aufgaben der Analyse ausschließen zu können, verzichtet. Die überarbeitete Version, die in allen weiteren Untersuchungen eingesetzt wurde, ist vollständig im Anhang B enthalten.

Um einschätzen zu können, inwieweit der finale Wissens- und Lerntest den drei wichtigsten Kriterien der klassischen Testtheorie (vgl. Rost, 2004) entsprach, wurden die Objektivität, Reliabilität und Validität des Instruments evaluiert. Die ermittelten Gütekriterien basieren auf den Antworten sämtlicher Teilnehmer der Studien, die in den folgenden Kapiteln noch ausführlich erläutert werden. Zunächst gilt eine hohe Durchführungs- und Auswertungsobjektivität als Voraussetzung dafür, dass ein Test reliabel und valide ist, weshalb dieses Gütekriterium zuerst überprüft wurde. Während die Durchführungsobjektivität gesichert werden konnte, indem eine ausführliche Bearbeitungsanweisung von dem Testleiter verlesen wurde, gestaltete sich die Gewährleistung der Auswertungsobjektivität etwas komplizierter. Die Auswertung der Multiple-Choice-Items stellte kein großes Problem dar, denn entweder hatte der Versuchsteilnehmer sein Kreuzchen an der richtigen Stelle gemacht oder nicht. Dagegen war die Beurteilung der Antworten auf die offenen Fragen mit sehr viel mehr Freiheitsgraden verbunden. Damit die Auswertung der offenen Formate auf einer objektiven Grundlage stand, wurde ein Kategorienschema mit Musterantworten entwickelt (Rost, 2004), anhand dessen die *Signierobjektivität* von zwei unabhängigen Beurteilern bestimmt werden konnte. Als Maß der Übereinstimmung wurde der Koeffizient κ von J. Cohen (1960) bestimmt. Tabelle 5.3 zeigt, dass die Kennwerte für die Auswertungsobjektivität aller Items eine hohe bis mittlere Ausprägung hatten (Bühner, 2006, S.140).

Tabelle 5.3: Auswertungsobjektivität aller Items mit offenem Format

Skala	Objektivität		
	niedrig $\kappa < .60$	mittel $\kappa = .60 - .90$	hoch $\kappa > .90$
freies Erinnern	–	1	11
Inferenz offenes Format	–	3	1

Da das κ bei keinem Item unter den Wert von .60 fiel, wurde das Gütekriterium der *Objektivität* als weitgehend erfüllt angesehen. Die exakten Kennwerte der Urteilsübereinstimmung sind für jedes einzelne (offene) Item tabellarisch im Anhang nachzulesen. Diese Tabelle enthält zudem alle weiteren Kennwerte, die im Rahmen der finalen Skalenkonstruktion bestimmt wurden. So bot sich vor allem die Trennschärfe als Kriterium bei

der Selektion *guter* Items an, doch ist dieser Kennwert nicht so einfach zu interpretieren, wie es auf den ersten Blick scheint (Bühner, 2006; Rost, 2004). Um zu verhindern, dass eine Auswahl nach hohen Trennschärfen zu einer Überrepräsentation von Fragen mittlerer Schwierigkeit führte, wurde mit S_{el} ein weiteres Selektionskriterium berechnet (vgl. Amelang & Zielinski, 1997). Durch diesen zusätzlichen Kennwert sollte der Umstand korrigiert werden, dass die Trennschärfe r_{it} für Item mit einer mittleren Schwierigkeit p , generell höher ist, als für Items mit extremen Schwierigkeiten. In Anbetracht der Tatsache, dass der Anzahl der richtig beantworteten Items von einer Vielzahl von Faktoren abhing, die neben den experimentell variierten Bedingungen, natürlich auch Personen und Situationsvariablen umfassten, wurde ein sehr liberaler cut-off-Wert bestimmt. Ein Item wurde erst dann aus der jeweiligen Skala und damit von allen weiteren statistischen Analysen ausgeschlossen, wenn $S_{el} < .10$ war. Tabelle 5.4 fasst die vier Skalen des Lern- und Wissenstests zusammen, die in den drei Untersuchungen zur Anwendung kamen.

Tabelle 5.4: Überblick über die Skalen des Wissenstests

Skalenbezeichnung	α	$\bar{r}_{ii'}$	\bar{x}	$s_{\bar{x}}$	N_I
Wiedererkennen	.50	.09	6.2	1.9	10 (12)
freies Erinnern	.62	.12	4.4	2.4	11 (12)
Inferenz MC	.56	.11	4.0	2.0	10 (12)
Inferenz offen	.38	.12	1.1	1.0	4 (4)
Gesamtskala	.76	.08	16.0	5.1	35 (40)

Anm.: α = Maß der internen Konsistenz nach Cronbach, $\bar{r}_{ii'}$ = Iteminterkorrelation, \bar{x} = Skalenmittelwert, $s_{\bar{x}}$ = Standardabweichung des Skalenmittelwerts, N_I = Anzahl der Items, die in der Skala verblieben sind (ursprüngliche Anzahl)

Wie man unschwer erkennen kann, führten die Reliabilitätsschätzungen mit Cronbachs- α auf den ersten Blick zu keinem guten Ergebnis. Dies betraf speziell die Skala *Inferenz offen*, deren α -Koeffizient mit .38 am niedrigsten war, wobei natürlich berücksichtigt werden musste, dass sie nur aus vier Items bestand. Aber auch die anderen Skalen würden nach den Beurteilungsrichtlinien gängiger Lehrbücher zur Testkonstruktion (Bühner, 2006) bzw. Forschungsmethodik (Bortz & Döring, 2006) als unreliabel und damit schlichtweg als schlecht gelten. In diesem Zusammenhang musste allerdings bedacht werden, dass sich diese Vorgaben vor allem auf homogene Konstrukte beziehen, die einen meist ziemlich eng definierten Verhaltensausschnitt umreißen. Dagegen stellte die Lernleistung, wie sie im Rahmen der Untersuchungen erfasst wurde, ein sehr viel heterogeneres Konstrukt dar. Jedes einzelne Item bezog sich auf einen anderen inhaltlichen Abschnitt der Texte bzw. Bilder, so dass der Zusammenhang zwischen den Items des Wissenstest folgerichtig schwächer war, als dies zum Beispiel bei Aufgaben zur Erfassung räumlicher Fähigkeiten der Fall ist. Diese inhaltliche Heterogenität der Skalen drückt

5 Untersuchungsmaterial und Instrumente

sich auch in deren mittleren Iteminterkorrelationen ($\bar{r}_{ii'}$) aus, welche die bei eindimensionalen Konstrukten empfohlene Mindestausprägung von $\bar{r}_{ii'} \geq .20$ (Bortz & Döring, 2006) klar unterschritten.

Besonders deutlich wird die Heterogenität des Konstrukts *Lernerfolg*, wenn man den Koeffizienten der *Gesamtskala* betrachtet. Neben den variierenden Inhalten, schränkten sehr wahrscheinlich die verschiedenen Aufgabenformate die Homogenität zusätzlich ein, so dass $\bar{r}_{ii'} = .08$ für die Gesamtskala am niedrigsten ausfiel. Im Grunde ist auch Cronbachs α ein Homogenitätsindex (vgl. Bortz & Döring, 2006), weshalb die in Tabelle 5.4 angeführten Werte keineswegs für eine *unzureichende* Qualität des Instruments sprechen müssen. Die Heterogenität der Skalen (und deren Items) war vielmehr ein Zeichen dafür, dass der Test die umfangreichen Inhalte des Stimulumsmaterials in seiner Bandbreite erfasste. Vor diesem Hintergrund wurden die Ergebnisse der Reliabilitätsschätzung als substantiell und letztlich auch als zufriedenstellend betrachtet.

Mit Blick auf die Skalenmittelwerte fiel auf, dass der Wissenstest in etwa durchschnittlich schwer war. Zwar besagt das arithmetische Mittel (\bar{x}) der *Gesamtskala*, dass mit 16 von 35 weniger als die Hälfte der Aufgaben gelöst wurden, doch lagen dieser Berechnung auch Werte von Personen zugrunde, die nicht das vollständige Lehrmaterial erhalten hatten. Alle in Tabelle 5.4 angeführten Kennwerte beruhen auf Items, die in die weiteren Analysen eingingen. Dabei mussten insgesamt fünf Items von der Berechnung der einzelnen Leistungsskalen ausgeschlossen werden. Bei zweien war die Trennschärfe negativ, während zwei andere Items das Kriterium $S_{el} > .10$ verfehlten. Das fünfte Item konnte nicht weiter verwendet werden, da es einen konzeptionellen Fehler enthielt. Darüber hinaus wäre es möglich gewesen, die Homogenität der Skalen durch die Eliminierung weiterer Items, deren Trennschärfen ziemlich gering $r_{itc} \leq .20$ waren, zu erhöhen. Dies hätte zwar zur Folge gehabt, dass die interne Konsistenz der Skalen und damit Cronbachs α gestiegen wäre, doch wurde dem Ziel, das Wissen der Versuchsteilnehmer möglichst breit zu erfassen, größere Bedeutung beigemessen.

Damit wurde bislang dargelegt, dass der Wissenstest unabhängig (*objektiv*) ist und (irgend) ein Konstrukt hinreichend genau (*reliabel*) erfasst, doch wurde noch nicht geklärt, was überhaupt gemessen wird. Um zu bestimmen, ob die Anzahl der richtig bearbeiteten Items tatsächlich dem Lernerfolg entsprachen, musste abschließend die Validität bzw. Gültigkeit des Wissenstests bestimmt werden. Per Definition (Amelang & Zielinski, 1997) ist ein Test dann valide, wenn das interessierende Merkmal, das er zu messen vorgibt, auch wirklich erfasst wird. Allerdings werden verschiedene Arten der Validität unterschieden, wobei die Unterteilung in *Inhaltsvalidität*, *Kriteriumsvalidität* und *Konstruktvalidität* am bekanntesten ist (vgl. Bühner, 2006; Amelang & Zielinski, 1997; Bortz & Döring, 1995). Murphy und Davidshofer (2001) weisen aber darauf hin, dass nur die *Inhaltsvalidität* der obigen Definition entspricht. Die inhaltliche Validität eines Instruments wird in der Regel aber nicht dadurch geschätzt, indem man einen Koeffizienten bzw. einen Kennwert errechnet. Vielmehr entscheiden logische, theoretische und fachliche Überlegungen darüber, ob der Test das als inhaltlich valide angesehen werden kann oder nicht. Auf dieser Grundlage konnte der Wissenstest in seiner vorliegenden Form durchaus als inhaltlich

valide angesehen werden.

Der Test sollte das Wissen erfassen, das zuvor auf der Grundlage verschiedener Texte und Bilder gelernt worden war. Da dieser aus Fragen bestand, die sich anhand der Inhalte des Lehrmaterials beantworten ließen, erfüllte der Test genau diesen Zweck. Legte man den Daten gleichzeitig das Binominalmodell (als Testmodell des klassischen Testtheorie) zugrunde (vgl. Rost, 2004), so konnte die Anzahl der richtig beantworteter Test-Items als ein valider Schätzer für den Lernerfolg des jeweiligen Versuchsteilnehmers angesehen werden. So wird niemand die Tatsache anzweifeln, dass jemand, der viel zum Thema Geodynamik weiss, mehr Fragen beantworten kann, als jemand, der sich nur schlecht in dieser Domäne auskennt. Allerdings geht man bei diesen Überlegungen von einem Repräsentationsschluss aus. So wird angenommen, dass die Items, die in dem Test enthalten sind, eine repräsentative Menge aller (theoretisch) möglichen Items darstellen, deren Gesamtheit das interessierende Merkmal vollständig erfassen. Der Nachweis eines gültigen Repräsentationsschlusses ist in der Praxis aber äußerst schwierig, wie Murphy und Davidshofer (2001) anhand eines Beispiels erläutern. Demnach kann nicht ausgeschlossen werden, dass selbst auf der Grundlage der *noch* recht übersichtlichen Lernmaterialien ganz andere Items hätten konstruiert werden können, die das Konstrukt repräsentativer messen würden.

Die Autoren schlagen stattdessen vor, die Inhaltsebene dieses Konstrukts zu beschreiben, festzulegen, welcher Inhaltsbereich durch welches Items erfasst wird und schließlich die Übereinstimmung von Teststruktur und der Struktur des Konstrukts zu überprüfen. Diese Prüfung erfolgt in der Regel anhand einer *konfirmatorischen Faktorenanalyse*, wobei auf eine Durchführung aufgrund des ungünstigen Verhältnisses zwischen Itemanzahl und Stichprobengröße verzichtet werden musste (Marsh, Hau, Balla & Grayson, 1998). Um dennoch einen Eindruck von der *Struktur* des Tests zu gewinnen, bot sich die Berechnung der Korrelationen zwischen den Skalen an, deren Ergebnis Tabelle 5.5 enthält.

Tabelle 5.5: Korrelationen der Skalen des Wissenstests untereinander

	Gedächtnis MC	Gedächtnis offen	Inferenz MC
Gedächtnis offen	.47**		
Inferenz MC	.29**	.46**	
Inferenz offen	.28**	.38**	.22**

Anm.: **. Die Korrelation ist auf einem Niveau von .01 einseitig signifikant

Beim ersten Blick auf die Korrelationskoeffizienten fällt auf, dass diese zwar alle signifikant waren, aber stellenweise geringer als erwartet ausfielen. Insbesondere erstaunt der niedrige Zusammenhang zwischen den beiden Inferenzskalen, für den es aber verschiedene Erklärungen gibt. Zum einen war gerade bei der Formulierung dieser Aufgaben darauf geachtet worden, dass sie inhaltlich möglichst wenige Überschneidungen aufwiesen. Solche Überschneidungen bestanden aber zwischen den Gedächtnis-Items und den

5 Untersuchungsmaterial und Instrumente

Inferenzaufgaben, weshalb hier die Korrelationen deutlich höher waren. Beispielsweise wurde in Item 40 nach den Gründen gefragt, weshalb Magma aus dem Erdinneren an die Oberfläche aufsteigt. Einzelne Informationen, die zur Beantwortung der Aufgabe erinnert und kombiniert werden mussten, wurden gleichzeitig aber mit Items der Gedächtnisskalen abgefragt. Ein weiterer Grund ist sicherlich der Tatsache geschuldet, dass die Skala *Inferenz offen* nur aus vier Items bestand, die sehr schwer waren. Beiden Umstände beschränkten möglicherweise die Varianz der Skala, was dann in niedrigen Korrelationskoeffizienten resultierte. Zwar wäre es möglich gewesen, zusätzlich die minderungskorrigierten Korrelationen (Amelang & Zielinski, 1997) als Schätzer der wahren Zusammenhänge zu berechnen, da aber die Reliabilitätskoeffizienten aller Skalen (vgl. 5.4 so niedrig waren, dass dies zu drastischen Erhöhungen der Koeffizienten geführt hätte (*Verdünnungsparadox*), wurde darauf verzichtet. Stattdessen stand die Überlegung im Raum, die vier Inferenzitems mit offenem Format von allen weiteren Analysen auszuschließen und damit die ganze Skala zu eliminieren. Dies hätte natürlich Informationsverluste nach sich gezogen, weshalb eine Ausschlussentscheidung an die weiteren Ergebnisse des Validierungsprozesses geknüpft werden sollte.

Aufgrund der bereits erwähnten Probleme mit dem Nachweis der Inhaltsvalidität, greifen Forscher in vielen Fällen auf externe Variablen zurück, um indirekt auf die Gültigkeit ihres Instruments zu schließen. Sowohl die Kriteriums- als auch Konstruktvalidität benötigen zu ihrer Bestimmung solche Außenvariablen, weshalb sie Rost (2004) unter dem Begriff *externer Validität* zusammenfasst. Der große Vorteil bei der Bestimmung der externen Validität besteht darin, dass man sie berechnen kann. Zu diesem Zweck wählt man *Kriteriumsvariablen* aus, mit denen man die Werte des eigenen Tests dann korreliert werden. Natürlich müssen diese Korrelationen bzw. Validitätskoeffizienten an Hypothesen über erwartete Zusammenhänge gebunden sein. So sollte ein Test mit Instrumenten, die ein ähnliches Konstrukt erfassen, möglichst hoch (*konvergent valide*) und mit Instrumenten, bei denen kein Zusammenhang vermutet wird, möglichst niedrig (*divergent valide*) korrelieren. Beispielsweise kann angenommen werden, dass Personen, die gut in einem Intelligenztest abschneiden, auch in anderen Intelligenztests erfolgreich sind.

Um die zweifellos vorhandene *Augenscheinvalidität* des Lern- und Wissenstests auch empirisch zu stützen, wurden dessen vier Skalen mit verschiedenen Kriteriumsvariablen korreliert. Dabei war es naheliegend Merkmale heranzuziehen, die zur statistischen Kontrolle vor oder nach den Sitzungen erhoben wurden und nicht unmittelbar die wissenschaftliche Fragestellung betrafen. Die Variablen *Interesse*, *Schulnote Erdkunde*, *Vorwissen*, *Motivation* sowie *räumliche* und *verbale Fähigkeiten* waren ausgewählt worden, da die begründete Vermutung bestand, dass sie den Wissenerwerb mit den Lehrmaterialien in ihrer bestehenden Form beeinflussen würden. Die vorhergesagten Zusammenhänge sollten allerdings ausschließlich positiver Natur sein, so dass keine Informationen zur divergenten Validität des Tests gesammelt werden konnten. Für die Variablen *Interesse*, *Schulnote Erdkunde*, *Motivation* sowie *räumliche* und *verbale Fähigkeiten* wurden dabei eher niedrige Korrelationskoeffizienten erwartet, da sie nur indirekt mit dem Konstrukt zusammenhängen. Im Gegensatz dazu, sollte das *Vorwissen* mindestens in mittlerer Aus-

prägung mit den Skalen des Wissenstests korrelieren und damit den stärksten Zusammenhang aufweisen. Tabelle 5.6 enthält die Ergebnisse der Prüfung der konvergenten Validität.

Tabelle 5.6: Korrelationskoeffizienten zur Prüfung der konvergenten Validität des Lern- und Wissenstest

	Interesse	Note Erdkunde	Vorwissen Geologie	aktuelle Motivation	Verbale Fähigkeiten	Räumliche Fähigkeiten
Gedächtnis MC	.21**	.05	.27**	.18**	.12**	.14*
Gedächtnis offen	.15*	.08	.35**	.14*	.20**	.22**
Inferenz MC	.27**	.14*	.26**	.33**	.20**	.21**
Inferenz offen	.21**	.12	.17*	.06	.11	.16*
Gesamtskala	.28**	.13*	.38**	.26**	.22*	.25*

Anm.: Die Korrelation ist auf einem Niveau von * .05 oder ** .01 einseitig signifikant

Wie man erkennen kann, korrelierten die Skalen des Wissenstests fast durchgehend mit den sechs Kriteriumsvariablen. Die Zusammenhänge waren zwar alle recht niedrig ausgeprägt, doch wurde dies auch weitgehend in dieser Form erwartet. Den Vermutungen entsprach zudem die Tatsache, dass Vorwissen das Kriterium war, mit dem der Wissenstest am höchsten korrelierte, was ein Vergleich der Validitätskoeffizienten der Gesamtskala verdeutlicht. Der geringe Zusammenhang der Skalen mit den Erdkundenoten mutet dagegen nur auf den ersten Blick seltsam an. Immerhin ist der innere Aufbau der Erde, neben Wirtschaftsgeographie und Länderkunde nur eines von vielen thematischen Gebieten, die unterrichtet werden. Zudem kann es von Bundesland zu Bundesland verschieden sein, ob und in welcher Klassenstufe Geologie auf dem Lehrplan steht. Darüber hinaus waren die Teilnehmer zum Zeitpunkt der Untersuchung bereits Studenten und hatten seit längerem keinen Geographie-Unterricht mehr besucht. Die Angaben zu den Erdkundenoten beziehen sich damit auf Ereignisse, die möglicherweise mehrere Jahre zurück liegen, was deren Brauchbarkeit natürlich einschränkt.

Viel wichtiger als die Erdkundenote war für den Wissenstest die Frage, was mit der Skala *Inferenz offen* geschehen sollte. Die Ergebnisse in Tabelle 5.6 zeigen, dass niedrige aber signifikante Korrelationen mit den Kriterien *Interesse*, *Vorwissen* und den *räumlichen Fähigkeiten* nachgewiesen werden konnte, was letztlich als Beleg für deren Brauchbarkeit angesehen wurde. Dies hatte schließlich zur Konsequenz, dass die Skala nicht eliminiert und stattdessen in den weiteren Analysen verwendet werden sollte. Die Prüfung der Gütekriterien der KTT kam damit abschließend zu dem Ergebnis, dass die überarbeitete Version des Lern- und Wissenstests ein hinreichend objektives, reliables und valides Instrument darstellte, um Wissen zu erfassen, das die Versuchsteilnehmer während der Präsentation des Lehrmaterials erworben hatten.

5.2.2 Mentale Belastung (post)

Ein Merkmal, von dem angenommen werden konnte, dass es unmittelbar mit dem Wissenserwerb zu tun hatte, ist die mentale Belastung (*cognitive load*), welche die Teilnehmer empfanden. Durch Berücksichtigung dieser Variable war es möglich, insbesondere die Leistungsergebnisse zusätzlich vor dem Hintergrund der CLT (Sweller & Chandler, 1994) zu interpretieren, wodurch die Studie auf eine breitere theoretische Basis gestellt wurde. Allerdings konnte die mentale Belastung im Rahmen der Untersuchung während der Lernens auftreten oder bei der Bearbeitung des Wissenstests, so dass es angebracht war, sie in beiden Fällen zu erheben. Um die Ursache der mentalen Belastung näher zu bestimmen, wurde zudem nach der empfundenen Schwierigkeit der Texte, Bilder oder Items gefragt. So erhielten die Teilnehmer direkt nach der Lernphase und am Ende des Wissenstests entsprechende Ratingskalen, auf denen sie retrospektiv angeben konnten, wie sehr sie sich angestrengt hatten und wie schwierig sie die Bilder, Texte oder die Fragen des Lerntests empfunden hatten. Abbildung 5.9 zeigt eines der Items, die in Anlehnung an Paas, van Merriënboer und Adam (1994) die mentale Belastung während des Wissenserwerbs erfassen sollten.

Wie stark haben Sie sich beim Lernen der Texte und Bilder angestrengt?

sehr wenig ○ ○ ○ ○ ○ ○ sehr stark

Abbildung 5.9: Item zur Erfassung der mentalen Belastung während der Lernphase

Diese Operationalisierung gründet auf der Annahme, dass Lernende in der Lage sind, kognitive Prozesse wahrzunehmen und über damit verbundene mentale Anstrengungen zu berichten (Paas & van Merriënboer, 1994). Beispielsweise berichten Gopher und Browne (1984) von Probanden, die ohne Probleme relevante kognitive Prozess während des Wissenserwerbs identifizieren und dem damit verbunden mentalen Aufwand einen numerischen Wert zuordnen konnten. Weitere Vorteile dieser Methode sind die einfache Umsetzung, der geringe Beanspruchung der Teilnehmer im Vergleich zu physiologischen Messungen und nicht zuletzt die hohe Augenscheinvalidität (O'Donnel & Eggemeier, 1986). Allerdings wurde die Anzahl der Skalenstufen von ursprünglich neun auf sechs reduziert, um die Items den anderen Instrumenten anzupassen, was die Brauchbarkeit des Verfahrens (Paas, 1992) sehr wahrscheinlich aber nicht geschadet haben dürfte (Mattell & Jacoby, 1971). Da immer nur ein Item für einen anderen Aufgabenbereich verwendet wurde, konnte keine Skalenanalyse vorgenommen werden. Wie die anderen adaptierten Instrument befinden sich auch die Fragen zur Erfassung der mentalen Anstrengung und empfundenen Schwierigkeit im Anhang B.

5.2.3 Vorwissen (prae)

Von allen Personeneigenschaften, die für den Wissenserwerb von Bedeutung sind, kommt dem Vorwissen eine besondere Rolle zu (Kalyuga, 2005). Dies gilt sicherlich nicht nur für den Bereich des multimedialen Lernens, sondern kann allgemein auf die Gestaltung von Lernumgebungen bzw. -materialien übertragen werden. Nun mag es banal erscheinen, dass Personen, die sich gut mit der Theorie der Plattentektonik auskennen, sehr wahrscheinlich auch erfolgreich in einem Test zu diesem Thema abschneiden werden. Dies gilt umso mehr, wenn sie vor diesem Test noch entsprechendes Lehrmaterial zur Auffrischung präsentiert bekamen. Ganz so einfach, wie es sich hier auf den ersten Blick darstellt, ist es aber nicht. Der Begriff *prior knowledge* bezeichnet kontextspezifisches Vorwissen, das Lernende mehr oder weniger bewusst nutzen können, um den Erwerb neuer Wissensinhalte effektiver und damit erfolgreicher zu gestalten. Dies geschieht nach Ansicht einer ganzen Reihe von Autoren über Elaborationsprozesse bei denen neue Informationen mit Hilfe vorhandener Wissenstrukturen organisiert, restrukturiert, verknüpft und integriert werden (vgl. Reigeluth, Merrill, Wilson & Spiller, 1980; Kalyuga, 2009).

Auch wenn verschiedene Studien zeigen, dass ein ausgeprägtes Vor- oder Hintergrundwissen den Lernprozess in der Regel erleichtert (Krajcik & Varelas, 2006; Kendeou & Broek, 2007), so kann unter bestimmten Umständen hohes domänenspezifische Vorwissen beim Wissenserwerb auch zu Nachteilen führen (Kalyuga, Ayres, Chandler & Sweller, 2003; D. S. McNamara, Kintsch, Songer & Kintsch, 1996). Erklärt werden diese Befunde mit der Annahme, dass die Struktur der Lernumgebung, die auch aus einem einfachen Text bestehen kann und die Wissensstruktur des Lernenden während des Lernprozesses miteinander in Interaktion treten. Solche Interaktionen können den Prozess des Wissenserwerbs fördern, aber auch behindern. Aus diesen Gründen war es im Vorfeld der Untersuchung wichtig, das Vorwissen als Kontrollvariable zu erheben.

Die ursprüngliche Version des Vorwissenstests umfasste insgesamt 15 Items, und bezog sich wie der Lern- und Wissenstest nur auf Informationen, die den Texten und Bildern des Lehrmaterials entnommen werden konnten. Jedes Item bestand aus einer Aussage zum Thema Geodynamik/Geologie und die Versuchsteilnehmer, mussten angeben ob diese Aussage zutrifft oder nicht. Auch hier wurde vor der Durchführung darum gebeten, das entsprechende Kästchen nur dann anzukreuzen, wenn die Antwort nicht geraten werden musste. Abbildung 5.10 zeigt eines der Items der Vorwissensskala, die vollständig im Anhang enthalten ist.

<p>Die modernsten Theorien gehen davon aus, dass die Konvektionsströme im Erdmantel für das Driften der Platten verantwortlich sind.</p>	<p>Richtig <input type="radio"/></p> <p>Falsch <input type="radio"/></p>
--	--

Abbildung 5.10: MC-Item zur Erfassung des domänenspezifischen Vorwissens

5 Untersuchungsmaterial und Instrumente

Die Aussagen überlappten sich inhaltlich mit den Fragen des Lern- und Wissenstests, wobei eine direkte Zuordnung einzelner Items aufgrund der unterschiedlichen Formulierung und Schwerpunktsetzung nicht möglich war. Eine *echte* Messwiederholung mit identischen Fragen wurde damit bewusst vermieden, da nicht klar sein konnte, ob mit Vor- und Nachtest tatsächlich dasselbe gemessen wurde (vgl. Rost, 2004). Anhand der Summe der richtig beantworteten Items sollte stattdessen eine Schätzung der vorhandenen Wissenstrukturen der Versuchsteilnehmer möglich sein, die zudem eine statistische Kontrolle dieser Personenvariablen zuließ. Hinsichtlich der Gütekriterien der Skala schien die *Objektivität* am unproblematischsten, da Anweisungen und das *multiple choice* Format diese weitgehend gewährleistet hatten. Alle weiteren Berechnungen zu den Merkmalen des Instruments fasst Tabelle 5.7 zusammen, wobei den Ergebnissen wiederum sämtliche Antworten der Versuchsteilnehmer zugrunde lagen.

Tabelle 5.7: Eigenschaften der Testskala Vorwissen

	α	$\bar{r}_{ii'}$	\bar{x}	$s_{\bar{x}}$	N_I
Skala Vorwissen	.73	.15	7.5	3.1	15 (15)

Anm.: α = Maß der internen Konsistenz nach Cronbach, $\bar{r}_{ii'}$ = Iteminterkorrelation, \bar{x} = Skalenmittelwert, $s_{\bar{x}}$ = Standardabweichung des Skalenmittelwerts, N_I = Anzahl der Items, die in der Skala verblieben sind (ursprüngliche Anzahl)

Die Tabelle zeigt, dass die Reliabilität des Instruments zwar niedrig war (vgl. Bühner, 2006), aber dennoch als ausreichend angesehen werden konnte. Insgesamt schien der Test mittelmäßig schwer gewesen zu sein, da durchschnittlich die Hälfte der Fragen richtig beantwortet wurden. Alle Items wiesen eine positive Trennschärfe von $r_{itc} > .15$ auf, so dass keines eliminiert werden musste und alle in der Skala verbleiben konnten. Auch die Validität als letztes und wichtigstes Gütekriterium wurde aufgrund inhaltlicher Überlegungen und der relativ hohen Korrelation mit dem Wissens- und Lerntest (siehe Tabelle 5.6) für gegeben erachtet. Der Vorwissenstest ist vollständig in Anhang B enthalten.

5.2.4 FAM Fragebogen zur aktuellen Motivation (post)

Jede Leistung erfordert neben einer grundsätzlichen motorischen bzw. intellektuellen Befähigung die grundsätzliche Bereitschaft sich anzustrengen. Das Vorhandensein dieser Anstrengung ist aber umso fraglicher, wenn das Abschneiden in einer Leistungssituation an keine Konsequenzen geknüpft ist, wie das bei Untersuchungen mit freiwilliger Teilnahme der Fall ist. Damit sollte den Teilnehmern kein mangelndes Engagement unterstellt werden, aber jemand, der sich für geodynamische Prozesse interessiert und/oder über eine generell hohe Leistungsmotivation verfügt, wird sich vermutlich sehr viel stärker anstrengen, als jemand, für den das nicht gilt. Um das Personenmerkmal Motivation in Abhängigkeit von den Inhalten der Lehrmaterialien zu erfassen, wurde der *Fragebogen zur*

aktuellen Motivation (FAM) (Rheinberg, Vollmeyer & Burns, 2001) an die Erfordernisse der Untersuchung adaptiert.

Insgesamt besteht der FAM aus 18 Items, die den vier Skalen/Dimensionen *Erfolgswahrscheinlichkeit* (4 Items), *Herausforderung* (4 Items), *Interesse* (5 Items) und *Misserfolgsbefürchtung* (5 Items) zugeordnet wurden. Rheinberg et al. (2001) zeigen, dass sich der Fragebogen bei verschiedenen computergestützten Lernumgebungen einsetzen ließ, doch konnten nur 10 der 18 Items auf die Inhalte der Untersuchung angepasst werden. Bei den restlichen Items war der Bezug zu Problemlöseaufgaben, für die der Fragebogen ursprünglich entwickelt worden war, zu stark, so dass eine *Adaption* nicht sinnvoll erschien. Abbildung 5.11 zeigt eines der eingesetzten Items, während das gesamte Instrument im Anhang enthalten ist.

Ich bin fest entschlossen, mich bei dem Lerntest voll anzustrengen.

trifft nicht zu trifft zu

Abbildung 5.11: MC-Item zur Erfassung der aktuellen Motivation

Die ursprünglich Version hatte eine Ratingskala mit sieben Abstufungen. Um jedoch neutrale Urteile zu vermeiden und damit eine zumindest tendentielle Richtung bei der Einschätzung der aktuellen Motivation der Teilnehmer zu erzwingen, wurde die Zahl der Skalenstufen auf sechs reduziert. Dieses Vorgehen war unproblematisch, zumal nicht erwartet werden musste, dass dieses Vorgehen eine Verminderung der Reliabilität des Instruments nach sich ziehen würde (Matell & Jacoby, 1971). Schwerwiegendere Konsequenzen hatte die Tatsache, dass die 10 *adaptierten* Items, sich nicht gleichmäßig auf die vier Dimensionen bzw. Skalen verteilten. Beispielsweise wurden drei Items der Dimension Interesse übernommen, während nur ein Item der Dimension Herausforderung in den Kontext der Untersuchung passte. Letztlich wurde die vierdimensionale Struktur aufgegeben und alle 10 Items auf einer Skala angeordnet. Ähnlich wie bei der Testskala zur Erfassung des Vorwissens, konnte das Gütekriterium der Objektivität aufgrund der Bearbeitungsanweisungen und des MC-Formats als weitgehend erfüllt erachtet werden. Dagegen zeigt Tabelle 5.8 die Ergebnisse der Item- und Skalenanalyse der *Kurzversion* des FAM.

Tabelle 5.8: Eigenschaften der Testskala aktuelle Motivation

	α	$\bar{r}_{ii'}$	\bar{x}	$s_{\bar{x}}$	N_I
FAM Kurzversion	.79	.30	27.5	7.1	09 (10)

Anm.: α = Maß der internen Konsistenz nach Cronbach, $\bar{r}_{ii'}$ = Iteminterkorrelation, \bar{x} = Skalenmittelwert, $s_{\bar{x}}$ = Standardabweichung des Skalenmittelwerts, N_I = Anzahl der Items, die in der Skala verblieben sind (ursprüngliche Anzahl)

5 Untersuchungsmaterial und Instrumente

Die Reliabilität lag zwar immer noch unter $r_\alpha = .80$ und musste damit als niedrig angesehen werden, doch war die Iteminterkorrelation $\bar{r}_{ii'}$ mit $.30$ recht hoch, was für die Homogenität der Skala sprach und eine Zusammenfassung der Items nachträglich rechtfertigte. Allerdings war es unumgänglich, ein Item zu entfernen, da dessen Trennschärfe mit $r_{itc} = .02$ einen zu niedrigen Wert hatte. Was die Validität als letztes und wichtigstes Gütekriterium anging, so konnte diese aufgrund inhaltlicher Überlegungen, der hohen Korrelation mit dem Score des Lern- und Wissenstests und nicht zuletzt aufgrund der Ergebnisse von Rheinberg et al. (2001) als ausreichend angesehen werden.

5.2.5 Verbale und räumliche Intelligenz (prae)

Wer intelligenter ist, dem fällt es auch leichter, sich neues Wissen anzueignen. Diese alltagspsychologische Grundannahme, findet sich auch in verschiedenen wissenschaftlichen Definitionen des Konstrukts *Intelligenz* wieder, obwohl eine einheitliche definitorische Fassung bislang nicht gelungen ist (Amelang & Zielinski, 1997). Der Unschärfe des Intelligenzkonstrukts steht jedoch die Tatsache entgegen, dass viele Tests gute Prädiktoren für die schulische und berufliche Leistung darstellen (Asendorpf, 1996). So konnte beispielsweise recht früh belegt werden, dass Leseverständnis und die Intelligenzleistung hoch miteinander korrelieren (Hage & Stroud, 1959). Vor diesem Hintergrund wurde es als notwendig erachtet, neben einer domänenspezifischen Personenvariablen wie dem Vorwissen, ein Maß für die eher *allgemeine* kognitive Leistungsfähigkeit bzw. Intelligenz zu erheben.

Als Instrument versprach der IST 2000 R (Liepmann, Beauducel, Brocke & Amthauer, 2007) Ökonomie, Überschaubarkeit und Anwendungsfreundlichkeit. Da der Test zusätzlich schulelevante Intelligenzleistungen valide erfasst, wie hohe Korrelationen mit schulischem Erfolg zeigen (Liepmann, Beauducel, Brocke & Amthauer, 2007), schien er besonders für Studenten im Grundstudium geeignet zu sein, die erst kürzlich das Gymnasium abgeschlossen hatten und einen Großteil der Stichprobe ausmachten. Die Durchführung des Basismoduls des IST 2000 R hätte allerdings 90 Minuten in Anspruch genommen, ein zeitlicher Umfang, der für die vollständige Sitzung mit Vortest, Präsentationsphase und Nachbefragung eingeplant worden war. Daher wurden nur zwei der neun Untertests des Grundmoduls für die Untersuchung ausgewählt. Abbildung 5.12 zeigt ein Item einer der ersten Skalen.

Ein Sport-Pokal ist immer ___?

- ein Andenken
- schwer
- wertvoll
- eine Anerkennung
- vergoldet

Abbildung 5.12: Item der Skala Satzergänzung (SE) des I-S-T 2000 R

Dabei handelt es sich um den Untertest *Satzergänzung (SE)*, der auf den ersten Blick

recht einfach erschien. Man musste lediglich entscheiden, welche(s) der vorgegebenen Wörter den dargestellten Satz sinnvoll ergänzte und die richtige Lösung ankreuzen. Die insgesamt 20 Items wurden aber zunehmend schwerer und eine richtige Beantwortung erforderte neben semantischem Wissen auch zunehmend logisches Denken. Mit Hilfe dieser Satzergänzungsaufgaben sollten in erster Linie die verbalen Fähigkeiten der Versuchsteilnehmer eingeschätzt werden, da ein großer Teil der Stimulusmaterials aus Texten bestand. Die Reliabilität dieser Skala ist mit $r_\alpha = .63$ relativ niedrig, was die Autoren auf eine heterogene Anforderungsstruktur zurückführen. Dennoch erwies sich der Unterstest bei einer Studie von Bühner, Ziegler, Krumm und Schmidt-Atzert (2006) als *raschskalierbar*, was letztlich für dessen hohe psychometrische Qualität spricht. Außerdem wurde es als Vorteil angesehen, dass die Satzergänzungsaufgaben aufgrund ihrer Heterogenität die verbale Intelligenz etwas breiter erfassten, da nur eine Skala aufgrund zeitlicher Beschränkungen eingesetzt werden konnte.

Anstelle eines weiteren Untertests zur Erfassung einer anderen Facette der verbalen Fähigkeiten, schien es sinnvoller zu sein, einen Intelligenzbereich zu messen, der speziell für die Verarbeitung von Bildern von Bedeutung war. Verschiedene Studien legen nahe, dass die Ausprägung räumlicher Fähigkeiten, den Wissenserwerb mit piktoralem Material maßgeblich beeinflussen kann. Ob bei der Bearbeitung von Karten (Sanchez & Branaghan, 2009), Diagrammen (Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997) oder Animationen (Mayer & Sims, 1994), stets konnten deutliche Unterschiede hinsichtlich des Lernergebnisses zwischen Personen mit hohen und niedrigen räumlichen Fähigkeiten festgestellt werden. Von den drei Aufgabengruppen, die der I-S-T 2000 R zur Erfassung figurativer Leistungen bereitstellt, fiel die Wahl auf die Würfelaufgaben. Auch dieser Untertest hatte sich mit geringen Einschränkungen als raschskalierbar erwiesen und war mit $r_\alpha = .80$ reliabler als die anderen figurativen Aufgaben. Das entscheidende Kriterium für diese Skala war jedoch deren dreidimensionaler Charakter, was Abbildung 5.13 verdeutlicht.

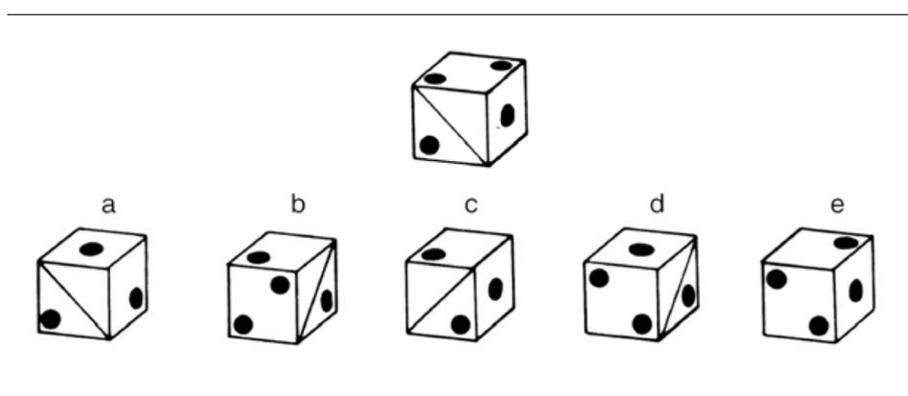


Abbildung 5.13: Item der Skala Würfelaufgaben (WÜ) des I-S-T 2000 R

Wie leicht zu erraten ist, besteht die Aufgabe darin, zu entscheiden, welchem der mit den Buchstaben *a* bis *e* gekennzeichneten Würfel, der Würfel in der ersten Zeile

5 Untersuchungsmaterial und Instrumente

entspricht. Zu diesem Zweck muss die Figur mental gedreht und/oder gekippt werden, was letztlich räumliches Vorstellungsvermögen voraussetzt. Dieses Vorstellungsvermögen schien gerade für die Arbeit mit den Bildern und Texten des Lehrmaterials bedeutsam zu sein, da man sich zentrale Konzepte der Geodynamik besonders gut räumlich vorstellen kann. Zwar haben Köller, Rost und Köller (1994) festgestellt, dass bei diesem Untertest eine Unterscheidung zwischen Flächen- und Raumwürfel getroffen werden muss, die klar differenzierbare Lösungsstrategien nach sich ziehen. Doch auch hier wurde die Heterogenität der Skala als ein Vorteil gesehen, da mit einer Skala neben dem räumlichen Vorstellungsvermögen noch weitere Aspekte der figurativen Intelligenz gemessen werden konnten. Tabelle 5.9 zeigt die Ergebnisse aller Probanden, welche an den drei Untersuchungen teilgenommen haben.

Tabelle 5.9: Stichprobenmerkmale der IST 2000 R Intelligenzsskalen

	α	$\bar{r}_{ii'}$	\bar{x}	$s_{\bar{x}}$
Satzergänzung (SE)	.62	.07	13.9	2.9
Würfelaufgaben (WÜ)	.83	.20	11.2	4.4

Anm.: α = Maß der internen Konsistenz nach Cronbach, $\bar{r}_{ii'}$ = Iteminterkorrelation, \bar{x} = Skalenmittelwert, $s_{\bar{x}}$ = Standardabweichung des Skalenmittelwerts

Die innerhalb der Stichprobe ermittelten internen Konsistenzen der beiden Untertests entsprachen dabei ziemlich genau den im Manual angegebenen Werten. Dort sind zwar keine Angaben zu den Iteminterkorrelationen der Skalen zu finden, doch bestätigt die relativ niedrige Ausprägung beider Koeffizienten (vgl. Bortz & Döring, 2006) die Einschätzung bezüglich des heterogenen Charakters der Aufgaben. Ein Vergleich der Skalenmittelwerte \bar{x} und deren Standardabweichung $s_{\bar{x}}$ mit den Angaben der verschiedenen Normtabellen für Gymnasiasten im Alter zwischen 19 und 30 Jahren ergab nur minimale Abweichungen. Damit entsprachen die Leistungen der Versuchsteilnehmer nahezu exakt den Leistungen der Normierungsstichprobe. DA Objektivität und Validität bei diesem bewährten Instrument als weitgehend gesichert angesehen werden können (Liepmann et al., 2007), müssen diese Merkmale des Tests nicht weiter diskutiert werden.

5.2.6 Lernstil (prae)

Neben den eher leistungsbezogenen Merkmalen bot es sich an, den *learning style* bzw. *cognitive style* der Studierenden zu erheben, wobei sich die beiden Begriffe in der Literatur nicht immer auf die gleiche Konzeption beziehen (Plass, Chun, Mayer & Leutner, 1998). Da die Lernmaterialien jeweils zur Hälfte aus Texten und Bildern bestanden, schien an dieser Stelle das *Visualizer-Verbalizer*-Konstrukt (Richardson, 1977) interessant zu sein. Zwar halten manche Autoren die Annahme, dass einige Menschen besonders gut Texte und andere besonders gut Bilder verarbeiten können, für einen festen Bestandteil

der erziehungswissenschaftlichen Folklore (Mayer & Massa, 2003), doch existieren auch Belege für einen Zusammenhang zwischen VV-maßen und Leistungsdaten (Kirby, Moore & Schofield, 1988). Zudem sprechen neuere Ergebnisse aus der Neuropsychologie (Kraemer, Rosenberg & Thompson-Schill, 2009) für das Vorhandensein einer verbalen bzw. visuellen Verarbeitungspräferenz.

Aufgrund verschiedener Schwierigkeiten mit dem ursprünglichen Verbalizer-Visualizer-Questionnaire (VVQ) von Richardson (1977), wurde eine überarbeitete und erweiterte Fassung von Kirby, Moore und Schofield (1988) verwendet. Abbildung 5.14 zeigt eines der 20 Items, die aus dem Englischen übersetzt wurden. Von den drei Skalen, die Kirby et

Ich lese sehr langsam	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
-----------------------	--

Abbildung 5.14: Item des Verbalizer-Visualizer-Questionnaires

al. (1988) eingesetzt haben, kamen in der Untersuchung nur zwei zum Einsatz. So bearbeiteten die Teilnehmer noch vor der Präsentation der Bilder und Texte die vollständigen *Visualizer-* und *Verbalizer-Skalen* mit jeweils zehn Items. Die dritte Skala *Traum* hatte inhaltlich kaum Bezug zu den Lernmaterialien, weshalb sie weggelassen wurde. Tabelle 5.10 enthält die Ergebnisse einer Skalenanalyse des VVQ.

Tabelle 5.10: Merkmale der VVQ-Skalen nach Kirby et al. (1988)

	α	$\bar{r}_{ij'}$	\bar{x}	$s_{\bar{x}}$	r_{vi}	r_{ri}
Verbalizer	.57	.11	6.4	1.9	.02	.02
Visualizer	.72	.20	7.9	2.0	.01	.22*

Anm.: α = Maß der internen Konsistenz nach Cronbach, $\bar{r}_{ij'}$ = Iteminterkorrelation, \bar{x} = Skalenmittelwert, $s_{\bar{x}}$ = Standardabweichung des Skalenmittelwerts, r_{vi} = Korrelation mit IST-SE, r_{ri} = Korrelation mit IST-WÜ

Die Reliabilität der beiden Skalen entspricht in etwa den internen Konsistenzen, welche auch Kirby et al. (1988) ermittelt haben. Die niedrigen α -Koeffizienten weisen aber auf deren Heterogenität hin, was die geringen Iteminterkorrelationen zusätzlich belegen. Allerdings ergab die Skalenanalyse für alle Items des VVQ durchweg positive Trennschärfen, so dass kein einziges entfernt werden musste. Um die Validität der aus dem englischen übertragenen VVQ-Skalen zu prüfen, wurden diese mit den beiden Untertests des IST korreliert. Dabei wurde jeweils positive Zusammenhänge zwischen der Verbalizer-Skala und den Satzergänzungsaufgaben, sowie der Visualizer-Skala und den Würfelaufgaben erwartet. Tabelle 5.10 zeigt aber, dass nur die Korrelation zwischen der Visualizer-Skala und dem räumlichen Vorstellungsvermögen signifikant korrelierte. Auch wenn die Validitätsprüfung nicht ganz geglückt war, wurden alle VVQ-Items, die in Anhang B enthalten

sind, bei den statistischen Analysen berücksichtigt.

5.2.7 Fragen zu Lerngewohnheiten (prae) & Fragen zum Lernen während der Präsentationsphase (post)

Prae- und Posttest enthielten noch weitere Fragen zum allgemeinen Umgang der Teilnehmer mit Text- und Bildinformationen und dem speziellen Vorgehen beim Wissenerwerb mit dem Stimulusmaterial. So war beispielsweise von Interesse, wie oft im Studium und/oder privat Bilder zum Wissenerwerb genutzt werden, oder ob sich die Probanden während der Präsentationsphase des Lernmaterials mehr auf die Texte oder mehr auf die Bild konzentriert haben. Insgesamt wurden den Teilnehmenden sieben solcher Items vorgelegt, die alle aus einer Frage und einer sechstufigen Skala bestanden, auf der eine Einschätzung vorgenommen werden konnte. Auch diese Fragen sind im Anhang B enthalten.

5.2.8 Soziodemographische Daten (prae)

Natürlich wurden im Rahmen der Untersuchung auch Fragen vorgelegt, mit denen soziodemografische Daten erhoben wurden. So sollten die Versuchsteilnehmer Angaben zum *Alter*, *Geschlecht*, *Studienfach*, der Anzahl absolvierter *Semester* und der durchschnittlichen *Erdkundenote* machen. Darüber hinaus wurde das *Interesse für Geologie* mit einer sechsstufigen Ratingsskala gemessen, die Werte zwischen 0 (=sehr gering) und 5 (sehr groß) annehmen konnte. Tabelle 5.11 fasst einige der allgemeinen Merkmale der Gesamtstichprobe zusammen. Der überwiegende Anteil der Versuchsteilnehmer hatte sich

Tabelle 5.11: Allgemeine Stichprobenmerkmale

	\bar{x}	$s_{\bar{x}}$
Alter (in Jahren)	22.9	4.0
Anzahl der Semester	3.5	2.9
Erdkundenote	2.3	0.9
Interesse für Geologie	1.7	1.2

Anm.: \bar{x} = Skalenmittelwert, $s_{\bar{x}}$ = Standardabweichung des Skalenmittelwerts

für das Studienfach Psychologie eingeschrieben (45%), während die anderen Probanden entweder eine Lehramtsausbildung (Grund-, Haupt-, Real- oder Sonderschulpädagogik) absolvierten (38%), Erziehungswissenschaften (9%) oder Sonstiges (8%) studierten. Wie bei dieser Fächerkombination erwartet werden konnte, nahmen mit rund 80 % deutlich mehr Frauen als Männer an den Studien teil. Die Gesamtstichprobe war mit 23 Jahren im Mittel recht jung und befand sich in der Regel noch im Grundstudium. Obwohl die Angaben zu den Erdkundenoten auf gute Zensuren hinwiesen, war das Interesse für Geologie

eher unterdurchschnittlich ausgeprägt.

Damit ist die Darstellung der Materialien und Instrumente an dieser Stelle abgeschlossen. Methodische Fragen, welche die Stichprobenziehung, Anordnung der Instrumente oder andere Aspekte der Versuchsplanung betreffen, werden in den folgenden Kapiteln zu den einzelnen Studien ausführlich dargestellt.

6 Studie I: Interferenzen versus Kapazität

Wie an anderer Stelle bereits erwähnt wurde, lässt sich ein überaus interessanter Effekt beobachten, wenn Texte und Bilder nacheinander gelernt werden. Demnach ist es besser, visuelle vor verbalen Informationen zu verarbeiten, da im umgekehrten Fall eine schlechtere Leistung zu befürchten ist. Natürlich liegt es nahe, einen solchen Befund auf die Praxis anzuwenden, doch mindestens ebenso faszinierend sind die damit verbundenen theoretischen Implikationen. Das gilt umso mehr, da gleich zwei Erklärungsansätze vorliegen, die den PTS-Effekt plausibel und konsistent herleiten können. Das Ziel der ersten Studie ist es daher, zu entscheiden, welche der beiden Modellvorstellungen tatsächlich zutrifft.

6.1 Fragestellung und Hypothesen

Die beiden Erklärungsansätze um die es geht, wurden in Kapitel 4 als Kapazitäts- und Interferenzhypothese bezeichnet. Theoretische Überlegungen legten dabei den Schluss nahe, dass diese Hypothesen unterschiedliche Vorhersagen machen, wenn neben der Reihenfolge das Informationsverhältnis zwischen Text und Bild variiert. Aus diesem Grund wurde kongruente sowie komplementäre Texte- und Bilder entworfen (siehe Kapitel 5.1.4, mit dem eine experimentelle Prüfung möglich sein sollte. Das setzt jedoch voraus, dass es auch gelingt, den PTS-Effekt tatsächlich zu replizieren, was nicht unbedingt als sicher gelten kann. Immerhin haben Kulhavy et al. (1994) ihre Studien nicht veröffentlicht, so dass mit der Arbeit von Verdi, Johnson, Stock, Kulhavy und Whitman-Ahern (1997) bislang nur eine einzige Untersuchung frei zugänglich ist, die den PTS-Effekt ausdrücklich belegt. Obwohl die berichteten statistischen Zusammenhänge beeindruckend sind, kann daher nicht ausgeschlossen werden, dass den Ergebnissen von Verdi et al. (1997) andere Ursachen zugrunde liegen und die Variation der Reihenfolge keine Bedeutung besitzt. So ist es durchaus möglich, dass der PTS-Effekt gar nicht existiert, womit sowohl die Inferenz- als auch die Kapazitätshypothese widerlegt würden.

Ein anderer Aspekt, der bei der Konkretisierung der Fragestellung berücksichtigt wurde, betraf weitere Personenmerkmale, von denen anzunehmen war, dass sie mit dem Lernprozess zusammen hingen. Alle bisherigen Ausführungen bezogen sich vor allem auf den *Lernerfolg*, wobei allein zwischen Behalten und Verstehen differenziert wurde. Es war durchaus anzunehmen, dass sich noch weitere Variablen in Abhängigkeit des Untersuchungsdesigns verändern würden, was vor allem im Hinblick auf eine umfassendere Deutung der Ergebnisse interessant sein sollte. Ein in diesem Zusammenhang besonders wichtiges Merkmal war die Belastung des Arbeitsgedächtnisses bzw. der *cognitive load*, der während der Lernphase oder der Berarbeitung des Wissenstests empfunden wurde.

6 Studie I: Interferenzen versus Kapazität

Auch hier ließen sich Hypothesen und prüfbare Vorhersagen ableiten.

So ist nach der Kapazitätsannahme immer dann mit einer gesteigerten Beanspruchung der Arbeitsgedächtnisses zu rechnen, wenn das Bild vor dem Text gelesen wird, da man davon ausgehen muss, dass die Etablierung von referentiellen Verknüpfungen kognitiv belastend ist. Der mentale Aufwand sollte deutlich geringer sein, falls der Lernende keine oder nur vereinzelte referentielle Verknüpfungen bilden kann. Dies beispielsweise dann der Fall, wenn Text und Bild sich inhaltlich kaum überlappen oder die Verarbeitungsreihenfolge es nicht zulässt. Folgt man hingegen dem Erklärungsansatz von Schnotz (2005), so wird eine negative mentale Belastung auftreten, wenn Interferenzen zwischen dem mentalen Modell und dem externen Bild entstehen. Wird das kongruente Bild zuerst verarbeitet oder es besteht ein komplementäres Informationsverhältnis zwischen den beiden Formaten sollte der empfundene *cognitive load* vergleichsweise gering ausfallen. Damit sagt die Interferenzhypothese eine Wechselwirkung zwischen dem Informationsverhältnis und der Reihenfolge von Text und Bild vorher.

Eine weiteres Merkmal, von dem angenommen wurde, dass es Aufschluss über den PTS-Effekt und die dahinter vermuteten Prozesse geben konnte, war die Lernzeit. Diese Überlegung gründete auf der einfachen Annahme, dass jede kognitive Operation ein gewisses Maß an Zeit beansprucht. Je komplexer ein Prozess ist, umso länger wird seine Durchführung dauern, obwohl dieser Zusammenhang je nach Vorwissen oder Fähigkeiten interindividuell variieren kann. Verdi et al. (1997) und sehr wahrscheinlich auch Kulhavy et al. (1994) haben in ihren Studien die Lernzeiten jeweils fixiert. Auf diese Weise konnten die Autoren die Lese- und Betrachtungsdauer experimentell kontrollieren, doch Auswirkungen der Reihenfolge ließen sich natürlich nicht mehr erfassen. Ob und welchen Einfluss der PTS-Effekt aus der Perspektive der Kapazitätshypothese auf die Lernzeit hat, kann daher wiederum nur geschlussfolgert werden. Ausgangspunkt der in erster Linie theoretischen Überlegungen sind die referentiellen Verknüpfungen zwischen visuellen und verbalen Informationen, die nur etabliert werden können, wenn das Bild vor dem Text gelernt wird. Nimmt man nun an, dass jede einzelne Verknüpfung visueller und verbaler Informationen zusätzlich Zeit kostet, dann müsste die Verarbeitung des Textes nach dem Bild länger dauern, als wenn der Text davor gelesen wird. Dieser Zusammenhang sollte selbst dann noch vorhanden sein, wenn Bild und Text in einem eher komplementären Verhältnis zueinander stehen. Zwar würden insgesamt weniger referentielle Verknüpfungen etabliert werden, doch allein der Versuch, dies zu tun, dürfte Zeit in Anspruch nehmen.

Auch anhand des Modells von (Schnotz, 2005) ließ sich ableiten, dass die Lernzeit mit der Präsentationsreihenfolge variieren wird. Dafür sind allerdings formatspezifische Effekte der beiden Repräsentationsformate verantwortlich, die nur indirekt mit Interferenzen zu tun haben. Demnach stellt es zunächst eine durchaus anspruchsvolle Aufgabe dar, ein mentales Modell allein anhand eines Textes zu konstruieren. Kann ein Lernender jedoch zuvor auf ein Bild zurückgreifen, das den Sachverhalt visualisiert, erhält er quasi einen Bauplan für den Konstruktionsprozess. Aus diesem Grund müsste der nachfolgenden Text schneller verarbeitet werden können, da die Umwandlung eines symbolischen in ein analoges Repräsentationsformat zumindest teilweise entfällt. Dieser Zusammen-

hang gilt jedoch nur dann, wenn Bild und Text sich inhaltlich überschneiden. Bei einem komplementären Informationsverhältnis dürfte die Lernzeit dagegen nicht von der Reihenfolge beeinflusst werden, da die Verarbeitung von Texten und Bildern ohne inhaltliche Überlappung zu relativ unabhängigen Repräsentationen führen müsste.

Diese Überlegungen betrafen allein die Lesezeit des verbalen Materials, da der Umgang mit den Bildern nur sehr schwer vorhergesagt werden kann. Nach Kulhavy et al. (1994) muss sich der Lernende in der Text-vor-Bild Bedingung nämlich entscheiden, ob er die verbalen Informationen replizieren oder lieber das Bild lernen will. Im Fall einer freien Wahl könnte das zur Folge haben, dass einzelne Versuchspersonen darauf verzichten, die visuellen Informationen zu verarbeiten. Ebenso lässt sich die Reaktion der Lernenden auf Interferenzen nur schwer vorhersagen. Es ist vorstellbar, dass einige Versuchspersonen trotz der kognitiven Konflikte versuchen werden, ihr internes Bild und die externe Abbildung zu integrieren, was wegen des zusätzlichen Aufwandes zu einer längeren Verarbeitungszeit führen sollte. Anderen hingegen dürfte eine solche Auseinandersetzung zu mühsam sein, so dass sie versuchen werden, den (zeitlichen) Aufwand möglichst klein zu halten. Daher beziehen sich sämtliche Vorhersagen zur Lernzeit zunächst auf die Betrachtungsdauer der Textseiten.

Zusammenfassend lässt sich festhalten, dass sich die Fragestellung mit dem *Lernerfolg*, der *kognitiven Belastung* und der *Lernzeit* auf insgesamt drei Personenmerkmale bezieht. Für jedes dieser Merkmale lassen sich anhand der Kapazitäts-, und Interferenzhypothese unterschiedliche Vorhersagen ableiten, wie die folgende Darstellung in übersichtlicher Weise zeigen soll.

Fragestellung 1: Welchen Einfluss haben das Informationsverhältnis und die Reihenfolge von Text und Bild auf den Lernerfolg?

1.1 Kapazitätshypothese: Der Lernerfolg wird größer sein, wenn das Bild vor dem Text gelernt wird. Dieser Zusammenhang gilt auch dann, wenn Bild und Text in einem komplementären Informationsverhältnis zueinander stehen.

1.2 Interferenzhypothese: Der Lernerfolg wird größer sein, wenn das Bild vor dem Text gelernt wird. Dieser Zusammenhang gilt allerdings nur, wenn Bild und Text in einem kongruenten Informationsverhältnis zueinander stehen.

Fragestellung 2: Welchen Einfluss haben das Informationsverhältnis und die Reihenfolge von Text und Bild auf die kognitive Belastung des Arbeitsgedächtnisses?

2.1 Kapazitätshypothese: Die mentale Belastung des Arbeitsgedächtnisses wird größer sein, wenn das Bild vor dem Text gelernt wird. Dieser Zusammenhang gilt auch dann, wenn Bild und Text in einem komplementären Informationsverhältnis zueinander stehen.

2.2 Interferenzhypothese: Die mentale Belastung wird größer sein, wenn der Text vor dem Bild gelernt wird. Dieser Zusammenhang gilt allerdings nur, wenn

6 Studie I: Interferenzen versus Kapazität

Bild und Text in einem kongruenten Informationsverhältnis zueinander stehen.

Fragestellung 3: Welchen Einfluss haben das Informationsverhältnis und die Reihenfolge von Text und Bild auf die Verarbeitungszeiten der einzelnen Formate?

3.1 Kapazitätshypothese: Die Verarbeitungszeit für den Text nimmt zu, wenn zuvor ein Bild betrachtet wurde. Dieser Zusammenhang gilt auch dann, wenn Bild und Text in einem komplementären Informationsverhältnis zueinander stehen.

3.2 Interferenzhypothese: Die Verarbeitungszeit für den Text nimmt ab, wenn zuvor ein Bild betrachtet wurde. Dieser Zusammenhang gilt nur, wenn Bild und Text in einem kongruenten Informationsverhältnis zueinander stehen.

6.2 Versuchsplanung

In Kapitel 5 wurde sowohl das verwendete Stimulusmaterial als auch alle Instrumente ausführlich beschrieben, die zur Erfassung der Kovariaten bzw. der abhängigen Personeneigenschaften entwickelt und/oder verwendet wurden. Die folgenden Abschnitte thematisieren daher vor allem den Ablauf der Studie, die Versuchsteilnehmer und die Durchführung.

6.2.1 Design

Das Untersuchungsmaterial bestand aus Texten (siehe Abschnitt 5.1.2) und Bildern (siehe Abschnitt 5.1.3), die für das bevorstehende Experiment so manipuliert wurden, dass letztlich zwei Versionen vorlagen (siehe Abschnitt 5.1.4). Beide Ausführungen bezogen sich auf die selbe Domäne (5.1.1) und thematisierten die selben Inhalte, wobei jedoch das Informationsverhältnis zwischen Text und Bild variierte. Während die kongruente Version gemeinsame Informationen enthielt, war bei der kongruenten Ausführung ein komplementäres Informationsverhältnis erzeugt worden. Um die Vorhersagen der verschiedenen Hypothesen zu prüfen, konnte daher eine Versuchsanordnung gewählt werden, in der jeweils zwei Bedingungen systematisch variiert werden konnten. Damit resultierte letztlich ein schlichtes 2 mal 2 Design mit den Faktoren Informationsverhältnis und Verarbeitungsreihenfolge, wie Abbildung 6.1 grafisch veranschaulicht.

Die Ziffern in den Klammern geben an, dass jeder Zelle 21 Versuchspersonen zugeteilt wurden, was einer Gesamtteilnehmerzahl von 84 entspricht. Dieser Stichprobenumfang beruht auf Überlegungen, die neben den statistischen Entscheidungsfehlern erster und zweiter Art die Stärke des PTS-Effekts berücksichtigte. Wie bereits erwähnt wurde, haben Kulhavy et al. (1994) ihre Experimente nicht veröffentlicht, weshalb mit der Studie von Verdi et al. (1997) nur eine empirische Arbeit vorliegt, anhand der sich die Höhe des vermuteten Zusammenhanges abschätzen lässt. Zwar geben die Autoren kein direktes

Informationsverhältnis	Reihenfolge	
	Text vor Bild (TvB)	Bild vor Text (BvT)
kongruent (<i>kon</i>)	kon_TvB (n= 21)	kon_BvT (n=21)
komplementär (<i>komp</i>)	komp_TvB (n=21)	komp_BvT (n=21)

Abbildung 6.1: Untersuchungsdesign der ersten Studie

Maß für die Effektstärke an, doch konnte diese anhand von F-Werten und Freiheitsgraden rekonstruiert werden. So war es möglich, mit einer Formel, die B. Rasch, Frieese, Hofmann und Naumann (2010) vorschlugen, *Chohen's f* zu bestimmen, wobei die ermittelten Werte je nach Experiment variierten. Die einzelnen Angaben schwankten zwischen einer mittleren Effektstärke von $f = 0.25$ ($\bar{\eta}_{part}^2 \approx 0.06$), und einem starken Zusammenhang von $f = 0.42$ ($\bar{\eta}_{part}^2 \approx 0.15$).

Für die weitere Planung wurde mit $f = 0.34$ ($\bar{\eta}_{part}^2 \approx 0.11$) die durchschnittliche Effektgröße berechnet. Da aber die Lernzeit als abhängige Variable erfasst werden sollte und nicht mehr unmittelbar einer experimentellen Kontrolle unterlag, war damit zu rechnen, dass der Einfluss der Verarbeitungskontrolle unter Umständen etwas geringer ausfallen könnte. Aus diesem Grund wurde bei den Überlegungen zur benötigten Stichprobengröße eine Effektsärke von $f = 0.31$ ($\bar{\eta}_{part}^2 \approx 0.087$) zugrunde gelegt. Das Signifikanzniveaus und die angestrebte Teststärke orientierte sich an dem in den Sozialwissenschaften üblichen Verhältnis von 1 zu 4 (vgl. Bortz & Döring, 2006). Hinzu kam, dass aus ökonomischen Gründen die Stichprobe und damit den Aufwand überschaubar gehalten werden sollte, weshalb mit einem α -Fehlerniveau von 5% und einem β -Fehlern von 20% geplant wurde. Damit waren letztlich alle Parameter vorhanden, um den optimalen Stichprobenumfang mithilfe des Software G*3 (Faul, Erdfelder, Lang & Buchner, 2007) zu berechnen, was eine Gesamteilnehmerzahl von 84 Personen ergab.

Die abhängigen Variablen, auf die sich die einzelnen Hypothesen und Fragestellungen bezogen, waren der *Lernerfolg*, die *kognitive Belastung* und schließlich die *Lernzeit*. Zusätzlich wurden eine Reihe von Kontrollvariablen erfasst, deren statistische Berücksichtigung mögliche Stichprobeneffekte neutralisieren sollte. Zwar war vorgesehen, die Versuchspersonen den Bedingungen randomisiert zuzuweisen, was letztlich die beste Kontrolle darstellt (Bortz & Döring, 2006), doch bei einem Stichprobenumfang von 84 Teilnehmern können mögliche Verzerrung nicht gänzlich ausgeschlossen werden. Aus diesem Grund wurden das *Interesse für Geologie/Geographie*, die *Erdkundenote*, das *Vorwissen*, die *aktuelle Motivation*, die *räumliche und verbale Intelligenz* sowie der *Lernstil* erhoben. Alle Instrumente, mit denen diese Personenmerkmale gemessen werden sollten, sind in Kapitel 5.2 ausführlich beschrieben.

6.2.2 Stichprobe

Die Untersuchung fand in den Räumen der Universität Koblenz-Landau / Campus Landau statt, so dass es nahe lag, die Versuchspersonen unter den Studierenden zu rekrutieren. Da zu dieser Zeit mehrere wissenschaftliche Untersuchungen um Probanden konkurrierten, war es nicht möglich, ausschließlich Teilnehmer des Studiengangs Psychologie für das Experiment zu gewinnen. Die Rekrutierungsprobleme führten letztlich auch dazu, dass sich die Durchführung der Studie über einen Zeitraum von fünfeneinhalb Wochen hinzog. Doch trotz dieser Schwierigkeiten konnten letztlich 84 Personen davon überzeugt werden, an der Untersuchung teilzunehmen, so dass der geplante Stichprobenumfang verwirklicht werden konnte. Die meisten Teilnehmer ($\approx 62\%$) hatten sich in den Studiengang Psychologie eingeschrieben, während der Rest Lehramt ($\approx 24\%$), Pädagogik ($\approx 8\%$) oder Sonstiges ($\approx 6\%$) studierte.

Obwohl die überwiegende Mehrheit ($\approx 79\%$) vier oder weniger Semester aufweisen konnte, handelte es sich bei rund 21% um eher erfahrenere Studierende. Dementsprechend betrug der Altersdurchschnitt etwa 23 Jahre ($M = 23.3, SD = 4$) mit einem Minimum von 18 und einem Maximum von 39 Jahren. Hinsichtlich des Geschlechts dominierten eindeutig die Frauen (79%), da männliche Studierende mit 21.4% in der Minderheit waren. Um eine Teilnahme möglichst reizvoll zu gestalten, wurden im Vorfeld der Studie mit verschiedenen Anreizen geworben. Während die Psychologiestudenten zwischen der Gutschrift von zwei Versuchspersonenstunden und 10,- € als Belohnung wählen konnten, erhielten die Studierenden anderer Fachrichtungen am Ende der Sitzung gleich das Geld.

Für eine abschließende Beurteilung der Stichprobe bieten sich mit der Homogenität und der Repräsentativität vor allem zwei Kriterien an. Demnach kann die Stichprobe auf den ersten Blick als relativ homogen angesehen werden, wenn man berücksichtigt, dass alle Teilnehmer Studenten waren und damit die Hochschulreife besaßen. Eine Einschränkung stellte jedoch die Tatsache dar, dass nicht ausschließlich Psychologiestudenten teilnahmen, da die verschiedenen Zulassungsvoraussetzungen der einzelnen Studiengänge eine Art Vorselektion bedeuteten. Ebenso dürfte es um die Motivation der Studierenden sehr unterschiedlich bestellt gewesen sein, denn ein Teil der Studierenden des Faches Psychologie *musste* quasi bei der Untersuchung mitmachen, um genügend Versuchspersonenstunden zu sammeln. In diesem Zusammenhang schien es durchaus angebracht zu sein, eine Reihe von Kontrollvariablen zu erheben, um möglichen Verzerrungen durch zufällige Effekte vorzubeugen. Hinsichtlich der Repräsentativität lässt sich festhalten, dass diese als eingeschränkt betrachtet werden musste. Die *TeilnehmerInnen* waren mehrheitlich junge, weibliche Studenten der Sozial- oder Erziehungswissenschaften, was beim Versuch einer Verallgemeinerung der Ergebnisse bedacht werden musste.

6.2.3 Ablauf und Durchführung der Untersuchung

Alle experimentellen Sitzungen wurden am Campus Landau der Universität Koblenz-Landau durchgeführt. Dem Versuchsleiter standen zu diesem Zweck zwei Räume zur Verfügung, die mit 8 bzw. mit 6 PCs ausgestattet waren. Jeder dieser Rechner besaß einen Zugang zum Internet, so dass über einen *Browser* die Texte und Bilder, die als HTML-Seiten vorlagen, präsentiert werden konnten. Um das Prinzip der Randomisierung zu verwirklichen, war im Vorfeld der Untersuchung eine Liste erstellt worden, mit deren Hilfe es möglich sein sollte, die Teilnehmer einer der Bedingungen zuzuweisen. Diese Liste bestand aus einer zufälligen Anordnung der Zahlen 1 bis 4, wobei jede Ziffer eine Experimentalgruppen repräsentierte und genau 21 mal enthalten war. Die Teilnehmer wurden dann nach der Reihenfolge ihres Erscheinens den einzelnen Versuchsbedingungen zugeordnet.

Die eigentliche Testsitzung begann, wenn alle Versuchspersonen, die sich angemeldet hatten, erschienen waren oder nach einer kurzen Wartezeit von 5 Minuten, so dass auch Nachzügler noch teilnehmen konnten. Nachdem der Versuchsleiter die Tür zum Untersuchungsraum geschlossen hatte, wies er den Teilnehmern einen Computerarbeitsplatz zu. Anschließend wurde den Studierenden mitgeteilt, was sie in der Sitzung zu erwarten hätten und dass ihre Daten vertraulich bzw. anonym behandelt würden. Bestanden keine weiteren Fragen, sollten die Versuchsteilnehmer den ersten Abschnitt des Fragebogens ausfüllen, der zuvor ausgeteilt worden war. Dabei ging es um Angaben zur Person, aber auch um Lerngewohnheiten, das Interesse für Geologie und die Erdkunde- bzw. Geographienote. Die Beantwortung der Fragen nahm in der Regel nicht mehr als fünf Minuten in Anspruch.

Im Anschluss erhielten die Studierenden die Untertest *Satzergänzung (SE)* aus dem IST 2000 R (Lipmann et al., 2007), zu dessen Bearbeitung 6 Minuten zur Verfügung standen (siehe 5.2.5). Zuvor hatte der Versuchsleiter, die im Manual angeführten Instruktionen verlesen und den Teilnehmern die Möglichkeit gegeben, alle bestehenden Unklarheiten zu klären. Die Bearbeitungszeit wurde mit einer Stoppuhr gemessen, wobei die meisten Versuchspersonen jedoch in der Regel keine 6 Minuten benötigten, um die 20 Items zu beantworten. Die restliche Zeit konnten die Teilnehmer dazu nutzen, ihre Antworten noch einmal zu überprüfen. Erst nachdem genau 6 Minuten verstrichen waren, wurden die Studierenden dazu aufgefordert, ihre Stifte beiseite zu legen. Anschließend gab es eine kurze Pause, bevor der Versuchsleiter die Anweisungen für die zweite Testskala aus dem IST 2000 R verlas. Diese bestand wiederum aus 20 Items, bei denen allerdings Würfel mental rotiert werden mussten, um eine Übereinstimmung zu erkennen (siehe Abbildung 5.13 auf Seite 125). Offenbar stellte die Bearbeitung der *Würfelaufgaben (WÜ)* eine sehr viel größere Herausforderung dar, denn in den meisten Fällen reichte die Bearbeitungszeit von 9 Minuten nicht aus, um alle Items zu lösen.

Nach der Erfassung der räumlichen Fähigkeiten, wurde erneut eine kurze Pause eingelegt, in der sich die Teilnehmer etwas erholen konnten, bevor es mit dem Vorwissenstest weiterging (siehe Abschnitt 5.2.3). Anders als bei den Skalen aus dem IST-2000R gab

6 Studie I: Interferenzen versus Kapazität

es für die Bearbeitung des Vorwissenstests keine zeitliche Vorgabe. Allerdings wurden die Studierenden gebeten, nur dann eine der beiden Antwortmöglichkeiten anzukreuzen, wenn sie sich einigermaßen sicher waren. Items, bei denen die Teilnehmer die Antwort wirklich nicht wussten, sollten lieber ausgelassen werden. Durch diese Instruktion sollte verhindert werden, dass die Teilnehmer die Antworten errieten, was letztlich zu einer Verzerrung der Ergebnisse geführt hätte. Der letzte Vortest bestand schließlich aus zwei Skalen einer überarbeiteten Fassung des Verbalizer-Visualizer-Questionnaires (siehe Kirby et al. (1988) bzw. Abschnitt 5.14), der zur Erfassung verbaler bzw. visueller Verarbeitungspräferenzen vorgelegt wurde. Auch hier gab es keine zeitliche Begrenzung, doch sollten die Teilnehmer laut Instruktion möglichst rasch und ohne langes Nachdenken antworten. Nachdem auch dieser Vortest abgeschlossen war, wurden die Fragebögen eingesammelt und das Experiment konnte beginnen.

Dabei war die experimentelle Intervention nicht viel mehr als eine Lernphase, bei der die Studierenden mit Hilfe von instruktionalen Bildern und Texten neues Wissen erwerben sollten. Allerdings unterlag der Prozess des Wissenserwerbs gewissen Einschränkungen, über die der Versuchsleiter die Teilnehmenden im Vorfeld aufgeklärt hatte. So wurden die betreffenden Bilder und Texte auf Computermonitoren präsentiert, die jeweils so eingestellt waren, dass beide Formate genau auf eine Bildschirmseite passten. Auf diese Weise waren die Studierenden nicht dazu gezwungen, zu scrollen, da alle Informationen auf einen Blick erfasst bzw. gelesen werden konnten. Die Teilnehmer mussten die zur Verfügung stehende Maus nur dann verwenden, wenn sie alle Informationen aufgenommen hatten und bereit waren, mit dem nächsten Bild bzw. Text fortzufahren. Dazu konnten sie auf einen kleinen *Vorwärtspfeil* klicken, der sich in der rechten Ecke des Bildschirms befand.

Das Lehr-Lernmaterial, dessen Eigenschaften bereits in Abschnitt 5.1.4 ausführlich dargestellt wurden, umfasste jeweils einen Text und ein Bild zu insgesamt vier thematischen Bereichen aus der Geologie. Die einzelnen Abschnitte bauten aufeinander auf, so dass unter allen experimentellen Bedingungen inhaltlich immer die selbe Reihenfolge eingehalten wurde. Jede Lernphase begann mit der *Theorie der Plattentektonik*, dann folgte eine Ausführung über die *Plattengrenzen*, anschließend die *Messungen im Erdinneren*, bevor sie mit Bildern und Texten zu den *Mechanismen der Plattenbewegung* endete. Neben dem Informationsverhältnis wurde lediglich die Sequenz variiert, in der die Teilnehmer Text und Bild präsentiert bekamen. Die Personen in der einen Gruppe lasen immer zuerst den Text, bevor sie das Bild betrachten konnten. Die Studierenden der anderen Gruppe betrachteten immer zuerst das Bild und lasen dann den Text. Dabei konnten die Teilnehmer selbständig bestimmen, wie lange sie auf den einzelnen HTML-Seite verweilen wollten. Somit gab es zwar keine Lernzeitbeschränkung, doch hatte sich die Versuchsperson entschieden, fortzufahren und den entsprechenden Vorwärtspfeil angeklickt, gab es kein Zurück mehr. Die Verarbeitungsreihenfolge war streng linear, so dass jeder Text und jedes Bild immer nur ein einziges Mal aufgerufen werden konnte.

Da das Lernmaterial rechnergestützt präsentiert wurde, stellte es kein großes Problem dar, die Lernzeit zu erheben. Zu diesem Zweck bot es sich an, die Verweildauer auf den

Seiten heranzuziehen, die sich aus der zeitlichen Differenz zwischen dem Aufruf und dem Verlassen der Seite bestimmen ließ. Hatten die Studierenden alle 8 HTML-Seiten betrachtet bzw. gelesen, erschien eine letzte Seite, die den Teilnehmenden dazu aufforderte sich an den Versuchsleiter zu wenden. Dieser gab dann einen zweiten Fragebogen aus, der neben Items zur aktuellen Motivation (Abschnitt 5.2.4) und der kognitiven Belastung (Abschnitt 5.2.2) vor allem den Test enthielt, mit dem der Erfolg des Wissenserwebs evaluiert werden sollte (siehe Kapitel 5.2.1). Aufgrund der unreglementierten Lernzeit, beendeten die Versuchspersonen die experimentelle Phase nicht gleichzeitig, so dass die Instruktionen in der Regel jedem Teilnehmer einzeln gegeben werden mussten. Während die Fragen zur aktuellen Motivation und der kognitiven Belastung keiner speziellen Anweisungen bedurften, wurden beim Lerntest erneut darum gebeten, die Antworten der *multiple choice*-Items möglichst nicht zu raten. Für die Bearbeitung des Abschlussfragebogens gab es keine zeitliche Beschränkungen, so dass die Teilnehmer alle Items ohne Druck beantworten konnten. Nachdem die Studierenden fertig waren, kamen sie einzeln zum Versuchsleiter und gaben den ausgefüllten Fragebogen ab. Zum Dank für ihre Teilnahme erhielten sie dann entweder eine Bestätigung über 2 Versuchspersonenstunden oder aber 10,- €, womit die Testsitzung beendet war.

6.2.4 Statistische Auswertung

Die Prüfung der Forschungshypothesen erfolgte anhand von varianzanalytischen Verfahren, die mit der Software SPSS 17.0 durchgeführt wurden. Da inferenzstatistische Berechnungen immer nur unter bestimmten Bedingungen gültige Schlüsse zulassen, wurden vor jeder Analyse zuerst die Voraussetzungen getestet. Dies betraf insbesondere die Annahme varianzhomogener Populationen und die Normalverteilung der abhängigen Variablen, was mit einem Test von (Levene, 1960) bzw. dem Kolmogorow-Smirnow-Anpassungstest überprüft wurde (vgl. Goodman, 1954). Vor der Verwendung multivariater Verfahren wurde zudem mit einem Verfahren nach Box (1949) getestet, ob die Kovarianzmatrizen zwischen den experimentellen Gruppen vergleichbar waren. Da der Box-Test gleichzeitig sehr sensibel auf die Verletzung der multivariaten Normalverteilung reagiert (Bortz, 1999), konnte bei einem nicht signifikanten Ergebnis auch diese Annahme als erfüllt gelten.

Allen Unterschiedshypothesen wurde ein Signifikanzniveau von $\alpha = .05$ zugrundegelegt, wobei die Angaben zu den empirischen Wahrscheinlichkeiten p immer einer zweiseitigen Testung entsprechen. Manche Fragestellungen machten es jedoch erforderlich, statistisch die Gültigkeit der Nullhypothese (H_0) abzusichern. Solche Äquivalenztests erforderten jedoch ein besonderes Vorgehen, das teilweise von den Überlegungen der sonst üblichen Testprozedur abweicht. Während bei einer *normalen* Testung mit der Festlegung der Irrtumswahrscheinlichkeit α vor allem der Fehler erster Art kontrolliert wird, findet der β -Fehler oft nur nachträglich im Rahmen von (post-hoc) Poweranalysen Beachtung. Für die Prüfung der Gültigkeit der H_0 ist es jedoch zwingend erforderlich, die Wahrscheinlichkeit für einen β -Fehler abzusenken bzw. die Teststärke ($1 - \beta$) zu steigern, da bei der

vorliegenden Fragestellung das Risiko einer fälschlichen Beibehaltung der Nullhypothese minimiert werden muss. Zur Kontrolle des β -Fehlers gibt es mehrere Möglichkeiten (Diehl & Arbinger, 1990), wobei die einfachste Methode darin besteht, das Signifikanzniveau α zu erhöhen. So empfehlen Bortz und Döring (2006) für die Testung von Äquivalenzannahmen ein α von 0.20, um die Teststärke auf ein vertretbares Niveau zu heben. Dieser Vorschlag wurde übernommen, wobei auch hier die Angaben zu den empirischen Wahrscheinlichkeiten p immer einer zweiseitigen Testung entsprachen.

Neben den üblichen Kennzahlen, wie den Prüfgrößen und den Freiheitsgraden, schien es angebracht zu sein, ein Maß für die praktische Bedeutsamkeit der gefundenen Zusammenhänge zu berichten. Um die Stärke der Effekte zu veranschaulichen, wurde η^2 gewählt, da dessen Berechnung keine besonderen Voraussetzungen erfordert und relativ leicht zu interpretieren ist (vgl. Diehl & Arbinger, 1990).

6.3 Ergebnisse

Die Ergebnisse einer empirischen Arbeit anschaulich und übersichtlich zu berichten, ist keine leichte Aufgabe, da mit vielen Zahlen und Kennwerten hantiert werden muss. Hinzu kommen Diskussionen, ob die verwendeten statistischen Verfahren, den Anforderungen der Stichprobe und damit den Daten genügen können oder nicht. Um die Resultate dennoch möglichst nachvollziehbar darzustellen, werden die Ergebnisse kapitelweise präsentiert.

6.3.1 Auswahl der der Kontrollvariablen

Es gibt verschiedene Möglichkeiten die interne Validität einer Untersuchung zu steigern, wobei sicherlich der eleganteste Weg die Randomisierung darstellt. Doch obwohl auf diese Weise der Einfluss personengebundener *Störvariablen* theoretisch nivelliert wird, bleibt gerade bei kleinen Stichproben immer ein gewisse Gefährdung bestehen. So kann beispielsweise die zufällige Häufung besonders intelligenter Personen in der einen Gruppe die tatsächlichen Zusammenhänge verzerren. Die offensichtliche Lösung dieses Problems ist eine Erhöhung des Stichprobenumfangs, was jedoch aus ökonomischen Gründen nicht immer als sinnvoll erachtet werden kann. Ein anderer Weg besteht darin, die potentiellen Störvariablen bei der Untersuchung mit zu erheben und statistisch zu kontrollieren. Indem diese *Kontrollvariablen* im Rahmen einer Kovarianzanalyse berücksichtigt werden, wird ihr Einfluss auf die eigentlich interessante Variable regressionstechnisch eliminiert.

Hinter diesem Vorgehen steht die Absicht die Fehlervarianz zu reduzieren, wodurch sich die Teststärke $(1 - \beta)$ erhöht, da die Wahrscheinlichkeit für einen β -Fehler abnimmt. Auf den ersten Blick wäre es daher sinnvoll, so viele potentielle Störgrößen wie möglich zu erheben, damit der Einfluss der experimentellen Bedingungen besonders *rein* gemessen werden kann. Allerdings reduziert sich mit jeder Kovariate, die in die statistische Analyse

eingeht, die Anzahl der Freiheitsgrade um 1, was bei der Signifikanztestung einer Verkleinerung der Stichprobe gleichkommt. Die Auswahl der Kontrollvariablen muss demnach mit Bedacht erfolgen, da eine Berücksichtigung von Kovariaten mit nur geringem Einfluss auf die Varianz der abhängigen Variablen letztlich zu einer Verringerung der Teststärke führt.

Auch im Rahmen der vorliegenden Arbeit wurden potentielle Störgrößen erhoben, wobei diese jedoch nicht zwingend in der Analyse berücksichtigt werden sollten. Die Entscheidung, eine Kontrollvariable aufzunehmen, machte nur dann Sinn, wenn sich deren durchschnittliche Ausprägung nicht gleichmäßig auf die einzelnen Experimentalgruppen verteilte. Aus diesem Grund wurden Äquivalenztests gegen eine α -Fehlerwahrscheinlichkeit von 20% durchgeführt. Als Prüfgröße dienten die F-Werte einer ANOVA mit den zweistufigen Faktoren *Reihenfolge* und *Informationsverhältnis*. Die Interpretation von F-Werten setzt jedoch voraus, dass die abhängige Variable normalverteilt ist und sich die Fehlervarianzen in den Gruppen nicht unterscheiden. Tabelle 6.1 zeigt das Ergebnis der Kolmogorow-Smirnow- und Levene-Tests für die einzelnen Kontrollvariablen.

Tabelle 6.1: Kolmogorow-Smirnow- und Levene-Tests für die Auswahl der Kontrollvariablen

Merkmal	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Interesse</i>	<i>Levene – Test</i>	$F = 0.04$	$df_1 = 3; df_2 = 80$	$p = .988$
	<i>K – S – Test</i>	$Z_{K-S} = 2,10$	–	$p < .001^{**}$
<i>Erdkundenote</i>	<i>Levene – Test</i>	$F = 1.44$	$df_1 = 3; df_2 = 80$	$p = .237$
	<i>K – S – Test</i>	$Z_{K-S} = 1,83$	–	$p = .002^{**}$
<i>Vorwissen</i>	<i>Levene – Test</i>	$F = 1.15$	$df_1 = 3; df_2 = 80$	$p = .333$
	<i>K – S – Test</i>	$Z_{K-S} = 1,05$	–	$p = .223$
<i>Motivation</i> <i>FAM</i>	<i>Leven – Test</i>	$F = 0.72$	$df_1 = 3; df_2 = 80$	$p = .544$
	<i>K – S – Test</i>	$Z_{K-S} = 0.56$	–	$p = .910$
<i>Verbale</i> <i>Intelligenz</i>	<i>Leven – Test</i>	$F = 1.48$	$df_1 = 3; df_2 = 80$	$p = .227$
	<i>K – S – Test</i>	$Z_{K-S} = 1,25$	–	$p = .089$
<i>Räumliche</i> <i>Intelligenz</i>	<i>Levene – Test</i>	$F = 0.55$	$df_1 = 3; df_2 = 80$	$p = .649$
	<i>K – S – Test</i>	$Z_{K-S} = 1,11$	–	$p = .167$
<i>Lernstil</i> <i>Verbalizer</i>	<i>Levene – Test</i>	$F = 0.44$	$df_1 = 3; df_2 = 80$	$p = .726$
	<i>K – S – Test</i>	$Z_{K-S} = 1,23$	–	$p = .098$
<i>Lernstil</i> <i>Visualizer</i>	<i>Levene – Test</i>	$F = 1.58$	$df_1 = 3; df_2 = 80$	$p = .200$
	<i>K – S – Test</i>	$Z_{K-S} = 2.04$	–	$p < .001^{**}$

Anm.: *K-S-Test* = Kolmogorow-Smirnow-Anpassungstest; *signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen (* < .05) bzw. durch zwei Sternchen (** < .01) hervorgehoben*

Es ist zu erkennen, dass die Fehlervarianzen bei allen Kontrollvariablen durchaus vergleichbar waren, weshalb die erste Voraussetzung als erfüllt angesehen werden konnte. Anders verhielt es sich dagegen mit der Normalverteilungsannahme, die nicht in jedem

6 Studie I: Interferenzen versus Kapazität

Fall zutraf. So überschritten die *Visualizer-Skala*, das *Interesse für Geologie* und die Angaben zu der *Erdkundenote* den kritischen Wert der Kolmogorow-Smirnow-Prüfstatistik deutlich, was es ziemlich unwahrscheinlich machte, dass diese Merkmale normalverteilt waren. Milligan, Wong und Thompson (1987) konnten in Simulationsstudien jedoch belegen, dass die F-Werte zweifaktorieller Varianzanalysen durchaus gültige Schlüsse zulassen, selbst wenn das Kriterium der Normalverteilung verletzt ist. Allerdings müssen dazu die Zellen gleichmäßig besetzt bzw. der Versuchsplan *orthogonal* sein, was im Rahmen der ersten Studie gegeben war. Aufgrund dieser relativen Robustheit konnte letztlich für alle Kontrollvariablen Varianzanalysen gerechnet werden, um deren Verteilung zu prüfen.

Für die Personenmerkmale *Interesse* und *Erdkundenote* ergab die ANOVA keine signifikanten Ergebnisse, weshalb sie nicht als Kovariaten berücksichtigt werden mussten. Etwas anders verhielt es sich mit dem *domänenspezifischen Vorwissen*, bei dem ein Unterschied gefunden wurde. Während der Faktor *Informationsverhältnis* ($F_{(1;80)} = 0.056$, $p = .813$, $\bar{\eta}^2 = .001$) und der Interaktionseffekt ($F_{(1;80)} = 0.182$, $p = .671$, $\bar{\eta}^2 = .002$) den kritischen Wert jeweils verfehlten, lag die Irrtumswahrscheinlichkeit beim Vergleich der *Reihenfolge* ($F_{(1;80)} = 2.748$, $p = .101$, $\bar{\eta}^2 = .033$) unter dem Wert des Signifikanzniveau von $\alpha = .20$. Damit galt die Äquivalenzannahme als verletzt, weshalb das *Vorwissen* als Kontrollvariable in die Analyse aufgenommen wurde. Auch das Merkmal der *Motivation*, das mit einer Kurzform des *FAM* (siehe Abschnitt 5.2.4) erhoben wurde, war offensichtlich nicht gleichmäßig verteilt. Zwar zeigten sich beim *Informationsverhältnis* ($F_{(1;80)} = 1.064$, $p = .305$, $\bar{\eta}^2 = .013$) und der Interaktion ($F_{(1;80)} = 0.898$, $p = .346$, $\bar{\eta}^2 = .011$) wiederum keine bedeutsamen Unterschiede. Doch überschritt die Prüfstatistik des Faktors *Reihenfolge* ($F_{(1;80)} = 2.101$, $p = .151$, $\bar{\eta}^2 = .026$) den kritischen Wert, weshalb auch die Motivation statistisch kontrolliert werden musste.

Für die beiden Intelligenz- und Lernstilskalen wurde MANOVAs berechnet, um die Äquivalenzannahme zu prüfen. Die betreffenden Skalen waren jeweils Teil eines übergeordneten Konstruktes, so dass zumindest theoretisch davon ausgegangen werden musste, dass die beiden Messwerte nicht unabhängig waren. Da die MANOVA aber besondere Anforderungen an die abhängigen Variablen und deren Verhältnis zu einander stellt, wurde vor der Berechnung der Box-Test durchgeführt, der jedoch weder im Fall der Intelligenz ($M_{Box} = 11.59$, $F = 1.229$, $df_1 = 9$, $df_2 = 73342.77$, $p = .271$), noch bei den Lernstilskalen ($M_{Box} = 5.98$, $F = .634$, $df_1 = 9$, $df_2 = 73342.77$, $p = .769$) signifikant wurde. Die multivariaten Tests bestätigten, dass die Leistung in den IST-Skalen als gleichverteilt gelten konnten, während die Ergebnisse zum Lernstil diese Deutung jedoch nicht zuließen. So zeigte der Kennwert Pillai's Spur (V) einen Unterschied zwischen den Gruppen des Faktors *Reihenfolge* an ($V = .057$, $F_{(2;79)} = 2.373$, $p = .100$, $\bar{\eta}^2 = .057$), der eine α -Fehlerwahrscheinlichkeit von 20% unterschritt. Das *Informationsverhältnis* ($V = .017$, $F_{(2;79)} = .689$, $p = .505$, $\bar{\eta}^2 = .017$) und der Interaktionseffekt ($V = .001$, $F_{(2;79)} = .055$, $p = .946$, $\bar{\eta}^2 = .001$) waren zwar nicht signifikant, doch musste die Äquivalenzannahme als verletzt angesehen werden.

Fasst man die Untersuchung zu den Kontrollvariablen zusammen, dann fällt auf, dass die Abweichungen in erster Linie den Faktor *Reihenfolge* betrafen. Offensichtlich unter-

schieden sich die Teilnehmer dieser Gruppen trotz der randomisierten Zuweisung, wie Tabelle 6.2 zeigt, die Zellenmittelwerte und Standardabweichung der berücksichtigten Kovariaten enthält.

Tabelle 6.2: Mittelwerte und Standardabweichung der Kontrollvariablen pro Versuchsbedingung

Merkmal	Versuchsbedingungen			
	Text vor Bild kongruent	Text vor Bild komplementär	Bild vor Text kongruent	Bild vor Text komplementär
<i>Vorwissen</i>	$\bar{X} = 6.76$ $s_{\bar{X}} = 2.16$	$\bar{X} = 6.43$ $s_{\bar{X}} = 2.20$	$\bar{X} = 5.71$ $s_{\bar{X}} = 1.95$	$\bar{X} = 5.81$ $s_{\bar{X}} = 2.80$
<i>Motivation</i> <i>FAM</i>	$\bar{X} = 3.16$ $s_{\bar{X}} = 0.82$	$\bar{X} = 3.18$ $s_{\bar{X}} = 0.70$	$\bar{X} = 2.76$ $s_{\bar{X}} = 0.90$	$\bar{X} = 3.01$ $s_{\bar{X}} = 0.68$
<i>Lernstil</i> <i>Verbalizer</i>	$\bar{X} = 6.24$ $s_{\bar{X}} = 1.76$	$\bar{X} = 6.19$ $s_{\bar{X}} = 2.32$	$\bar{X} = 7.10$ $s_{\bar{X}} = 2.05$	$\bar{X} = 6.95$ $s_{\bar{X}} = 1.94$
<i>Lernstil</i> <i>Visualizer</i>	$\bar{X} = 7.76$ $s_{\bar{X}} = 2.07$	$\bar{X} = 8.47$ $s_{\bar{X}} = 1.69$	$\bar{X} = 7.19$ $s_{\bar{X}} = 2.44$	$\bar{X} = 7.62$ $s_{\bar{X}} = 2.56$

Anm.: \bar{X} = Mittelwert; $s_{\bar{X}}$ = Standardabweichung, in jeder Zelle des 2x2 Versuchsplans befanden sich n=21 Personen

Das betrifft vor allem das *Vorwissen*, aber auch die *Motivation*. Demnach verfügten die Personen in den Text-vor-Bild-Gruppen nicht nur über mehr domänenspezifische Kenntnisse, sondern waren gleichzeitig auch motivierter. Es erschien also durchaus angebracht zu sein, diese Unterschiede regressionstechnisch zu bereinigen, um einer Verzerrung der Ergebnisse entgegen zu wirken. Ob auch die Kontrolle des Lernstils die Teststärke erhöhen würde, war nicht ganz so sicher, da nur ein schwacher Zusammenhang des Merkmals mit den abhängigen Variablen vermutet werden konnte. Allerdings zeigt Tabelle 6.2, dass durchaus Unterschiede hinsichtlich der Verarbeitungspräferenz bestanden. Während in den Text-vor-Bild-Gruppen tendenziell mehr Visualizer zu finden waren, bevorzugten die Teilnehmer in den Bild-vor-Text-Gruppen eher verbales Material für den Wissenserwerb. Die Zusammenhänge waren zwar schwach, doch da sich die Anzahl der Kovariaten im Rahmen hielt, sollte auch der Lernstil statistisch kontrolliert werden. Allerdings wurde aus der Visualizer- und Verbalizerskala ein Wert gebildet (vgl. Kirby et al., 1988), womit sich die Reduzierung der Freiheitsgrade in Grenzen hielt.

6.3.2 Forschungsfrage 1: Behalten & Verstehen

Einer der zentralen Aspekte des *picture-text-sequencing*-Effekts betrifft die Frage nach dem *Lernerfolg*. Um zu erfassen wie viel bzw. oder wie gut die Versuchsteilnehmer in den verschiedenen Experimentalgruppen gelernt haben, wurde den Studierenden nach der Treatmentphase ein Wissenstest vorgelegt, der in Kapitel 5 detailliert beschrieben ist. An dieser Stelle sei nur darauf hingewiesen, dass die zugrunde liegende Operationalisierung des Konzepts *Lernerfolg* insgesamt vier Skalen umfasste, von denen 2 die Verständnis-

6 Studie I: Interferenzen versus Kapazität

leistung und 2 die Behaltensleistung messen sollten. Inwieweit die einzelnen Skalen alle Voraussetzungen erfüllten, um im Rahmen von varianzanalytischen Verfahren verwendet werden zu können, zeigt Tabelle 6.3

Tabelle 6.3: Kolmogorow-Smirnow- und Levene-Tests für die Skalen des Lerntests

Skala	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Wiedererkennen</i>	<i>Levene – Test</i>	$F = 0.02$	$df_1 = 3; df_2 = 80$	$p = .996$
	<i>K – S – Test</i>	$Z_{K-S} = 1.33$	–	$p = .056$
<i>freies Erinnern</i>	<i>Levene – Test</i>	$F = 0.52$	$df_1 = 3; df_2 = 80$	$p = .668$
	<i>K – S – Test</i>	$Z_{K-S} = 1.20$	–	$p = .112$
<i>Inferenz MC</i>	<i>Levene – Test</i>	$F = 1.69$	$df_1 = 3; df_2 = 80$	$p = .176$
	<i>K – S – Test</i>	$Z_{K-S} = 1.142$	–	$p = .148$
<i>Inferenz offen</i>	<i>Levene – Test</i>	$F = 0.21$	$df_1 = 3; df_2 = 80$	$p = .889$
	<i>K – S – Test</i>	$Z_{K-S} = 2.05$	–	$p < .001^{**}$
<i>Gesamtscore</i>	<i>Levene – Test</i>	$F = 0.75$	$df_1 = 3; df_2 = 80$	$p = .527$
	<i>K – S – Test</i>	$Z_{K-S} = 0.73$	–	$p = .659$

Anm.: K-S-Test = Kolmogorow-Smirnow-Anpassungstest; signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen (* < .05) bzw. durch zwei Sternchen (** < .01) hervorgehoben

Wie unschwer zu erkennen ist, konnten die Zellenpopulationen als varianzhomogen angesehen werden, da für keine der Skalen die Prüfstatistik des Levene-Tests einen kritischen Wert erreichte. Anders verhielt es sich mit der Normalverteilungsannahme, die für die Skala *offene Inferenzfragen* nicht zu halten war. Im Zusammenhang mit der Testung potentieller Kovariaten wurde jedoch bereits darauf hingewiesen (Seite 6.3.1), dass sich varianzanalytische Verfahren gegenüber einer Verletzung der Normalverteilungsannahme als relativ robust erwiesen haben. Demnach war nicht damit zu rechnen, dass die Berechnung einer 2X2 MANCOVA mit den Kontrollvariablen *Vorwissen*, *Motivation* und dem *Verbalizer-Visualizer-Score* und den Faktoren *Reihenfolge* und *Informationsverhältnis* zu Problemen führen sollte. Dies galt umso mehr, da auch der Box-Test keine signifikantes Ergebnis ergab ($M_{Box} = 30.00$, $F = .909$, $df_1 = 30$, $df_2 = 17596.18$, $p = .609$), so dass die Teststatistiken ohne allzu große Bedenken interpretiert werden konnten.

Bei den Ergebnissen der Analyse fiel zunächst auf, dass alle drei Kovariaten signifikant wurden. Demnach trugen das *Vorwissen* ($V = .161$, $F_{(4;74)} = 3.561$, $p = .010$, $\bar{\eta}^2 = .161$), die *Motivation* ($V = .153$, $F_{(4;74)} = 3.350$, $p = .014$, $\bar{\eta}^2 = .153$) und selbst die *Verarbeitungspräferenz* ($V = .201$, $F_{(4;74)} = 4.649$, $p = .002$, $\bar{\eta}^2 = .201$) dazu bei, die Fehlervarianzen der vier abhängigen Variablen bedeutsam zu reduzieren. Womöglich ist es auf diese *Bereinigung* zurückzuführen, dass neben dem *Informationsverhältnis* ($V = .287$, $F_{(4;74)} = 7.455$, $p < .001$, $\bar{\eta}^2 = .287$) auch der Faktor *Reihenfolge* ($V = .146$, $F_{(4;74)} = 3.153$, $p = .019$, $\bar{\eta}^2 = .146$) signifikant wurde. Die Interaktion ($V = .003$, $F_{(4;74)} = 0.049$, $p = .995$, $\bar{\eta}^2 = .003$) zwischen den beiden Faktoren erwies sich dagegen als unbedeutsam. Das Ergebnis die-

ser Analyse zeigte zwar recht klar, dass Unterschiede zwischen den Experimentalgruppen vorhanden waren, doch welche Richtung diese hatten, ließ sich erst anhand der Zellenmittelwerte bestimmen, die Tabelle 6.4 enthält.

Tabelle 6.4: Adjustierte Mittelwerte und deren Standardfehler für die Skalen des Lerntests

Lerntestskala	Versuchsbedingungen			
	Text vor Bild kongruent	Text vor Bild komplementär	Bild vor Text kongruent	Bild vor Text komplementär
<i>Wiedererkennen</i>	$M = 7.10$ $SE = 0.36$	$M = 6.02$ $SE = 0.36$	$M = 7.56$ $SE = 0.37$	$M = 6.18$ $SE = 0.36$
<i>freies Erinnern</i>	$M = 5.21$ $SE = 0.39$	$M = 3.48$ $SE = 0.39$	$M = 6.11$ $SE = 0.40$	$M = 4.24$ $SE = 0.39$
<i>Inferenz MC</i>	$M = 4.83$ $SE = 0.36$	$M = 3.45$ $SE = 0.36$	$M = 5.79$ $SE = 0.36$	$M = 4.46$ $SE = 0.36$
<i>Inferenz MC</i>	$M = 0.96$ $SE = 0.21$	$M = 0.93$ $SE = 0.21$	$M = 1.58$ $SE = 0.21$	$M = 1.48$ $SE = 0.20$

Anm.: M = Mittelwert; SE = Standardfehler des Mittelwerts, adjustiert nach den Kovariaten Vorwissen, Motivation und Verarbeitungspräferenz

Wie zu erkennen ist, war es durchweg vorteilhafter, wenn die Versuchsteilnehmer beim Wissenserwerb kongruente Lehrmaterialien nutzen konnten und/oder das Bild vor dem Text präsentiert wurde. Dieser Zusammenhang betraf jedoch nicht alle Skalen gleichermaßen. Während die Mittelwerte der Skalen *freies Erinnern* und *Inferenzen MC* mit den Haupteffekten der MANCOVA ohne Probleme zu vereinbaren waren, ließen die Ergebnisse der MC-Gedächtnisaufgaben bzw. den offenen Inferenzitems eine solch einfache Deutung nicht zu. Eine univariate Analyse ergab im Fall der Skala *Wiedererkennen* einen signifikanten Haupteffekt für den Faktor *Informationsverhältnis* ($F_{(1;77)} = 11.631$, $p < .001$, $\bar{\eta}^2 = .131$), doch blieb der Einfluss der *Reihenfolge* ($F_{(1;77)} = 0.704$, $p = .404$, $\bar{\eta}^2 = .009$) ohne Bedeutung. Die ANCOVA für die offenen Inferenzitems hatte das gegenteilige Resultat zur Folge. Hier wurde der Faktor *Reihenfolge* ($F_{(1;77)} = 7.774$, $p = .007$, $\bar{\eta}^2 = .092$) signifikant, während die Mittelwertsunterschiede zwischen den Gruppen mit komplementären und kongruenten Lernmaterialien ($F_{(1;77)} = 0.088$, $p = .768$, $\bar{\eta}^2 = .001$) den kritischen F-Wert deutlich verfehlt hatten.

Die Rangreihe der Zellenmittelwerte war jedoch mit nur einer Ausnahme für die abhängigen Variablen immer die gleiche. Das durchgehend beste Ergebnis erzielte die Bild-vor-Text-Gruppe der kongruenten Versuchsbedingung, während die Teilnehmer, die zum Wissenserwerb das komplementäre Material nutzen mussten und den Text vor dem Bild lasen, auf allen vier Skalen am schlechtesten abschnitten. In drei von vier Fällen erreichte die Text-vor-Bild-Gruppe mit den kongruenten Lernmaterialien das zweitbeste Ergebnis, während die komplementäre Bild-vor-Text-Bedingung hinsichtlich der Leistungsvariablen in der Regel auf dem dritten Rang landete. Die einzige Ausnahme betraf die Skala *Inferenz offen*, da hier beide Bild-vor-Textgruppen deutlich mehr Items lösen konnten, als die

6 Studie I: Interferenzen versus Kapazität

anderen beiden Bedingungen.

Um letztlich aber zu entscheiden, welche der drei Hypothesen der ersten Forschungsfrage am besten mit diesen Ergebnissen übereinstimmte, war es notwendig, eine kleine Vereinfachung vorzunehmen. Da die MANCOVA eindeutig zwei signifikante Haupteffekte ergeben hatte und deren Richtung aufgrund der Rangreihe der adjustierten Mittelwerte ziemlich sicher bestimmt werden konnte, wurde über alle vier Skalen der *Gesamtscore* berechnet. Das Ergebnis dieser Datenaggregation zeigt Abbildung 6.2.

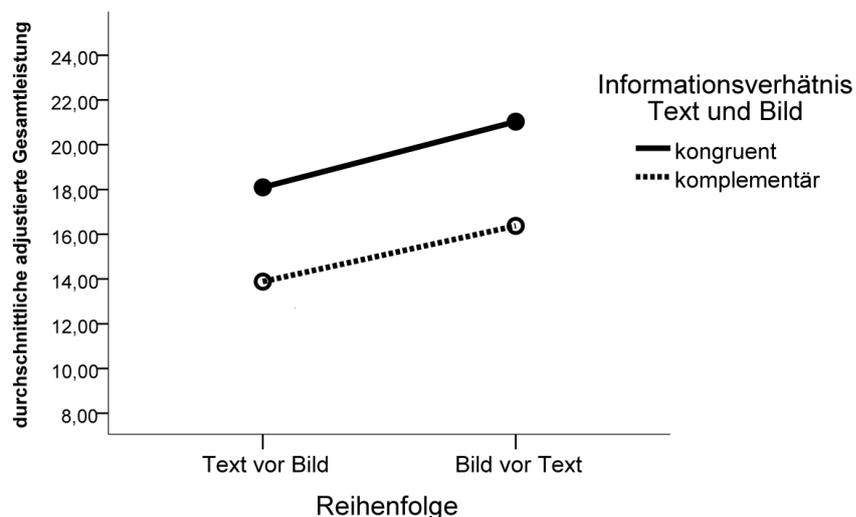


Abbildung 6.2: Mittlere adjustierte Gesamtleistung im Wissenstest

In dem Diagramm ist gut zu erkennen, dass sowohl die *Reihenfolge* als auch das *Informationsverhältnis* zwischen Text und Bild den Lernerfolg beeinflusst hatten. So veranschaulicht die Steigung der beiden Linien, dass der *picture-text-sequencing*-Effekt tatsächlich nachgewiesen werden konnte. Versuchsteilnehmer, die zuerst das Bild für den Wissenserwerb nutzten, erzielten demnach einen höheren Gesamtscore, als Studierende, die zuerst den Text gelesen hatten. Zudem deutet der Abstand zwischen den Linien darauf hin, dass ein kongruentes Informationsverhältnis zwischen dem visuellen und verbalen Format sich durchgehend als vorteilhaft erwies. Gleichzeitig ist der parallele Verlauf ein deutliches Zeichen dafür, dass die beiden Faktoren nicht miteinander interagieren. Dieses Ergebnis entspricht genau den Vorhersagen, die anhand des Modells von (Kulhavy et al., 1994) abgeleitet wurden. Obwohl die Gruppen mit kongruentem Lernmaterial besser abschnitt, als die Teilnehmern mit komplementärem Lernmaterial, war in beiden Fällen die Bild-vor-Text-Bedingung überlegen. Damit konnte die *Kapazitätshypothese* als gültig angesehen werden, während die Interferenzhypothese zurückgewiesen werden mussten.

6.3.3 Forschungsfrage 2: Mentale Belastung (*cognitive load*)

Das zweite Merkmal, von dem anzunehmen war, dass es durch die Verarbeitungsreihenfolge und das Informationsverhältnis beeinflusst wurde, war die mentale Belastung bzw. der *cognitive load*. Dieser konnte eine Folge des Lernprozesses sein oder entstehen, wenn die Versuchsteilnehmer ihr Wissen abriefen, um die Testitems zu lösen. Aus diesem Grund war die mentale Belastung sowohl nach der Präsentationsphase als auch nach dem Wissenstest erhoben worden. Um ihre Ursache näher zu bestimmen, wurden die Studierenden zusätzlich um eine Einschätzung der Text-, Bild- und Testschwierigkeit gebeten. Vor der Prüfung der Forschungshypothesen, musste jedoch erneut untersucht werden, ob die abhängigen Variablen alle Voraussetzungen der Varianzanalyse erfüllten. Tabelle 6.5 gibt die Ergebnisse wieder.

Tabelle 6.5: Kolmogorow-Smirnow- und Levene-Tests für empfundene Schwierigkeit und mentale Belastung

Skala	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Schwierigkeit Bild</i>	<i>Levene – Test</i>	$F = 0.67$	$df_1 = 3; df_2 = 80$	$p = .571$
	<i>K – S – Test</i>	$Z_{K-S} = 2.01$	–	$p < .001^{**}$
<i>Schwierigkeit Text</i>	<i>Levene – Test</i>	$F = 0.54$	$df_1 = 3; df_2 = 80$	$p = .654$
	<i>K – S – Test</i>	$Z_{K-S} = 1.61$	–	$p = .011^*$
<i>Schwierigkeit Test</i>	<i>Levene – Test</i>	$F = 1.71$	$df_1 = 3; df_2 = 80$	$p = .172$
	<i>K – S – Test</i>	$Z_{K-S} = 2.47$	–	$p < .001^{**}$
<i>Belastung Lernen</i>	<i>Levene – Test</i>	$F = 0.81$	$df_1 = 3; df_2 = 80$	$p = .494$
	<i>K – S – Test</i>	$Z_{K-S} = 3.10$	–	$p < .001^{**}$
<i>Belastung Test</i>	<i>Levene – Test</i>	$F = 0.38$	$df_1 = 3; df_2 = 80$	$p = .770$
	<i>K – S – Test</i>	$Z_{K-S} = 2.10$	–	$p < .001^{**}$

Anm.: *K-S-Test* = Kolmogorow-Smirnow-Anpassungstest; signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen ($* < .05$) bzw. durch zwei Sternchen ($** < .01$) hervorgehoben

Es ist gut zu erkennen, dass die Angaben der Versuchsteilnehmer zur empfundenen Schwierigkeit und der eigenen Anstrengungen allesamt nicht normalverteilt waren. Gleichzeitig konnte aber von homogenen Fehlervarianzen ausgegangen werden, da der Levene-Test für keine der Skalen signifikant wurde. Vor diesem Hintergrund und bei gleichmäßig besetzten Zellen erschien die Verwendung varianzanalytischer Verfahren zur Prüfung der Forschungshypothesen unbedenklich zu sein. Allerdings stellte sich die Frage, ob die Testung uni- oder multivariat erfolgen sollte, da die Variablen nicht unabhängig waren. So ist anzunehmen, dass die Schwierigkeit einer Aufgabe und die Anstrengung zu ihrer Bewältigung zusammenhängen. Und je stärker jemand sich beim Lernen anstrengt, umso leichter wird ihm die Bearbeitung der Fragen im Wissenstest fallen. Aufgrund dieser Überlegungen wurde entschieden, über alle fünf Messwerten eine 2x2 MANCOVA mit den Kovariaten *Vorwissen*, *Motivation* und *Verarbeitungspräferenz* zu rechnen. Da

6 Studie I: Interferenzen versus Kapazität

der der Box-Test ($M_{Box} = 45.94$, $F = .91$, $df_1 = 45$, $df_2 = 15841.87$, $p = .654$) kein signifikantes Resultat ergab, konnte die Voraussetzungen für eine multivariate Testung als erfüllt angesehen werden.

Die eigentlichen Ergebnisse der Analyse erwiesen sich jedoch weitgehend als *unbedeutsam*. Weder die *Reihenfolge* ($V = .030$, $F_{(5;73)} = 0.445$, $p = .816$, $\bar{\eta}^2 = .030$), noch das *Informationsverhältnis* ($V = .117$, $F_{(5;73)} = 1.928$, $p = .100$, $\bar{\eta}^2 = .117$) oder die Interaktion ($V = .005$, $F_{(5;73)} = 0.074$, $p = .996$, $\bar{\eta}^2 = .005$) der beiden Faktoren deuteten auf einen Zusammenhang mit den abhängigen Variablen hin. Von den Kovariaten wurde lediglich die *Motivation* signifikant ($V = .509$, $F_{(5;73)} = 15.122$, $p < .001$, $\bar{\eta}^2 = .509$), während die Prüfgrößen des *Vorwissens* ($V = .093$, $F_{(5;73)} = 1.505$, $p = .093$, $\bar{\eta}^2 = .093$) und der *Verarbeitungspräferenz* ($V = .125$, $F_{(5;73)} = 2.092$, $p = .076$, $\bar{\eta}^2 = .125$) den kritischen Wert verfehlt hatten.

Aufgrund dieser globalen Ergebnisse erstaunt es nur wenige, dass nachträglich berechnete univariate Analysen lediglich zwei *marginal* signifikante Effekte ergaben. So lag die Irrtumswahrscheinlichkeit des Faktors *Informationsverhältnis* ($F_{(1;77)} = 3.583$, $p = .062$, $\bar{\eta}^2 = .044$) für die mentale Belastung beim Wissenserwerb nur knapp über dem dem Signifikanzniveau von 5%. Gleiches gilt für die Schwierigkeitseinschätzung beim Wissenstest, wo der kritische F-Wert für das *Informationsverhältnis* ($F_{(1;77)} = 3.826$, $p = .054$, $\bar{\eta}^2 = .047$) ebenfalls nur knapp verfehlt wurde. Streng genommen konnte aber damit für keine der abhängigen Variablen ein signifikanter Effekt nachgewiesen werden, was offensichtlich nicht zu den Vorhersagen der *Kapazitäts-* und *Interferenzhypothese* passen wollte.

6.3.4 Forschungsfrage 3: Lernzeiten

Das dritte Hypothesenbündel, welches im Rahmen der ersten Studie geprüft werden sollte, betraf die Lernzeit. Operationalisiert wurde diese abhängige Variable, indem die Verweildauer der Versuchsteilnehmer auf den einzelnen HTML-Seiten gemessen wurde. Damit ergaben sich jeweils vier Bild- und vier Text-Messwert, wobei sich die theoretischen Vorhersagen vor allem auf die Verarbeitungszeit des verbalen Materials bezogen. Wie in den vorangegangenen Abschnitten sollte die Hypothesentestung mit varianzanalytischen Verfahren erfolgen, so dass zuerst die Voraussetzungen geprüft werden mussten. Tabelle 6.8 fasst die Ergebnisse für die Textseiten zusammen.

Wie zu erkennen ist, konnte man die Normalverteilungsannahme akzeptieren und von homogenen Fehlervarianzen ausgehen, da überzufällige Abweichungen nicht zu erkennen waren. Auch ergab der Test von Box (1949), mit dem die Kovarianzmatrizen der abhängigen Variablen unter den verschiedenen Bedingungen auf Gleichheit geprüft werden, kein signifikantes Ergebnis ($M_{Box} = 44.178$, $F = 1.34$, $df_1 = 30$, $df_2 = 17596.18$, $p = .102$). Die Ergebnisse der MANCOVA ließen sich also ohne Bedenken interpretieren. Dabei zeigte sich, dass keine der Kontrollvariablen einen bedeutsamen Zusammenhang mit den Lesezeiten der Texte aufwies. Die Prüfstatistiken der *Motivation* ($V = .049$, $F_{(4;74)} = 0.963$, $p = .433$, $\bar{\eta}^2 = .049$), des *Vorwissens* ($V = .055$, $F_{(4;74)} = 1.079$, $p = .373$, $\bar{\eta}^2 = .055$) und der *Verarbeitungspräferenz* ($V = .017$, $F_{(4;74)} = 0.326$, $p = .860$, $\bar{\eta}^2 = .017$) hatten den kri-

Tabelle 6.6: Kolmogorow-Smirnow- und Levene-Tests für die Lesezeiten der Texte

	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Lesezeit Text I</i>	<i>Levene – Test</i>	$F = 0.52$	$df_1 = 3; df_2 = 80$	$p = .671$
	<i>K – S – Test</i>	$Z_{K-S} = 1.17$	–	$p = .127$
<i>Lesezeit Text II</i>	<i>Levene – Test</i>	$F = 1.52$	$df_1 = 3; df_2 = 80$	$p = .215$
	<i>K – S – Test</i>	$Z_{K-S} = 1.01$	–	$p = .258$
<i>Lesezeit Text III</i>	<i>Levene – Test</i>	$F = 1.32$	$df_1 = 3; df_2 = 80$	$p = .273$
	<i>K – S – Test</i>	$Z_{K-S} = .74$	–	$p = .648$
<i>Lesezeit Text IV</i>	<i>Levene – Test</i>	$F = .99$	$df_1 = 3; df_2 = 80$	$p = .403$
	<i>K – S – Test</i>	$Z_{K-S} = 1.20$	–	$p = .114$

Anm.: *K-S-Test* = Kolmogorow-Smirnow-Anpassungstest; *signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen (* < .05) bzw. durch zwei Sternchen (** < .01) hervorgehoben*

tischen Wert jeweils verfehlt. Anders verhielt es sich mit den Haupteffekten *Reihenfolge* ($V = .121$, $F_{(4;74)} = 2.547$, $p = .046$, $\bar{\eta}^2 = .121$) und dem *Informationsverhältnis* ($V = .396$, $F_{(4;74)} = 12.141$, $p < .001$, $\bar{\eta}^2 = .396$), die beide signifikant wurden. Die Interaktion der Faktoren ($V = .029$, $F_{(4;74)} = 0.545$, $p = .703$, $\bar{\eta}^2 = .029$) erwies sich dagegen als unbedeutsam. Die Richtung der gefundenen Unterschiede lässt sich anhand Tabelle 6.7 ablesen.

Tabelle 6.7: Adjustierte Mittelwerte und deren Standardfehler für die Lesezeiten der Texte

	Versuchsbedingungen			
	Text vor Bild kongruent	Text vor Bild komplementär	Bild vor Text kongruent	Bild vor Text komplementär
<i>Text I</i>	$M = 200.05$ $SE = 12.50$	$M = 179.43$ $SE = 12.56$	$M = 163.73$ $SE = 12.70$	$M = 136.98$ $SE = 12.46$
<i>Text II</i>	$M = 175.03$ $SE = 9.76$	$M = 115.98$ $SE = 9.81$	$M = 141.33$ $SE = 9.92$	$M = 90.73$ $SE = 9.73$
<i>Text III</i>	$M = 144.74$ $SE = 9.39$	$M = 112.77$ $SE = 9.22$	$M = 132.46$ $SE = 8.62$	$M = 103.41$ $SE = 8.67$
<i>Text IV</i>	$M = 145.83$ $SE = 8.62$	$M = 97.91$ $SE = 8.67$	$M = 126.35$ $SE = 8.76$	$M = 94.53$ $SE = 8.60$

Anm.: M = Mittelwert; SE = Standardfehler des Mittelwerts, adjustiert nach den Kovariaten Vorwissen, Motivation und Verarbeitungspräferenz

Es ist gut zu erkennen, dass sämtliche Mittelwerte in die gleiche Richtung deuteten. Ganz offensichtlich dauerte das Lesen der kongruenten Texte länger, was jedoch nicht wirklich überraschte. Sehr viel interessanter war dagegen der Effekt, den die *Reihenfolge* hatte. Hier zeigte sich, dass Versuchspersonen, die den Text zuerst erhielten, mehr Zeit in die Verarbeitung der verbalen Informationen investierten, als die Bild-vor-Text-Gruppen.

6 Studie I: Interferenzen versus Kapazität

Vergleicht man die Ergebnisse in Tabelle 6.7 mit den Vorhersagen der beiden Hypothesen, dann ergibt sich wiederum kein eindeutiges Bild. Zunächst widersprach die kürzere Verarbeitungszeit der Bild-vor-Text-Gruppen eindeutig den Annahmen der Kapazitätshypothese, deren Erklärungsansatz damit wenig plausibel erschien. Aber auch die Überlegungen auf der Basis des Modells von Schnotz (2005) trafen nur für das kongruente Material zu. Bei einem komplementären Informationsverhältnis hätte es keine Unterschiede bei der Verarbeitungsdauer der Textseiten geben dürfen, was jedoch der Fall war.

Um eine bessere Grundlage für die Deutung der Ergebnisse zu erhalten, wurde *post hoc* ein weitere Analyse durchgeführt, bei der auch die Betrachtungszeiten der Bilder berücksichtigt wurden. Dabei interessierte es weniger, welche Text- oder Bildseite die Teilnehmer am längsten betrachtet hatten. Von sehr viel größerer Bedeutung waren mögliche Interaktionen zwischen den formatspezifischen Lernzeiten und den experimentellen Bedingungen. Daher wurden die insgesamt 8 Messwerte zu einer *Gesamtbetrachtungszeit Text* und einer *Gesamtbetrachtungszeit Bild* aggregiert und aus Gründen der Vergleichbarkeit z-standardisiert (Bortz & Döring, 2006). Diese beiden Werte wurden dann wie zwei Messzeitpunkte einer Variablen varianzanalytisch ausgewertet. Zuvor mussten jedoch die Voraussetzungen geprüft werden, was in Tabelle 6.8 dargestellt ist.

Tabelle 6.8: Kolmogorow-Smirnow- und Levene-Tests für die formatspezifischen Lernzeiten

	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Gesamtlernzeit Text</i>	<i>Levene – Test</i>	$F = 0.87$	$df_1 = 3; df_2 = 80$	$p = .462$
	<i>K – S – Test</i>	$Z_{K-S} = .94$	–	$p = .335$
<i>Gesamtlernzeit Bild</i>	<i>Levene – Test</i>	$F = 1.23$	$df_1 = 3; df_2 = 80$	$p = .303$
	<i>K – S – Test</i>	$Z_{K-S} = .86$	–	$p = .447$
<i>Gesamtlernzeit (Text + Bild)</i>	<i>Levene – Test</i>	$F = 1.47$	$df_1 = 3; df_2 = 80$	$p = .230$
	<i>K – S – Test</i>	$Z_{K-S} = 1.27$	–	$p = .081$

Anm.: *K-S-Test = Kolmogorow-Smirnow-Anpassungstest; signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen (* < .05) bzw. durch zwei Sternchen (** < .01) hervorgehoben*

Die Ergebnisse zeigen, dass die Normalverteilungsannahme gegeben und die Fehlervarianzen homogen waren, da keiner der entsprechenden Tests signifikant wurde. Außerdem entfiel bei nur einem Messwertpaar die Prüfung der Sphärizität, so dass sich die resultierenden F-Werte ohne Einschränkungen interpretiert ließen. Im Grunde handelte es bei dem verwendeten Verfahren um eine dreifaktorielle Varianzanalyse mit zwei *normalen* und einem Messwiederholungsfaktor. Demnach konnte man zwischen drei Haupteffekten, drei Interaktionseffekten erster Ordnung und einem Interaktionseffekt zweiter Ordnung unterscheiden. Hinzu kamen die Kovariaten, die auch bei dieser Analyse berücksichtigt wurden. Dabei zeigte sich, dass die *Motivation* ($F_{(1;77)} = .262, p = .610, \bar{\eta}^2 = .003$) offensichtlich nicht dazu beigetragen hatte, die Fehlervarianz zu reduzieren. Anders verhielt es sich mit den Kontrollvariablen *Vorwissen* ($F_{(1;77)} = 7.136, p = .009, \bar{\eta}^2 = .085$) und der

Verarbeitungspräferenz ($F_{(1;77)} = 3.991$, $p = .049$, $\bar{\eta}^2 = .049$), deren Prüfgröße jeweils über dem kritischen Wert lagen.

Betrachtete man den Messwiederholungsfaktor ($F_{(1;77)} = 3.607$, $p = .061$, $\bar{\eta}^2 = .045$), dann fiel auf, dass es keinen bedeutsamen Unterschied zwischen der standardisierten Betrachtungszeiten der Texte und Bilder gab. Der Faktor *Reihenfolge* ($F_{(1;77)} = 1.669$, $p = .200$, $\bar{\eta}^2 = .021$) wurde ebenfalls nicht signifikant, während für das *Informationsverhältnis* ($F_{(1;77)} = 25.431$, $p < .001$, $\bar{\eta}^2 = .248$) ein starker Effekt nachgewiesen werden konnte. Von den drei möglichen Interaktionen erster Ordnung erwies sich allein die Wechselwirkung der Faktoren *Zeit* und *Reihenfolge* ($F_{(1;77)} = 58.206$, $p < .001$, $\bar{\eta}^2 = .430$) als bedeutsam. Dementsprechend blieben die Prüfgrößen der Interaktionen zwischen den Faktoren *Zeit* und dem *Informationsverhältnis* ($F_{(1;77)} = .336$, $p = .564$, $\bar{\eta}^2 = .004$) sowie zwischen der *Reihenfolge* und dem *Informationsverhältnis* ($F_{(1;77)} = .695$, $p = .407$, $\bar{\eta}^2 = .009$) jeweils unter dem kritischen Wert. Besonders interessant ist vor allem der Befund, dass die Interaktion zweiter Ordnung ($F_{(1;77)} = 4.879$, $p = .030$, $\bar{\eta}^2 = .060$) signifikant wurde. Um die einzelnen Zusammenhänge besser deuten zu können, erschien es hilfreich, die einzelnen Effekte grafisch zu veranschaulichen. So zeigt Abbildung 6.3 zunächst den signifikanten Zusammenhang zwischen dem Informationsverhältnis und der durchschnittlichen Gesamtlesezeit.

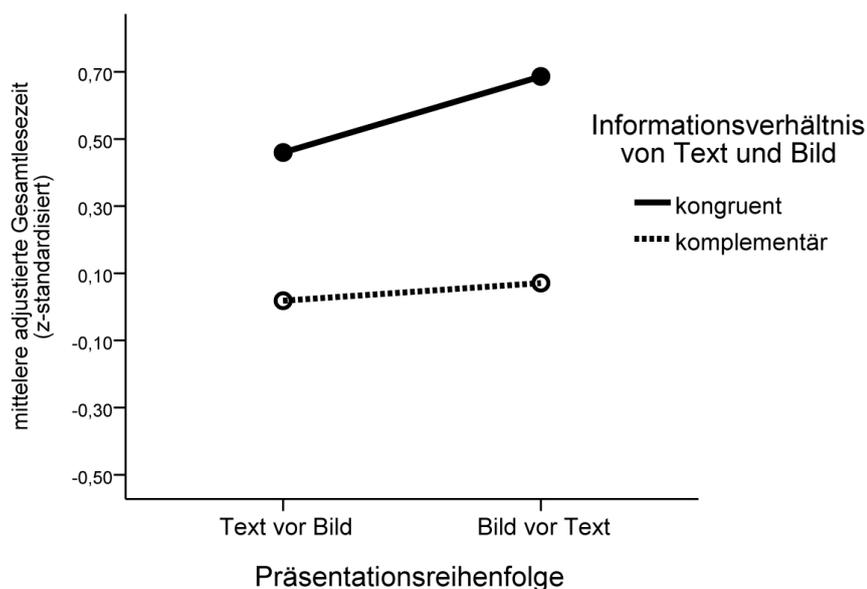


Abbildung 6.3: Mittlere adjustierte Gesamtlesezeit

Es ist sehr deutlich zu erkennen, dass die kongruenten Formate sehr viel länger betrachtet bzw. gelesen wurden, als die komplementären Text oder Bilder. Allerdings ist dieser Zusammenhang äußerst trivial, da die Versuchsteilnehmer mit den kongruenten Formaten deutlich mehr Worte und Bildelemente zu verarbeiten hatten.

Etwas weniger vorhersehbar war dagegen der Interaktionseffekt zwischen der Reihenfolge und dem Messwiederholungsfaktor, wie Abbildung 6.4 für die kongruenten Bedin-

6 Studie I: Interferenzen versus Kapazität

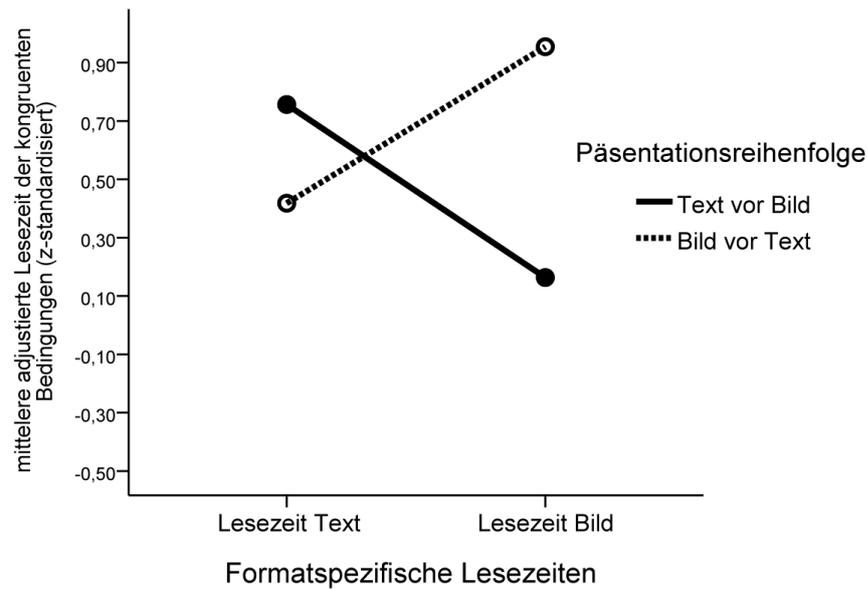


Abbildung 6.4: Darstellung der Interaktion Zeit*Reihenfolge für die kongruente Bedingung

ungen veranschaulicht. Offensichtlich haben sich die Studierenden mit dem Format, das Ihnen als erstes präsentiert wurde, am intensivsten auseinandergesetzt, was sich in den Betrachtungszeiten und der dargestellten Wechselwirkung widerspiegelt. So haben Versuchsteilnehmer, die zuerst den Text lasen, vergleichsweise wenig Zeit darauf verwendet das nachfolgende Bild zu verarbeiten. Studierende, die sich zuerst mit dem Bild auseinandersetzen mussten, benötigten zwar auch weniger Zeit, um den entsprechenden Text zu lesen, doch ist der Unterschied bei weitem nicht so deutlich. Stellt man den gleichen Zusammenhang für die komplementären Bedingungen dar, ergibt sich auf den ersten Blick ein ähnliches Bild, wie Abbildung 6.5 zeigt.

Wenn man diese Darstellung mit Abbildung 6.4 vergleicht, fällt jedoch auf, dass beide Linien etwas flacher ausfallen. Der Interaktionseffekt variiert demnach mit dem Informationsverhältnis bzw. schwächt sich unter der komplementären Bedingung ab. Aus diesem Grund liegt es nahe, dass die Wechselwirkung zwischen dem Messwiederholungsfaktor, der Reihenfolge und dem Informationsverhältnis signifikant wurde. Hätte man diese Interaktion für die kongruente ($F_{(1;37)} = 43.595$, $p < .001$, $\bar{\eta}^2 = .541$) und komplementäre Bedingungen ($F_{(1;37)} = 17.320$, $p < .001$, $\bar{\eta}^2 = .319$) getrennt berechnet, so wäre der Effekt zwar deutlich geringer ausgefallen, aber immer noch sehr stark gewesen. Zusammenfassend ließ sich festhalten, dass immer das Format, dessen zuerst erfolgt war, zumindest zeitlich am intensivsten verarbeitet wurde. Dieser Effekt konnte selbst dann beobachtet werden, wenn Text und Bild in einem komplementären Informationsverhältnis zueinander standen.

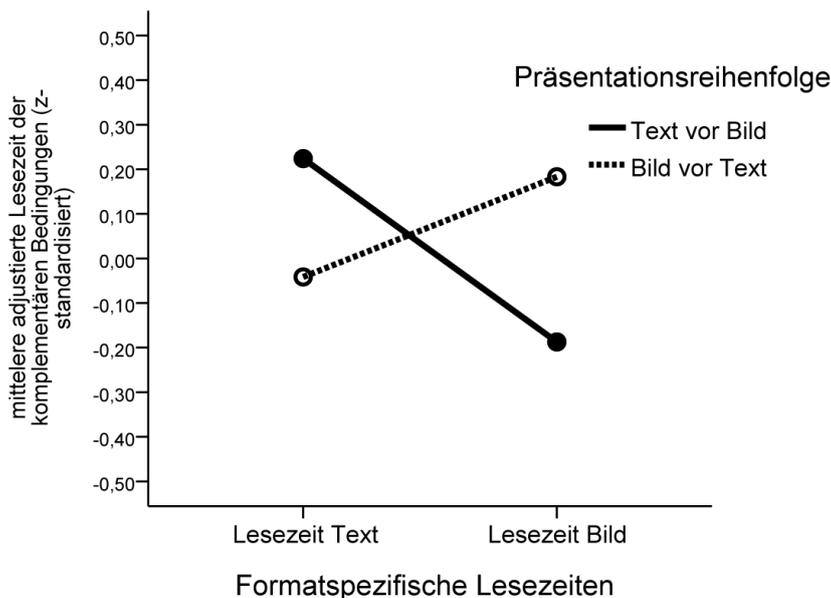


Abbildung 6.5: Darstellung der Interaktion Zeit*Reihenfolge für die komplementäre Bedingung

6.3.5 Weitere Analysen

Zusätzlich zu den bisher behandelten Personenvariablen waren direkt nach der Präsentationsphase vier Fragen zum Lernen erhoben worden. Dadurch sollte untersucht werden, wie die *Reihenfolge* und das *Informationsverhältnis* zwischen Text und Bild das Vorgehen der Teilnehmer beim Wissenswerb möglichst beeinflussen. Jedes dieser Items umfasste eine sechsstufige Ratingskala, auf der die Studierenden ihre Einschätzung abgeben konnten. Die Auswertung der Antworten erfolgte anhand von univariaten Kovarianzanalysen, in deren Rahmen erneut das *Vorwissen*, die *Motivation* und die *Verarbeitungspräferenz* statistisch kontrolliert wurden. Die folgenden vier Abschnitte fassen die Ergebnisse zusammen.

Details oder Zusammenhänge?

Bei der ersten Frage sollten die Teilnehmer einschätzen, ob sie sich beim Lernen eher auf Details oder auf übergeordnete Zusammenhänge konzentriert haben. Die Antwort konnten einen Wert zwischen 0 (nur Details) und 5 (nur Zusammenhänge) annehmen, wobei der Gesamtmittelwert von $\bar{X} = 3.2$ dafür sprach, dass die meisten Versuchspersonen dazu tendiert hatten, eher die Zusammenhänge verstehen zu wollen. Vor der Durchführung der ANCOVA wurde zunächst geprüft, ob die Fehlervarianz homogen und die abhängige Variable normalverteilt waren. Während der Levene-Test nicht signifikant wurde ($F = 0.84$, $df_1 = 3$; $df_2 = 80$, $p = .475$), ergab der Kolmogorow-Smirnow-Test einen Wert, bei dem die Normalverteilungsannahme als verletzt gelten musste ($Z_{K-S} = 2.07$, $p < .001$). Doch bei gleichmäßig besetzten Zellen, war eine Verletzung dieser Annahme zu vernachlässigen, weshalb die Ergebnisse einer Kovarianzanalyse ohne Einschränkungen interpretiert

6 Studie I: Interferenzen versus Kapazität

werden konnte.

Es zeigte sich, dass weder das *Vorwissen* ($F_{(1;77)} = 0.070$, $p = .792$, $\bar{\eta}^2 = .001$), noch die *Motivation* ($F_{(1;77)} = 0.025$, $p = .875$, $\bar{\eta}^2 = .000$) oder die *Verarbeitungspräferenz* ($F_{(1;77)} = 0.416$, $p = .521$, $\bar{\eta}^2 = .005$) einen bedeutsamen Zusammenhang mit den Einschätzungen der Teilnehmer aufwiesen. Interessant war jedoch, dass der Faktor *Reihenfolge* ($F_{(1;77)} = 4.084$, $p < .047$, $\bar{\eta}^2 = .050$) signifikant wurde, während das *Informationsverhältnis* ($F_{(1;77)} = 0.161$, $p = .689$, $\bar{\eta}^2 = .002$) und die Interaktion ($F_{(1;77)} = 0.000$, $p = .997$, $\bar{\eta}^2 = .000$) jeweils deutlich über dem α -Fehlerniveau von 5% lagen. Abbildung 6.6 veranschaulicht, dass Teilnehmer, die zuerst das Bild gesehen hatten, es offenbar vorzogen, sich mehr auf die Zusammenhänge zu konzentrieren, als die Versuchspersonen in den Text-vor-Bild-Gruppen.

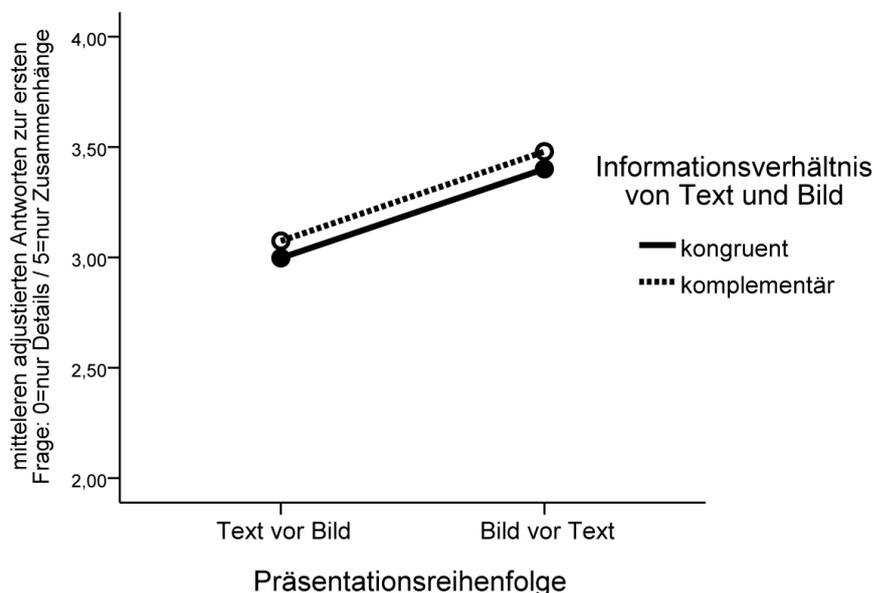


Abbildung 6.6: Fragen zum Lernen 1: Details oder Zusammenhänge?

Symbolisches oder anschauliches Memorieren

Die zweite Frage bezog sich darauf, ob die Studierenden versucht haben, sich die Inhalte eher bildlich oder eher sprachlich zu merken. Wiederum konnten die Einschätzungen der Teilnehmer Werte zwischen 0 (nur sprachlich) und 5 (nur bildlich) annehmen. Das arithmetische Mittel über alle Gruppen lag bei $\bar{X} = 2.9$, was einem tendenziell ausgewogenen Vorgehen entsprach. Die Prüfung der Voraussetzungen für die ANCOVA ergab ein ähnliches Bild, wie bei der ersten Frage. Während die Varianzhomogenität als gegeben betrachtet werden konnte ($F = 1.83$, $df_1 = 3$; $df_2 = 80$, $p = .149$), musste die Normalverteilungsannahme ($Z_{K-S} = 2.35$, $p < .001$) als verletzt gelten.

Die Ergebnisse ließen sich jedoch trotzdem interpretieren, auch wenn erneut nur ein signifikanter Effekt berichtet werden konnte. So gab es offensichtlich keine bedeutsamen Unterschiede zwischen den vier Gruppen, da die F-Werte der Faktoren *Reihenfolge*

ge($F_{(1;77)} = 0.254$, $p = .616$, $\bar{\eta}^2 = .003$), *Informationsverhältnis* ($F_{(1;77)} = 1.629$, $p = .206$, $\bar{\eta}^2 = .021$) und deren Interaktion ($F_{(1;77)} = 3.505$, $p = .065$, $\bar{\eta}^2 = .044$) das kritische Niveau jeweils verfehlt hatten. Das gleich traf auf die Kovariaten *Vorwissen* ($F_{(1;77)} = 0.054$, $p = .816$, $\bar{\eta}^2 = .001$) und *Motivation* ($F_{(1;77)} = 1.546$, $p = .218$, $\bar{\eta}^2 = .020$) zu. Allein die *Verarbeitungspräferenz* ($F_{(1;77)} = 6.597$, $p = .012$, $\bar{\eta}^2 = .079$) korrelierte signifikant mit den Einschätzungen der Studierenden.

Text oder Bild?

Mit dem dritten Item sollte erfasst werden, ob sich die Teilnehmer in der Lernphase mehr auf den Text oder mehr auf das Bild konzentriert haben. Die Skala war dabei so gepolt, dass ein Wert von 0 anzeigte, dass der Lernende ausschließlich die Texte zum Wissenserwerb genutzt hatte. Dagegen entsprach ein Wert 5 einer alleinigen Verwendung des Bildes. Das arithmetische Mittel über alle Gruppen lag bei $\bar{X} = 2.3$, was dafür sprach, dass die Teilnehmer in der Regel beide Formate ausgewogen nutzten. Die Prüfung der Voraussetzungen ergab erneut ein einen nicht signifikanten Levene-Test ($F = 1.77$, $df_1 = 3$; $df_2 = 80$, $p = .159$), während die Annahme einer normalverteilten abhängigen Variablen verworfen werden musste ($Z_{K-S} = 1.87$, $p = .002$).

Dies stellte jedoch kein allzu großes Problem dar, so dass die ANCOVA ganz normal berechnet und interpretiert wurde. Bei den Ergebnissen fiel auf, dass von den Kovariaten weder das *Vorwissen* ($F_{(1;77)} = 1.484$, $p = .227$, $\bar{\eta}^2 = .019$) noch die *Motivation* ($F_{(1;77)} = 1.633$, $p = .205$, $\bar{\eta}^2 = .021$) oder die *Verarbeitungspräferenz* ($F_{(1;77)} = 2.419$, $p = .124$, $\bar{\eta}^2 = .030$) signifikant wurde. Allerdings ergab die Analyse für die *Reihenfolge* ($F_{(1;77)} = 7.397$, $p = .008$, $\bar{\eta}^2 = .088$) und das *Informationsverhältnis* ($F_{(1;77)} = 4.429$, $p = .039$, $\bar{\eta}^2 = .054$) bedeutsame Unterschiede. Die Interaktion wurde dagegen nicht signifikant ($F_{(1;77)} = 0.173$, $p = .678$, $\bar{\eta}^2 = .002$). Abbildung 6.7 zeigt, dass sich die Teilnehmer immer auf das Format stärker konzentriert haben, das ihnen zuerst präsentiert wurde. Besonders interessant ist dabei, dass sich die komplementären Gruppen besonders stark auf den Text konzentriert haben.

6 Studie I: Interferenzen versus Kapazität

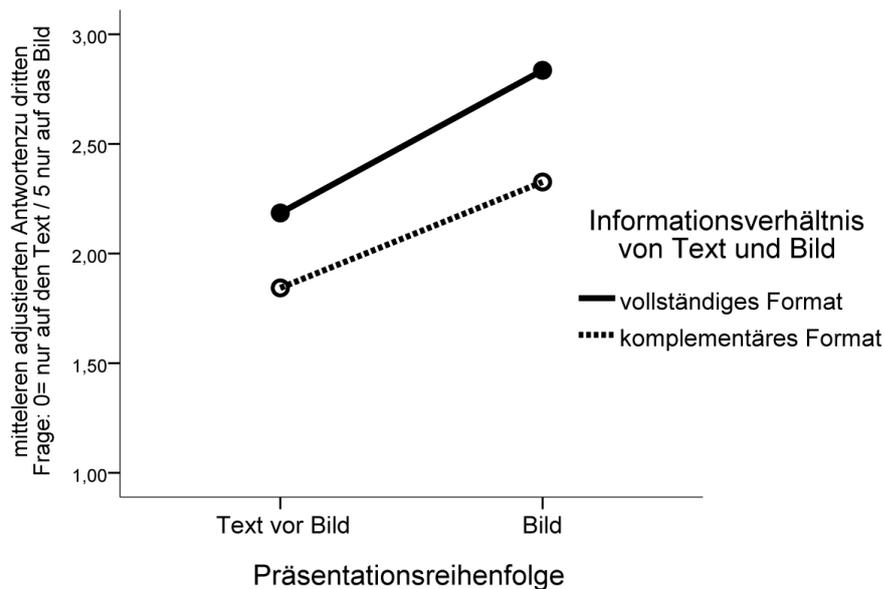


Abbildung 6.7: Fragen zum Lernen 3: Texte oder Bilder?

Beurteilung der Präsentation am Computer

Bei der letzten Frage sollten die Teilnehmer angeben, ob es eher positiv oder eher negativ war, dass die Texte und Bilder an einem Computerbildschirm präsentiert wurden. Ein niedriger Wert entsprach einer positiven Einschätzung, während ein hoher Wert eine negative Haltung anzeigte. Das arithmetische Mittel über alle Gruppen lag bei $\bar{X} = 2.7$, was zum Ausdruck brachte, dass die computergestützte Präsentation nicht nur auf Zustimmung stieß. Bei der Prüfung der Voraussetzungen zeigt sich das gleiche Bild, wie bei den drei Items zuvor. Die Varianzhomogenität konnte als gegeben gelten ($F = 1.40$, $df_1 = 3$; $df_2 = 80$, $p = .250$), während die Normalverteilungsannahme verletzt war ($Z_{K-S} = 1.52$, $p = .020$).

Für die Durchführung der ANCOVA hatte dieses Ergebnis erneut keine Konsequenzen, auch wenn die Faktoren *Reihenfolge* ($F_{(1;77)} = 2.187$, $p = .143$, $\bar{\eta}^2 = .028$), *Informationsverhältnis* ($F_{(1;77)} = 1.051$, $p = .309$, $\bar{\eta}^2 = .013$) und deren Interaktion ($F_{(1;77)} = 0.035$, $p = .852$, $\bar{\eta}^2 = .000$) nicht signifikant wurden. Auch die Kovariaten *Vorwissen* ($F_{(1;77)} = 0.119$, $p = .732$, $\bar{\eta}^2 = .002$) und *Verarbeitungspräferenz* ($F_{(1;77)} = 0.982$, $p = .325$, $\bar{\eta}^2 = .013$) erwiesen sich als unbedeutsam. Offenbar korrelierte allein die *Motivation* ($F_{(1;77)} = 16.446$, $p = .001$, $\bar{\eta}^2 = .127$) der Teilnehmer mit ihrer Einstellung gegenüber dem Lernen am Computer.

6.4 Diskussion

Bisher wurde der PTS-Effekt vor allem im Zusammenhang mit Leistungsmaßen beobachtet, die sich in erster Linie auf die freie Wiedergabe von zuvor gelernten Informationen konzentrierten (Kulhavy et al., 1994; Verdi et al., 1997). Bedeutungsvolles Lernen, wie es beispielsweise (Mayer, 2005) definiert, ist jedoch mehr als das Aufnehmen und Spei-

chern bestimmter Inhalte, weshalb neben der Behaltens- auch die Verstehensleistung (siehe Kapitel 5.2.1) im Rahmen der ersten Studie erfasst wurde. Die Ergebnisse der statistischen Analysen zeigten recht eindeutig, dass Teilnehmer, die das Bilder vor dem Text lernten, sowohl mehr Informationen wiedergeben konnten, als auch in der Lage waren, mehr Fragen zu beantworten, die schlussfolgerndes Denken erforderten. Durch diesen Nachweis des PTS-Effekts wurde die *Parallelitätshypothese* recht eindeutig falsifiziert, da die Reihenfolge in der Texte und Bilder zum Wissenserwerb genutzt werden, den Prozess der Informationsverarbeitung offensichtlich beeinflusst.

Der Umstand, dass dieser Effekt selbst dann zu beobachten war, wenn die Bilder und Texte ein komplementäres Informationsverhältnis aufwiesen, sprach zunächst eindeutig für die *Kapazitätshypothese* und den Erklärungsansatz von Kulhavy et al. (1994). Demnach konnten die Lernenden zwar erheblich weniger referentielle Verknüpfungen etablieren, wenn sich die visuell und verbal präsentierten Informationen inhaltlich wenig überlappen. Der Verarbeitungsvorteil, den der Aufbau eines intakten *images* mit sich brachte, blieb jedoch bestehen, was sich an den beiden signifikanten Haupteffekten und dem parallelen Verlauf der beiden Linien in Abbildung 6.2 ablesen lässt. Vor diesem Hintergrund erscheint die *Interferenzhypothese*, die auf einem Erklärungsansatz von (Schnotz, 2002, 2005) beruht, sehr viel weniger plausibel zu sein. Da unter der komplementären Bedingung letztlich kaum gemeinsame Inhalte zwischen Bild und Text vorhanden waren, die zu Interferenzen zwischen der mentalen Repräsentation und dem nachfolgenden Bild hätte führen können, ist das Ergebnis der ersten Analyse kaum zu erklären. Daran ändern auch die etwas uneinheitlichen Resultate der einzelnen Leistungsskalen nichts.

Die Ergebnisse zur mentalen Belastung bzw. dem *cognitive load* sind dagegen nicht so einfach zu interpretieren, wie die Befunde zur Lernleistung. Letztlich musste festgestellt werden, dass selbst bei univariaten Analysen kein statistisch bedeutsamer Haupt- oder Interaktionseffekt nachgewiesen werden konnte. Zwar verfehlten die entsprechenden Prüfstatistiken bei den Skalen *Schwierigkeit des Wissenstests* und *mentale Belastung beim Wissenserwerb* nur jeweils knapp den kritischen Wert für einen α -Fehler von 5%, doch sprachen inferenzstatistische Prinzipien gegen eine Deutung dieser *marginal* signifikanten Mittelwertsunterschiede. Natürlich kann man darüber spekulieren, weshalb für die empfundene Schwierigkeit und die mentale Belastung kein PTS-Effekt nachgewiesen werden konnte. Die einfachste Antwort auf diese Frage besteht darin, dass ein solcher Zusammenhang gar nicht existiert. Geht man jedoch davon aus, dass zumindest einzelne Annahmen der *cognitive load theory* von Chandler und Sweller (1991) zutreffen und die Lernleistung zu einem gewissen Teil von der mentalen Belastung abhängt, dann lassen sich auch alternative Erklärungen für die nicht signifikanten Ergebnisse finden. So ist es durchaus möglich, dass die Teststärke nicht ausreichend war, da die Planung des Stichprobenumfangs ausschließlich auf Angaben zur Leistungsmessung beruhte. Bei einer Gesamtteilnehmerzahl von 84 Personen und einem globalen α -Fehler von 5% betrug die Wahrscheinlichkeit für einen β -Fehler bei einem vorhandenen mittleren Effekt ($\bar{\eta}^2 = .10$) 38% und bei einem vorhandenen kleinen Effekt ($\bar{\eta}^2 = .01$) sogar 85%. Demnach hätte grundsätzlich Glück dazugehört, tatsächliche bestehende kleine und mittlere

Zusammenhänge nachzuweisen.

Ein weiterer Punkt betrifft die Operationalisierung und Erfassung der mentalen Belastung und der empfundenen Schwierigkeit. Obwohl ein bereits bewährtes Verfahren verwendet wurde (Paas et al., 1994), gibt es Autoren, die eine Messung des *cognitive loads* über Skalen zur Selbsteinschätzung kritisch betrachten (Schnotz & Kürschner, 2007). Außerdem erfolgte die Erfassung der mentalen Belastung retrospektiv und aggregiert, so dass die Teilnehmer hinsichtlich ihrer Einschätzung nicht zwischen einzelnen Texten, Bildern oder bestimmten Aufgaben differenzieren konnten. Unter Umständen hätte eine Befragung nach jeder Bild- und Textseite zu einer besseren Messung geführt. Hinzu kommt, dass in der Formulierung der Belastungs- bzw. Anstrengungsitems (siehe Abschnitt 5.2.2) eine starke motivationale Komponente enthalten war, was sich nicht zuletzt darin äußerte, dass der FAM-Score als einzige Kovariate signifikant wurde. Möglicherweise wurde gar nicht die Anstrengung im Sinne der *cognitive load theory* von Chandler und Sweller (1991) gemessen, sondern die Motivation der Teilnehmer, sich anzustrengen (Schnotz, Fries & Horz, 2009). Insgesamt bleibt für Untersuchung der empfundenen Schwierigkeit und mentalen Belastung jedoch festzuhalten, dass keine der Hypothesen imstande war, die empirischen Befunde wirklich zu erklären.

Ein ähnlich Resultat ergaben die Analysen zu den Lesezeiten der Texte, die zwar statistisch bedeutsame Ergebnisse lieferten, aber den theoretischen Vorhersagen nur teilweise zugeordnet werden konnten. Der signifikante Haupteffekt für den Faktor *Informationsverhältnis* stellte zunächst keine große Überraschung dar. Die kongruenten Versionen umfassten einfach mehr Worte, so dass die Teilnehmer mehr Zeit benötigten, um die Texte ganz zu lesen. Sehr viel interessanter war dagegen der zweite Faktor, der einen bedeutsamen Einfluss der Präsentationsreihenfolge auf die Lesezeit der Texte anzeigte. Die Mittelwerte in Tabelle 6.7 veranschaulichten, dass Studierende, die zuerst den Text erhielten, diesen durchschnittlich länger verarbeiteten, als Teilnehmer, die zuvor das Bild betrachtet hatten. Dieser Befund verhielt sich entgegengesetzt zu den Vorhersagen der *Kapazitätshypothese*. Auf der Grundlage des Modells von Kulhavy et al. (1994) war erwartet worden, dass die Bild-zuerst-Gruppen für die Verarbeitung nachfolgender Texte länger brauchen würden. Hinter dieser Einschätzung stand die Vermutung, dass die Etablierung referentieller Verknüpfungen zwischen verbalen und piktorialen Informationen einen weiteren kognitiven Prozess darstellt, der zur *normalen* Verarbeitungszeit der Texte hinzu gerechnet werden muss. Dass zusätzliche kognitive Operationen zu einer geringeren Lernzeit führen, war aber sehr unwahrscheinlich. Damit sprachen die Befunde zur Lernzeit gegen das Knüpfen referentieller Verbindungen als Erklärung für den PTS-Effekt. Aber auch die Vorhersagen der *Interferenzhypothese* stimmten nur teilweise mit den empirischen Ergebnissen überein. So war für die kongruenten Bedingungen die Einschätzung richtig, dass Studierende, die das Bild zuerst betrachtet hatten, weniger Zeit in die Verarbeitung des nachfolgenden Textes investierten, als die Gruppe, die den Text vor dem Bild gelesen hatte. Dieser Zusammenhang blieb jedoch unter einem komplementären Informationsverhältnis zwischen Bild und Text erhalten, was die Gültigkeit des Erklärungsansatzes von (Schnotz, 2005) wiederum einschränkte.

Versucht man auf eine abschließende Antwort auf die Frage zu finden, welche der beiden Forschungshypothesen nun am ehesten auf die Resultate passte, dann fällt die Antwort weniger eindeutig aus, als im Vorfeld der ersten Studie vermutet worden war. Immerhin bleibt festzuhalten, dass die Ergebnisse für die Lernleistung und -zeit für einen *picture-text-sequencing*-Effekt sprechen. Daran ändert auch die Tatsache nichts, dass die Untersuchung zur empfundenen Schwierigkeit und der mentale Belastung keine bedeutsame Zusammenhänge erbrachte. Der Vergleich zwischen der *Kapazitäts-* und der *Interferenzhypothese* ist dagegen sehr viel schwieriger, da die Leistungsvergleiche beim Wissenstest zunächst den Erklärungsansatz von (Kulhavy et al., 1994) unterstützten. Einer solcher Deutung widersprachen jedoch die Ergebnisse zur Lernzeit, was die Interpretation der Befunde nicht einfacher machte. Allerdings ist die Annahme von referentiellen Verknüpfungen, die zur *normalen* Textverarbeitung hinzu kommen, mit einer kürzeren Leszeit nach der Präsentation der Bilder gänzlich unvereinbar. Das gilt umso mehr, da nach dem Modell von (Kulhavy et al., 1994) während der Rezeption der verbalen Informationen nicht der vollständige Arbeitsspeicher zur Verfügung steht. Ein Teil der begrenzten Kapazitäten sollte eigentlich durch die Aufrechterhaltung des *image* in Anspruch genommen, mit dessen Elementen die sequentiell eingehenden verbalen Informationen verknüpft werden, was zusätzlich Zeit kosten müsste. Trotz der korrekten Vorhersagen zur Lernleistung ist die *Kapazitätshypothese* als Erklärungsmodell für den PTS-Effekt damit letztlich zu verwerfen.

Ähnliches gilt für die *Interferenzhypothese*, deren Vorhersagen bei keiner der drei Fragestellungen vollständig zutrafen. Sieht man von den Ergebnissen zur mentalen Belastung ab, dann waren vor allem die Einschätzungen zu den komplementären Gruppen falsch. Um dennoch zu einer einheitlichen Deutung der Resultate zu gelangen, schien es hilfreich zu sein, die *post hoc* Analyse zu berücksichtigen, bei der die Lernzeit der Bilder und Texte in einem Messwiederholungsdesign gegenübergestellt wurden. Hier zeigte sich, dass die Teilnehmer immer dann länger auf den Text- bzw. Bildseiten verweilten, wenn diese zuerst präsentiert wurden. Dieses Ergebnis legte die Vermutung nahe, dass sich die beiden Formate in Abhängigkeit der Reihenfolge gegenseitig kompensierten bzw. ersetzen. Ein solcher Zusammenhang ist im Fall der kongruenten Bedingung durchaus nachvollziehbar, da einige Informationen beiden Formaten entnommen werden konnten. Hatte der Leser eines Textes den Eindruck gewonnen, alle wichtigen Inhalte erfasst und verstanden zu haben, hielt er es unter Umständen nicht mehr für notwendig, das nachfolgende Bild genauso intensiv zu verarbeiten.

Erstaunlich ist jedoch, dass sich die gleich Interaktion zwischen der Reihenfolge und den Text- bzw. Bildlesezeiten auch unter der komplementären Bedingung nachweisen ließ. Zwar war der Effekt deutlich schwächer, doch wurde das erste Format auch dann intensiver zum Wissenserwerb genutzt, wenn Bild und Text keine oder nur wenige gemeinsame Informationen enthielten. Dieses schwer *nachvollziehbare* Verhalten spiegelt sich auch in Abbildung 6.3.5 wieder, wonach sich die komplementäre Gruppe, die zuerst die verbalen Informationen erhalten hatte, sich ganz besonders auf den Text konzentrierte. Die Teilnehmer waren aber auf beide Formate angewiesen, weshalb die Kompensation

oder das Ersetzen von Inhalten hier keinen Sinn machte. Vielleicht hatten die Lernenden den Informationsgehalt und die Bedeutung der Illustrationen einfach unterschätzt, so dass sie weniger Zeit investierten (Weidenmann, 2006). Ähnlich wie bei Salomon (1984) wurden die Bilder möglicherweise als ein leicht verständliches Medium angesehen, was in der Konsequenz zu einer geringeren Verarbeitungstiefe und -dauer geführt hat. Allerdings hätte dieser Effekt auch bei den Teilnehmern auftreten müssen, denen die piktorialen Informationen zuerst präsentiert wurden.

Eine schlüssigere Erklärung lässt sich anhand des theoretischen Modells von Schnotz und Bannert (1999) herleiten, das von formatspezifischen Pfaden beim Aufbau mentaler Repräsentationen ausgeht. Dabei spielt es zunächst keine Rolle, ob ein Bild oder ein Text verarbeitet wird, da in beiden Fällen das Ziel die Konstruktion eines mentalen Modells ist. Allerdings unterscheiden sich je nach Format die Prozesse, die zum Aufbau der Repräsentation erforderlich sind. Dienen piktoriale Informationen als Ausgangspunkt für den Konstruktionsprozess, dann fällt dieser vergleichsweise leicht, da Bild und mentales Modell auf analogen Repräsentationsprinzipien beruhen. Nachträglich aufgenommene verbale Informationen können vor allem zum Abgleich, Ausbau oder zur Verbesserung des ersten *Modellentwurfes* verwendet werden. Falls die Inhalte des Textes der piktorialen Darstellung nicht widersprechen, ist daher mit einer zügigen Verarbeitung zu rechnen. Erfolgt der Aufbau des mentalen Modells jedoch anhand verbaler Informationen, wird der Aufwand größer sein. Das Verstehen von Texten beinhaltet zunächst die Repräsentation der Textoberflächenstruktur, auf deren Grundlage Propositionen extrahiert werden. Die propositionale Repräsentation ist wiederum die Ausgangsbasis für die Konstruktion des mentalen Modells, das Charakteristika oder Funktionen von Sachverhalten anschaulich abbildet. Verglichen mit der Bildverarbeitung ist das nicht nur ein zusätzlicher Prozess, vielmehr muss ein symbolisches in ein analoges Repräsentationsformat konvertiert werden, was unweigerlich zu einer längeren Verarbeitungsdauer führt. Zu diesen Überlegungen passt auch, dass die Bild-zuerst-Gruppenangaben, sich während der Lernphase mehr auf Zusammenhänge und weniger auf einzelne Details konzentriert zu haben (siehe Abschnitt 6.3.5). Man könnte diesen Befund auch dahin gehend deuten, dass es diesen Teilnehmern leichter fiel, ein mentales Modell zu konstruieren.

Dieser Erklärungsansatz ist dabei weitgehend unabhängig vom Informationsverhältnis zwischen den verbalen und piktorialen Informationen. Betrachtet beispielsweise ein Studierender ein Bild, das die Eigenschaften konservativen Plattengrenzen veranschaulicht, wird er ein mentales Modell der Darstellung konstruieren. Enthält der nachfolgende Text dann die verbale Beschreibung einer anderen Plattengrenze, kann die zuvor erzeugte Repräsentation die Verarbeitung durchaus erleichtern. Das mentale Modell lässt sich auch als bei einer komplementären Relation zwischen Text und Bild als Entwurf verwenden, der solange interpoliert wird, bis er auf die Textinformationen passt. Invariante strukturelle Gemeinsamkeiten können übernommen, abweichende Merkmale müssen geändert werden, damit letztlich eine konsistente Repräsentation der im Text beschriebenen Sachverhalte entsteht. Dieser Vorteil entfällt natürlich, wenn die Konstruktion des mentalen Modells zuerst anhand verbaler Informationen erfolgen muss. Dem Lernenden steht dann keine Vor-

lage zur Verfügung, so dass gezwungen ist, auf sein Vorwissen zurückgreifen, um sich anschauliche Inhalte vorstellen bzw. verstehen zu können. Ohne *piktoriale Anleitung* durchläuft der Konstruktionsprozess sämtliche Stufen der kognitiven Textverarbeitung, was die Umwandlung *symbolischer* Sprachinformationen in ein analoges Repräsentationsformat mit einschließt und zu längeren Verarbeitungszeiten führt.

Das integrierte Modell des Text- und Bildverstehens bietet aber noch eine weitere Erklärungsmöglichkeit, die an die ursprüngliche Hypothese anknüpft. Demnach könnte man den Einfluss der Reihenfolge unter den komplementären Bedingungen auch als Hinweis für das Auftreten von Interferenzen deuten. Das setzt jedoch die Annahme voraus, dass die Konfrontation eines mentalen Modells mit einem nachfolgenden Bild selbst dann zu einem kognitiven Konflikt führt, wenn sich deren Informationen inhaltlich kaum überlappen. Der resultierenden Interferenzen würden dabei weniger auf die abweichende Repräsentation des selben Sachverhaltes, sondern auf strukturelle oder funktionale Differenzen zwischen dem mentalen Modell und dem externen Bild zurückzuführen sein. Stellt sich beispielsweise ein Lernender auf der Grundlage des Textes zunächst eine *konser-vative* Plattengrenze vor und betrachtet im Anschluss die Abbildung einer *destruktiven* Plattengrenze, handelt es sich rein inhaltlich, um zwei verschiedene Dinge. Die mentale Modellkonstruktion anhand des Textes erfordert jedoch die Repräsentation räumlicher Strukturen, die sich mehr oder weniger auf jeden Grenztypus anwenden lassen. Es kann daher nicht ausgeschlossen werden, dass der Abgleich struktureller Eigenschaften zu Interferenzen führte, die es den Lernenden erschwerten, den nachfolgenden Bilder Informationen zu entnehmen. Die Teilnehmer der Text-vor-Bild-Gruppen mussten sich möglicherweise beim Wissenserwerb allein auf die Texte stützen, was die bessere Leistung der kongruenten gegenüber der komplementären Bedingung erklärt.

Als Fazit lässt sich festhalten, dass sich die teils widersprüchlichen Resultate anhand des Modells des integrierten Text- und Bildverstehens deuten lassen, obwohl die Vorhersagen der ursprünglichen *Interferenzhypothese* nicht zutrafen. Um wirklich abschätzen zu können, welche kognitiven Prozesse dem PTS-Effekt zugrunde liegen, musste eine zusätzliche Untersuchung durchgeführt werden, die im folgenden Kapitel ausführlich erläutert wird.

7 Studie II: Interferenzen versus Konstruktion

Die Ergebnisse der ersten Studie hatten gezeigt, dass ein picture-text-sequencing-Effekt vorhanden war, wenn die Teilnehmer piktoriale und verbale Informationen sequentiell zum Aufbau von Wissensstrukturen nutzten. Allerdings konnte keine der beiden Hypothesen die verschiedenen Befunde befriedigend erklären, weshalb anhand des Modells des integrierten Text und Bildverstehens (Schnotz & Bannert, 2003) zwei alternative Deutungsansätze abgeleitet wurden. Während der Erste von Vorteilen beim Aufbau des mentalen Modells ausging, griff der zweite Ansatz die ursprüngliche Annahme von Interferenzen auf. Da beide Deutungen im Rahmen des theoretischen Modells durchaus plausibel erschienen, wurden sie in einer zweiten Untersuchung gegenübergestellt.

7.1 Fragestellung und Hypothesen

Das Ziel der zweiten Studie war es, die beiden Erklärungsansätze zu vergleichen, die sich aus der Interpretation der ersten Studie ergeben hatten. So ließ sich der Umstand, dass selbst unter der komplementären Bedingung der PTS-Effekt bei der Lernleistung und -zeit beobachtet wurde, nur durch zusätzliche Annahmen deuten. Obwohl die Gruppen mit dem kongruenten Material im Wissenstest besser abgeschnitten hatten, war das Informationsverhältnis von Texten und Bildern für den Einfluss der Verarbeitungsreihenfolge offenbar unerheblich. Demnach spielte es keine Rolle, ob sich die verbalen und piktorialen Informationen auf die identischen Inhalte bezogen oder nicht. Es müssen also andere Eigenschaften der Texte und Bilder bzw. ihrer Repräsentation gewesen sein, die den PTS-Effekt verursacht haben. Nach Johnson-Laird (1983) zeichnen sich mentale Modelle durch strukturelle Übereinstimmungen mit den konkreten Sachverhalten aus, weshalb sich nicht ausschließen lässt, dass Analogiebeziehungen auch abstrahiert werden. Daher ist es grundsätzlich möglich, dass sich nacheinander verarbeitete Informationen auch dann beeinflussen, wenn sie sich nicht auf identische Sachverhalte beziehen. Abstrahierte Strukturen und funktionale Eigenschaften des ersten mentalen Modells können beispielsweise als Referenz, Entwurf oder auch als Bausteine bei nachfolgenden Konstruktionsprozessen verwendet werden. Eine solche Nutzung bereits vorhandener Elemente macht vor allem dann Sinn, wenn sich die verschiedenen Modelle auf Sachverhalte beziehen, die Teil einer gemeinsamen inhaltlichen Domäne sind.

Diese theoretischen Überlegungen lassen sich auch auf die Verarbeitung verbaler und piktorialer Informationen übertragen, wobei es in diesem Fall formatspezifische Eigenschaften zu berücksichtigen gilt. Fest steht, dass die Reihenfolge, in der Text und Bild zur Modellkonstruktion herangezogen wurden, Lernerfolg und -zeit offenbar beeinflusst haben. Allerdings stellt sich die Frage, welche Richtung und damit welche Ursache dieser

7 Studie II: Interferenzen versus Konstruktion

Einfluss hat. Argumentiert man auf der Grundlage des Modells des integrierten Text und Bildverstehens (Schnotz & Bannert, 1999, 2003), dann lassen sich zwei verschiedene Ansätze ableiten, die den PTS-Effekt jeweils auf die Wechselwirkung zwischen Reihenfolge und formatspezifischen Verarbeitungsprozessen zurückführen.

Der erste Ansatz geht davon aus, dass es Lernenden grundsätzlich leichter fällt, ein mentales Modell anhand eines Bildes zu konstruieren, da es sich in beiden Fällen um ein analoges Format handelt. Erhält ein Lernender das Bild vor dem Text, so wird er die piktorialen Informationen nutzen, um eine Repräsentation der dargestellten Sachverhalte aufzubauen. Dabei ist nicht nur der Konstruktionsprozess einfacher, das resultierende mentale Modell erleichtert auch die nachträgliche Verarbeitung des Textes. Es kann entweder als Vorlage oder Entwurf genutzt werden oder der Lernende verwendet im Fall eines komplementären Informationsverhältnisses einzelne Bausteine und Elemente beim Textverstehen. Auf diese Weise wird die beschränkte Kapazität des Arbeitsgedächtnisses optimal genutzt und sämtliche mentalen Ressourcen lassen sich für den Auf- und Ausbau von Wissensstrukturen einsetzen. Erfolgt die Präsentation des Textes jedoch vor dem Bild, muss der Lernende vergleichsweise mühsam das mentale Modell anhand einer propositionalen Repräsentation konstruieren. Die Überführung eines prinzipiell symbolischen Formates in ein analoges Format ist schon an sich eine anspruchsvolle Aufgabe. Hinzu kommt, dass der Lernende keine Möglichkeit hat auf eine Modellvorlage oder einzelne Bausteine zurückzugreifen, weshalb die mentalen Ressourcen nicht optimal eingesetzt werden können. Die Präsentation des Bildes vor dem Text hat damit vor allem eine positive Wirkung, da die richtige Anordnung den Aufbau von Wissensstrukturen fördert.

Im Gegensatz dazu, knüpft der zweite Ansatz an die ursprüngliche *Interferenzhypothese* aus Studie I an, die von einem negativen Einfluss der Reihenfolge auf den Prozess des Wissenserwerbs ausgeht. Demnach ist der PTS-Effekt vor allem darauf zurückzuführen, dass es unvorteilhaft ist, den Text vor dem Bild zu lernen. Schnotz (2002, 2005) macht dafür Interferenzen verantwortlich, die entstehen, wenn ein Lernender versucht, ein mentales Modell, das auf der Grundlage eines Textes konstruiert wurde, und eine nachfolgendes Bild zu integrieren. So wird das Bild zwangsläufig von dem mentalen Modell abweichen, da die interne Repräsentation nicht zuletzt zu einem großen Teil das individuelle Vorwissen des Lernenden umfasst. Im Rahmen der ersten Untersuchung war noch davon ausgegangen worden, dass diese Abweichungen zwischen Bild und Modell nur dann negative Konsequenzen haben würden, wenn die zugrunde liegenden Formate sich auf die gleichen Sachverhalte beziehen. Allerdings konnte nicht ausgeschlossen werden, dass Interferenzen auch selbst dann auftreten, wenn Text und Bild in einem komplementären Informationsverhältnis zueinander stehen. In diesem Fall wären es allgemeine Strukturen oder funktionale Merkmale, die das mentale Modell und das Bild gemeinsam haben, aber die sehr wahrscheinlich voneinander abweichen werden. Ein Abgleich der beiden Repräsentationen würde zu einem kognitiven Konflikt führen, der es den Lernenden erschwert, den nachfolgenden Bilder Informationen zu entnehmen und in das bestehende mentale Modell zu integrieren.

Aus dem Blickwinkel der *cognitive load theory* sind Interferenzen aber nichts anderes

als *extraneous load* (Sweller & Chandler, 1994), der durch eine spezielle Anordnung von Text und Bild verursacht wird. Dass die erste Studie keinen bedeutsamen Zusammenhang zwischen der Verarbeitungsreihenfolge und der mentalen Belastung ergab, mag daran gelegen haben, dass sich die verschiedenen *load*-Formen anhand von Selbsteinschätzungen faktisch nicht unterscheiden lassen (J. van Merriënboer et al., 2002; Schnotz & Kürschner, 2007). So ist es durchaus vorstellbar, dass die mentale Belastung, welche die Teilnehmer während des Lernens empfanden, der Höhe nach zwar vergleichbar war, aber unterschiedliche Ursachen und damit auch Konsequenzen für den Wissenserwerb hatte. Die Anstrengung, die von den Bild-vor-Text Gruppen empfunden wurde, lässt sich möglicherweise auf *germane load* (Sweller et al., 1998) zurückführen, der beim Aufbau neuer Wissenstrukturen entsteht. Im Gegensatz dazu, könnte die empfundenen Belastung der Teilnehmer, die zuerst den Text gelesen hatten, eine Folge von Interferenzen bzw. *extraneous load*) gewesen sein. Diese Vermutungen ließen sich allerdings nicht allein anhand der Daten belegen, die im Rahmen der ersten Studie erhoben wurden. Wirklich aussagekräftig war der Vergleich mit Personen, die entweder nur die Bilder oder nur die Texte zum Wissenserwerb nutzen konnten. Dabei sollten Teilnehmer, denen beide Formate und damit auch mehr Informationen zur Verfügung standen, beim Lerntest deutlich besser abschneiden. Dies würde jedoch nicht gelten, wenn Interferenzen den Lernprozess beeinflussten. Demnach ist vorstellbar, dass Versuchspersonen der Text-vor-Bild Gruppen sich zwar mehr anstrengen, aber kein besseres Testergebnis erzielen. Würden die Teilnehmer, die nur den Text gelesen hatten, dagegen eine geringere mentale Belastung empfinden aber ein vergleichbares Resultat im Wissenstest erzielen, dann spräche dies für das Auftreten von *extraneous load* und damit für die *Interferenzhypothese*.

Auf der anderen Seite ist zu erwarten, dass sich auch die Bild-vor-Text Gruppen mehr anstrengen müssen, als Personen, denen allein die Bilder vorgelegt wurden. Die nachträglich verarbeiteten verbalen Informationen müssen in das vorhandene mentale Modell integriert werden, was den nicht selten Rückgriff auf schematisches Vorwissen erforderlich macht. Die resultierende Anstrengung sollte aber mit einer tieferen Verarbeitung und besseren Lernergebnissen einhergehen, weshalb diese Form der mentalen Belastung im Sinne der CLT als *germane load* anzusehen wäre. Damit lassen sich erneut zwei alternative Erklärungen für den PTS-Effekt formulieren, die sich aber nicht unbedingt widersprechen müssen. Grundsätzlich gilt es jedoch zu bedenken, dass auch der Umfang des Lernmaterials als eine Quelle der Belastung empfunden werden kann. Schnotz (2010) spricht in diesem Zusammenhang auch von *mentaler Energie*, die im Verlauf eines Lernprozesses aufgebraucht wird. Die Fähigkeit und Bereitschaft eines Studierenden sich über längere Zeit intensiv mit einem Thema auseinander zu setzen, ist demnach eine endliche Ressource. Lernende, die nur vier Texte zu lesen haben, werden daher in der Regel weniger Energie aufbringen müssen, als Personen, die sich zusätzlich mit vier weiteren Bildern konfrontiert sehen. Konsequenterweise wird die empfundene mentale Belastung in den Gruppen mit zwei Formaten höher sein, ohne dass dies zwingend ein Hinweis auf Interferenzen oder die besonders elaborierte Konstruktion eines mentalen Modells ist. Eine solche Verzerrung kann umgangen werden, indem man die Lernzeit statistisch

7 Studie II: Interferenzen versus Konstruktion

kontrolliert und ihren Einfluss auf die abhängigen Variablen herauspartialisiert. Treffen die theoretischen Vermutungen zu, müssten sich die vorhergesagten Zusammenhänge dennoch nachweisen lassen. Dementsprechend lauten Forschungsfrage und Hypothesen:

Fragestellung: Ist der PTS-Effekt das Ergebnis von Interferenzen oder auf Verarbeitungsvorteile bei der Konstruktion eines mentalen Modells zurückzuführen?

Interferenz-Hypothese: Für den PTS-Effekt sind negative Einflüsse auf den Lernprozess verantwortlich, die auf Interferenzen zurückzuführen sind. Bei kontrollierter Lernzeit strengen sich Teilnehmer der Text-vor-Bild Gruppen beim Wissenserwerb stärker an, als Personen, die nur die Texte lesen. Trotz der größeren Anstrengung schneiden die Teilnehmer der Text-vor-Bild Gruppen im Lerntest nicht besser ab.

Konstruktions-Hypothese: Für den PTS-Effekt sind positive Einflüsse auf den Lernprozess verantwortlich, die auf Unterstützungsfunktionen bei der Konstruktion mentaler Modelle zurückzuführen sind. Trotz kontrollierter Lernzeit strengen sich Teilnehmer der Bild-vor-Text Gruppen beim Lernen stärker an, als Personen, denen nur die Bilder vorgelegt wurden. Dank der größeren Anstrengung schneiden die Teilnehmer der Bild-vor-Text Gruppen im Lerntest aber besser ab.

7.2 Versuchsplanung

Um Vergleichbarkeit zu gewährleisten, orientierten sich die Überlegungen zur zweiten Studie weitgehend an den Vorgaben der ersten Studie. Das betraf vor allem die Messinstrumente, das Design und die Durchführung. Auf Abweichungen und Besonderheiten der zweiten Studie wird in den folgenden Abschnitten näher eingegangen.

7.2.1 Design

Der offensichtlichste Weg die Fragestellung der zweiten Untersuchung experimentell umzusetzen, bestand darin, das Design der ersten Studie um eine weitere Bedingung bzw. einen weiteren Faktor zu ergänzen. Dies hätte ein 2x2x2 Design mit insgesamt 8 Zellen ergeben, wobei es notwendig gewesen wäre, die Faktoren im Vergleich zur ersten Studie umzubenennen. Sehr viel problematischer war dagegen der Aufwand, der mit einem vollständig gekreuzten, dreifaktoriellen Design einherging. Bei 21 Personen pro Versuchsbedingung hätte eine vollkommen neue Stichprobe die Rekrutierung von insgesamt 168 Versuchspersonen erfordert. Da aus zeitlichen und ökonomischen Gründen dieser Stichprobenumfang schlicht zu umfangreich, wurde lediglich ein 2x2 Design umgesetzt, das Abbildung 7.1 zeigt.

Informationsverhältnis	Repräsentationsformat	
	Text (T)	Bild (B)
<i>kongruent (kon)</i>	kon_T (n= 21)	kon_B (n=21)
<i>komplementär (komp)</i>	komp_T (n=21)	komp_B (n=21)

Abbildung 7.1: Untersuchungsdesign der zweiten Studie

Anstelle von Texten und Bildern erhielten die Versuchsteilnehmer der zweiten Studie immer nur eines der beiden Repräsentationsformate. Gleichzeitig wurde erneut das Informationsverhältnis variiert, wobei es sich hier eigentlich um vollständige und unvollständige Texte und Bilder handelte. Die ursprünglichen Bezeichnungen kongruent bzw. komplementär wurde jedoch beibehalten, um den inhaltlichen Bezug zur ersten Studie zu erleichtern. So sollte mit der zweiten Erhebung in erster Linie eine Grundlage geschaffen werden, anhand der sich die Ergebnisse der ersten Untersuchung interpretieren ließen. Ein solches Vorgehen war aus forschungsmethodischer Sicht allerdings nicht ganz unproblematisch. Obwohl auch die Teilnehmer der zweiten Studie zufällig auf die einzelnen Versuchsbedingungen verteilt werden sollten, konnte das Prinzip der Randomisierung bei Mittelwertvergleichen zwischen Studie I und Studie II nicht als erfüllt gelten.

Faktisch hatten die Teilnehmer der ersten Untersuchung keine Möglichkeit den Versuchsbedingungen der zweiten Untersuchung zugelost zu werden, so dass alle Analysen, die beide Stichproben miteinbezogen, nur in einem quasiexperimentellen Rahmen erfolgen konnten. Ohne Randomisierung war es jedoch nicht mehr möglich, systematische Verzerrungen auszuschließen, die auf eine ungleichmäßige Verteilung von Störgrößen zurückzuführen waren. In diesem Zusammenhang erschien es besonders angebracht zu sein, potentielle Störvariablen statistisch zu kontrollieren und als Kovariaten in den einzelnen Analysen zu berücksichtigen. Daher wurden erneut das *Interesse für Geologie/Geographie*, die *Erdkundenote*, das *domänenspezifische Vorwissen*, die *aktuelle Motivation*, die *räumliche und verbale Intelligenz* sowie der *Lernstil* erhoben. Neben den Kontrollvariablen wurden darüber hinaus alle abhängigen Variablen und Maße erfasst, die bereits während der ersten Untersuchung gemessen worden waren.

7.2.2 Stichprobe

Die Untersuchung fand in den selben Räumen der Universität Koblenz-Landau / Campus Landau statt, wie die erste Studie und erstreckte sich über einen Zeitraum von insgesamt sieben Wochen. Alle 84 Teilnehmer waren Studenten, von denen $\approx 39\%$ Lehramtsfächer, $\approx 37\%$ Psychologie, $\approx 10\%$ Pädagogik oder $\approx 14\%$ Sontiges studierten.

7 Studie II: Interferenzen versus Konstruktion

Die Mehrheit ($\approx 69\%$) befand sich mit vier oder weniger Semestern noch im Grundstudium, während 31% weiter fortgeschritten waren. Der Altersdurchschnitt betrug etwa 23 Jahre ($M = 23.2, SD = 3.6$) mit einem Minimum von 18 und einem Maximum von 35 Jahren. Hinsichtlich des Geschlechts dominierten eindeutig die Frauen (79%), was die männlichen Studierenden mit zu einer Minderheit machte. Erneut wurde eine Belohnung für die Teilnehmer in Aussicht gestellt, die entweder aus der Gutschrift von 2 Versuchspersonenstunden oder 10,-€ bestand. Vergleicht man die Stichproben der ersten und zweiten Studie miteinander (siehe Seite 136), dann gab es vor allem hinsichtlich des Studienwahl der Teilnehmer Unterschiede. An der ersten Untersuchung nahmen deutlich mehr Psychologiestudenten teil, was insofern ein Problem darstellen konnte, da dieser Studiengang einer Zulassungsbegrenzung unterliegt. Allerdings war damit zu rechnen, dass die Berücksichtigung leistungsrelevanter Kontrollvariablen stichprobenbedingte Verzerrungen zumindest teilweise ausgleichen würde. Hinsichtlich der Homogenität und Repräsentativität der Stichprobe galten ähnliche Bedenken, wie bei der ersten Studie.

7.2.3 Ablauf und Durchführung der Untersuchung

Die Durchführung der zweiten Studie erfolgte nach dem selben Muster, wie im Abschnitt 6.2.3 bereits beschrieben wurde. Alle Instruktionen, der Ablauf und die Reihenfolge der Testungen waren identisch mit dem Vorgehen bei der ersten Untersuchung, bis auf zwei Ausnahmen. Erstens erhielten die Teilnehmer ausschließlich Bilder oder ausschließlich Texte, was die Lernzeit erheblich verkürzte. Zweitens wurden Items, die sich auf die Schwierigkeit des Repräsentationsformats bezogen, das die Studierenden nicht erhalten hatten, aus der Nachbefragung entfernt.

7.2.4 Statistische Auswertung

Zur Beantwortung der zentralen Forschungsfrage wurden auch bei der zweiten Untersuchung vor allem varianzanalytische Verfahren verwendet, deren Berechnung mithilfe der Software SPSS 17.0 erfolgte. Da der Stichprobenumfang überschaubar war und bei einigen Vergleichen ein quasiexperimentelles Design vorlag, erschien es erneut angebracht zu sein, mögliche Störvariablen in die Analysen miteinzubeziehen. Zu diesem Zweck wurde vor jeder Hypothesentestung untersucht, ob die einzelnen Versuchsgruppen hinsichtlich der potentiellen Kovariaten Unterschiede aufwiesen. Zudem erfolgte im Vorfeld eine Prüfung der Voraussetzungen, unter denen die Varianzanalyse interpretierbare Ergebnisse lieferte. Die Annahme varianzhomogener Populationen wurde mit einem Test von (Levene, 1960) und die Annahme gleicher Kovarianzmatritzen im Fall der MANOVA mit einem Test von Box (1949) kontrolliert. Ob die Voraussetzung einer normalverteilten abhängiger Variablen erfüllt war, sollte darüber hinaus der Kolmogorow-Smirnow-Anpassungstest (vgl. Goodman, 1954) zeigen. Je nach Fragestellung wurde für vorhergesagte Unterschiede ein Signifikanzniveau von $\alpha = .05$ und bei der Prüfung von Äquivalenzannahmen ein $\alpha = .20$ zugrunde gelegt. Neben den üblichen Angaben, wie

den Prüfstatistiken und den Freiheitsgraden, wurde erneut $\bar{\eta}^2$ als Maß für die praktische Bedeutsamkeit der gefundenen Zusammenhänge berichtet.

7.3 Ergebnisse

Die Ergebnisdarstellung beginnt mit einem Vergleich der Texte und Bilder, den auch Abbildung 7.1 zeigt. Zwar war es für die eigentliche Fragestellung nicht von unmittelbarer Bedeutung, ob die Teilnehmer besser mit verbalen oder piktorialen Informationen lernten, doch sollten die Ergebnisse einen Eindruck davon geben, welches der Repräsentationsformate für sich genommen mehr Nutzen versprach.

7.3.1 Text versus Bild

Vor der Berechnung der Kovarianzanalyse sollte zuerst geprüft werden, welche Kontrollvariablen sinnvoller Weise aufgenommen werden mussten. Aus diesem Grund wurden für jede potentielle Störgröße ein Äquivalenztest durchgeführt, um sicherzustellen, dass sich diese möglichst gleichmäßig auf die einzelnen Experimentalgruppen verteilten. Eine zutreffende Interpretation der resultierenden F-Werten setzt jedoch voraus, dass die betreffende Variable normalverteilt und die Fehlervarianzen in den Gruppen vergleichbar waren. Tabelle 7.1 zeigt das Ergebnis der Prüfung dieser Voraussetzungen, die anhand des Kolmogorow-Smirnow-Tests (Goodman, 1954) und einem Verfahren nach Levene (1960) erfolgte.

Die Ergebnisse zeigten, dass eine ganze Reihe von Variablen entweder nicht normalverteilt waren oder keine homogenen Fehlervarianzen aufwiesen. Da die Zellen jedoch gleichmäßig besetzt und der Versuchsplan vollständig gekreuzt war, stellte dies im Fall der verletzten Normalverteilungsannahme kein Problem dar (Milligan et al., 1987). Sehr viel schwerer wog dagegen die Varianzheterogenität beim *Vorwissen*, was sich nicht so einfach vernachlässigen ließ. Bei einer Verletzung der Homogenitätsannahme raten Diehl und Arbinger (1990) daher, auf die ANOVA zu verzichten und stattdessen ein alternatives Verfahren zu verwenden. Aus diesem Grund wurden für das *Vorwissen* eine Prüfstatistik nach Brown und Forsythe (1974) berechnet, während die Untersuchung der anderen potentiellen Störgrößen mittels zweifaktorieller Varianzanalysen erfolgte. Da getestet werden sollte, ob sich die einzelnen Merkmale gleichmäßig auf die verschiedenen Bedingungen verteilten, wurde ein α -Fehler von 20% zugrunde gelegt.

Gleich die erste ANOVA für das Variable *Interesse*, ergab einen bedeutsamen Unterschied. Zwar wurden weder der Faktor *Repräsentationsformat* ($F_{(1;80)} = 0.09$, $p = .925$, $\bar{\eta}^2 = .000$), noch der Interaktionseffekt ($F_{(1;80)} = 0.09$, $p = .302$, $\bar{\eta}^2 = .013$) signifikant, doch deutete das Ergebnis für den Faktor *Informationsverhältnis* ($F_{(1;80)} = 3.23$, $p = .079$, $\bar{\eta}^2 = .039$) auf einen Unterschied zwischen den Gruppen hin. Aus diesem Grund wurde das Interesse für Geologie als Kovariate aufgenommen. Die Analyse für die *Erdkundennote* hatte keine signifikanten Effekte ergeben, weshalb dieses Merkmal nicht berücksichtigt

7 Studie II: Interferenzen versus Konstruktion

Tabelle 7.1: Kolmogorow-Smirnow- und Levene-Tests für die Auswahl der Kontrollvariablen der zweiten Untersuchung

Merkmal	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Interesse</i>	<i>Levene – Test</i>	$F = 0.93$	$df_1 = 3; df_2 = 80$	$p = .964$
	<i>K – S – Test</i>	$Z_{K-S} = 2.13$	–	$p < .001^{**}$
<i>Erdkundenote</i>	<i>Levene – Test</i>	$F = 1.66$	$df_1 = 3; df_2 = 80$	$p = .183$
	<i>K – S – Test</i>	$Z_{K-S} = 2.09$	–	$p < .001^{**}$
<i>Vorwissen</i>	<i>Levene – Test</i>	$F = 3.57$	$df_1 = 3; df_2 = 80$	$p = .018^*$
	<i>K – S – Test</i>	$Z_{K-S} = 1.12$	–	$p = .165$
<i>Motivation</i> <i>FAM</i>	<i>Leven – Test</i>	$F = 2.28$	$df_1 = 3; df_2 = 80$	$p = .086$
	<i>K – S – Test</i>	$Z_{K-S} = 1.01$	–	$p = .257$
<i>Verbale</i> <i>Intelligenz</i>	<i>Leven – Test</i>	$F = 1.39$	$df_1 = 3; df_2 = 80$	$p = .252$
	<i>K – S – Test</i>	$Z_{K-S} = 1.43$	–	$p = .033^*$
<i>Räumliche</i> <i>Intelligenz</i>	<i>Levene – Test</i>	$F = 1.57$	$df_1 = 3; df_2 = 80$	$p = .203$
	<i>K – S – Test</i>	$Z_{K-S} = 0.65$	–	$p = .795$
<i>Lernstil</i> <i>Verbalizer</i>	<i>Levene – Test</i>	$F = 0.32$	$df_1 = 3; df_2 = 80$	$p = .809$
	<i>K – S – Test</i>	$Z_{K-S} = 1.02$	–	$p = .252$
<i>Lernstil</i> <i>Visualizer</i>	<i>Levene – Test</i>	$F = 0.31$	$df_1 = 3; df_2 = 80$	$p = .814$
	<i>K – S – Test</i>	$Z_{K-S} = 1.67$	–	$p = .007^{**}$

Anm.: K-S-Test = Kolmogorow-Smirnow-Anpassungstest; signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen (* < .05) bzw. durch zwei Sternchen (** < .01) hervorgehoben

wurde. Zur Testung des *Vorwissens* konnte nicht auf eine ANOVA zurückgegriffen werden, da die Varianzen in den Gruppen heterogen waren. Daher wurde ein Verfahren nach Brown und Forsythe (1974) verwendet, das bei gleich stark besetzten Zellen zu identischen F-Werten führt wie eine *normale* Varianzanalyse, aber robuster ist. Die eigentliche Korrektur fand über die Freiheitsgrade statt, was einer verzerrungsbedingten Erhöhung des α -Fehlers entgegen wirken sollte. Da die Statistiksoftware SPSS 17 keine exakte Irrtumswahrscheinlichkeit für die zweifaktorielle Variante des Verfahrens ausgibt, werden im Folgenden die p-Werte berichtet, die nach einer Tabelle von Diehl und Arbinger (1990) ermittelt wurden. Es zeigte sich, dass der Prüfwert für den Faktor *Repräsentationsformat* den kritischen Wert für ein α von 20 % überschritt ($F_{(1;73)} = 1.72, p < .20, \bar{\eta}^2 = .021$). Das gleiche traf auf den Interaktionseffekt zu ($F_{(1;73)} = 5.10, p < .05, \bar{\eta}^2 = .060$), während der Faktor *Informationsverhältnis* nicht signifikant wurde ($F_{(1;73)} = 1.53, p > .20, \bar{\eta}^2 = .019$). Aufgrund dieser Unterschiede zwischen den Gruppen wurde das Vorwissen in die Analyse aufgenommen.

Als ähnlich ungleichmäßig verteilt erwies sich das Merkmal *Motivation*, das mit einer Kurzform des FAM erhoben worden war. Die ANOVA ergab bedeutsame Unterschiede für das *Repräsentationsformat* ($F_{(1;80)} = 3.26, p = .075, \bar{\eta}^2 = .039$), das *Informationsverhältnis*

($F_{(1;80)} = 4.94$, $p = .029$, $\bar{\eta}^2 = .058$) und eine signifikante Interaktion ($F_{(1;80)} = 4.79$, $p = .032$, $\bar{\eta}^2 = .056$). Damit musste die Motivation in die Analysen miteinbezogen werden.

Für die beiden Untertests des IST 2000 R und die Lerstilskalen wurden MANOVA's berechnet, da ein Vergleich der Kovarianzen in den zu keinem signifikanten Ergebnis geführt hatte, wie der Box-Tests für die Intelligenz- ($M_{Box} = 11.69$, $F = 1.240$, $df_1 = 9$, $df_2 = 73342.77$, $p = .265$) und Verbalizer-Visualizer-Skalen ($M_{Box} = 6.82$, $F = 0.72$, $df_1 = 9$, $df_2 = 73342.77$, $p = .688$) zeigte. Das multivariate Verfahren ergab für die Untertests des IST 2000 R eine Prüfstatistik die über dem kritischen Wert für ein α von 20% lag. Offenbar gab es Unterschiede zwischen den Gruppen, die verschiedene *Repräsentationsformate* erhalten hatten ($V = .065$, $F_{(2;79)} = 2.741$, $p = .070$, $\bar{\eta}^2 = .065$). Der Faktor *Informationsverhältnis* ($V = .026$, $F_{(2;79)} = 1.036$, $p = .360$, $\bar{\eta}^2 = .026$) und die Interaktion ($V = .014$, $F_{(2;79)} = 0.577$, $p = .564$, $\bar{\eta}^2 = .014$) wurden nicht signifikant. Dagegen ergab die Analyse der beiden Lerntilskalen keine Hinweise auf nennenswerte Gruppenunterschiede, weshalb neben dem *Interesse*, dem *Vorwissen* und der *Motivation* nur die Messwerte für die *verbale* und *räumliche Intelligenz* als Kovariaten berücksichtigt werden mussten.

Die erste Studie hatte ein ganze Reihe von statistischen Analysen beinhaltet, die Forschungshypothesen zur Lernleistung, der mentalen Belastung und der Lernzeit überprüften. Für das Design, das Abbildung 7.1 veranschaulicht, gab es jedoch keine hypothetischen Vorhersagen, da alle inhaltlichen Fragen einen Vergleich zwischen den Gruppen der ersten und der zweiten Untersuchung hinausliefen. Aus Gründen der Übersichtlichkeit wurde daher lediglich für den Lernerfolg eine zweifaktorielle Kovarianzanalyse gerechnet. Zudem hatten im Rahmen der ersten Studie mehr oder weniger alle vier Skalen des Wissenstests in die gleiche Richtung gedeutet, so dass es sich der Einfachheit halber anbot, den Gesamtscore als abhängige Variable zu verwenden. Die Prüfung der Bedingungen, unter denen die Varianzanalyse gültige Schlüsse zulässt, ergab dabei durchweg positive Resultate. Weder der Kolmogorow-Smirnow-Anpassungstest ($Z_{K-S} = 1.12$, $p = .162$) noch der Levene-Test ($F_{(3;80)} = 1.37$, $p = .257$) wiesen auf eine Verletzung zentraler Annahmen hin, weshalb nichts gegen die Berechnung einer zweifaktoriellen ANCOVA mit fünf Kovariaten sprach. Abbildung 7.2 gibt das Ergebnis der Analyse grafisch wieder.

Die Zahlen, die dieser Grafik zugrunde lagen, zeigten zunächst, dass nicht alle Kovariaten signifikant wurden. Während die Merkmale *Interesse* ($F_{(1;75)} = 8.641$, $p = .004$, $\bar{\eta}^2 = .103$), *Motivation* ($F_{(1;75)} = 5.971$, $p = .017$, $\bar{\eta}^2 = .074$) und *verbale Intelligenz* ($F_{(1;75)} = 5.350$, $p = .023$, $\bar{\eta}^2 = .067$) einen bedeutsamen Zusammenhang mit der abhängigen Variablen aufwiesen, war der Einfluss des *Vorwissens* ($F_{(1;75)} = 0.537$, $p = .466$, $\bar{\eta}^2 = .007$) und der *räumlichen Intelligenz* ($F_{(1;75)} = 1.885$, $p = .164$, $\bar{\eta}^2 = .025$) zu vernachlässigen. Davon abgesehen, ergab die ANCOVA zwei signifikante Haupteffekte. Das *Repräsentationsformat* ($F_{(1;75)} = 8.534$, $p = .005$, $\bar{\eta}^2 = .102$) wie auch das *Informationsverhältnis* ($F_{(1;75)} = 10.043$, $p = .002$, $\bar{\eta}^2 = .118$) wirkten sich demnach überzufällig stark auf die Lernleistung der Versuchsteilnehmer aus. Dagegen blieb die Interaktion ($F_{(1;75)} = 1.051$, $p = .309$, $\bar{\eta}^2 = .014$) zwischen den beiden Faktoren ohne Bedeutung. Abbildung 7.2 veranschaulicht, dass die Studierenden mehr Wissen erwarben, wenn sie dazu einen Text verwenden konnten. Ebenso war es von Vorteil, *vollständige* bzw. kongruente Versionen zu verwenden, da

7 Studie II: Interferenzen versus Konstruktion

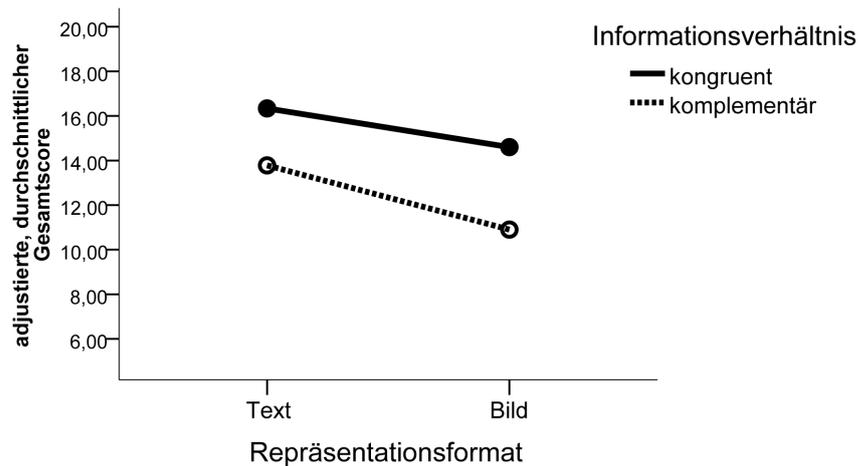


Abbildung 7.2: Mittlere adjustierte Gesamtleistung in der zweiten Studie

diese einfach mehr Informationen enthielten.

7.3.2 Interferenz-Hypothese

Der inhaltliche Schwerpunkt der zweiten Untersuchung betraf weniger die Frage, ob man besser mit Texten oder Bildern lernt. Aus theoretischer Sicht war es sehr viel interessanter herauszufinden, wodurch der PTS-Effekt möglicherweise verursacht wird. Die erste Annahme, die in diesem Abschnitt untersucht werden soll, machte Interferenzen für die Vorteile der Bild-vor-Text-Gruppen beim Wissenserwerb verantwortlich. Es handelt sich hierbei um einen negativen Effekt, der in Anlehnung an die *CLT* auf eine starke mentale Belastung als Folge von *extraneous load* zurückgeführt werden kann. Der Nachweis dieser Annahme erforderte den Vergleich zwischen Lernenden, die zuerst einen Text und dann ein Bild verarbeitet hatten, und Lernenden, denen nur der Text vorlag. Um den vermuteten Zusammenhang zu untersuchen, wurden die Bild-vor-Text- und die nur-Text-Gruppen in einem zweifaktoriellen Design angeordnet, das Abbildung 7.3 zeigt.

Informationsverhältnis	Zusätzliches Format	
	<i>ja</i>	<i>nein</i>
<i>kongruent</i>	kongruente Texte + Bilder (n= 21)	kongruente Texte (n=21)
<i>komplementär</i>	komplementäre Texte +Bilder (n=21)	komplementäre Texte (n=21)

Abbildung 7.3: Untersuchungsdesign zur Prüfung der Interferenz-Hypothese

Mit Blick auf den quasiexperimentellen Charakter der Untersuchung und die unterschiedliche Zusammensetzung der beiden Stichproben (siehe Abschnitt 6.2.2 bzw. 7.2.2, sollte die Auswertung kovarianzanalytisch erfolgen. Zu diesem Zweck wurden erneut Äquivalenztests durchgeführt, was zunächst die Prüfung einiger Voraussetzungen erforderlich machte. Tabelle 7.2 gibt die Ergebnisse der Voraussetzungsprüfung wieder.

Tabelle 7.2: Kolmogorow-Smirnow- und Levene-Tests für die Auswahl der Kontrollvariablen zur Testung der Interferenz-Hypothese

Merkmal	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Interesse</i>	<i>Levene – Test</i>	$F = 0.19$	$df_1 = 3; df_2 = 80$	$p = .901$
	<i>K – S – Test</i>	$Z_{K-S} = 2.15$	–	$p < .001^{**}$
<i>Erdkundenote</i>	<i>Levene – Test</i>	$F = 2.1$	$df_1 = 3; df_2 = 80$	$p = .107$
	<i>K – S – Test</i>	$Z_{K-S} = 2.14$	–	$p < .001^{**}$
<i>Vorwissen</i>	<i>Levene – Test</i>	$F = 0.57$	$df_1 = 3; df_2 = 80$	$p = .638$
	<i>K – S – Test</i>	$Z_{K-S} = 1.07$	–	$p = .203$
<i>Motivation FAM</i>	<i>Leven – Test</i>	$F = 0.384$	$df_1 = 3; df_2 = 80$	$p = .764$
	<i>K – S – Test</i>	$Z_{K-S} = 0.70$	–	$p = .715$
<i>Verbale Intelligenz</i>	<i>Leven – Test</i>	$F = 1.71$	$df_1 = 3; df_2 = 80$	$p = .172$
	<i>K – S – Test</i>	$Z_{K-S} = 1.35$	–	$p = .053$
<i>Räumliche Intelligenz</i>	<i>Levene – Test</i>	$F = 0.576$	$df_1 = 3; df_2 = 80$	$p = .633$
	<i>K – S – Test</i>	$Z_{K-S} = 0.91$	–	$p = .373$
<i>Lernstil Verbalizer</i>	<i>Levene – Test</i>	$F = 0.377$	$df_1 = 3; df_2 = 80$	$p = .770$
	<i>K – S – Test</i>	$Z_{K-S} = 1.11$	–	$p = .168$
<i>Lernstil Visualizer</i>	<i>Levene – Test</i>	$F = 0.688$	$df_1 = 3; df_2 = 80$	$p = .562$
	<i>K – S – Test</i>	$Z_{K-S} = 2.13$	–	$p < .001^{**}$

Anm.: K-S-Test = Kolmogorow-Smirnow-Anpassungstest; signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen ($* < .05$) bzw. durch zwei Sternchen ($** < .01$) hervorgehoben

Es ist zu erkennen, dass für die Merkmale *Interesse*, *Erdkundenote* und die *Visualizer-Skala* die Normalverteilungsannahme nicht zu halten war. Da jedoch der Levene-Test bei keiner der Variablen signifikant wurde und die Zellen gleichmäßig besetzt waren, erschien die Berechnung einer zweifaktoriellen ANOVA unbedenklich zu sein. Die Äquivalenztestung erfolgte erneut auf der Grundlage einer α -Fehlerwahrscheinlichkeit von 20% und begann mit dem *Interesse* für Geologie, wobei die Ergebnisse auf eine gleichmäßige Verteilung des Merkmals hinwies. Dagegen deutete sich bei der *Erdkundenote* ein kleiner Gruppenunterschied für den Faktor *zusätzliches Format* ($F_{(1;80)} = 2.277$, $p = .135$, $\bar{\eta}^2 = .028$) an, da die ermittelte Irrtumswahrscheinlichkeit unter dem Signifikanzniveau von $\alpha = .20$ lag. Obwohl das *Informationsverhältnis* ($F_{(1;80)} = 0.307$, $p = .581$, $\bar{\eta}^2 = .004$) und der Interaktionseffekt ($F_{(1;80)} = 0.529$, $p = .469$, $\bar{\eta}^2 = .007$) den kritischen F-Wert deutlich verfehlten, wurde die Erdkundenote als Kovariate aufgenommen.

7 Studie II: Interferenzen versus Konstruktion

Ähnliches traf auf das *Vorwissen* zu, auch wenn hier die ungleichmäßige Verteilung des Merkmals sehr viel deutlicher war. Demnach unterschritt sowohl der Effekt für den Faktor *Informationsverhältnis* ($F_{(1;80)} = 3.870$, $p = .053$, $\bar{\eta}^2 = .046$), als auch die Interaktion ($F_{(1;80)} = 6.964$, $p = .010$, $\bar{\eta}^2 = .080$) das α -Fehlerniveau von 20%. Allerdings gab es offenbar keine Unterschiede hinsichtlich des *zusätzlichen Formats* ($F_{(1;80)} = 0.389$, $p = .535$, $\bar{\eta}^2 = .005$), was für die Entscheidung, das Vorwissen als Kontrollvariable aufzunehmen, aber keine Rolle mehr spielte.

Die Ergebnisse zur *Motivation* wiesen ebenfalls auf eine Verletzung der Äquivalenzannahme hin, da der Faktor *zusätzliches Format* ($F_{(1;80)} = 4.615$, $p = .035$, $\bar{\eta}^2 = .055$) signifikant wurde. Zwar lagen die Prüfgrößen für das *Informationsverhältnis* ($F_{(1;80)} = 0.004$, $p = .953$, $\bar{\eta}^2 = .000$) und den Interaktionseffekt ($F_{(1;80)} = 0.001$, $p = .976$, $\bar{\eta}^2 = .000$) klar unter dem kritischen Wert, doch schien es auch in diesem Fall angeraten zu sein, das Merkmal *Motivation* in den weiteren Analysen miteinzubeziehen.

Etwas anders verhielt es sich dagegen bei die Skalen zur Erfassung der *verbalen* und *räumlichen Intelligenz*, deren Verteilung mit einer MANOVA geprüft wurde. Aufgrund des Box-Tests ($M_{Box} = 9.292$, $F = 0.99$, $df_1 = 9$, $df_2 = 73342.77$, $p = .449$) erschien die Durchführung einer multivariaten Varianzanalyse unproblematisch zu sein, doch verfehlten alle F-Werte das kritische Niveau, weshalb die beiden IST-Skalen nicht berücksichtigt werden mussten.

Die letzte Äquivalenztestung erfolgte für die *Verbalizer-* und *Visualizer-Skalen*, die den *Lernstil* bzw. die *Verarbeitungspräferenzen* erfassen sollten. Auch hier war die Kovarianz der beiden abhängigen Variablen über alle Versuchsgruppen vergleichbar ($M_{Box} = 3.80$, $F = 0.40$, $df_1 = 9$, $df_2 = 73342.77$, $p = .934$), weshalb eine MANOVA ohne Bedenken gerechnet werden konnte. Dabei zeigte sich, dass für den Effekt des Faktors *Informationsverhältnis* ($V = .046$, $F_{(2;79)} = 1.919$, $p = .153$, $\bar{\eta}^2 = .046$) eine Irrtumswahrscheinlichkeit ermittelt wurde, die unter dem Signifikanzniveau von 20% lag. Für das *zusätzliche Format* ($V = .008$, $F_{(2;79)} = 0.316$, $p = .730$, $\bar{\eta}^2 = .008$) und die Interaktion ($V = .001$, $F_{(2;79)} = 0.047$, $p = .954$, $\bar{\eta}^2 = .001$) konnte dagegen kein bedeutsamer Zusammenhang mit dem Merkmal nachgewiesen werden. Letztlich war das Ergebnis knapp, doch wurde auch der *Lernstil* als Kontrollvariable berücksichtigt.

Die Prüfung der eigentlichen Hypothesen erfolgte anhand einer 2x2 ANCOVA mit den Faktoren *Informationsverhältnis* und *zusätzliches Format*. Um die Fehlervarianz möglichst klein zu halten, wurden die *Erdkundenote*, das *Vorwissen*, die *Motivation* und der *Lernstil* als Kovariaten berücksichtigt. Neben diesen vier Merkmalen, war es aufgrund der Fragestellung notwendig, mit der *Lernzeit* eine fünfte Kontrollvariable in die Analyse miteinzubeziehen. Da die Hypothese zwei abhängige Variablen betraf, die das quasiexperimentelle Design aber auf unterschiedliche Weise beeinflussen sollte, wurden zwei getrennte ANCOVAs durchgeführt. Die erste Analyse bezog sich auf die *mentale Belastung* beim Lernen, wobei aufgrund theoretischer Überlegungen ein signifikanter Effekt für den Faktor *zusätzliches Format* vorhergesagt wurde. Die Prüfung der Voraussetzungen für die Verwendung varianzanalytischer Verfahren ergab, dass die Selbsteinschätzung zur mentalen Belastung offensichtlich nicht normalverteilt war ($Z_{K-S} = 2.81$, $p < .001$). Auf-

grund des Levene-Tests ($F_{(3;80)} = 0.843$, $p = .475$) konnte jedoch von Varianzhomogenität ausgegangen werden, so dass die Berechnung der ANCOVA bei gleich stark besetzten Zellen unbedenklich erschien.

Als Ergebnis zeigte sich, dass nur die Kovariaten *Motivation* ($F_{(1;75)} = 5.723$, $p = .019$, $\bar{\eta}^2 = .071$) und *Lernzeit* ($F_{(1;75)} = 4.197$, $p = .044$, $\bar{\eta}^2 = .053$) einen bedeutsamen Zusammenhang mit der abhängigen Variablen aufwiesen. Die *Erdkundenote* ($F_{(1;75)} = 2.895$, $p = .093$, $\bar{\eta}^2 = .037$), das *Vorwissen* ($F_{(1;75)} = 0.025$, $p = .876$, $\bar{\eta}^2 = .000$) und der *Lernstil* ($F_{(1;75)} = 0.036$, $p = .850$, $\bar{\eta}^2 = .000$) leisteten dagegen keinen signifikanten Beitrag zur Kontrolle der Fehlervarianz. In ähnlicher Weise verfehlten die Prüfgrößen des Faktors *Informationsverhältnis* ($F_{(1;75)} = 1.222$, $p = .273$, $\bar{\eta}^2 = .016$) und der Interaktion ($F_{(1;75)} = 0.473$, $p = .494$, $\bar{\eta}^2 = .006$) den kritischen Wert. Allerdings ergab die Analyse einen signifikanten Haupteffekt für den Faktor *zusätzliches Format* ($F_{(1;75)} = 4.047$, $p = .048$, $\bar{\eta}^2 = .051$), dessen Irrtumswahrscheinlichkeit knapp unter der Signifikanzniveau von 5% lag. Die Richtung des Effekts lässt sich anhand von Abbildung 7.4 ablesen, wo zu erkennen ist, dass mentale Belastung trotz kontrollierter Lernzeit bei den Text-vor-Bild-Gruppen deutlich höher war.

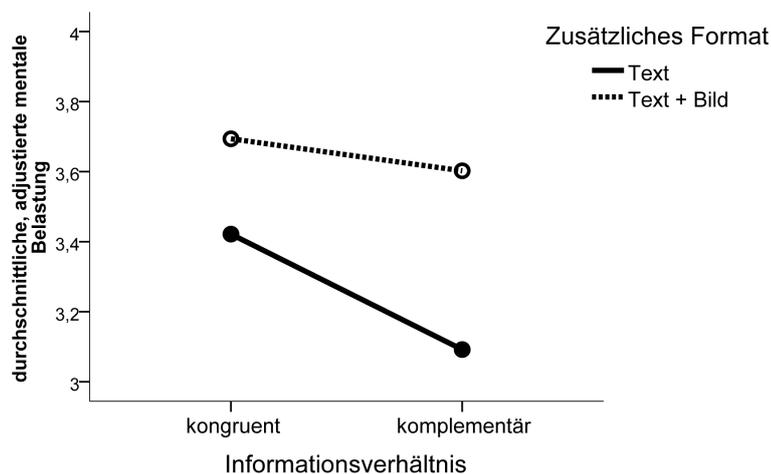


Abbildung 7.4: Interferenz-Hypothese: mentale Belastung

Bei der zweiten abhängigen Variablen, die im Rahmen der *Interferenz-Hypothese* einer Testung unterzogen werden sollte, handelte es sich um den *Lernerfolg*. Zu diesem Zweck wurde mit dem *Gesamtscore* des Wissenstests eine zweite ANCOVA gerechnet, wobei in diesem Fall jedoch kein Unterschied zwischen den Gruppen des Faktors *zusätzliches Format* zu erwarten war. Damit zielte diese Analyse auf die Bestätigung der Nullhypothese ab, was eine Anhebung des α -Fehlers auf 20% notwendig machte. Vor der Berechnung musste zunächst geprüft werden, ob auch der Gesamtscore ohne Einschränkungen als abhängige Variable verwendet werden konnte. Da sowohl der Kolmogorow-Smirnow-Anpassungstest ($Z_{K-S} = 0.88$, $p = .421$), als auch der Levene-Test ($F_{(3;80)} = 0.382$, $p = .767$) kein signifikantes Ergebnis ergaben, war die Gefahr einer Verzerrung der statistischen Prüfgrößen gering.

7 Studie II: Interferenzen versus Konstruktion

Ähnlich wie bei der ersten ANCOVA, wurde auch bei zweiten Analyse nur ein Teil der Kovariaten signifikant. Während für das *Vorwissen* ($F_{(1;75)} = 8.485, p = .005, \bar{\eta}^2 = .102$) und die *Motivation* ($F_{(1;75)} = 13.202, p = .001, \bar{\eta}^2 = .150$) bedeutsame Zusammenhänge nachgewiesen werden konnten, blieben die Prüfgrößen der Merkmale *Erdkundenote* ($F_{(1;75)} = 2.134, p = .148, \bar{\eta}^2 = .028$), *Lernstil* ($F_{(1;75)} = .364, p = .548, \bar{\eta}^2 = .005$) und *Lernzeit* ($F_{(1;75)} = 3.265, p = .075, \bar{\eta}^2 = .042$) unter dem kritischen Wert. Zudem ergab die Analyse einen signifikanten Haupteffekt für der Faktor *Informationsverhältnis* ($F_{(1;75)} = 6.177, p = .015, \bar{\eta}^2 = .076$). Die Effekte der Interaktion ($F_{(1;75)} = 0.234, p = .630, \bar{\eta}^2 = .003$) und auch die Irrtumswahrscheinlichkeit für den Faktor *zusätzliches Format* ($F_{(1;75)} = 0.003, p = .957, \bar{\eta}^2 = .000$) lagen deutlich über dem α von 20%, so dass die Äquivalenzannahme als bestätigt gelten konnte. Abbildung 7.5 veranschaulicht die Ergebnisse der zweiten Analyse grafisch.

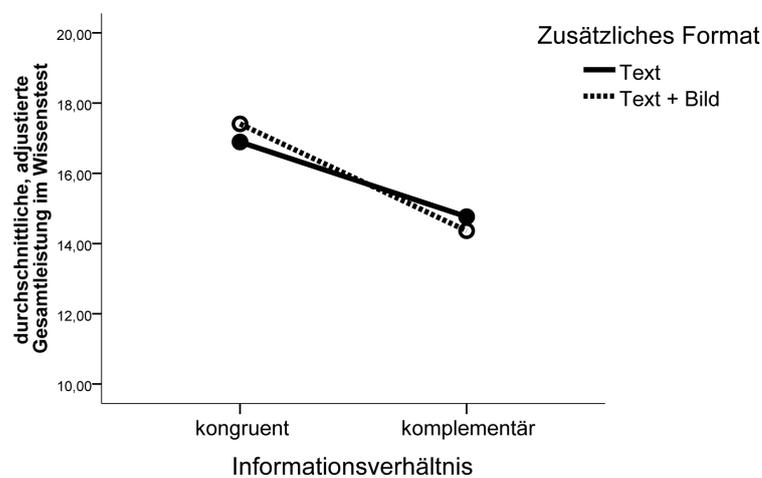


Abbildung 7.5: Interferenz-Hypothese: Gesamtleistung im Wissenstest

Fasst man die Resultate zusammen und vergleicht sie mit den Vorhersagen der Interferenzhypothese (siehe Seite 166), dann lässt sich eine sehr gute Übereinstimmung feststellen. So zeigt Abbildung 7.4, dass der signifikante Unterschied zwischen den Textvor-Bild und den nur-Text Gruppen hinsichtlich der mentalen Belastung beim Lernen in die vorhergesagte Richtung wies. Offensichtlich waren die nachfolgenden Bilder trotz kontrollierter Lernzeit eine Quelle zusätzlicher Beanspruchung. Zu diesem Befund passt es, dass die Verarbeitung der piktorialen Informationen den Lernenden keinen Vorteile beim Wissenserwerb brachte. Zwar waren die Teilnehmer mit den kongruenten bzw. vollständigen Materialien im Test erfolgreicher (vgl. Abbildung 7.5), aber von der Nutzung eines zusätzlichen Repräsentationsformates profitierten auch sie nicht. Die höhere mentale Belastung bei einer vergleichbaren Lernleistungen spricht recht deutlich für das Auftreten von *extraneous load* als Folge von Interferenzen. Allerdings kann auf dieser Grundlage nicht ausgeschlossen werden, dass womöglich auch positive Effekte für den PTS-Effekt verantwortlich sind. Daher wurde eine weitere Analyse durchgeführt, was in dem folgenden Abschnitt dargestellt ist.

7.3.3 Konstruktions-Hypothese

Die bisherigen Ergebnisse unterstützten die Vermutung, dass der PTS-Effekt durch das Auftreten von Interferenzen verursacht wird, die als Folge des Abgleichs einer mentalen (analogen) Repräsentation mit einem externen Bild entstehen. Gleichzeitig ist aber durchaus vorstellbar, dass es den Wissenserwerb positiv beeinflussen kann, wenn Lernende das Bild vor dem Text verarbeiten. Um letztlich zu ergründen, ob die mentale Belastung der Bild-zuerst-Gruppen beim Aufbau neuer Wissenstrukturen entstehen, wurde folgendes Untersuchungsdesign aus Gruppen der ersten und der zweiten Studie kombiniert.

		Zusätzliches Format	
		<i>ja</i>	<i>nein</i>
Informationsverhältnis	<i>kongruent</i>	kongruente Bilder + Texte (n= 21)	kongruente Bilder (n=21)
	<i>komplementär</i>	komplementäre Bilder + Texte (n=21)	komplementäre Bilder (n=21)

Abbildung 7.6: Untersuchungsdesign zur Prüfung der Konstruktions-Hypothese

Wie man in Abbildung 7.6 erkennen kann, sollten die Bild-vor-Text-Bedingungen mit den nur Bild-Bedingungen verglichen werden. Vor der eigentlichen Prüfung der *Konstruktions-Hypothese* wurde jedoch erneut die Verteilung der potentieller Störgrößen unter den verschiedenen Bedingungen untersucht, um potentielle Kontrollvariablen auszuwählen. Da zu diesem Zweck Äquivalenztestungen auf der Grundlage einer zweifaktoriellen Varianzanalyse durchgeführt werden sollten, war es wieder notwendig, zuerst die Voraussetzungen für den Einsatz dieses Verfahrens zu prüfen. Tabelle 7.3 fasst die Ergebnisse der Kolmogorow-Smirnow- und des Levene-Tests für die verschiedenen Merkmale zusammen.

Wie im Abschnitt zuvor, konnte die Normalverteilungsannahme für die Merkmale *Interesse*, *Erdkundenote* und die *Visualizer-Skala* nicht aufrecht erhalten werden. Da bei den betroffenen Variablen jedoch von homogenen Fehlervarianzen ausgegangen werden konnte und die Zellen gleichmäßig besetzt waren, bestand nur eine geringen Gefahr, eventuell verzerrte Ergebnisse zu interpretieren. Sehr viel schwerer wog dagegen die Varianzheterogenität beim *Vorwissen* und der *Motivation*, die sich nicht so einfach vernachlässigen ließ (Diehl & Arbinger, 1990). Aus diesem Grund wurden für diese beiden Merkmale eine Prüfstatistik nach Brown und Forsythe (1974) berechnet, während die Untersuchung der anderen potentiellen Störgrößen mittels zweifaktorieller Varianzanalysen erfolgte. Da getestet werden sollte, ob sich die einzelnen Merkmale gleichmäßig auf die verschiedenen Bedingungen verteilten, wurde wieder ein α -Fehler von 20% zugrunde gelegt.

7 Studie II: Interferenzen versus Konstruktion

Tabelle 7.3: Kolmogorow-Smirnow- und Levene-Tests für die Auswahl der Kontrollvariablen zur Testung der Konstruktions-Hypothese

Merkmal	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Interesse</i>	<i>Levene – Test</i>	$F = 0.24$	$df_1 = 3; df_2 = 80$	$p = .869$
	<i>K – S – Test</i>	$Z_{K-S} = 2.07$	–	$p < .001^{**}$
<i>Erdkundenote</i>	<i>Levene – Test</i>	$F = 0.43$	$df_1 = 3; df_2 = 80$	$p = .734$
	<i>K – S – Test</i>	$Z_{K-S} = 1.89$	–	$p = .002^{**}$
<i>Vorwissen</i>	<i>Levene – Test</i>	$F = 4.91$	$df_1 = 3; df_2 = 80$	$p = .003^{**}$
	<i>K – S – Test</i>	$Z_{K-S} = 0.99$	–	$p = .286$
<i>Motivation</i> <i>FAM</i>	<i>Levene – Test</i>	$F = 3.46$	$df_1 = 3; df_2 = 80$	$p = .020^*$
	<i>K – S – Test</i>	$Z_{K-S} = 1.08$	–	$p = .195$
<i>Verbale</i> <i>Intelligenz</i>	<i>Levene – Test</i>	$F = 1.28$	$df_1 = 3; df_2 = 80$	$p = .288$
	<i>K – S – Test</i>	$Z_{K-S} = 1.33$	–	$p = .058$
<i>Räumliche</i> <i>Intelligenz</i>	<i>Levene – Test</i>	$F = 0.03$	$df_1 = 3; df_2 = 80$	$p = .994$
	<i>K – S – Test</i>	$Z_{K-S} = 1.12$	–	$p = .164$
<i>Lernstil</i> <i>Verbalizer</i>	<i>Levene – Test</i>	$F = 0.11$	$df_1 = 3; df_2 = 80$	$p = .954$
	<i>K – S – Test</i>	$Z_{K-S} = 1.21$	–	$p = .106$
<i>Lernstil</i> <i>Visualizer</i>	<i>Levene – Test</i>	$F = 1.59$	$df_1 = 3; df_2 = 80$	$p = .199$
	<i>K – S – Test</i>	$Z_{K-S} = 2.07$	–	$p < .001^{**}$

Anm.: K-S-Test = Kolmogorow-Smirnow-Anpassungstest; signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen (* < .05) bzw. durch zwei Sternchen (** < .01) hervorgehoben

Die ANOVA für das *Interesse* ergab einen signifikanten Interaktionseffekt ($F_{(1;80)} = 2.502$, $p = .118$, $\bar{\eta}^2 = .030$), während die F-Werte der Faktoren *Informationsverhältnis* ($F_{(1;80)} = 1.463$, $p = .230$, $\bar{\eta}^2 = .018$) und *zusätzliches Format* ($F_{(1;80)} = 0.216$, $p = .643$, $\bar{\eta}^2 = .003$) keinen Zusammenhang anzeigten. Dagegen konnte bei der *Erdkundenote* von einer gleichmäßigen Verteilung ausgegangen werden. Aus diesem Grund war es nur erforderlich, das *Interesse* als Kontrollvariable zu verwenden.

Der Test nach Brown und Forsythe (1974) ergab für das *Vorwissen* keine signifikanten Ergebnisse. So überschritten weder die Prüfgrößen der Haupt-, noch des Interaktionseffekts den kritische Wert von $F_{(1;68)} = 1.679$ für ein α von 20%, weshalb das *Vorwissen* nicht berücksichtigt werden musste. Etwas anders verhielt es sich mit der *Motivation*, für die ebenfalls das Verfahren nach Brown und Forsythe (1974) verwendet wurde. Hier deutete ein signifikanter Effekt für den Faktor *Informationsverhältnis* ($F_{(1;77)} = 9.875$, $p < .005$, $\bar{\eta}^2 = .110$) eine Verletzung der Äquivalenzannahme an, während das zusätzliche Format ($F_{(1;77)} = 0.313$, $p > .200$, $\bar{\eta}^2 = .013$) und die Interaktion ($F_{(1;77)} = 0.377$, $p > .200$, $\bar{\eta}^2 = .010$) keine bedeutsamen Zusammenhang aufwiesen. Aus diesem Grund wurde die *Motivation* als Kontrollvariable in die nachfolgenden Analysen aufgenommen.

Zur Prüfung der Äquivalenzannahme für die beiden Untertests des IST 2000 R wur-

de eine MONOVA gerechnet. Dabei war das Ergebnis des Box-Tests nicht signifikant ($M_{Box} = 10.82$, $F = 1.15$, $df_1 = 9$, $df_2 = 73342.77$, $p = .324$), so dass man von gleichen Kovarianzen zwischen den Gruppen ausgehen konnte. Die Analyse ergab einen Effekt für den Faktor *zusätzliches Format* ($V = .060$, $F_{(2;79)} = 2.541$, $p = .085$, $\eta^2 = .060$), der über dem kritischen Wert für eine Irrtumswahrscheinlichkeit von 20% lag. Der Faktor *Informationsverhältnis* ($V = .000$, $F_{(2;79)} = 0.008$, $p = .992$, $\eta^2 = .000$) und die Interaktion ($V = .006$, $F_{(2;79)} = 0.230$, $p = .0795$, $\eta^2 = .006$) blieben dagegen ohne Bedeutung. Aufgrund des Ergebnisses wurden die verbale und die räumliche Intelligenz als Kovariaten aufgenommen.

Die letzte Äquivalenzprüfung betraf die *Verbalizer-* und *Visualizer-*Skalen, die zur Erfassung des *Lernstils* eingesetzt wurden. Nach dem positiven Ergebnis des Box-Tests ($M_{Box} = 9.12$, $F = 0.97$, $df_1 = 9$, $df_2 = 73342.77$, $p = .465$) wurde auch in diesem Fall eine MANOVA gerechnet. Diese ergab einen signifikanten Haupteffekt für den Faktor *zusätzliches Format* ($V = .092$, $F_{(2;79)} = 3.989$, $p = .022$, $\eta^2 = .092$), während die Prüfgrößen für das *Informationsverhältnis* ($V = .009$, $F_{(2;79)} = 0.343$, $p = .711$, $\eta^2 = .009$) und die Interaktion ($V = .003$, $F_{(2;79)} = 0.100$, $p = .905$, $\eta^2 = .003$) den kritischen Wert jeweils verfehlten. Damit musste auch der Lernstil als Kontrollvariable bei der nachfolgenden Analyse berücksichtigt werden.

Zu Untersuchung der *Konstruktions-Hypothese* sollte erneut eine 2x2 ANCOVA mit den Faktoren *Informationsverhältnis* und *zusätzliches Format* berechnet werden. Als Kovariaten wurden neben der obligatorischen *Lernzeit*, das *Interesse*, die *Motivation*, die *verbale* und *räumliche Intelligenz* sowie der *Lernstil* in die Analyse mit aufgenommen. Erneut sollten für die *mentale Belastung* und den Lernerfolg zwei getrennte ANCOVAs durchgeführt werden, wobei zuerst zu prüfen war, ob die beiden abhängigen Variablen die Voraussetzungen für eine Verwendung dieses Verfahrens erfüllten. Dabei zeigte sich, dass bei der *mentalen Belastung* die Normalverteilungsannahme verletzt wurde (Kolmogorow-Smirnow-Anpassungstest: $Z_{K-S} = 2.37$, $p < .001$). Da jedoch die Varianzen homogen zu sein schienen (Levene-Test: ($F_{(3;80)} = 0.148$, $p = .913$), konnte die ANCOVA berechnet werden.

Von den Kontrollvariablen ließ sich allein für die *Motivation* ($F_{(1;74)} = 8.114$, $p = .006$, $\eta^2 = .099$) ein bedeutsames Ergebnis nachweisen. Die Prüfgrößen der Merkmale *Lernzeit* ($F_{(1;74)} = 3.656$, $p = .060$, $\eta^2 = .047$), *Interesse* ($F_{(1;74)} = 0.585$, $p = .447$, $\eta^2 = .008$), *verbale Intelligenz* ($F_{(1;74)} = 0.714$, $p = .401$, $\eta^2 = .010$), *räumliche Intelligenz* ($F_{(1;74)} = 1.956$, $p = .166$, $\eta^2 = .026$) und *Lernstil* ($F_{(1;74)} = 0.634$, $p = .428$, $\eta^2 = .008$) hatten dagegen den kritischen Wert nicht überschritten. Zudem ergab die Analyse weder signifikante Effekte für die Faktoren *Informationsverhältnis* ($F_{(1;74)} = 2.584$, $p = .112$, $\eta^2 = .034$) oder das *zusätzliche Format* ($F_{(1;74)} = 0.223$, $p = .638$, $\eta^2 = .003$), noch erwies sich die Interaktion ($F_{(1;74)} = 3.687$, $p = .057$, $\eta^2 = .047$) als statistisch bedeutsam. Demnach gab es keine Unterschiede zwischen den vier Versuchsbedingungen hinsichtlich der mentalen Belastung.

Bevor die ANCOVA für den *Lernerfolg* berechnet werden konnte, mussten auch hier zunächst die Voraussetzungen geprüft werden. Der Kolmogorow-Smirnow-Anpassungstest ergab kein signifikantes Ergebnis ($Z_{K-S} = 1.138$, $p < .150$), so dass die Normalverteilungsannahme als erfüllt gelten konnte. Auch der Levene-Test wurde nicht signifikant

7 Studie II: Interferenzen versus Konstruktion

($F_{(3;80)} = 2.343$, $p = .079$), allerdings verfehlte die Prüfgröße nur knapp den kritischen Wert, was zumindest auf mäßig heterogene Varianzen schließen ließ. Dennoch wurde die Entscheidung getroffen, die ANCOVA zu verwenden und nur bei sehr knappen Ergebnissen für oder wider die Hypothese mit dem Verfahren von Brown und Forsythe (1974) zu vergleichen. Allerdings zeigte sich, dass ein solches Vorgehen nicht notwendig war, da eine Korrektur der Nennerfreiheitsgrade von 74 auf 71 keine nennenswerten Folgen für den F-Wert hatte, der in beiden Fällen für eine α -Fehlerwahrscheinlichkeit von 5% bei $F \approx 3.978$ lag.

Zunächst ergab die Analyse, dass die einbezogenen Kovariaten nur teilweise ein bedeutsamen Zusammenhang mit der abhängigen Variablen aufwiesen. So wurden das *Interesse* ($F_{(1;74)} = 8.588$, $p = .004$, $\bar{\eta}^2 = .104$), die verbale Intelligenz ($F_{(1;74)} = 13.807$, $p < .001$, $\bar{\eta}^2 = .157$), die räumliche Intelligenz ($F_{(1;74)} = 15.413$, $p < .001$, $\bar{\eta}^2 = .172$) und der *Lernstil* ($F_{(1;76)} = 4.229$, $p = .043$, $\bar{\eta}^2 = .054$) signifikant, während der Beitrag der *Motivation* ($F_{(1;74)} = 1.889$, $p = .173$, $\bar{\eta}^2 = .025$) und der *Lernzeit* ($F_{(1;74)} = 1.127$, $p = .292$, $\bar{\eta}^2 = .015$) zur Varianzaufklärung offenbar gering war. Mit Blick auf die Fragestellung konnte festgestellt werden, dass die Faktoren *Informationsverhältnis* ($F_{(1;74)} = 13.616$, $p < .001$, $\bar{\eta}^2 = .155$) und das *zusätzliche Format* ($F_{(1;74)} = 14.342$, $p < .001$, $\bar{\eta}^2 = .162$) einen bedeutsamen Einfluss auf den Wissenserwerb besaßen. Die Interaktionseffekt ($F_{(1;74)} = 0.099$, $p = .754$, $\bar{\eta}^2 = .001$) wurde dagegen nicht signifikant. Dieser Zusammenhang lässt sich auch gut in Abbildung 7.7 ablesen, da der parallele Verlauf der beiden Linien eine Wechselwirkung der beiden Faktoren sehr unwahrscheinlich macht. Die Darstellung verdeutlicht, dass es für den Lernerfolg sehr viel vorteilhafter war, wenn die Teilnehmer auf kongruentes bzw. vollständiges Material und auf einen zusätzlichen Text zurückgreifen konnten.

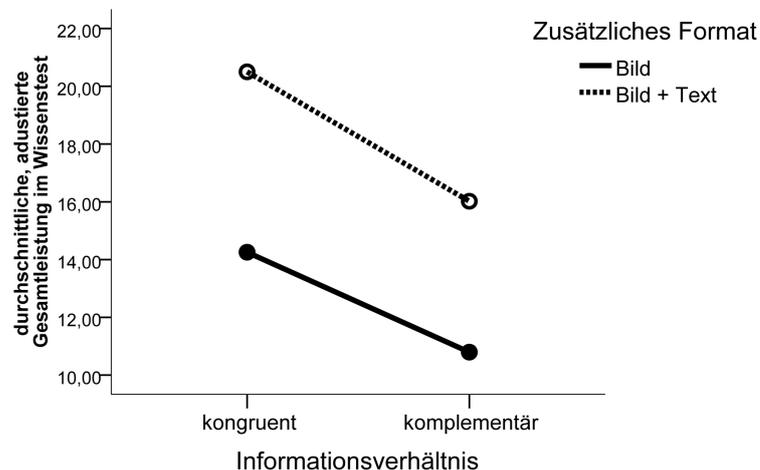


Abbildung 7.7: Konstruktions-Hypothese: Gesamtleistung im Wissenstest

Vergleicht man nun diese Ergebnisse und das Resultat der Analyse zur mentalen Belastung mit den den theoretischen Vorhersagen auf Seite 166, dann wird deutlich, dass nur ein Teil der *Konstruktions-Hypothese* zutrif. So strengten sich die Versuchsteilnehmer, die Bild und Text erhalten hatten, beim Lernen nicht stärker an, als Studierende,

denen nur die Bilder präsentiert wurden. Hypothesenkonform sind dagegen die Ergebnisse zum Wissenstest, da der Text trotz kontrollierter Lernzeit einen überaus positiven Einfluss auf den Wissenserwerb hatte. Welche Schlussfolgerungen sich aus diesen Befunden ziehen lassen, wird im folgenden Abschnitt diskutiert.

7.4 Diskussion

Das zentrale Anliegen der zweiten Studie war es, Hinweise auf die Ursachen des *picture-text-sequencing*-Effekts zu finden, der im Rahmen der ersten Untersuchung nachgewiesen werden konnte. Zu diesem Zweck erhielten die Teilnehmer einer zweiten Stichprobe entweder Texte oder Bilder in den entsprechenden komplementären bzw. kongruenten Versionen. Obwohl die insgesamt vier Bedingungen in erster Linie als Interpretationsgrundlage für die erste Studie dienen sollten, lag es nahe, die beiden Formate zu vergleichen. Zwar waren im Vorfeld der Analyse keine Hypothesen formuliert worden, doch erschienen die in Abschnitt 7.3.1 dargestellten Resultate interessant genug zu sein, um sie etwas näher zu betrachten. So ist es zuerst einmal trivial, dass Personen mit den vollständigen Formaten mehr Informationen lernen konnten, als die Studierenden, die auf die komplementären Versionen zurückgreifen mussten. Sehr viel interessanter war die Gegenüberstellung der Texte und Bilder, bei der eindeutig das verbale Format besser abschnitt. In diesem Zusammenhang stellte sich die Frage, weshalb die Versuchsteilnehmer mit den Texten eine höhere Gesamtleistung erzielten.

Eine Erklärung ist, dass die verbalen Formate einfach mehr testrelevante Informationen und Details enthielten, die in den Bildern nicht zur Verfügung stand. Dafür sprechen auch die Ergebnisse der Überprüfung des Lehrmaterials (siehe Kapitel 5.1.5), die ergeben hatte, dass die kongruenten Bild- und Textversionen sich inhaltlich in Regel nicht einmal zu 50% überlappten. Zudem war bei der Konstruktion des Wissenstests nicht explizit darauf geachtet worden, dass die Items immer auf der Grundlage beider Formate gelöst werden konnten. Dies war mit Blick auf die Fragestellung der ersten Studie aber auch gar nicht notwendig gewesen. Neben der Anzahl der enthaltenen Informationen, könnte eine weitere Ursache formatspezifische Merkmale sein. Demnach sind verbal dargestellte Sachverhalte relativ überschaubar, da ein Text eine bestimmte Länge besitzt, die durch die Anzahl der verwendeten Worte vorgegeben wird. Natürlich ist es möglich, dass bestimmte Inhalte umständlich aufbereitet wurden, oder sehr viel Vorwissen eingebracht werden muss, um die inhaltlichen Zusammenhänge zu verstehen. Handelt es sich aber um einen expositorischen Text, der eine Aneinanderreihung von Informationen darstellt, lassen sich diese in der vom Autor intendierten Reihenfolge recht einfach ablesen.

Bei Bildern ist eine vergleichbar reglementierte Verarbeitung häufig nicht möglich. Hier ist es sehr viel schwieriger relevante von irrelevanten Inhalten zu unterscheiden, da die Anzahl piktorial repräsentierter Informationen oftmals gar nicht erfasst werden kann. Die Herausforderung für die Bildgruppen bestand demnach vor allem darin, die richtigen Informationen zu identifizieren, wobei die ausgegebene Instruktion, möglichst viele der dargestellten Sachverhalte zu lernen, keine wirklich Orientierung bot. Beispielsweise konnte die

7 Studie II: Interferenzen versus Konstruktion

Anzahl der großen Lithosphärenplatten auch dem Bild entnommen werden, indem man den eingezeichneten Grenzen folgte und die größten Flächen einfach abzählte. Doch neben der Tatsache, dass dieses Vorgehen mit der Entscheidung verbunden war, ob es sich im Einzelfall tatsächlich um eine *große Platte* handelte, mussten die Lernenden erst einmal auf die Idee kommen, die Platten zu zählen. Im Text wurde die gleiche Information mit wenigen Worten und der Ziffer 9 repräsentiert, was einen Verarbeitungsvorteil darstellte.

Hinzu kommt, dass trotz der propagierten visuellen Wende (vgl. Mitchell, 1995), in einem schulischen und universitären Kontext das Leitmedium immer noch der Text ist. Da sich die Stichprobe aus Studierenden zusammensetzte, konnte angenommen werden, dass es sich bei den Teilnehmern um zumindest einigermaßen geübte Leser handelte. Ob dies auch auf die Verarbeitung von piktorialen Informationen zutrifft, lässt sich durchaus bezweifeln. Statische oder bewegte Bilder sind mittlerweile ein fester Bestandteil verschiedener Lehr-Lernmaterialien, doch ihre Nutzung will gelernt sein (Shah, Mayer & Hegarty, 1999; Mautone & Mayer, 2007). Zudem werden nach Ansicht von Weidenmann (2006) Bilder von Lernenden zwar gerne betrachtet und auf einer Textseite zuerst angeschaut, aber häufig auch unterschätzt. Die Lernenden verfügen oft gar nicht über die Erfahrung, um eine angemessenen Verarbeitungstiefe beim Bildverstehen anzustreben. Deshalb bedarf es unterstützender und anregender Texte, die den Lernenden beim Wissenserwerb mit dem Bild anleiten und die Entnahme der wesentlichen Informationen durch Hinweise fördern (Shah et al., 1999). Den Lernenden in den Bildgruppen standen jedoch nur die verbalen Informationen zur Verfügung, die in den grafischen Darstellungen selbst enthalten waren, wie Legenden oder Beschriftungen. Diese Textfragmente kennzeichneten zwar bedeutungsvolle Bildelemente, eine anleitende Funktion besaßen sie jedoch nicht. Fasst man letztlich alle möglichen Ursachen zusammen, ist es kein Wunder, dass die Teilnehmer mit den verbalen Informationen besser abschnitten.

Obwohl diese Ergebnisse recht interessant waren, bestand der eigentliche Zweck der zweiten Studie darin, zwei Hypothesen im Zusammenhang mit den Ursachen des PTS-Effekts zu prüfen. Gleich bei der ersten Testung zeigte sich, dass die Ergebnisse der statistischen Analysen sehr gut mit den Vorhersagen der *Interferenz-Hypothese* übereinstimmten. So hatte die Gegenüberstellung der Text-vor-Bild-Gruppen mit den nur-Text-Bedingungen einen bedeutsamen Unterschied hinsichtlich der mentalen Belastung beim Lernen ergeben. Da der Einfluss der Lernzeit herauspartialisiert worden war, konnte ausgeschlossen werden, dass die verstärkte Anstrengung allein durch die zusätzliche Verarbeitung der piktorialen Informationen verursacht wurde. Gleichzeitig waren die Leistungen im Wissenstest aber vergleichbar, ganz unabhängig davon, ob die Teilnehmer zusätzlich auf Bilder zurückgreifen konnten oder nicht. Diese Ergebnisse lassen sich mithilfe der *cognitive load theory* (Chandler & Sweller, 1991) recht eindeutig interpretieren. Demnach empfanden die Lernenden, denen nach dem Text ein Bild präsentiert wurde, eine zusätzliche mentale Belastung, die nicht durch die längere Lernzeit erklärt werden kann. Gleichzeitig führte die stärkere Anstrengung zu keiner besseren Leistung, was den Schluss nahe legt, dass die nachträgliche Präsentation des Bildes *extraneous load* (Sweller & Chandler, 1994) ausgelöst hat. Zieht man nun die Modellannahmen von Schnotz

(2002, 2005) hinzu, dann entsteht dieser *extraneous load*, wenn der Lernende versucht, sein mentales Modell mit dem externen Bild abzugleichen. Dementsprechend kann man anhand dieser Ergebnisse folgern, dass dem PTS-Effekt schließlich doch Interferenzen zugrunde liegen.

Diese Deutung hatte jedoch nicht zwangsläufig die Ablehnung der *Konstruktionshypothese* zur Folge. So war es weiterhin möglich, dass neben den negativen auch günstige Einflüsse für die Verarbeitungsvorteile der Bild-vor-Text-Gruppen verantwortlich waren. Durch die Präsentation der Bilder vor den Texten wurden Interferenzen offenbar vermieden, so dass sich die Frage stellte, ob diese Anordnung noch weitere Effekte nach sich zog. Trotz kontrollierter Lernzeit konnte allerdings hinsichtlich der mentalen Belastung kein bedeutsamer Unterschied für den Faktor *zusätzliches Format* festgestellt werden. Auch die anderen Effekte wurde nicht signifikant, selbst wenn die Prüfstatistik der Interaktion den kritischen Wert nur knapp verfehlt hatte. Möglicherweise ist dieses Ergebnis die Folge einer zufälligen Verzerrung anzusehen, die durch die Kovariaten nicht aufgefangen werden konnte. So hatten sich in der Gruppe, denen nur die komplementären Bilder präsentiert wurden, offenbar erschöpfte, für Belastungen anfällige oder auch engagierte Personen konzentriert.

Immerhin zeigte sich, dass trotz kontrollierter Lernzeiten die Bedingungen, die zusätzlich Texte erhalten hatten, deutlich besser im Wissenstest abschnitten. Der Vorteil, den die zusätzlichen Texte mit sich brachten, war jedoch kein rein additiver Effekt, sonst wäre eine bedeutsamer Zusammenhang zwischen der Lernzeit und der Gesamtleistung zu beobachten gewesen. Plausibler ist es, die Unterschiede auf die Qualität der mentalen Repräsentationen als Folge einer integrativen Verarbeitung zurückzuführen. Dennoch musste festgestellt werden, dass die bessere Leistung im Wissenstest nicht mit einer erhöhten Anstrengung in Verbindung zu bringen war. Die Präsentation des Textes nach dem Bild verursachte demnach keine zusätzliche mentale Belastung, die sich als *germane load* infolge einer tieferen Verarbeitung deuten ließ. Es mag zwar sein, dass der PTS-Effekt auch durch positive Effekte der Formatreihenfolge verursacht wird, doch ließen sich keine Anhaltspunkte für diese Annahme finden. Damit musste die *Konstruktions-Hypothese* zurückgewiesen werden.

Aber selbst wenn die Resultate gut zur *Interferenz-Hypothese* passten, musste man bei der Verallgemeinerung der Ergebnisse vorsichtig sein. Immerhin handelte es sich bei der entsprechenden Analyse um ein quasiexperimentelles Design, was trotz der Berücksichtigung von Kovariaten immer mit einer verminderten internen Validität verbunden ist. Hinzu kommt, dass im Fall von relativ kleinen Stichproben selbst die randomisierte Zuweisung der Versuchsteilnehmer zu unrepräsentativen Merkmalsverteilungen führen kann, die sich statistische Kontrollen nicht bereinigen lassen. Führt man diesen Gedankengang jedoch zu Ende, dann können auch die Ergebnisse, die zur Annahme der *Interferenz-Hypothese* geführt haben, das Resultat einer zufälligen Häufung bestimmter Merkmalsausprägungen sein. Die Gefahr, ein Artefakt produziert zu haben, ist umso größer, da die nachträgliche Selbsteinschätzung der mentalen Belastung nicht unumstritten ist (J. van Merriënboer et al., 2002; Schnotz & Kürschner, 2007). Außerdem bleibt aus theoretischen

7 Studie II: Interferenzen versus Konstruktion

schers Sicht der Zusammenhang zwischen Lernzeit und den verschiedenen load-Formen weitgehend unklar. In der ursprünglichen Konzeption der CLT spielt die Dauer der Lernprozesse praktisch keine Rolle, selbst wenn Ermüdungseffekte durchaus plausibel sind und die Wahrnehmung der mentalen Belastung beeinflussen dürften. In der Studie wurde die Lernzeit zwar herausgerechnet, doch ist der lineare Zusammenhang zwischen Bearbeitungsdauer und empfundener Anstrengung eine vereinfachende Annahme.

Aus diesem Grund war es kein Fehler, nach weiteren Anzeichen für die *Interferenz-Hypothese* zu suchen. Ob es diese Belege auch auf einer physiologischen Ebene gab, wurde in einer dritten und letzten Studie geprüft, die im folgenden Kapitel vorgestellt wird.

8 Studie III: Interferenzen & Blickbewegungen

Die Ergebnisse der zweiten Untersuchung stützten die Vermutung, dass Interferenzen den *picture-text-sequencing*-Effekt verursachten. Von besonderem Interesse waren in diesem Zusammenhang die kognitiven Prozesse, die bislang nur indirekt über den Lernerfolg oder die mentale Belastung erschlossen werden konnten. Eine Möglichkeit, die Verarbeitung der Texte und Bilder *online* zu verfolgen, bestand darin, die Blickbewegungen zu erfassen.

8.1 Fragestellung

Fasst man die Resultate der ersten und zweiten Studie zusammen, dann sprachen die Befunde in ihrer Tendenz für die *Interferenzhypothese*. Insbesondere der Vergleich zwischen den beiden Untersuchungen stützte diese Annahme. Allerdings handelte es sich bei den erfassten abhängigen Variablen um Merkmale, die zeitlich erst nach der Phase erhoben wurden, um die es eigentlich ging. So war es das Ziel sämtlicher Untersuchungen, Modellannahmen über kognitive Prozesse zu untersuchen, die während des Wissenserwerbs auftraten. Diese Prozesse ließen sich jedoch nicht direkt beobachten, weshalb über die Manipulation experimenteller Bedingungen mögliche Zusammenhänge inferiert werden mussten. Vereinfacht gesagt basierten damit sämtliche Interpretationen zum PTS-Effekt und seinen Ursachen auf indirekten Schlussfolgerungen. Daran ist natürlich grundsätzlich nichts auszusetzen, immerhin verdankt die experimentelle Psychologie diesem Vorgehen viele zentrale Ergebnisse. Dennoch ist die Frage nach einer direkteren Methode zur Beobachtung kognitiver Prozesse berechtigt, insbesondere wenn sich auf diese Weise bisherige Annahmen untermauern lassen.

Eine Möglichkeit, einer solch unmittelbaren Beobachtung mentaler Vorgänge besteht darin, die Blickbewegungen während der Informationsaufnahme zu erfassen. Grundlage hierfür sind kurze, ruckartige Bewegungen des okulomotorischen Apparates, die mehrere Male in der Sekunde stattfinden und als *Sakkaden* bezeichnet werden. Allgemein vermutet man, dass während dieser Sprünge keine Informationen aufgenommen werden, weshalb das Interesse der Forscher insbesondere den Pausen zwischen den Sakkaden gilt. Hier werden bestimmte Areale des Gesichtsfeldes, einzelne Worte oder Buchstaben bis zu 600 ms lang fixiert, so dass in diesem kurzen Zeitraum das Auge mehr oder weniger still hält. Viele Wissenschaftler gehen davon aus, dass während einer solchen *Fixation*, die eigentliche Informationsaufnahme erfolgt.

Orientiert man sich an Just und Carpenter (1980), dann lassen sich über Blickbewegungen mentale Vorgänge um einiges unmittelbarer messen, als dies durch Leistungstests möglich wäre. Die beiden Autoren untersuchten den Zusammenhang zwischen

Aufmerksamkeit, Blickrichtung und Dauer beim Lesen, was sie zu der Feststellung veranlasste, dass ein Wort so lange fixiert wird, bis es kognitiv verarbeitet ist. Mit ihrer *eye mind assumption* setzen Just und Carpenter (1980) im Grunde Wahrnehmungs- und mentale Prozesse gleich, wofür sie auch empirische Belege finden konnten. So stieg die Fixationsdauer bei selten verwendeten Worten oder an Stellen, die eine Integration verschiedener Satzabschnitte erforderten, deutlich an. Einigen Autoren geht die ursprüngliche *eye mind assumption* hingegen zu weit, da auch bereits gelesene Worte die Betrachtungsdauer beeinflussen und somit den Zusammenhang zwischen Fixation und Kognition zumindest einschränken (Underwood & Everatt, 1992). Doch trotz vereinzelter Einwände belegen zahlreiche Studien den unmittelbaren Zusammenhang zwischen Blickbewegungen und aktuellen kognitiven Prozessen (Rayner, 1998). Kompetente Leser, die gute Textverständnisse erzielen, machen demnach weniger Fixationen, die kürzer und durch lange Sakkaden verbunden sind, wie beispielsweise Everatt und Underwood (1994) zeigen konnten. Dagegen benötigen schlechte Leser mehr Fixationen, die zudem länger und durch vergleichsweise kurze Sakkaden verbunden sind (Eden, Stein, Wood & Wood, 1994).

Bei solch eindeutigen Befunden zum Textverstehen lag es nahe, die Methode der Blickbewegungsmessung auch bei der Verarbeitung von Bildern einzusetzen. Allerdings galt es zu bedenken, dass sich Erkenntnisse aus der Leseforschung nicht ohne weiteres auf das Verstehen piktorialer Inhalte übertragen lassen. Nach Ansicht von Ballstaedt (2005) gehorchen die Augen beim Lesen von Texten anderen Regeln als beim Betrachten von Bildern. Während die visuelle Verarbeitung verbaler Informationen von links nach rechts, Wort für Wort und Zeile für Zeile erfolgt, verfügt ein Lernender bei der Konfrontation mit einem Bild über sehr viel mehr Freiheitsgrade. So bleibt es letztlich ihm überlassen, welches Objekt bzw. Areal er zuerst betrachtet und auch die Reihenfolge, in der die piktorialen Informationen aufgenommen werden sollen, ist in den meisten Fällen nicht vorgegeben. Entweder wird der Blick durch visuelle Merkmale des Bildes mehr oder weniger angezogen, oder durch das Interesse bzw. die Zielsetzung des Betrachters gesteuert.

Gemessen an der Anzahl von Studien zum Leseverstehen, gibt es sehr viel weniger blickbewegungsbasierte Untersuchungen zur Verarbeitung piktorialer Informationen (Rayner, 1998). Die variable Reihenfolge, in der Bildelemente betrachtet werden können, und die oft komplexe Abgrenzung von Bedeutungseinheiten erschweren die Gleichsetzung von Perzeption und Kognition bei der Verarbeitung von bildlichen Darstellungen. Dennoch lassen sich auch hier Zusammenhänge zwischen Wahrnehmungs- und Verarbeitungsprozessen feststellen, selbst wenn sie etwas unschärfer sind. So konnten Christianson, Loftus, Hoffman und Loftus (1991) nachweisen, dass die Anzahl von Fixationen ein guter Prädiktor für Gedächtnisleistungen ist, bei denen das Bild oder einzelne Elemente wiedererkannt werden müssen. Je häufiger der Lernende ein Areal oder Bildausschnitt fixiert hat, um so wahrscheinlicher ist es, dass die entsprechende Information auch memoriert wurde. Diese Feststellung ist zwar etwas banal, aber sie verdeutlicht, dass zumindest einfache Zusammenhänge zwischen Blickbewegungsdaten und kognitiven Prozessen bestehen. Ähnlich wie Merkmale von Worten und Texten das Lesever-

halten beeinflussen, wirken sich auch die Eigenschaften von Bildern auf Sakkadenlänge und Fixationsdauer aus (Rayner & Pollatsek, 1992). Enthält die grafische Darstellung beispielweise viele Informationen, die dicht nebeneinander angeordnet sind, dann steigt die durchschnittliche Fixationsdauer, während die Sakkadenlänge abnimmt. Bei Bildern mit einer eher weitflächigen Informationsverteilung sinkt dagegen die Fixationsdauer und die Sakkadenlänge nimmt zu.

Vor diesem Hintergrund war es durchaus plausibel, dass die Präsentationsreihenfolge und das Informationsverhältnis, wie sie in den Studien zuvor untersucht wurden, das Blickbewegungsverhalten beeinflussen würden. Das Forschungsinteresse richtete sich dabei in erster Linie auf die Verarbeitung der Bilder, da sich hier theoretisch die entscheidenden Prozesse abspielten. Allerdings musste damit gerechnet werden, dass die erfassten Parameter zum Teil stark schwankten (Rayner & Pollatsek, 1992). So gibt es insbesondere bei der Betrachtung von Bildern hinsichtlich der Fixationsdauer und Sakkadenlänge große inter- und intraindividuelle Unterschiede, die sich nicht immer eindeutig erklären lassen. Trotz dieser Einschränkung sollte der Versuch unternommen werden, Hinweise in den Blickbewegungsdaten zu finden, die den Erklärungsansatz der *Interferenzhypothese* stützten. Da jedoch kaum vorherzusehen war, wie sich die einzelnen Parameter verändern würden, schien es das Beste zu sein, auf konkrete Hypothesen zu verzichten und die Forschungsfrage explorativ zu formulieren:

Welchen Einfluss besitzen die Präsentationsreihenfolge und das Informationsverhältnis von Texten und Bildern auf das Blickbewegungsverhalten bei der Verarbeitung piktorialer Informationen?

8.2 Versuchsplanung

8.2.1 Design

Im Grunde stellte die dritte Untersuchung eine Replikation der ersten Studie dar. Die Versuchsteilnehmer bearbeiteten die gleichen Tests und erhielten die gleichen Lernmaterialien. Der entscheidende Unterschied bestand jedoch darin, dass während der Präsentation der Texte und Bilder die Blickbewegungen der Versuchsteilnehmer erfasst wurden. Das erforderte jedoch Einzelsitzungen, die für den Versuchsleiter wie auch die Probanden mit einem hohen zeitlichen Aufwand verbunden waren. Aus diesem Grund wurde der Stichprobenumfang gegenüber der ersten Studie deutlich reduziert, was zu einem 2x2 Design mit folgender Zellenbesetzung führte.

Abbildung 8.1 zeigt, dass insgesamt 24 Personen teilnahmen, die sich gleichmäßig auf die vier experimentellen Bedingungen verteilten. Der ziemlich geringe Stichprobenumfang brachte es jedoch mit sich, dass kleine ($1 - \beta = 0.05$) und mittlere Effekte ($1 - \beta = 0.23$), selbst wenn sie vorhanden waren, sehr wahrscheinlich unentdeckt blieben. Um bei einem α -Fehler eine Teststärke von .80 zu erreichen, musste der statistische Zusammenhang mindestens $\eta^2 = .275$ betragen, damit ein Unterschied signifikant wurde. Als abhängige

Informationsverhältnis	Reihenfolge	
	Text vor Bild (TvB)	Bild vor Text (BvT)
kongruent (<i>kon</i>)	kon_TvB (n= 6)	kon_BvT (n=6)
komplementär (<i>komp</i>)	komp_TvB (n=6)	komp_BvT (n=6)

Abbildung 8.1: Untersuchungsdesign der dritten Studie

Variablen, die im Mittelpunkt Analyse stehen sollten, kamen eine ganze Reihe von Kennwerten infrage. Der Einfachheit halber wurden mit der *Anzahl der Fixationen*, der *Fixationsdauer* und der *mittleren Sakkadenamplitude* letztlich drei Parameter gewählt, die in den meisten vergleichbaren Studien erhoben werden und sich relativ leicht interpretieren lassen.

8.2.2 Stichprobe

Die dritte Untersuchung fand in einem Raum der Universität Koblenz-Landau / Campus Landau statt, der speziell für Experimente mit Blickbewegungsmessungen eingerichtet worden war. Da es sich als ziemlich schwierig erwiesen hatte, geeignete Versuchspersonen zu finden, erstreckte sich die Rekrutierung der Teilnehmer über einen Zeitraum von 5 Wochen. Um letztlich alle Zellen gleichmäßig auffüllen zu können, wurden zusätzlich Schüler der 12ten Klasse eines Gymnasiums in Landau angeworben. Aus diesem Grund nahmen neben 19 Studierenden 5 Oberstufenschüler an der Untersuchung teil, die im folgenden Schuljahr Abitur machen wollten. Von den studentischen Versuchspersonen studierte elf eine Kombination aus Lehramtsfächern, drei studierten Pädagogik und zwei waren in Psychologie eingeschrieben, während die Übrigen der Kategorie Sonstiges zugeordnet wurden. Die Mehrheit ($\approx 63\%$) befand sich mit vier oder weniger Semestern noch im Grundstudium, während 36% weiter fortgeschritten waren. Aufgrund der teilnehmenden Schüler lag der Altersdurchschnitt mit 21 Jahren $M = 20.7, SD = 2.9$ etwas unter dem der ersten beiden Studien. Hinsichtlich des Geschlechts dominierten erneut eindeutig die Frauen (87,5%), da insgesamt nur 3 männliche Personen teilgenommen hatten. Nach Abschluss der Untersuchung erhielt jeder Schüler bzw. Student 10 € als Aufwandsentschädigung, während die Möglichkeit, Versuchspersonenstunden gut geschrieben zu bekommen, kein einzige Mal genutzt wurde.

8.2.3 Durchführung

Die dritte Untersuchung war im Grunde eine Wiederholung der ersten Studie, auch wenn es einzelne *Besonderheiten* gab. So erfolgte die Durchführung in einem speziellen Raum,

in dem die Apparatur zur Erfassung der Blickbewegungsmessungen stand. Alle Instruktionen, die verwendeten Instrumente und die Reihenfolge der Vortests waren jedoch identisch mit dem Procedere der ersten Untersuchung. Ebenso wurden die selben Texte und Bilder verwendet, doch erforderte das verwendete Gerät (*Eyelink II*) der Firma *SR Research* eine Kalibrierungsphase, bevor die eigentliche Präsentation des Stimulusmaterials beginnen konnte. Die Messung der Blickbewegungen erfolgte über zwei Infrarotkameras, die an einem *head set* befestigt waren, das einer Art Helm glich und etwa 420g wog. Dieses *head set* wurde jedem Teilnehmer vorsichtig auf den Kopf gesetzt und über zwei Gurte grob angepasst. Die beiden Kameras waren über eine Halterung von Oben schräg auf die Pupillen gerichtet, so dass sich über zusätzliche Marker am Helm und dem Präsentationsbildschirm Blickrichtung und Fixationspunkt berechnen ließen. Nach Angaben des Herstellers war es auf diese Weise möglich, die Blickbewegung mit einer Abweichung 0.5° zu erfassen, was es jedoch erforderlich machte, den Kopf möglichst still zu halten. Daher wurden die Versuchspersonen, die in ca. 80 cm Entfernung zum Bildschirm saßen, instruiert sich möglichst wenig zu bewegen.

Individuelle Unterschiede, die Kopfform, Größe oder Merkmale der Augen (z.B. Hornhautkrümmung, Pupillenkontrast, etc.) betrafen, verlangten eine individuelle Feinkalibrierung der Kameras, deren Güte in verschiedenen Validierungsdurchgängen geprüft wurde. Obwohl die Aufzeichnung immer anhand eines Auges erfolgte und daher bei Problemen auf das zweite Auge ausgewichen werden konnten, kam es in insgesamt vier Fällen vor, dass keine ausreichende Messgenauigkeit erreicht wurde. Dem Versuchsleiter blieb dann nichts anderes übrig, als den Versuch abubrechen und die betreffenden Teilnehmer ohne Aufwandsentschädigung zu entlassen. Verließ die Kalibrierung hingegen zufriedenstellend, konnte die Präsentation der Texte und Bilder beginnen. Anstelle einer Maus erhielten die Schüler bzw. Studierenden ein *Gamepad*. Wollten die Teilnehmer mit dem nächsten Bild oder Text fortfahren, dann mussten sie eine bestimmte Taste drücken, was ihnen aber vom Versuchsleiter zuvor erklärt worden war.

Das ungewöhnliche Eingabegerät wurde notwendig, da die Präsentation der Stimulusmaterials nicht über *normale* HTML-Seiten erfolgen konnte. Die Aufzeichnung der Blickbewegungsdaten sowie der gesamte Versuch wurden über die Software *Eyelink II* des Herstellers *SR Research* gesteuert, die ein solches Format nicht erkannte. Daher mussten die Bild- und Textseiten im Vorfeld der Untersuchung zuerst in JPEG-Dateien umgewandelt werden, damit sie sich für den Versuch verwenden ließen. Die Darstellung des Lehrmaterials erfolgte dann über einen Display-Rechner, wobei während der Skalierung darauf geachtet wurde, möglichst nah an den ursprünglichen Formaten zu bleiben. Zu der Versuchsanordnung zählte zudem ein Host-Computer, der sowohl mit der Display-Rechner als auch mit dem *head set* verbunden war und die Blickbewegungsdaten mit einer Frequenz von 250 Hz aufzeichnete. Nach Abschluss der Lernphase wurde den Teilnehmern der Helm abgenommen und die selben Instrumente vorgelegt, die auch in den ersten Studie verwendet worden waren.

8.2.4 Auswertung

Die umfangreichen Blickbewegungsdaten erforderten zunächst eine Bearbeitung im Data Viewer der *Eyelink II*-Software, bevor sie für statistische Analysen verwendet werden konnten. Beispielsweise erfolgte die Festlegung der Mindestdauer für eine gültige Fixation auf 80 ms, so dass kürzere Fixationen nicht gezählt und von allen weiteren Berechnungen ausgeschlossen wurden. Anschließend mussten die Daten ausgelesen und transformiert werden, damit eine Auswertung mit SPSS 17.0 überhaupt möglich war. Zur Analyse der Daten wurden vor allem MANOVAs, ANOVAs und einfache Produkt-Moment-Korrelationen berechnet. Ob die abhängigen Variablen die Voraussetzungen für das jeweilige Verfahren erfüllten, sollte erneut mit dem Levene-Test, dem Kolmogorow-Smirnow-Anpassungstest und dem Box-Test geprüft werden. Aufgrund des geringen Stichprobenumfangs wäre der Verlust von Freiheitsgraden, den die Berücksichtigung von Kontrollvariablen zur Folge gehabt hätte, sehr viel stärker ins Gewicht gefallen, als bei den ersten beiden Studien. Zudem war nicht sicher, ob es einen Zusammenhang zwischen den Blickbewegungsparametern und den erhobenen Kovariaten gab, weshalb letztlich auf eine statistische Kontrolle potentieller Störgrößen verzichtet wurde. Allen Analysen, die im Rahmen der explorativen Fragestellung erfolgten, wurde ein Signifikanzniveau von $\alpha = .05$ bei zweiseitiger Testung zugrunde gelegt. Die folgende Ergebnisdarstellung umfasst für jede Berechnung neben den Prüfstatistiken und Freiheitsgraden auch $\bar{\eta}^2$, das der Beurteilung der praktischen Bedeutsamkeit dienen sollte.

8.3 Ergebnisse

Da auf die Berücksichtigung von Kovariaten verzichtet wurde, beginnt die Ergebnisdarstellung nicht mit der Äquivalenztestungen zur Auswahl potentieller Störgrößen. Stattdessen werden zunächst die Resultate berichtet, die sich auf die Leistungen der Teilnehmer im Wissenstest, der mentalen Belastung beim Lernen und die Betrachtungszeit bezogen. Der zweite Abschnitt enthält die Ergebnisse zu den Blickbewegungsmessungen, wobei ausschließlich die Parameter analysiert wurden, deren Aufzeichnung bei der Präsentation der Bildseiten erfolgte.

8.3.1 Wissenstest, mentale Belastung und Lernzeit

Wie in der ersten Studie wurde eine MANOVA zur Prüfung von Gruppenunterschieden hinsichtlich der Lernleistung berechnet. Dabei waren die vier Skalen des Wissenstests die abhängigen Variablen, während die *Repräsentationsreihenfolge* und das *Informationsverhältnis* die Faktoren des 2x2 Designs darstellten. Vor der eigentlichen Analyse wurden zunächst die Voraussetzungen geprüft, unter denen sich die Ergebnisse einer MANOVA interpretieren lassen. Tabelle 8.1 zeigt, dass die Annahme homogener Fehlervarianzen und univariater Normalverteilung als erfüllt angesehen werden konnten.

Dagegen wies das signifikante Ergebnis des Tests von Box (1949) ($M_{Box} = 81.94$, $F =$

Tabelle 8.1: Kolmogorow-Smirnow- und Levene-Tests für die Skalen des Lerntests in der dritten Untersuchung

Skala	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
Wiedererkennen	<i>Levene – Test</i>	$F = 0.49$	$df_1 = 3; df_2 = 20$	$p = .692$
	<i>K – S – Test</i>	$Z_{K-S} = 0.54$	–	$p = .929$
freies Erinnern	<i>Levene – Test</i>	$F = 0.42$	$df_1 = 3; df_2 = 20$	$p = .743$
	<i>K – S – Test</i>	$Z_{K-S} = 0.90$	–	$p = .392$
InferenzMC	<i>Levene – Test</i>	$F = 1.97$	$df_1 = 3; df_2 = 20$	$p = .151$
	<i>K – S – Test</i>	$Z_{K-S} = 0.817$	–	$p = .516$
Inferenz offen	<i>Levene – Test</i>	$F = 1.06$	$df_1 = 3; df_2 = 20$	$p = .388$
	<i>K – S – Test</i>	$Z_{K-S} = 0.889$	–	$p = .408$

Anm.: *K-S-Test = Kolmogorow-Smirnow-Anpassungstest; signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen (* < .05) bzw. durch zwei Sternchen (** < .01) hervorgehoben*

1.67, $df_1 = 30$, $df_2 = 1099.76$, $p = .013$) auf ungleiche Kovarianzmatritzen und/oder eine Verletzung der multivariaten Normalverteilungsannahme hin. Field (2009) geht jedoch davon aus, dass bei gleicher Zellenbesetzung die Teststatistik *Pillai's trace* (V) recht robust gegenüber heterogenen Kovarianzmatritzen ist. Da der Box-Test nach Meinung von Bortz (1999) zudem sehr sensitiv reagiert, erschien es vertretbar zu sein, die Prüfgrößen der MANOVA dennoch zu interpretieren. Allerdings gab es nicht viel zu deuten, da weder die Faktoren *Reihenfolge* ($V = .238$, $F_{(4;17)} = 1.326$, $p = .301$, $\bar{\eta}^2 = .238$) und *Informationsverhältnis* ($V = .220$, $F_{(4;17)} = 1.202$, $p = .346$, $\bar{\eta}^2 = .220$), noch deren Interaktion ($V = .290$, $F_{(4;17)} = 1.739$, $p = .188$, $\bar{\eta}^2 = .290$) einen bedeutsamen Zusammenhang anzeigten.

Zur Analyse der mentalen Belastung wurde lediglich eine ANOVA mit dem Item berechnet, das die empfundene Belastung beim Lernen erfassen sollte. Bei der Prüfung der Voraussetzungen zeigte der Levene-Test ($F = 1.28$, $df_1 = 3$; $df_2 = 20$ $p = .310$) kein signifikantes Ergebnis, doch musste die Normalverteilungsannahme verworfen werden ($Z_{K-S} = 1.625$, $p = .010$), was bei der gleichmäßigen Besetzung der Zellen jedoch unproblematisch war. Ähnlich wie bei der ersten Studie konnten auch hier keine bedeutsamen Zusammenhänge zwischen der abhängigen Variablen und den experimentellen Bedingungen ermittelt werden. Die *Reihenfolge* ($F_{(1;20)} = 0.303$, $p = .588$, $\bar{\eta}^2 = .015$), das *Informationsverhältnis* ($F_{(1;20)} = 0.217$, $p = .646$, $\bar{\eta}^2 = .011$) und deren Wechselwirkung ($F_{(1;20)} = 1.212$, $p = .284$, $\bar{\eta}^2 = .057$) zeigten keine systematische Gruppenunterschiede an.

Die statistische Analyse der Lernzeit folgte dem selben Vorgehen, wie bei der ersten Studie. So wurden Betrachtungszeiten der einzelnen Bild- und Textseiten formatspezifisch aufaddiert und z-transformiert. Die beiden *neuen* Variablen wurden dann wie zwei unterschiedliche Messzeitpunkte behandelt, so dass ein Design mit 2x2x2 Faktoren gerechnet werden konnte. Vor der eigentlichen Durchführung erfolgte zunächst die Prüfung

8 Studie III: Interferenzen & Blickbewegungen

der Voraussetzungen, die Tabelle 8.2 zusammenfasst.

Tabelle 8.2: Kolmogorow-Smirnow- und Levene-Tests für die Lernzeiten in der dritten Untersuchung

	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Gesamtlernzeit Text</i>	<i>Levene – Test</i>	$F = 1.79$	$df_1 = 3; df_2 = 20$	$p = .182$
	<i>K – S – Test</i>	$Z_{K-S} = .85$	–	$p = .468$
<i>Gesamtlernzeit Bild</i>	<i>Levene – Test</i>	$F = 1.45$	$df_1 = 3; df_2 = 20$	$p = .258$
	<i>K – S – Test</i>	$Z_{K-S} = 1.24$	–	$p = .092$

Anm.: K-S-Test = Kolmogorow-Smirnow-Anpassungstest; signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen ($* < .05$) bzw. durch zwei Sternchen ($** < .01$) hervorgehoben

Wie man erkennen kann, ließen sich die zentralen Annahmen als erfüllt ansehen. Da auch der Box-Test ($M_{Box} = 10.87$, $F = 0.99$, $df_1 = 9$, $df_2 = 4583.9.76$, $p = .448$) den kritischen Wert verfehlt hatte, war nicht damit zu rechnen, dass die Varianzanalyse zu verzerrten Ergebnissen führen würde. Es zeigte sich jedoch, dass keiner der drei möglichen Haupteffekte auf einen bedeutsamen Zusammenhang hinwies. Weder die Messwiederholung ($F_{(1;20)} = 0.000$, $p = .000$, $\bar{\eta}^2 = .000$) noch die *Reihenfolge* ($F_{(1;20)} = 0.043$, $p = .838$, $\bar{\eta}^2 = .002$) oder das *Informationsverhältnis* ($F_{(1;20)} = 2.752$, $p = .113$, $\bar{\eta}^2 = .121$) deuteten auf überzufällige Mittelwertsunterschiede hin. Auch von den Interaktionen erster Ordnung wurden die Wechselwirkungen zwischen der *Reihenfolge* und dem *Informationsverhältnis* ($F_{(1;20)} = 0.058$, $p = .813$, $\bar{\eta}^2 = .003$) sowie zwischen dem Messwiederholungsfaktor und dem *Informationsverhältnis* ($F_{(1;20)} = 0.119$, $p = .734$, $\bar{\eta}^2 = .006$) nicht signifikant. Allerdings überschritt die Prüfgröße für die Interaktion zwischen der *Reihenfolge* und dem Messwiederholungsfaktor ($F_{(1;20)} = 6.818$, $p = .017$, $\bar{\eta}^2 = .254$) den kritischen F-Wert und auch die Interaktion zweiter Ordnung ($F_{(1;20)} = 4.382$, $p = .049$, $\bar{\eta}^2 = .180$) wies auf einen bedeutsamen Effekt hin.

Die Bedeutung dieses Ergebnisses wird etwas klarer, wenn man die Analyse für die komplementären und kongruenten Bedingungen getrennt durchführt. So zeigt sich beim Vergleich der standardisierten Lesezeiten der komplementären Gruppen zunächst kein signifikanter Interaktionseffekt ($F_{(1;10)} = 0.218$, $p = .651$, $\bar{\eta}^2 = .021$) zwischen dem Messwiederholungsfaktor und der *Reihenfolge*. Diese Wechselwirkung wird umso deutlicher unter der kongruenten Bedingung ($F_{(1;10)} = 8.255$, $p = .017$, $\bar{\eta}^2 = .452$), wie auch Abbildung 8.2 veranschaulicht.

Es ist gut zu erkennen, dass sich die Teilnehmer der kongruenten Gruppen mit dem Format länger auseinander gesetzt haben, das Ihnen zuerst präsentiert wurde. Da dieser Effekt nur unter einer Bedingung des Faktors Informationsverhältnis beobachtet werden konnte, ist es nicht verwunderlich, dass auch die Prüfgröße für die Interaktion zweiter Ordnung den kritischen Wert überschritt.

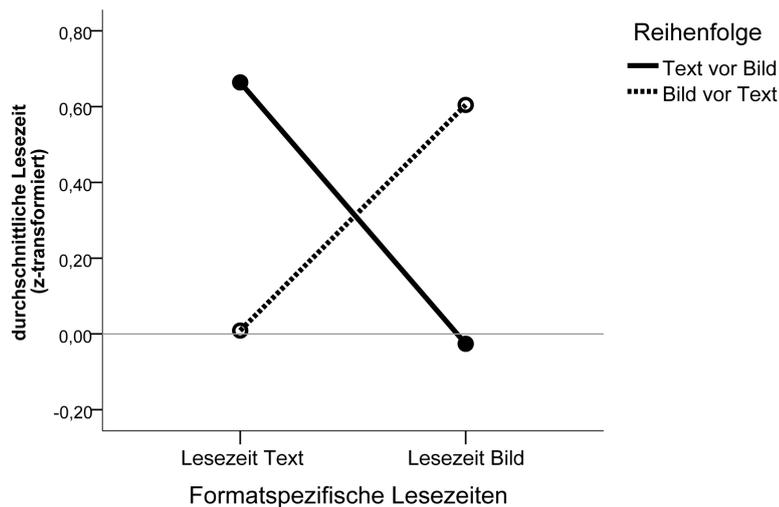


Abbildung 8.2: Interaktion zwischen formatspezifischer Lesezeit und Reihenfolge unter kongruenter Bedingung

8.3.2 Blickbewegungen

Die Darstellung der Analyseergebnisse zu den Augenbewegungen konzentrierte sich ausschließlich auf Daten, die während der Betrachtung der Bildseiten erhoben wurden. Dem lag die Überlegung zugrunde, dass, wenn es tatsächlich Hinweise auf Interferenzen gab, diese vor allem bei der Betrachtung der Bilder auftraten. Was auf den Textseiten geschah, war somit nur von nachgeordnetem Interesse, zumal die Darbietung der verbalen Informationen auf einer Seite, die Worte vergleichsweise *klein* erscheinen ließ. Eine Präsentation in einer größeren Schriftgröße war nicht vorgesehen, da unbedingt vermieden werden sollte, dass die Teilnehmer *scrollen* mussten. Letztlich hatten Studierende wie Schüler zwar keinerlei Probleme beim Lesen, doch erwiesen sich die Messungen über die Kameras am Helm als zu ungenau, um beispielsweise bestimmte Fixationen stets zuverlässig einzelnen Worten zuordnen zu können.

Der Umstand, dass die Parameter für jede der Bildseite erhoben worden waren, machte es möglich, für die die drei Blickbewegungsparameter jeweils eine MANOVA durchzuführen. Auf die Weise gingen keine Informationen verloren, was etwa bei einer Aggregation der vier Messwerte zu einem Gesamtmittelwert der Fall gewesen wäre. Außerdem konnte man davon ausgehen, dass verglichen mit Einzeltestungen der β -Fehler kleiner sein würde, da nicht mit einer Inflation des α -Fehlers gerechnet werden musste. Der erste Blickbewegungsparameter, der einer Analyse unterzogen wurde, war die *Fixationsanzahl*. Natürlich mussten auch hier zuerst die Voraussetzungen und Annahmen geprüft werden, die eine Interpretation der Ergebnisse erst sinnvoll machte. Tabelle 8.3 gibt die Ergebnisse dieser Prüfung wieder.

Wie unschwer zu erkennen ist, konnten für alle vier Messwerte die Normalverteilungsannahme beibehalten werden. Ebenso erwiesen sich die Fehlervarianzen in den Gruppen als weitgehend homogen, da nicht einer der Levene-Tests zu einem signifikanten Ergeb-

Tabelle 8.3: Kolmogorow-Smirnow- und Levene-Tests für die Anzahl valider Fixationen

	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Bild I : Die Theorie der Plattentektonik</i>	<i>Levene – Test</i>	$F = 1.25$	$df_1 = 3; df_2 = 20$	$p = .318$
	<i>K – S – Test</i>	$Z_{K-S} = 0.89$	–	$p = .402$
<i>Bild II : Die Bedeutung der Plattengrenzen</i>	<i>Levene – Test</i>	$F = 2.17$	$df_1 = 3; df_2 = 20$	$p = .123$
	<i>K – S – Test</i>	$Z_{K-S} = 0.54$	–	$p = .931$
<i>Bild III : Messungen im Erdinneren</i>	<i>Levene – Test</i>	$F = 2.16$	$df_1 = 3; df_2 = 20$	$p = .124$
	<i>K – S – Test</i>	$Z_{K-S} = 1.02$	–	$p = .246$
<i>Bild IV : Mechanismen der Plattenbewegung</i>	<i>Levene – Test</i>	$F = 1.10$	$df_1 = 3; df_2 = 20$	$p = .372$
	<i>K – S – Test</i>	$Z_{K-S} = 0.72$	–	$p = .674$

Anm.: K-S-Test = Kolmogorow-Smirnow-Anpassungstest; signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen (* < .05) bzw. durch zwei Sternchen (** < .01) hervorgehoben

nis führte. Aus diesem Grund stellte es kein besonderes Problem dar, dass der Box-Test ($M_{Box} = 74.06$, $F = 1.52$, $df_1 = 30$, $df_2 = 1099.8$, $p = .037$) auf eine Verletzung der multivariaten Normalverteilungannahme bzw. ungleiche Kovarianzmatrizen hinwies. Als Teststatistik wurde wie zuvor *Pillai's trace V* verwendet, das bei gleichmäßig besetzten Zellen sehr robust ist. Allerdings zeigte sich, dass weder die *Reihenfolge* ($V = .189$, $F_{(4;17)} = 0.988$, $p = .441$, $\bar{\eta}^2 = .189$) noch das *Informationsverhältnis* ($V = .286$, $F_{(4;17)} = 1.702$, $p = .196$, $\bar{\eta}^2 = .286$) einen bedeutsamen Einfluss auf die Anzahl der Fixationen pro Bild hatte. Auch die Interaktion der beiden Faktoren ($V = .071$, $F_{(4;17)} = 0.325$, $p = .858$, $\bar{\eta}^2 = .071$) wurde nicht signifikant, weshalb potentielle Wechselwirkungen ausgeschlossen werden konnten.

Nachdem die Anzahl der Fixationen keinen Zusammenhang ergeben hatte, konzentrierte sich die zweite Analyse auf deren Dauer. Zunächst wurden jedoch wieder die Voraussetzungen geprüft, wie Tabelle 8.4 zeigt.

Erneut konnte von erfüllten Voraussetzungen ausgegangen werden, da keiner der Tests zu einem signifikanten Ergebnis geführt hatte. Allerdings überschritt die Prüfstatistik des Box-Tests ($M_{Box} = 84.5$, $F = 1.73$, $df_1 = 30$, $df_2 = 1099.8$, $p = .009$) erneut den kritischen Wert, was bei der Interpretation der MANOVA-Ergebnisse jedoch vernachlässigt werden konnte. Diese ergaben, dass der Faktor *Reihenfolge* ($V = .115$, $F_{(4;17)} = 0.554$, $p = .699$, $\bar{\eta}^2 = .115$) keinen bedeutsamen Einfluss auf die vier Messungen der Fixationsdauer besaß und auch der Interaktionseffekt verfehlte das kritische Niveau klar ($V = .081$, $F_{(4;17)} = 0.375$, $p = .823$, $\bar{\eta}^2 = .081$). Dagegen deutete das signifikante Ergebnis ($V = .484$, $F_{(4;17)} = 3.993$, $p = .018$, $\bar{\eta}^2 = .484$) für den Faktor *Informationsverhältnis* auf einen Unterschied zwischen den Gruppen mit komplementären oder kongruenten Texten und Bildern hin. Die Richtung des Zusammenhangs wird deutlich, wenn man die Zellenmittelwerte betrachtet, die Tabelle 8.5 zeigt.

So ist gut zu erkennen, dass die mittlere Fixationsdauer unter der komplementären Be-

Tabelle 8.4: Kolmogorow-Smirnow- und Levene-Tests für die durchschnittliche Fixationsdauer

	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Bild I : Die Theorie der Plattentektonik</i>	<i>Levene – Test</i>	$F = 0.25$	$df_1 = 3; df_2 = 20$	$p = .864$
	<i>K – S – Test</i>	$Z_{K-S} = 0.52$	–	$p = .952$
<i>Bild II : Die Bedeutung der Plattengrenzen</i>	<i>Levene – Test</i>	$F = 0.77$	$df_1 = 3; df_2 = 20$	$p = .523$
	<i>K – S – Test</i>	$Z_{K-S} = 1.05$	–	$p = .217$
<i>Bild III : Messungen im Erdinneren</i>	<i>Levene – Test</i>	$F = 2.69$	$df_1 = 3; df_2 = 20$	$p = .074$
	<i>K – S – Test</i>	$Z_{K-S} = 1.07$	–	$p = .206$
<i>Bild IV : Mechanismen der Plattenbewegung</i>	<i>Levene – Test</i>	$F = 1.79$	$df_1 = 3; df_2 = 20$	$p = .181$
	<i>K – S – Test</i>	$Z_{K-S} = 0.91$	–	$p = .379$

Anm.: K-S-Test = Kolmogorow-Smirnow-Anpassungstest; signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen ($* < .05$) bzw. durch zwei Sternchen ($** < .01$) hervorgehoben

Tabelle 8.5: Mittelwerte und Standardabweichung der Fixationsdauer in Millisekunden

	Versuchsbedingungen			
	Text vor Bild kongruent	Text vor Bild komplementär	Bild vor Text kongruent	Bild vor Text komplementär
<i>Bild I : Die Theorie der Plattentektonik</i>	$M = 288.1$ $SD = 48.9$	$M = 278.1$ $SD = 33.8$	$M = 307.56$ $SD = 45.7$	$M = 293.8$ $SD = 50.7$
<i>Bild II : Die Bedeutung der Plattengrenzen</i>	$M = 299.1$ $SD = 48.2$	$M = 285.0$ $SD = 41.2$	$M = 306.5$ $SD = 69.5$	$M = 296.5$ $SD = 34.4$
<i>Bild III : Messungen im Erdinneren</i>	$M = 287.3$ $SD = 54.1$	$M = 252.7$ $SD = 40.3$	$M = 311.6$ $SD = 110.9$	$M = 255.0$ $SD = 29.3$
<i>Bild IV : Mechanismen der Plattenbewegung</i>	$M = 277.3$ $SD = 36.3$	$M = 231.7$ $SD = 38.9$	$M = 312.4$ $SD = 96.1$	$M = 235.6$ $SD = 28.9$

Anm.: M = Mittelwert; SD = Standardabweichung

dingung durchgehend niedriger ausfiel. Obwohl dies für die Messwerte aller vier Bilder zutraf, war der Unterschied nur bei einer Messung signifikant, wie nachträglich berechnete univariate Testungen ergaben. Die Prüfstatistiken für Bild I ($F_{(1;20)} = 0.424$, $p = .522$, $\bar{\eta}^2 = .021$), Bild II ($F_{(1;20)} = 0.350$, $p = .561$, $\bar{\eta}^2 = .017$) und Bild III ($F_{(1;20)} = 2.936$, $p = .102$, $\bar{\eta}^2 = .128$) verfehlten alle den kritischen Wert, selbst wenn es in einem Fall relativ knapp war. Lediglich bei Bild IV ($F_{(1;20)} = 6.528$, $p = .019$, $\bar{\eta}^2 = .246$) war der Unterschied groß genug, um das α -Fehlerniveau von 5% zu unterschreiten. Beim Blick auf Tabelle 8.5 fällt zudem auf, dass die Standardabweichungen der einzelnen Gruppen-Messwerte erheblich schwankte. Insbesondere bei der Bild-vor-Text-Gruppe mit kongruentem Material variierte die mittlere Fixationsdauer offensichtlich stark, was sehr wahrscheinlich auf den Einfluss eines Ausreißers zurückzuführen ist, wie die Analyse des mit 24 Personen noch recht übersichtlichen Datensatzes ergab. Von einem Ausschluss des betreffenden Falls

8 Studie III: Interferenzen & Blickbewegungen

wurde jedoch abgesehen, da eine erhöhte Fehlervarianz tendenziell eher zu konservativen Urteilen führt. So handelte es sich bei dem gefundenen Zusammenhang also sehr wahrscheinlich nicht um ein *Artefakt*, das aufgrund eines extremen Wertes zustande kam.

Die dritte und letzte Analyse betraf die Sakkadenweite, die als mittlere Amplitude in Grad° erfasst worden war. Vor der Durchführung der MANOVA wurden erneut deren Voraussetzungen geprüft, unter denen das Verfahren gültige Schlüsse zulässt. Die Ergebnisse in Tabelle 8.6 veranschaulichen, dass jede der vier Messungen als normalverteilt gelten konnte und die Varianzen in den einzelnen Zellen nicht wesentlich voneinander abwichen.

Tabelle 8.6: Kolmogorow-Smirnow- und Levene-Tests für die mittlere Sakkadenamplitude

	Verfahren	Prüfgröße	Freiheitsgrade	Irrtumswahrscheinlichkeit
<i>Bild I : Die Theorie der Plattentektonik</i>	<i>Levene – Test</i>	$F = 0.51$	$df_1 = 3; df_2 = 20$	$p = .683$
	<i>K – S – Test</i>	$Z_{K-S} = 0.70$	–	$p = .708$
<i>Bild II : Die Bedeutung der Plattengrenzen</i>	<i>Levene – Test</i>	$F = 0.80$	$df_1 = 3; df_2 = 20$	$p = .507$
	<i>K – S – Test</i>	$Z_{K-S} = 0.41$	–	$p = .996$
<i>Bild III : Messungen im Erdinneren</i>	<i>Levene – Test</i>	$F = 0.66$	$df_1 = 3; df_2 = 20$	$p = .585$
	<i>K – S – Test</i>	$Z_{K-S} = 0.60$	–	$p = .864$
<i>Bild IV : Mechanismen der Plattenbewegung</i>	<i>Levene – Test</i>	$F = 1.11$	$df_1 = 3; df_2 = 20$	$p = .367$
	<i>K – S – Test</i>	$Z_{K-S} = 0.76$	–	$p = .608$

Anm.: K-S-Test = Kolmogorow-Smirnow-Anpassungstest; signifikante Abweichungen bzw. Unterschiede werden durch ein Sternchen (* < .05) bzw. durch zwei Sternchen (** < .01) hervorgehoben

Selbst der sensitive Box-Test ($M_{Box} = 57.21$, $F = 1.17$, $df_1 = 30$, $df_2 = 1099.8$, $p = .241$) wurde nicht signifikant, so dass auch die Gleichheit der Kovarianzmatritzen gegeben war. Bei den Prüfung der faktoriellen Effekte zeigte sich ein ähnliches Bild, wie bei der mittleren Fixationsdauer. Die *Reihenfolge* ($V = .252$, $F_{(4;17)} = 1.429$, $p = .252$, $\bar{\eta}^2 = .252$) und die *Interaktion* ($V = .270$, $F_{(4;17)} = 1.576$, $p = .226$, $\bar{\eta}^2 = .270$) besaßen keinen bedeutsamen Einfluss auf die Sakkadenamplitude. Doch wurde erneut der Faktor *Informationsverhältnis* ($V = .600$, $F_{(4;17)} = 6.376$, $p = .003$, $\bar{\eta}^2 = .600$) signifikant. Allerdings gestaltete sich die Zuordnung des durchaus großen Effekts nicht ganz so einfach, wie Tabelle 8.7 verdeutlicht.

Zuerst fällt auf, dass die Unterschiede zwischen den Gruppen keine einheitliche Richtung aufwiesen, wie das bei der mittleren Fixationsdauer noch der Fall gewesen war. Während beim ersten und vierten Bild durchschnittlich kürzere Sakkadenamplituden beobachtet werden konnten, wenn die Teilnehmer kongruente Materialien verwendet hatten, war es beim zweiten Bild umgekehrt. Hier wurden kürzere Sakkadenamplituden unter der komplementären Bedingung gemessen. Ein gänzlich uneinheitliches Bild stellten die Werte des dritten Bildes dar, weil hier die Richtung des Unterschieds zusätzlich über die

Tabelle 8.7: Mittelwerte und Standardabweichung der Sakkadenamplituden

	Versuchsbedingungen			
	Text vor Bild kongruent	Text vor Bild komplementär	Bild vor Text kongruent	Bild vor Text komplementär
<i>Bild I : Die Theorie der Plattentektonik</i>	$M = 2.81$ $SD = 0.42$	$M = 2.95$ $SD = 0.69$	$M = 2.59$ $SD = 0.62$	$M = 2.63$ $SD = 0.69$
<i>Bild II : Die Bedeutung der Plattengrenzen</i>	$M = 2.43$ $SD = 0.41$	$M = 2.38$ $SD = 0.27$	$M = 2.75$ $SD = 0.45$	$M = 2.29$ $SD = 0.43$
<i>Bild III : Messungen im Erdinneren</i>	$M = 2.68$ $SD = 0.45$	$M = 2.55$ $SD = 0.65$	$M = 2.49$ $SD = 0.54$	$M = 2.68$ $SD = 0.59$
<i>Bild IV : Mechanismen der Plattenbewegung</i>	$M = 2.32$ $SD = 0.35$	$M = 2.98$ $SD = 0.64$	$M = 2.21$ $SD = 0.51$	$M = 3.24$ $SD = 0.85$

Anm.: M = Mittelwert; SD = Standardabweichung

Reihenfolge variierte. Aus diesem Grund erstaunt es wenig, dass nachträglich berechnete ANOVA's lediglich einen bedeutsamen Effekt für den Faktor *Informationsverhältnis* ergeben hatten. So erwiesen sich die Unterschiede bei Bild I ($F_{(1;20)} = 0.135$, $p = .717$, $\bar{\eta}^2 = .007$), Bild II ($F_{(1;20)} = 2.386$, $p = .138$, $\bar{\eta}^2 = .107$) und Bild III ($F_{(1;20)} = 0.200$, $p = .659$, $\bar{\eta}^2 = .010$) als nicht signifikant. Allein für das vierte Bild überschritt der F-Wert des Faktors *Informationsverhältnis* das kritische Niveau deutlich ($F_{(1;20)} = 11.310$, $p = .003$, $\bar{\eta}^2 = .361$), was möglicherweise auch zu dem großen Effekt bei der MANOVA geführt hat.

Um letztlich die Ergebnisse der dritten Studie besser in den Gesamtzusammenhang einordnen zu können, sollte zum Abschluss noch geklärt werden, in welcher Beziehung die Blickbewegungsparameter zu den anderen zentralen aV's der Untersuchung standen. Aus diesem Grund bot es sich an, die *Anzahl* und *mittlere Dauer* der Fixationen sowie die *mittlere Sakkadenamplitude* mit der *Gesamtleistung* im Wissenstest, der *mentalenn Belastung* während des Lernens und der *Lernzeit* zu korrelieren. Tabelle 8.8 veranschaulicht die Ergebnisse.

Tabelle 8.8: Korrelationen der Blickbewegungsparameter mit Lernleistung, mentaler Belastung und Lernzeit

	Lernleistung	mentale Belastung	Lernzeit Text	Lernzeit Bild
Anzahl der Fixation	.52**	.40*	.55**	.94**
Dauer der Fixationen	.10	.08	.51**	.53**
Sakkadenamplitude	.15	-.39*	-.21	-.22

Anm.: Die Korrelation ist auf einem Niveau von * $<.05$ bzw. ** $<.01$ einseitig signifikant

Es fällt auf, dass die Anzahl der Fixationen der einzige Blickbewegungsparameter ist,

der durchgängig mit den anderen Variablen korrelierte. Je häufiger die Teilnehmer einzelne Elemente fixierten, umso mehr Items konnten sie im Wissenstest beantworten, was jedoch gleichzeitig mit einer erhöhten mentalen Belastung verbunden war. Der Zusammenhang zwischen Lernzeit und der Anzahl bzw. der Dauer der Fixationen erstaunte dagegen wenig, da beide Größen in die Variable eingingen. Sehr viel interessanter war der Umstand, dass die Leistung im Wissenstest nicht von der durchschnittlichen Fixationsdauer abhing und offensichtlich auch keinen Einfluss auf die empfundene Anstrengung besaß. Das galt jedoch nicht für die mittlere Sakkadenamplitude, die negativ mit der empfundenen Anstrengung korrelierte, während mit der Lernleistung und -zeit kein bedeutsamer Zusammenhang nachgewiesen werden konnte. Offensichtlich gingen durchschnittlich kürzere Sakkaden mit einer stärker empfundenen mentalen Belastung einher. Welche Schlussfolgerungen sich aus diesem und den anderen Ergebnissen ziehen lassen, ist jedoch das Thema des folgenden Kapitels.

8.4 Diskussion

Mit der dritten und letzten Studie war die Absicht verbunden gewesen, anhand von Blickbewegungsparametern Hinweise auf Interferenzen als mögliche Ursache für den *picture-text-sequencing*-Effekt zu finden. Zu diesem Zweck wurde die erste Untersuchung wiederholt, wobei die Stichprobe deutlich reduziert werden musste, da sich die Messung der Augenbewegungen während der Lernphase als sehr aufwändig erwiesen hatte. Die vergleichsweise geringe Teilnehmerzahl wirkte sich natürlich auf die Teststärke des Designs aus, so dass Zusammenhänge, die in der ersten Studie noch signifikant wurden, nicht mehr nachweisbar waren. Das betraf vor allem die beiden Haupteffekte für die Faktoren *Repräsentationsreihenfolge* und *Informationsverhältnis*. So ergab eine *post hoc* Analyse für den PTS-Effekt hinsichtlich der Lernleistung, dessen Stärke in der ersten Studie bei $\bar{\eta}^2 = .146$ gelegen hatte, eine β -Fehler-Wahrscheinlichkeit von $\approx 15\%$. Die Reduktion der Teilnehmerzahl von 84 auf 24 in der zweiten Untersuchung hatte zur Folge, dass diese Wahrscheinlichkeit auf $\approx 75\%$ anstieg. Damit war es deutlich wahrscheinlicher, einen tatsächlich vorhandenen Effekt mittlerer Größe nicht finden. Gleiches gilt für den Faktor *Informationsverhältnis*, wobei hier die β -Fehler-Wahrscheinlichkeit von $\approx 1\%$ auf $\approx 44\%$ zunahm. Außerdem war die Stichprobe durch die Aufnahme von Schülern heterogener, was trotz Randomisierung zu einer erhöhten Fehlervarianz geführt haben dürfte. Da aufgrund der wenigen Teilnehmer keine Kovariaten berücksichtigt werden konnten, verwunderte es nicht, dass die Analysen zum Lernerfolg keine bedeutsamen Effekte ergab. Ähnliches lässt sich für die mentale Belastung feststellen, da die vermuteten Gruppen-Unterschiede hier noch kleiner sein dürften.

Allerdings konnten ungeachtet der geringen Teststärke des Designs auch Zusammenhänge nachgewiesen werden, die bereits in der ersten Studie vorhanden waren. Dies betraf insbesondere Interaktionseffekte erster und zweiter Ordnung, die bei den formatspezifischen Lernzeiten beobachtet wurden. Demnach nutzten die Teilnehmer der kongruenten Bedingungen stets das Format am längsten, welches sie zuerst erhalten hatten. Dieser

Effekt verschwand, wenn die Versuchspersonen auf komplementäre Texte und Bilder zurückgreifen mussten, was offensichtlich der Grund für den signifikanten Interaktionseffekt zweiter Ordnung war. Zusammengefasst bestärken die Ergebnisse zur Lernzeit die Annahme, dass bei den kongruenten Gruppen tatsächlich eine Art *Ersetzungseffekt* auftrat. Informationen, die dem ersten Format entnommen wurden, haben dazu geführt, dass vergleichbare Inhalte der nachfolgenden Texte oder Bilder weniger intensiv verarbeitet wurden. Aus einer lernökonomischen Perspektive ist ein solches Vorgehen durchaus plausibel, da Aufmerksamkeit und Konzentration nicht beliebig lange aufrecht erhalten werden können. Die begrenzten mentalen Ressourcen sollten daher möglichst optimal eingesetzt werden, was die Verarbeitung von Informationen, die der Lernende als redundant ansieht, natürlich nicht mit einschließt.

Erstaunlich war jedoch, dass die Analyse zur Lernzeit keinen signifikanten Zwischensubjektfaktor *Informationsverhältnis* ergeben hatte. Immerhin war der Unterschied zwischen den komplementären und kongruenten Gruppen mit einem $\eta^2 = .248$ noch relativ groß gewesen und auch die Höhe der Teststärke ($1 - \beta = .69$) für eine Design mit 24 Personen noch akzeptabel. Dass trotzdem kein bedeutsamer Zusammenhang gefunden werden konnte, lag unter Umständen an der vergleichsweise hohen Fehlervarianz, die auf Größe und Heterogenität der Stichprobe zurückzuführen ist. Ebenso kann es sein, dass Unterschiede im Ablauf einen gewissen Einfluss besaßen. Beispielsweise hatten es die Blickbewegungsmessungen erforderlich gemacht, dass die Untersuchung in Einzelsitzungen stattfand, bei denen der Versuchsleiter die ganze Zeit über anwesend war. In Studie I erfolgten die Lern- und Testphasen überwiegend in Gruppen, was die Interaktion zwischen Versuchsteilnehmer und -leiter deutlich abschwächte. Über die konkreten Auswirkungen der Einzelsitzungen auf die Verhaltens- und Leistungsdaten kann nachträglich nur spekuliert werden, dennoch stellen die Unterschiede bei der Durchführung eine potentielle Fehlerquelle dar, die zumindest einen Vergleich der beiden Untersuchungen erschwert.

Dieses Problem gab es bei der Analyse und Einordnung der Ergebnisse zu der Blickbewegungsmessung nicht, da im Grunde kaum Annahmen im Voraus gemacht werden konnten. Der explorative Charakter des Vorgehens war mit der schlichten Absicht verbunden, Hinweise auf Interferenzen zu finden, die direkter und verhaltensnäher waren, als beispielsweise nachträglicher Selbsteinschätzungen oder ein Lerntest. Die Ergebnisse der durchgeführten Berechnungen legten jedoch nahe, dass ein solcher Zusammenhang zwischen der Repräsentationsreihenfolge und den ausgewählten Blickbewegungsparametern nicht bestand. Natürlich kann auch hier die geringe Teststärke des Design als mögliche Erklärung für statistisch unbedeutende Gruppenunterschiede herhalten, doch erschien diese Überlegung in erster Linie für die Anzahl der Fixationen plausibel zu sein.

Das wird vor allem deutlich, wenn man bei der Interpretation die Ergebnisse berücksichtigt, die in Tabelle 8.8 dargestellt sind. So ist zu erkennen, dass die Anzahl der Fixationen und die Betrachtungszeit auf den Bildseiten sehr hoch miteinander korrelierten. Gleiches gilt für die Textseiten, wobei hier die Korrelation, die nicht in der Tabelle enthalten ist, mit $r = .97$ noch höher war. Bei einem so starken Zusammenhang liegt es

nahe, die Anzahl der Fixationen mit der Lernzeit einfach gleichzusetzen. Demnach hätte eine Gegenüberstellung der formatspezifischen Fixationshäufigkeiten in einem Messwiederholungsdesign sehr wahrscheinlich ähnliche Interaktionseffekte mit dem Faktor *Repräsentationsreihenfolge* ergeben. Da stattdessen aber eine MANOVA berechnet worden war, bei der keine Effekte ermittelt werden konnten, liegt der Schluss nahe, dass es nur etwas mehr *power* bedurft hätte, um die vermuteten Unterschiede auch statistisch abzusichern.

Ganz anders verhielt es sich hingegen mit den anderen beiden Blickbewegungsparametern, die offensichtlich nicht von der *Repräsentationsreihenfolge* beeinflusst wurden. Hier ergaben die Analysen jeweils einen Haupteffekt für den Faktor *Informationsverhältnis*, was auf den ersten Blick recht gut zu den Befunden passte, die Rayner und Pollatsek (1992) berichten. Die kongruenten Bilder enthielten letztlich mehr piktoriale Elemente, so dass die enthaltenen Informationen dichter angeordnet waren. Vereinfacht ausgedrückt, erforderte die höhere Anzahl an Inhalten pro Quadratzentimeter eine (zeitlich) intensivere Verarbeitung, während die Sprünge zwischen einzelnen Informationseinheiten weniger weit sein durften, wenn nicht etwas ausgelassen werden sollte. Dementsprechend stieg im Vergleich zu der komplementären Bedingung die mittlere Fixationsdauer und die durchschnittliche Sakkadenamplitude nahm ab. In beiden Fällen waren die Effektstärken der MANOVA ziemlich groß, wobei der Unterschied vor allem auf die kongruente und komplementäre Versionen des vierten Bildes zurückzuführen waren, wie nachträglich berechnete univariate Analysen ergeben hatten.

Bei der Fixationsdauer stellte dieser Befund im Grunde kein Problem dar, denn alle Zellmittelwerte bewegten sich in dieselbe Richtung, so dass die übrigen Einzelvergleiche sehr wahrscheinlich wegen der geringen Teststärke nicht signifikant wurden. Diese Ergebnisdeutung fällt im Zusammenhang mit den Sakkadenamplituden etwas schwerer, da hier die Mittelwertsstruktur kein einheitliches Bild ergab. In 3 von 8 Fällen, waren die Sakkaden durchschnittlich länger, wenn kongruentes Material präsentiert wurde, was den Befunden von Rayner und Pollatsek (1992) jedoch widerspricht. Umso erstaunlicher ist es, dass gerade hier die Berechnungen einen besonders starken Effekt anzeigten, obwohl die Zellenmittelwerte teilweise in verschiedene Richtungen deuteten. Ein möglicher Erklärungsansatz bietet eine Studie von Cole, Maxwell, Arvey und Salas (1994), die herausgefunden haben, dass die Teststärke einer MANOVA sehr unterschiedlich ausfallen kann. Dafür verantwortlich sind in erster Linie komplexe Wechselwirkungen zwischen den Korrelationen der abhängigen Variablen untereinander, der Richtung der Mittelwerte und der Größe der Effekte. Auch wenn die Autoren bei ihrer Simulationsstudie nur zwei Gruppen miteinander verglichen haben und die Schätzung des β -Fehlers bei mehrfaktoriellen Designs noch komplizierter sind, kann das Ergebnis zumindest teilweise ein Beiprodukt des statistischen Verfahrens sein.

Allerdings erklärt diese Annahme nicht, weshalb gerade bei Bild IV besonders große Unterschiede auftraten. Die plausibelste Erklärung hierfür war, dass diese Unterschiede vor allem durch die Präsentation verursacht wurden. Betrachtet man die beiden Versionen von Bild IV (siehe Anhang), dann fällt auf, dass die komplementäre Version etwa

halb so groß war, wie die kongruente Ausführung. Bei der Präsentation mit der Software *Eyelink 2.0* wurde das Stimulusmaterial aber stets auf die volle Bildschirmgröße skaliert, weshalb kleinere Bilder von den Betrachtern als deutlich größer wahrgenommen werden mussten. Dementsprechend nahm die Anzahl von Informationen pro Flächeneinheit ab, was letztlich zu den kürzeren Fixationen und längeren Sakkaden führte. Die signifikanten Effekte für den Faktor *Informationsverhältnis* wurde damit allem Anschein nach durch die Skalierung während der Präsentation verursacht und hatten demnach nichts mit irgendwelchen kognitiven Prozessen zu tun. Dafür sprechen auch die niedrigen Korrelationen zwischen Fixationsdauer, Sakkadenamplitude und der Leistung im Wissenstest. Offensichtlich ließen die beiden Blickbewegungsparameter keine Rückschlüsse auf Ablauf und Güte der Lernprozesse zu.

Anders verhielt es sich mit der Anzahl der Fixationen, die recht hoch mit der Leistung im Wissenstest korrelierte. Einen ähnlichen Befund haben bereits Christianson et al. (1991) berichtet, was jedoch aufgrund des engen Zusammenhangs mit der Lernzeit nicht wirklich überraschte. So ist es ebenfalls nicht ungewöhnlich, dass die Korrelation der Fixationszahl mit der mentalen Belastung hoch genug war, um selbst bei der kleinen Stichprobe signifikant zu werden. Mehr Fixationen bedeuten längeres Lernen und längeres Lernen geht mit einer vermehrten Anstrengung einher, vor allem wenn der Teilnehmer selbst bestimmt, wann er den Lernprozess beendet. Wirklich überraschend war hingegen der negative Zusammenhang zwischen den Sakkadenlänge und der mentalen Belastung. Offenbar haben Teilnehmer, die sich während der Lernphase vergleichsweise wenig anstrengten, ihre Blicke auf den Bildern hin und her *schweifen* lassen. Ob diese Korrelation auf (feste) interindividuelle Unterschiede, volitional gesteuertes Verhalten oder eine zufällige Verteilung zurückzuführen ist, konnte aufgrund der vorhandenen Daten nicht geklärt werden.

Abschließend bleibt festzuhalten, dass die Analyse der Blickbewegungsparameter keine neuen Erkenntnisse zum PTS-Effekt oder Anzeichen für Interferenzen ergeben hat. Die signifikanten Ergebnisse stimmten zwar teilweise mit den Befunden anderer Autoren überein, Rückschlüsse auf zugrunde liegende kognitive Prozesse waren jedoch nicht möglich.

Teil III

Gesamtdiskussion

9 Fazit

Alle experimentellen Studien dieser Arbeit wurden mit dem Ziel durchgeführt, mehr über den Wissenserwerb mit Texten und Bildern zu erfahren. Im Zentrum stand dabei ein Phänomen, das Schnotz (2005) als *picture-text-sequencing*-Effekt bezeichnet und immer dann zu beobachten ist, wenn verbale und piktoriale Informationen nacheinander verarbeitet werden müssen. Demnach ist es für den Wissenserwerb zuträglicher, wenn die Präsentation des Bildes vor dem Text erfolgt. Als Ursache für diesen Effekt machen Kulhavy et al. (1994) und Schnotz (2002, 2005) jeweils unterschiedliche Prozesse verantwortlich. Durch eine Gegenüberstellung von Hypothesen, die anhand der theoretischen Modelle beider Autoren hergeleitet wurden, sollten schließlich Rückschlüsse auf die Beschaffenheit der kognitiven Architektur gewonnen werden. Die empirischen Befunde waren jedoch weniger eindeutig, als erwartet, so dass es sinnvoll erschien, zwei weitere Untersuchungen durchzuführen. Bevor nun der Versuch unternommen wird, die verschiedenen Studien in einen größeren Zusammenhang einzuordnen, veranschaulicht die folgende Auflistung zunächst die drei wichtigsten Ergebnisse.

1. Der *picture-text-sequencing*-Effekt konnte reproduziert werden
2. Die Ergebnisse von Studie II stützen die *Interferenzhypothese*
3. Bei allen Ergebnissen kommt der Lernzeit eine besondere Bedeutung zu

9.1 Integrierte Deutung der Resultate

Es gibt eine Vielzahl von robusten Effekten, die bislang beim Wissenserwerb mit multiplen Repräsentationsformaten nachgewiesen wurden (vgl. Mayer & Moreno, 2003; Mayer, 2005). Der wohl bekannteste dieser Befunde bezieht sich auf die einfache Feststellung, dass Lernende in vielen Fällen davon profitieren, wenn sie neben einem Text zusätzlich auf Bilder (Diagramme, schematische Zeichnungen, etc.) zurückgreifen können (Levie & Lentz, 1982; J. Levin, Anglin & Carney, 1987; Carney & Levin, 2002). Dass die Reihenfolge, in der verbale und piktoriale Informationen verarbeitet werden, das Ergebnis des Lernprozesses beeinflusst, ist dagegen eher unbekannt. So wird zunächst von (Kulhavy et al., 1994) die Bedeutung der Präsentationsreihenfolge für den Wissenserwerb erwähnt, allerdings ohne auf veröffentlichte Studien hinzuweisen. Die ausführliche Darstellung einer Untersuchung zum Einfluss der Sequenz auf das Lernen mit Texten und Bildern findet sich schließlich bei Verdi et al. (1997). Die Autoren stellten in einer Reihe von Experimenten fest, dass Schüler immer dann mehr Wissen erwarben, wenn das Bild vor dem Text gelernt wurde. Die Stärke der statistischen Zusammenhänge variierte dabei

9 Fazit

zwischen $\bar{\eta}_{part}^2 \approx 0.06$) und $\bar{\eta}_{part}^2 \approx 0.15$) was einem mittleren bzw. einem großen Effekt entsprach. Obwohl es anhand dieser Befunde nahe liegt, Rückschlüsse auf die Architektur des kognitiven Systems zu ziehen, ließen sich keine weiteren Studien zum *picture-text-sequence-Effekt* finden. Mögliche Gründe für die Vernachlässigung der Reihenfolge beim Wissenserwerb mit Bildern und Texten gibt es verschiedene.

Zum einen ist der PTS-Effekt aus einer rein pragmatischen Perspektive nicht wirklich interessant. So sollten nach dem Prinzip der *temporal contiguity* verbale und piktoriale Informationen möglichst gleichzeitig präsentiert werden, da es eine zusätzliche Belastung der Kapazität des Arbeitsgedächtnisses darstellt, wenn Lernende dazu gezwungen sind, die beiden Formate sequentiell zu lernen. Die Gestalter multimedialer Lernumgebung werden demnach darauf achten, Bilder und Texte räumlich und zeitlich integriert anzuordnen, so dass eine sequenzierte Präsentation von Bild und Text erst gar nicht notwendig wird. Ein möglicher weiterer Grund könnte die CTML von Mayer (2005) sein, die sicherlich die bekannteste Theorie zum Wissenserwerb mit multiplen Repräsentationsformaten darstellt und eine Vielzahl empirischer Arbeiten angeregt hat. Innerhalb der CTML werden zwar verbale und ein bildliche mentale Modelle unterschieden, aber Annahmen zu formatspezifischen Merkmalen beider Formate und ihrer Bedeutung für den Lernprozess macht der Autor nicht. Theoretisch ist es daher unerheblich, ob Bild oder Text zuerst verarbeitet werden, so dass die Präsentationsreihenfolge ohne Konsequenzen für den Wissenserwerb bleiben müsste. Spezifische Merkmale verbaler und piktorialer Repräsentationsformate und ihre Folgen für den Lernprozess liegen damit im toten Winkel der CTML.

Natürlich besteht auch die Möglichkeit, dass es sich im Zusammenhang mit den wenigen Veröffentlichungen zum PTS-Effekt um ein *file drawer problem* handelt (Rosenthal, 1979). Dabei ist die Kernannahme, dass empirische Befunde, die in verhaltenswissenschaftlichen Kontexten berichtet und veröffentlicht werden, grundsätzlich eine Verzerrung der tatsächlichen Sachverhalte darstellen. Verantwortlich ist die Tendenz, nur Arbeiten mit signifikanten Effekten bei Fachzeitschriften einzureichen, während nicht signifikante Studien in der Schublade verschwinden. Es kann also sein, dass es den PTS-Effekt gar nicht gibt und den Untersuchungen von (Verdi et al., 1997) und (Kulhavy et al., 1994) eine Vielzahl von Forschungsarbeiten gegenüber stehen, die nicht publiziert wurden, weil kein Einfluss der Verarbeitungsreihenfolge auf den Lernerfolg nachzuweisen war. Diesen Überlegungen widersprechen jedoch die Ergebnisse der ersten Studie, wonach die Reihenfolge, in der Bilder und Texte präsentiert wurden, die Leistungen beim Wissenserwerb sehr wohl beeinflusst. Berücksichtigt man zudem die ermittelte Effektstärke von $\bar{\eta}^2 = .146$, dann handelt es sich um einen vergleichsweise starken Zusammenhang, was einen α -Fehler recht unwahrscheinlich machte. Es ist daher auch zu vernachlässigen, dass im Rahmen der Blickbewegungsstudie kein Effekt der Reihenfolge auf die Leistung im Wissenstest gefunden werden konnte, da die Teststärke $1 - \beta = .15$ des Designs sehr gering war.

Dieses Ergebnis war durchaus positiv zu bewerten, auch wenn die Interpretation sämtlicher Resultate sich um einiges schwieriger erwies als erwartet. Das lag weniger an den

Befunden zum Lernerfolg, die für sich allein genommen eindeutig für die *Kapazitätshypothese* sprachen. Auch die Analysen zur mentalen Belastung, die keine Unterschiede ergeben hatten, stellten den Erklärungsansatz von Kulhavy et al. (1994) nicht wirklich in Frage, da mit niedrigen Effektstärken argumentiert werden konnte. Als vollkommen unvereinbar mit der *Kapazitätshypothese* erwiesen sich allerdings die Ergebnisse zur Lernzeit. Demnach ließ sich eine theoretisch besonders intensive Phase kognitiver Operationen nur sehr schwer mit vergleichsweise kurzen Verarbeitungszeiten vereinbaren. Um die Deutungskraft der Interferenzhypothese war es aber nicht sehr viel besser bestellt, da deren Vorhersagen immer nur teilweise zugetroffen hatten. Insbesondere die Annahmen zum Bedeutung des Informationsverhältnisses waren falsch gewesen. Die Ergebnisse zeigten eindeutig, dass die Teilnehmer mit Bildern und Texten sehr viel mehr Wissen erwerben konnten, wenn diese sich inhaltlich stark überlappten. Das Besondere war jedoch, dass auch zwischen den komplementären Gruppen diejenigen besser abgeschnitten hatten, denen das Bild zuerst präsentiert worden war.

Eine große Stärke des Modells von Schnotz (2005) ist es, dass es gleichermaßen positive und negative Effekte beim Wissenserwerb mit multiplen Repräsentationen theoretisch erklären kann. Die Überlegungen im Vorfeld der ersten Studie hatten sich vor allem auf kognitive Konflikte konzentriert, die entstehen sollten, wenn ein Lernender versucht, sein mentales Modell und eine externe grafische Darstellung des gleichen Sachverhaltes zu integrieren. Dabei wurde übersehen, dass es eigentlich nahe liegt, dass die Präsentation der piktorialen vor den verbalen Informationen auch positive Auswirkungen auf den Wissenserwerb haben kann. Schnotz und Bannert (1999, 2003) selbst gehen davon aus, dass es den Lernenden leichter fällt, ein mentales Modell anhand eines Bildes zu konstruieren, als anhand eines Textes. Als Grund sehen die Autoren den Umstand an, dass Bild und Modell auf den selben analogen Repräsentationsprinzipien beruhen. Ein Text ist dagegen ein symbolisches Format, das für die Modellkonstruktion Umwandlungsprozesse erforderlich macht, deren Erfolg stark von dem Vorwissen des Lernenden abhängt. Nimmt man beispielsweise an, dass sich beide Formate auf den gleichen Sachverhalt beziehen, dann stellt die Präsentation des Bildes vor dem Text ein erhebliche Erleichterung dar. Der Lernende nutzt gezwungenermaßen zuerst die piktorialen Informationen zur Modellkonstruktion, was ihm aber vergleichsweise leicht fällt. Nach den Vorstellungen von Schnotz (2005) schließt die Verarbeitung eines Bildes den Aufbau einer propositionalen Struktur mit ein, die Elemente und Bedeutungen Modells benennt. Wendet sich der Lernende nach der Verarbeitung des Bildes dem Text zu, dann ist das mentale Modell in seinen Grundzügen bereits vorhanden. Die verbalen Informationen werden dann vor allem genutzt, um die propositionale Struktur zu ergänzen, die Grundlage der Modellinspektion und -interpolation ist.

Diese Annahme bildete die Grundlage für eine Erklärung, weshalb auch unter der komplementären Bedingung ein PTS-Effekt zu beobachten war. Demnach blieb der Verarbeitungsvorteil selbst dann bestehen, wenn beide Formate nur wenige gemeinsame Informationen enthielten. Auch anhand eines komplementären Bildes ließ sich ein mentales Modell konstruieren, welches das Textverstehen erleichtern konnte. Zwar erhielt der Ler-

nende keinen unmittelbaren Entwurf der verbal umschriebenen Sachverhalte, was die eigenständige Konstruktion einer analogen Repräsentation anhand verbaler Informationen weitgehend überflüssig gemacht hätte. Dafür ließen sich aber strukturelle Komponenten oder Elemente des mentalen Modells verwenden, die bei der Verarbeitung des nachfolgenden Textes hilfreich waren. Beispielsweise hatten in der ersten Studie Lernende ein komplementäres Bild betrachtet, das mögliche Ursachen für die Verschiebungen der tektonischen Platten veranschaulichte. Kern der beiden Abbildung war eine schematische Zeichnung, in der drei Lithosphären dargestellt wurden, die sich voneinander weg bewegten oder zusammenprallten. Der dazugehörige Text thematisierte ebenfalls theoretische Mechanismen der Plattenbewegung, die jedoch andere Kräfte für den Kontinentaldrift verantwortlich machten. Hatte der Teilnehmer nun das komplementäre Bild vor dem Text gesehen, dann konnte er Elemente der mentalen Repräsentation der Abbildung für die Modellkonstruktion der verbale umschriebenen Ursachen nutzen. Das mentale Modell des Bildes musste zwar anhand der Informationen im Text interpoliert werden, damit es die Kausalzusammenhänge kohärent wiedergab. Die Anordnung der drei tektonischen Platten musste jedoch nicht eigens konstruiert werden, was die Verarbeitung der verbalen Inhalte erleichtert haben dürfte. Allerdings verlangt ein solches Vorgehen, dass der Lernende in der Lage ist, strukturelle Gemeinsamkeiten zwischen Bild und Text zu erkennen und auch zu übertragen. Der PTS-Effekt zwischen den komplementären Gruppen legte jedoch nahe, dass dies den Teilnehmer zumindest ansatzweise gelungen war.

Lernende, die dagegen den Text zuerst erhielten, hatten keine Vorgaben für die Modellkonstruktion. Sie waren dazu gezwungen, auf ihr Vorwissen zurückgreifen, um sich anschauliche Inhalte vorstellen bzw. verstehen zu können. Der Aufbau des mentalen Modells erforderte das Durchlaufen sämtlicher Stufen der kognitiven Textverarbeitung, was auch die Umwandlung *symbolischer* Informationen in ein analoges Repräsentationsformat beinhaltete. Dabei ließ sich nicht ausschließen, dass auch die Verarbeitung des Textes vor dem Bild Auswirkungen auf den Lernprozess hatte. Aufgrund der Ergebnisse musste allerdings davon ausgegangen werden, dass dieser Einfluss negativer Natur war. Möglicherweise wurde der PTS-Effekt doch durch Interferenzen verursacht, die als Folge struktureller oder funktionaler Differenzen zwischen dem mentalen Modell und dem externen Bild entstanden. Auf diese Weise konnten anhand des theoretischen Modells von Schnotz (2005) gleich zwei Interpretationen abgeleitet werden, welche die Befunde der ersten Studie zu erklären vermochten. Um aber letztlich eine Antwort zu finden, ob nun positive oder negative Effekte der Reihenfolge für die besseren Ergebnisse der Gruppen verantwortlich waren, die zuerst mit den Bildern gelernt hatten, mussten weitere Daten erhoben werden.

Durch den Vergleich von Teilnehmern mit Bildern und Texten und Teilnehmern, die jeweils nur ein Format für den Wissenserwerb nutzen konnten, sollte Klarheit geschaffen werden. Dabei war der Hintergedanke, dass Interferenzen das Resultat eines kognitiven Konfliktes sind, der entsteht, wenn ein Lernender versucht sein mentales Modell mit einem externen Bild abzugleichen bzw. zu integrieren. Dieser Abgleich kann auch irrelevante Details betreffen, jedenfalls ist damit zu rechnen, dass es dem Individuum aufgrund

der Interferenzen schwerer fallen wird, die piktorialen Informationen angemessen zu verarbeiten. Argumentiert man auf der Grundlage der *cognitive load theory* dann entsteht in dieser Phase *extraneous load* (Sweller & Chandler, 1994), der auf die Anordnung des Textes vor dem Bild zurückzuführen ist. Bei einem negativen Einfluss der Reihenfolge sollten die Teilnehmer, die den Text vor dem Bild erhalten hatten, daher im Wissenstest nicht besser abschneiden, als Personen, denen allein der Text zur Verfügung stand. Zudem war davon auszugehen, dass die Interferenzen trotz kontrollierter Lernzeit eine zusätzliche Quelle der mentalen Belastung darstellte, was in den Selbsteinschätzungen der Studierenden zum Ausdruck kommen würde.

Die Ergebnisse der Analysen waren ganz im Sinne der Interferenzhypothese, da der Vergleich der Text-vor-Bild und nur Text Gruppen keine Unterschiede hinsichtlich der Lernleistung ergeben hatte. Zudem zeigten die Selbstauskünfte der Teilnehmer, die beide Formate erhalten hatten, dass diese sich beim Wissenserwerb stärker angestrengen mussten. Der Effektgröße des Unterschiedes zu der nur-Text Gruppe war mit einem $\bar{\eta}^2 = .051$ zwar klein bis mittel, aber dennoch bedeutsam. Obwohl die Befunde zuerst einmal die Interferenzhypothese stützten, war nicht auszuschließen, dass die Reihenfolge von Bild und Text dem Wissenserwerb auch förderlich sein konnte. In diesem Zusammenhang wurde angenommen, dass sich die Teilnehmer, denen die Bilder vor dem Text präsentiert wurden, trotz kontrollierter Lernzeit mehr anstrengen würden, als die Studierenden, die allein das Bild erhalten hatten. Die zusätzliche mentale Belastung sollte aber mit einer tieferen Verarbeitung und besseren Lernergebnissen verbunden sein, weshalb sie im Sinne der CLT als *germane load* anzusehen wäre (Sweller et al., 1998). Die Ergebnisse der zweiten Analyse zeigten zwar, dass die Gruppen mit den Bildern und Texten deutlich besser im Wissenstest abschnitten, doch konnte kein Unterschied hinsichtlich der mentalen Belastung festgestellt werden.

Natürlich war es möglich, dass die positiven Einflüsse der Reihenfolge auf den Wissenserwerb kleiner waren als die negativen und dass daher die Teststärke des Designs zu gering war. Ebenso könnte eine ungünstige Verteilung von personenbezogenen Störgrößen die Ergebnisse verzerrt haben, worauf die *marginal* signifikante Interaktion hinweist, welche die α -Fehlerwahrscheinlichkeit von 5% nur knapp verfehlt hatte. Fest stand aber auch, dass der Vergleich zwischen den Text-vor-Bild und nur-Text Gruppen, genau auf die Vorhersagen der Interferenzhypothese passte. Insbesondere der Umstand, dass die Teilnehmer nicht von dem zusätzlichen Bild profitieren konnten, ließ sich als Beleg für das Auftreten von *extraneous load* ansehen. Für die Überlegenheit der Gruppen, die zu den Bildern noch Texte erhalten hatten, gab es dagegen mehrere Erklärungsmöglichkeiten, die nicht unbedingt auf eine besonders elaborierte Verarbeitung zurückgeführt werden musste. So hatte die Analyse zu Beginn der zweiten Studie gezeigt, dass durchschnittlich mehr Wissen erworben wurde, wenn die Teilnehmer auf die Texte anstatt auf die Bilder zurückgreifen konnten. Offenbar enthielten die Texte mehr Informationen bzw. diese Information konnten im Vergleich zu den Bildern relativ einfach abgelesen werden. Die Verarbeitung piktorialer Inhalte stellte ohne Text eine sehr viel größere Herausforderung dar, was bislang vor allem bei Verstehen logischer nachgewiesen wurde (vgl. Shah, Mayer

& Hegarty, 1999; Mautone & Mayer, 2007). Diese Überlegungen änderten nichts an der Tatsache, dass die Befunde der zweiten Studie allein die Interferenzhypothese stützten, während die Konstruktionshypothese zu verwerfen war.

Dass mit der Untersuchung von Blickbewegungsparametern im Rahmen der dritten Studie keinerlei *direkte* Hinweise auf Interferenzen gefunden wurden, ist dagegen keine sehr große Überraschung, da physiologische Parameter zur Erfassung von *cognitive load* allgemein recht schwierig sind (vgl. Schnotz & Kürschner, 2007). Möglicherweise konnten aufgrund der geringen Teststärke des Designs keine Effekte nachgewiesen werden, immerhin hatten nur 24 Personen an der Studie teilgenommen. Doch trotz dessen zeigte sich auch hier, dass die Reihenfolge, in der Text und Bilder gelernt wurden, die formatspezifische Lernzeit maßgeblich beeinflusst hatte. Demnach nutzten vor allem die Teilnehmer der kongruenten Bedingungen das Format am längsten, welches ihnen zuerst präsentiert wurde. Dieser Effekt verschwand (Studie III) bzw. war deutlich schwächer (Studie I), wenn die Versuchspersonen auf komplementäre Texte und Bilder zurückgreifen mussten, was die signifikanten Interaktionseffekte zweiter Ordnung erklärte. Diese Ergebnisse legen die Deutung nahe, dass zumindest bei den kongruenten Gruppen tatsächlich ein *Ersetzungseffekt* auftrat. Demnach mussten Informationen, die dem ersten Format entnommen werden konnten, nicht noch einmal dem zweiten Format entnommen werden, was in der Konsequenz zu einer kürzeren Lernzeit führte. Geht man davon aus, dass ein Lernender aufgrund der beschränkten Kapazität seines Arbeitsgedächtnisse (Baddeley, 1992) mehr oder weniger bewusst ökonomischen handelt, dann ist ein solches Vorgehen durchaus plausibel. Viele Menschen wissen aus eigener Erfahrung, dass Aufmerksamkeit und Konzentration mentale Ressourcen sind, die nicht unbegrenzt zur Verfügung stehen. Daher müssen die vorhandenen Kapazitäten möglichst optimal genutzt werden, weshalb die Verarbeitung redundanter Informationen vermieden werden sollte.

Damit lassen die Befunde zu den formatspezifischen Lernzeiten eine weitere und sehr viel einfachere Interpretation des PTS-Effekts zu. So könnte anstelle von Interferenzen oder *extraneous load* eine simple Fehleinschätzung der Lernenden für den Einfluss der Reihenfolge auf Lernzeit und Wissenserwerb verantwortlich sein. Lernende die den Text zuerst lasen, hatten möglicherweise den Eindruck gewonnen, alle wichtigen Informationen bereits zu kennen, weshalb sie das nachfolgende Bild weniger intensiv verarbeiteten, selbst wenn es wichtige Informationen für das Verständnis der Sachverhalte enthielt. Zu dieser Argumentation passt auch die Einschätzung von Weidenmann (2006), nach der Lernende Bilder sehr gerne anschauen, diese hinsichtlich ihrer Informationstiefe häufig aber unterschätzten. Wurde das Bild zuerst präsentiert, dann wusste der Teilnehmer jedoch nicht, welche Informationen der nachfolgende Text enthielt und welche nicht. Damit war es für den Wissenserwerb erforderlich, die piktorialen Informationen möglichst genau zu analysieren, bevor der Text gelesen wurde. Dabei ist nicht davon auszugehen, dass die Teilnehmer diese Anforderung unterschätzten, selbst wenn die Lesezeiten signifikant kürzer waren. So sind trotz der piktorialen Wende (vgl. Mitchell, 1995) Studierende weiterhin daran gewöhnt, neues Wissen mit Texten zu erwerben. Möglicherweise hatte die Präsentation des Bildes vor dem Text dessen Verarbeitung erleichtert, indem relevantes

Vorwissen aktiviert oder die Konstruktion eines mentalen Modells unterstützt wurde, was zu einer kürzeren Lesezeit geführt haben könnte.

Endgültig belegen ließ sich diese Interpretation der Ergebnisse anhand der gesammelten Daten allerdings nicht, zumal die Resultate der zweiten Studie eindeutig für die Interferenzhypothese sprachen. Demnach hatte die nachträgliche Präsentation eines Bildes bei kontrollierter Lernzeit zu einer erhöhten mentalen Belastung geführt, während gleichzeitig nicht mehr Wissen erworben werden konnte. Der berechneten Kovarianzanalyse lag jedoch die Annahme eines linearen Zusammenhangs zwischen Lernzeit und mentaler Belastung zugrunde, was nicht unbedingt zutreffen muss. Denkbar wäre auch, dass die beiden Variablen über eine Exponentialfunktion miteinander verknüpft sind. In diesem Fall würde mit fortlaufender Dauer die mentale Belastung sehr viel stärker zunehmen als zu Beginn der Präsentationsphase, weshalb der Einfluss der Lernzeit womöglich nur unzureichend kontrolliert wurde. Diese Überlegungen machen deutlich, dass es zwar Belege für die Interferenzhypothese gibt, aber empirisch wirklich abgesichert ist sie nicht. Welche methodischen Vorgehensweisen unter Umständen zu einem eindeutigeren Ergebnis geführt hätten, wird in dem folgenden Kapitel diskutiert.

9.2 Methodische Kritik

Nach einer Untersuchung weiß man in der Regel mehr als zuvor. Das betrifft nicht nur den zu erwartenden inhaltlichen Erkenntnisgewinn, sondern häufig auch das methodische Vorgehen. Hier sind es allerdings Fehler oder Schwächen des eigenen Ansatzes, die in der Regel zwar ärgerlich sind, letztlich aber der Verbesserung zukünftiger Forschungsbemühungen dienen. Viele der getroffenen Entscheidungen hätte man anders umsetzen können, wobei alles seine Vor- und Nachteile hat. Den *cognitive load* bzw. die mentale Belastung in der Lernphase retrospektiv mit nur einer Skala zu erfassen, hatte sicherlich den Vorzug, dass es einfach und schnell ging. Es gab keine Unterbrechung der Lernphase und die Teilnehmer hatten nicht das Gefühl, ständig das Gleiche gefragt zu werden, was unter Umständen ihrer *compliance* zugute kam. Problematisch an dieser Methode den *cognitive load* zu messen war jedoch, dass es sich um nur einen Werte handelte, der einen Zustand über einen Zeitraum von 15 bis 25 Minuten erfassen sollte. Mehrere Messungen, zum Beispiel nach jeder Text- und Bildseite, hätten die Interferenzhypothese empirisch sicherlich auf ein solideres Fundament gestellt. Hinzu kommt, dass die Formulierung nach Paas et al. (1994) eine starke motivationale Komponente umfasste. Vielleicht wurde gar nicht die mentale Belastung gemessen, sondern die Motivation der Teilnehmer, sich bei dem Experiment einzubringen. Aus diesem Grund hätte es sicherlich nicht geschadet, neben der Selbsteinschätzung noch weitere Indikatoren für den *cognitive load* beim Wissenserwerb zu erheben, was jedoch einfacher klingt als es ist (Schnotz & Kürschner, 2007).

Eine weitere tendenziell problematische Entscheidung betrifft die Tatsache, dass die interessantesten Ergebnisse von Studie II auf einem quasiexperimentellen Design beruhen. Demnach wurden Gruppen miteinander verglichen, die nicht durch eine rando-

misierte Zuteilung der Versuchspersonen entstanden, so dass die interne Validität als eingeschränkt gelten musste. Die Berücksichtigung verschiedener Merkmalen als potentielle Störgrößen hatte sicherlich dazu beigetragen, die Fehlervarianz zu reduzieren, die Randomisierung ersetzen konnte ein solches Vorgehen sehr wahrscheinlich aber nicht. Der Grund, weshalb die Stichproben der beiden Untersuchungen getrennt gezogen wurden, liegt in der Planung der ersten Studie, die sich auf den Vergleich der Interferenz- und Kapazitätshypothese konzentriert hatte. So war fest damit gerechnet worden, dass die Ergebnisse eindeutig für die eine oder andere Hypothese sprechen würden. Die Mehrdeutigkeit der Ergebnisse und ihre Interpretation erforderte jedoch ein Design, in dessen Rahmen Gruppen mit einem Format Gruppen mit beiden Formaten gegenübergestellt wurden. Die zusätzliche Erhebung von 168 Personen, stellte jedoch einen zu großen Aufwand dar, weshalb auf ein experimentelles Design verzichtet werden musste. Vielleicht hätte eine bessere Planungsarbeit im Rahmen der ersten Studie zu einem echten Experiment geführt, doch blieb fraglich, ob es überhaupt möglich war, die Resultate so zu antizipieren.

Die letzten der hier angeführten methodenkritischen Reflexionen betreffen die dritte Studie, bei der kein Einfluss der Verarbeitungsreihenfolge auf verschiedene Blickbewegungsparameter gefunden werden konnte. Zwar lag der Verdacht nahe, dass es einen solchen Einfluss gar nicht gibt, doch möglicherweise wurde auch eine ungünstige Versuchsanordnung gewählt. Beispielsweise weist Rayner (1998) darauf hin, dass Blickbewegungsparameter interindividuell sehr unterschiedlich ausfallen können, ohne dabei im Zusammenhang mit kognitiven Prozessen zu stehen. Eine Möglichkeit diese individuellen Messwerte zu kontrollieren besteht darin, anstelle eines between-Designs ein within-Design als Versuchsanordnung zu wählen. Jeder Teilnehmer hätte dann Bilder und Texte in variierender Reihenfolge und mit variierendem Informationsverhältnis erhalten, wobei die Umsetzung einer Versuchsanordnung mit mehr als einer Person pro Kombination sehr viel mehr als 24 Teilnehmer erfordern würde. Zudem wäre die Vergleichbarkeit mit Studie I erheblich eingeschränkt gewesen, was letztlich zu dem Verzicht auf ein within-Design geführt hatte. Ob wirklich die Versuchsanordnung für die nicht signifikanten Effekte verantwortlich ist, wird vielleicht in einer zukünftigen Forschungsarbeit untersucht werden.

9.3 Praktische Implikationen

Obwohl sich die Arbeit in erster Linie auf eine Fragestellung konzentrierte, anhand der zwei theoretische Modelle miteinander verglichen werden sollten, lassen sich anhand der Ergebnisse auch Schlussfolgerungen für die Praxis ziehen. So ist es für Lernende und Lehrende zwar interessant aber nicht wirklich wichtig, ob nun Interferenzen oder eine andere Ursache für den PTS-Effekt verantwortlich sind. Was dagegen zählt, sind Anregungen und Methoden, mit denen Wissen möglichst effektiv vermittelt wird. Bevor man jedoch pauschal dazu rät, Bilder immer vor Texten zu repräsentieren, ist es angebracht, die Ergebnisse der drei Studien unter praktischen Gesichtspunkten zu betrachten.

Man sollte auf jeden Fall bedenken, dass der *picture-text-sequencing*-Effekt durch ei-

ne experimentelle Anordnung erzeugt wurde, die in einer realistischen Lernumgebung so sicherlich nur sehr selten anzutreffen ist. Das betrifft vor allem die Einschränkung, dass es nicht gestattet war, frei zwischen den Bild und Textseiten zu navigieren. In einer realen Situation können Lernende aber nach Belieben in Büchern oder auf Webseiten vor und zurück zu blättern. Es ist sogar ein wichtiger Bestandteil der selbstgesteuerten Lernens, den eigenen Wissenserwerb kontinuierlich zu überwachen, um bei Zweifeln und Unklarheiten bestimmte Bilder oder Textabschnitte zu wiederholen. Über diese Freiheiten, verfügten die Teilnehmer der drei Studien nicht. Hätte man ihnen erlaubt, frei zu navigieren, dann wäre vermutlich keine Unterschiede zwischen den einzelnen Gruppen festzustellen gewesen, da die Reihenfolge den individuellen Bedürfnissen und Vorlieben beim Wissenserwerb angepasst worden wäre.

Allgemein lassen sich die Ergebnisse der Untersuchungen nur mit Vorbehalt generalisieren, zumal die Zusammensetzung der Stichprobe nicht repräsentativ war. Bei nahezu allen Versuchspersonen handelte es sich um Studierende, die sich in Psychologie, einem Lehramtsstudiengang oder in Erziehungswissenschaften eingeschrieben hatten. Gemessen an den Voraussetzungen für die allgemeine Hochschulreife handelte es sich damit um erfolgreiche Lerner, deren kognitiven Fähigkeiten über dem Bevölkerungsdurchschnitt gelegen haben dürften. Wie eine andere Zielgruppe auf die Variation der Präsentationsreihenfolge reagiert hätte, kann daher nur vermutet werden. Möglicherweise wäre bei inhaltlichen Experten andere Effekte zu beobachten gewesen als bei Lernenden mit einer schwach ausgeprägten Befähigung zum selbstgesteuerten Lernen. Es ist daher letztlich nicht leicht, aufgrund der Ergebnisse dieser Arbeit Ratschläge für die Praxis zu geben, da neben den experimentellen Einschränkungen die Merkmale der Zielgruppe berücksichtigt werden müssen.

Einige wenige Empfehlungen lassen sich jedoch aussprechen, auch wenn sie sich nur indirekt auf den PTS-Effekt beziehen. So hatte sich gezeigt, dass Redundanzen bzw. eine kongruentes Informationsverhältnis nicht generell schlecht für den Wissenserwerb mit multiplen Repräsentationen sind. Das gilt insbesondere dann, wenn Bild und Text nicht integriert oder auf der einer Seite präsentiert werden können. Offenbar hatte die Wiederholung von Informationen keine gesteigerte mentale Belastung zur Folge, sondern zur Etablierung stabiler Wissensstrukturen geführt. Natürlich können Richtung und Ausmaß dieses Effektes mit den Merkmalen und Vorlieben der Lernenden variieren, doch scheint es durchaus Anwendungskontexte zu geben, in denen Redundanzen sich sinnvoll einsetzen lassen. Darüber hinaus stellt es sicherlich keinen Schaden dar, wenn Schüler und Studierende strukturiert lernen, Bilder als Informationsmedium sinnvoll für den Wissenserwerb zu gebrauchen. Auf diese Weise könnte verhindert werden, dass Lernende die Verarbeitungstiefe piktorialer Inhalte unterschätzen, was die Effektivität multimedialer Lernumgebungen steigern würde.

9.4 Ausblick

Es kommt häufig vor, dass Forschungsarbeiten einige Fragen beantworten, aber gleichzeitig neue Fragen aufwerfen. Das ist auch im Zusammenhang mit dem *picture-text-sequencing*-Effekt der Fall, der nach der Literatur zu urteilen einen vergleichsweise unerforschten Aspekt des Wissenserwerbs mit multiplen Repräsentationen darstellt. Daher wäre es kein Fehler, zuerst den Versuch zu unternehmen, die vorgefundenen Gruppenunterschiede zu replizieren, da es nicht endgültig als gesichert gelten kann, dass für den Einfluss der Verarbeitungsreihenfolge Interferenzen verantwortlich sind. Die neue Studie sollte dabei auf einem experimentellen Design mit Randomisierung beruhen und die mentale Belastung der Teilnehmer möglichst anhand mehrerer Indikatoren zu erheben. Denkbar wäre auch, über *lautes Denken* die kognitiven Prozesse während Lernpause näher zu untersuchen, obwohl diese Methode geschulte Versuchspersonen erfordert.

Ein weiterer interessanter Forschungsansatz, der in eine etwas andere Richtung zielt, würde die mögliche Wechselwirkung zwischen der Verarbeitungsreihenfolge und der Expertise der Lernenden betreffen. In den drei Studien war das Vorwissen als Störgröße behandelt und dessen Einfluss daher herauspartialisiert worden. Ein alternatives Vorgehen bestünde darin, die Expertise der Teilnehmer als Faktor zu deklarieren und entsprechende Stichproben zu ziehen. Natürlich wäre in diesem Zusammenhang die Frage zu stellen, welche inhaltliche Domäne und damit welche Bildarten man für eine neue Untersuchung auswählt. Neben Karten und schematischen Zeichnungen böten sich auch logische Bilder an, da sie häufig im schulisch-akademischen Kontext verwendet werden. Die ersten Arbeiten zu diesem Phänomen konzentrierten sich auf geographische Karten bzw. räumlich organisierte Abbildungen, weshalb sich hier die Gelegenheit bietet, den theoretischen auf weitere Bildarten auszuweiten.

Die Beschaffenheit des piktoralen Materials ist allerdings nur ein Aspekt des sequenzierten Lernens mit Texten und Bildern. Ebenso interessant wäre es, den Einfluss von Modalität und Verarbeitungreihenfolge auf den Wissenserwerb und die mentale Belastung zu untersuchen. Möglicherweise ist es für den Lernprozess hilfreich, beide Sinneskanäle zu nutzen, damit sich die vorhandenen kognitiven Ressourcen voll ausschöpfen lassen. Vielleicht wird eine rein akustische Darbietung des Textes aber auch zu einer schlechteren Lernleistung führen, insbesondere wenn Bild und verbale Informationen nicht gleichzeitig präsentiert werden. Lohnenswert erscheint es zudem, den Forschungsansatz mit den Blickbewegungsdaten weiterzuverfolgen, selbst wenn die Analysen der dritten Studie keine bedeutsamen Effekte ergeben hatte. Das könnte sich jedoch ändern, wenn bei einer neuen Untersuchung ein within-Design umgesetzt würde. Außerdem wäre es eventuell hilfreich sich bei der Auswertung der Daten auf Blickpfade und nicht auf globale Parameter zu konzentrieren, wobei in diesem Fall das Stimulusmaterial vereinfacht werden sollte. Die Ausführungen verdeutlichen, dass es noch einige offene Fragen und Forschungsmöglichkeiten hinsichtlich der Verarbeitungsreihenfolge von Texten und Bildern gibt. So bleibt zu hoffen, dass die vorliegende Arbeit als Anregung verstanden wird, sich weiterhin mit diesem Thema zu beschäftigen.

Literaturverzeichnis

- Aaronson, D. & Scarborough, H. S. (1977). Performance theories for sentence coding: Some quantitative models. *Journal of Verbal Learning and Verbal Behavior*, 16, 277-303.
- Alba, J. W. & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, 93, 203-231.
- Amelang, M. & Zielinski, W. (1997). *Psychologische Diagnostik und Intervention*. Berlin: Springer.
- Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. (1983a). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (1983b). A spreading theory of activation. *Journal of Verbal Learning and Verbal Behavior*, 22, 261-295.
- Anderson, J. R. (1996). Act: A simple theory of complex cognition. *American psychologist*, 51, 355-365.
- Anderson, J. R. (2001). *Kognitive Psychologie* (3rd Aufl.). Heidelberg: Spektrum.
- Anderson, J. R., Bothell, D., Douglas, S., Lebiere, C., Qin, Y. & Byrne, M. D. (2004). An integrated theory of mind. *Psychological Review*, 111, 1036-1060.
- Anderson, J. R. & Brown, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, 79(2), 97-123.
- Anderson, R. C., Reynolds, R. E., Schallert, D. L. & Geotz, E. T. (1977). Frameworks for comprehending discourse. *American Educational Research Journal*, 14, 637-381.
- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In J. Baumert et al. (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69-137). Opladen: Leske + Budrich.
- Asendorpf, J. B. (1996). *Psychologie der Persönlichkeit*. Berlin: Springer.
- Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, p. 89-195). New York: Academic Press.
- Baddeley, A. D. (1992). Working memory. *Science*, 255, 556-559.
- Baddeley, A. D. (2000). The episodic buffer in working memory. *Trends in Cognitive Sciences*, 4, 417-423.

Literaturverzeichnis

- Baddeley, A. D. (2001). Is working memory still working? *American Psychologist*, 56, 851-864.
- Baddeley, A. D. & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, p. 47-89). New York: Academic Press.
- Ballstaedt, S.-P. (1997). *Wissensvermittlung. Die Gestaltung von Lernmaterial*. Weinheim: Beltz.
- Ballstaedt, S.-P. (2005). Text-Bild-Kompositionen im Unterrichtsmaterial. *Der Deutschunterricht*, 4(4), 61-70.
- Bartlett, F. C. (1932/1995). *Remembering. a study in experimental and social psychology*. New York: Cambridge University Press. (Original work published 1932)
- Bellezza, F. S. (1981). Mnemonic devices: Classification, characteristics and criteria. *Review of Educational Research*, 51, 247-275.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Bühner, M., Ziegler, M., Krumm, S. & Schmidt-Atzert, L. (2006). Ist der I-S-T 2000 R Rasch-skalierbar? *Diagnostica*, 52, 119-130.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5 Aufl.). Berlin: Springer.
- Bortz, J. & Döring, N. (1995). *Forschungsmethoden und Evaluation für Sozialwissenschaftler*. Berlin: Springer.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Berlin: Springer.
- Bortz, J., Lienert, G. A. & Boehnke, K. (1990). *Verteilungsfreie Methoden in der Biostatistik*. Berlin: Springer.
- Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika*, 36, 317-346.
- Bransford, J. D., Barclay, J. R. & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, 3, 193-209.
- Bransford, J. D. & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, 2, 331-350.
- Brewer, W. F. & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, 13, 207-230.
- Brünken, R., Steinbacher, S., Plass, J. L. & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual task methodology. *Experimental Psychology*, 49, 109-119.
- Brown, M. B. & Forsythe, A. (1974). The anova and multiple comparisons for data with heterogeneous variances. *Biometrics*, 30, 719-724.

- Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology, 98*, 182-197.
- Carney, R. N. & Levin, J. R. (2002). Pictorial illustrations still improves students' learning from text. *Educational Psychology Review, 14*(1), 5-26.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1-22.
- Chafe, W. (1994). *Discourse, consciousness, and time*. Chicago: University of Chicago Press.
- Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293-332.
- Christianson, S. A., Loftus, E. F., Hoffman, H. & Loftus, G. R. (1991). Eye fixations and memory for emotional events. *Journal of Experimental Psychology: Learning, Memory & Cognition, 17*, 693-701.
- Clark, J. M. & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*, 149-210.
- Cohen, J. (1960). A coefficient for agreement of nominal scales. *Journal of Educational and Psychological Measurement, 20*(1), 37 - 46.
- Cohen, N. J. & Squire, L. R. (1980). Preserved learning and retention of pattern analyzing skills in amnesia; dissociation of knowing how and knowing that. *Science, 210*, 207-210.
- Cole, D. A., Maxwell, S. E., Arvey, R. & Salas, E. (1994). How the power of manova can both increase and decrease as a function of the intercorrelations among the dependent variables. *Psychological Bulletin, 115*, 465-474.
- Cooper, G. & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79*, 347-362.
- Craik, F. I. M. (1979). Human memory. *Annual Review of Psychology, 63*-102.
- Craik, F. I. M. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.
- Craik, F. I. M. & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*, 268-294.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.
- Davis, O. I. & Hunkins, F. (1968). The usefulness of a map with geographic text. *Journal of Geography, 67*, 362-366.
- Dean, R. S. & Kulhavy, R. W. (1981). Influence of spatial organization in prose learning. *Journal of Educational Psychology, 73*, 57-64.

Literaturverzeichnis

- Diehl, J. & Arbinger, R. (1990). *Einführung in die Inferenzstatistik*. Frankfurt: Verlag Dietmar Klotz.
- Dixon, D. & Saltz, E. (1977). The role of imagery on concept acquisition in lower-SES children. *Child Development*, 48(1), 288-291.
- Dutke, S. (1999). Der Crossover-Effekt von propositionaler Textrepräsentation und mentalen Modellen: Zur Rolle interindividueller Fähigkeitunterschiede. *Zeitschrift für Experimentelle Psychologie*, 46, 164-176.
- Dutke, S. (2003). Anaphor resolution as a function of spatial distance and priming: Exploring the spatial distance effect in situation models. *Experimental Psychology*, 54, 270-284.
- Eden, G. F., Stein, J. F., Wood, H. M. & Wood, F. B. (1994). Differences in eye movement and reading problems in dyslexic and normal children. *Visions*, 34, 1345-1358.
- Engelkamp, J. (1991). *Das menschliche Gedächtnis*. Göttingen: Hogrefe.
- Engelkamp, J. (1994). Episodisches Gedächtnis: Von Speichern zu Prozessen und Informationen. *Psychologische Rundschau*, 45, 195-210.
- Ericsson, K. A. (1988). Concurrent verbal reports on text comprehension: A review. *Text*, 8, 295-326.
- Ericsson, K. A. & Chase, W. G. (1982). Exceptional memory. *American Scientist*, 70, 607-615.
- Ericsson, K. A. & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211-245.
- Ericsson, K. A., Kintsch, W. & Patel, V. (2000). How experts' adaptations to representative task demands account for the expertise effect in memory recall: Comment on Vicente and Wang (1998). *Psychological Review*, 107, 578-592.
- Everatt, J. & Underwood, G. (1994). Individual differences in reading subprocesses: Relationships between reading ability, lexical access, eye movement control. *Language and Speech*, 37, 283-297.
- Eysenck, M. W. (1978). Levels of processing: A critique. *British Journal of Psychology*, 69, 157-169.
- Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- Feldman, J. A. & Ballard, D. H. (1982). Connectionist models and their properties. *Cognitive Science*, 6, 205-254.
- Field, A. (2009). *Discovering statistics using spss* (3 ed.). London: Sage.
- Fischer, B. & Glanzer, M. (1986). Short-term storage and the processing of cohesion during reading. *The Quarterly Journal of Experimental Psychology*, 38A, 431-460.

- Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
- Frederiksen, C. H. (1975). Representing logical and semantical structure of knowledge acquired from discourse. *Cognitive Psychology*, 7, 371-458.
- Garnham, A. (1980). Default values, criteria and constructivism. *Cognitive Science*, 4, 427-433.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.
- Glanzer, M., Fischer, B. & Dorfman, D. (1984). Short-term storage in reading. *Journal of Verbal Learning and Verbal Behavior*, 23, 467-486.
- Glenberg, A. M. & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, 9, 558-565.
- Glenberg, A. M. & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language*, 31, 129-151.
- Glenberg, A. M., Meyer, M. & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, 26, 69-83.
- Glenberg, A. M., Wilkinson, A. C. & Epstein, W. (1982). The illusion of knowing: Failure in self-assessment of comprehension. *Memory & Cognition*, 10, 597-602.
- Goodman, L. A. (1954). Komogorow-smirnow-tests for psychological research. *Psychological Bulletin*, 51, 160-168.
- Gopher, D. & Browne, R. (1984). On the psychophysics of workload: Why bother with subjective measures? *Human Factors*, 26, 519-532.
- Graesser, A., Hoffman, N. & Clark, L. (1980). Structural components of reading time. *Journal of Verbal Learning and Behavior*, 19, 135-151.
- Greeno, J. G. (1983). Conceptual entities. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (p. 227-252). Hillsdale, N.J: Lawrence Erlbaum.
- Gyselinck, V. & Tardieu, H. (1994). Illustrations, mental models and comprehension of instructional text. In W. Schnotz & R. W. Kulhavy (Eds.), *Comprehension of graphics* (p. 139-153). Amsterdam: North Holland.
- Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *The Behavioral and Brain Sciences*, 6, 1-11.
- Hage, D. S. & Stroud, J. B. (1959). Reading proficiency and intelligence scores, verbal and nonverbal. *Journal of Educational Research*, 52, 258-262.
- Haist, F., Shimamura, A. P. & Squire, L. R. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18(4), 691-702.

Literaturverzeichnis

- Hamrodi, D. & Lieber, G. (2008). Was Schulbücher aus modernen Bilderbüchern lernen können... In G. Lieber (Hrsg.), *Lehren und Lernen mit Bildern* (S. 183-197). Schneider Verlag Hohengehren: Baltmannsweiler.
- Hauck, G. (2004). *Effekte von Modalitätspräferenzen beim Wissenserwerb mit multimedialen Lernsystemen*. Berlin: Logos.
- Hegarty, M., Carpenter, P. A. & Just, M. A. (1996). Diagrams in the comprehension of scientific texts. In R. Barr, M. L. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research. vol. ii* (p. 641-668). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hegarty, M. & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory & Cognition*, 22, 411-430.
- Hegarty, M. & Steinhoff, K. (1997). Individual differences in use of diagrams as external memory in mechanical reasoning. *Learning and Individual Differences*, 9, 19-42.
- Hess, D. J., Foss, D. J. & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General*, 124, 62-82.
- Horowitz, L. M., Inouye, D. & Siegelman, E. Y. (1979). On averaging judges' ratings to increase their correlation with an external criterion. *Journal of Consulting and Clinical Psychology*, 47, 453-458.
- Issing, L. J. (1994). Wissenserwerb mit bildlichen Analogien. In B. Weidenmann (Hrsg.), *Wissenserwerb mit Bildern* (S. 149-176). Bern: Huber.
- Johnson-Laird, P. (1980). Mental models in cognitive science. *Cognitive Science*, 4, 71-115.
- Johnson-Laird, P. (1983). *Mental models. Towards a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press.
- Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.
- Kade, J. & Seitter, W. (2007). Lebenslanges Lernen. In M. Göhlich, C. Wulf & J. Zirfas (Hrsg.), *Pädagogische Theorien des Lernens*. Beltz: Weinheim.
- Kalyuga, S. (2005). Prior knowledge principle. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (p. 325-337). New York: Cambridge University Press.
- Kalyuga, S. (2009). Knowledge elaboration: A cognitive load perspective. *Learning and Instruction*, 19, 402-410.
- Kalyuga, S., Ayres, P., Chandler, P. & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38, 23-31.
- Kellogg, R. T. (2001). Long-term working memory in text production. *Memory & Cognition*, 29, 43-52.

- Kendall, M. G. & Smith, B. B. (1939). The problem of m rankings. *Annals of Mathematical Statistics*, 10, 275-287.
- Kendeou, P. & Broek, P. van den. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, 35(7), 1567-1577.
- Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, N.J: Lawrence Erlbaum.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 9, 163-182.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: University Press.
- Kintsch, W. & Keenan, J. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology*, 5, 257-274.
- Kintsch, W., Kozminsky, E., Streby, W., McKoon, G. & Keenan, J. M. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, 14, 196-214.
- Kintsch, W. & van Dijk, T. (1978). Toward a model of text and production. *Psychological Review*, 85, 363-394.
- Kirby, J. R. (1994). Comprehending and using maps: Are there two modes of map processing. In W. Schnotz & R. Kulhavy (Eds.), *Comprehension of graphics* (p. 63-76). Amsterdam: Elsevier Science B.V.
- Kirby, J. R., Moore, P. J. & Schofield, N. J. (1988). Verbal and visual learning styles. *Contemporary Educational Psychology*, 13, 169-184.
- Köller, O., Rost, J. & Köller, M. (1994). Individuelle Unterschiede beim Lösen von Raumvorstellungsaufgaben aus dem IST- bzw. IST-70-Untertest "Würfelaufgaben". *Zeitschrift für Psychologie*, 202, 65-85.
- Kluwe, R. H. (1992). Gedächtnis und Wissen. In H. Spada (Hrsg.), *Allgemeine Psychologie* (Vol. 2, S. 115-187). Bern: Huber.
- Knowlton, J. Q. (1966). On the definition of picture. *Audiovisual Communication Review*, 14(2), 157-183.
- Kosslyn, S. M. (1981). The medium and the message in mental imagery: An theory. *Psychological Review*, 88, 46-66.
- Kosslyn, S. M. (1994). *Image and brain*. Cambridge, MA: MIT Press.
- Kosslyn, S. M. & Pomerantz, J. R. (1977). Imagery, propositions, and the form of internal representations. *Cognitive Psychology*, 9, 52-76.
- Kraemer, D. J. M., Rosenberg, L. M. & Thompson-Schill, S. L. (2009). The neural correlates of visual and verbal cognitive styles. *The Journal of Neuroscience*, 29, 3792-3798.

- Krajcik, J. & Varelas, M. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, 90(6), 1073-1091.
- Kulhavy, R. W., Lee, B. & Caterino, L. C. (1985). Conjoint retention of maps and related discourse. *Contemporary Educational Psychology*, 10, 28-37.
- Kulhavy, R. W., Stock, W. A. & Caterino, L. (1994). Reference maps as framework for remembering text. In W. Schnotz & R. W. Kulhavy (Eds.), *Comprehension of graphics* (p. 153-162). Amsterdam: North-Holland.
- Kulhavy, R. W., Stock, W. A. & Kealy, W. A. (1993). How geographical maps increase recall of instructional text. *Educational Technology Research and Development*, 41, 47-62.
- Kulhavy, R. W., Stock, W. A., Verdi, M. P., Rittschof, K. A. & Savenye, W. (1993). Why maps improve memory for text: The influence of structural information on working memory operations. *European Journal of Cognitive Psychology*, 5, 375-392.
- Langer, I., Schulz von Thun, F. & Tausch, R. (2002). *Sich verständlich ausdrücken*. München: Reinhardt.
- Larkin, J. H. (1983). The role of problem representation in physics. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (p. 75-98). Hillsdale, N.J: Lawrence Erlbaum.
- Larkin, J. H. & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65-99.
- Lenzner, A. (2009). *Visuelle Wissenskommunikation: Effekte von Bildern beim Lernen*. Hamburg: Verlag Dr. Kovač.
- Levene, H. (1960). Robust tests for equality of variances. In I. Olkin, S. G. Ghurye, W. H. W. G. Madow & H. B. Mann (Eds.), *Contributions to probability and statistics: Essays in honor of Harold Hotelling* (p. 278-292).
- Levie, W. H. & Lentz, R. (1982). Effects of text illustration: A review of research. *Educational Communication and Technology Journal*, 30, 195-232.
- Levin, J., Anglin, G. & Carney, R. (1987). On empirically validating functions of pictures in prose. In M. Willow & H. Houghton (Eds.), *The psychology of illustration* (p. 51-85). New York: Springer.
- Levin, J. R., Shriberg, L. K. & Berry, J. K. (1983). A concrete strategy for remembering abstract prose. *American Educational Research Journal*, 20, 277-290.
- Liepmann, D., Beauducel, A., Brocke, B. & Amthauer, R. (2007). *I-S-T 2000 R Intelligenz-Struktur-Test 2000 R*. Göttingen: Hogrefe.
- Lowe, R. K. (1994). Selectivity in diagrams: reading beyond the lines. *Educational Psychology*, 14, 467-491.
- Mani, K. & Johnson-Laird, P. N. (1982). The mental representation of spatial description. *Memory & Cognition*, 10, 181-187.

- Marsh, H. W., Hau, K.-T., Balla, J. R. & Grayson, D. (1998). Is more ever too much? the number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, 33(2), 181-220.
- Matell, M. S. & Jacoby, J. (1971). Is there an optimal number of alternatives for likert scales items? *Educational and Psychological Measurement*, 31, 657-674.
- Mautone, P. D. & Mayer, R. E. (2007). Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology*, 99, 640-652.
- Mayer, R. E. (1996). Learning strategies for making sense out of expository text: The SOI model for guiding three cognitive processes in knowledge construction. *Educational Psychological Review*, 8, 357-371.
- Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, 32, 1-19.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (p. 31-48). New York: Cambridge University Press.
- Mayer, R. E. & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84, 444-452.
- Mayer, R. E., Heiser, J. & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, 93, 187-198.
- Mayer, R. E. & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style and learning preference. *Journal of Educational Psychology*, 95, 833-846.
- 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, 312-320.
- Mayer, R. E. & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38, 43-52.
- Mayer, R. E., Moreno, R., Boire, M. & Vagge, S. (1999). Maximizing constructivistic learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology*, 91, 638-643.
- Mayer, R. E. & Sims, V. K. (1994). For whom is a picture worth a thousand words? extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, 86, 389-401.
- McNamara, D. S., Kintsch, E., Songer, N. B. & Kintsch, W. (1996). Are good texts always better? interaction of text coherence, background knowledge and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1-43.

- McNamara, T. N., Miller, D. L. & Bransford, J. D. (1996). Mental model and reading comprehension. In R. Barr, M. L. Kamil, P. Mosenthal & P. Pearson (Eds.), *Handbook of reading research*. (Vol. 2, p. 490-511). Mahwah, NJ: Lawrence Erlbaum Associates.
- McVee, M. B., Dunsmore, K. & Gavelek, J. R. (2005). Schema theory revisited. *Review of Educational Research*, 75, 531-566.
- Mielke, R. (2001). *Psychologie des Lernens - Eine Einführung*. Stuttgart: Kohlhammer.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Milligan, G. W., Wong, D. S. & Thompson, P. A. (1987). Robustness properties of nonorthogonal analysis of variance. *Psychological Bulletin*, 101, 464-470.
- Mitchell, T. W. J. (1995). *Picture theory: Essays on verbal and visual representations*. Chicago: University of Chicago Press.
- Mokros, J. R. & Tinker, R. F. (1987). The impact of microcomputer-based labs on childrens abilities to interpret graphs. *Journal of Research in Science Teaching*, 24, 369-383.
- Murphy, K. R. & Davidshofer, C. O. (2001). *Psychological testing: Principles and applications* (5 ed.). Upper Saddle River, NJ: Prentice Hall.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353-383.
- Neisser, U. (1983). The rise and fall of the sensory register. *The Behavioral and Brain Sciences*, 6, 35.
- Nelson, D. L., Reed, V. S. & Walling, J. R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 523-528.
- Nelson, T. O. (1977). Repetition and depth of processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 151-177.
- Oberauer, K., Süß, H.-M., Wilhelm, O. & Wittmann, W. W. (2000). Working memory capacity - facets of a cognitive ability construct. *Personality and Individual Differences*, 29, 1017-1045.
- Oberauer, K., Süß, H.-M., Wilhelm, O. & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, 31, 167-193.
- O'Donnel, R. & Eggemeier, F. T. (1986). Workload assessment methodology. In K. Boff, L. Kaufman & J. Thomas (Eds.), *handbook of perception and human performance* (p. 1-49). New York: Wiley.
- Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology*, 84, 429-434.

- Paas, F. G. W. C., Tuovinen, J. E. & Darabi, J. J. G. v. A. A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *Education Technology Research and Development*, 53, 25-34.
- 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, 122-133.
- Paas, F. G. W. C., van Merriënboer, J. J. G. & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, 79, 419-430.
- Paivio, A. (1967). Paired associative learning and free recall of nouns as a function of concreteness. *Psychological Reports*, 20, 239-245.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart and Winston. (Reprinted 1979, Hillsdale, NJ: Erlbaum)
- Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45, 255-287.
- Paivio, A. & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding. *Cognitive Psychology*, 5, 176-206.
- Paivio, A., Rogers, T. B. & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, 11, 137-138.
- Palmer, S. E. (1978). Fundamental aspects of cognitive representations. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (p. 259-303). Hillsdale, N.J: Lawrence Erlbaum.
- Peeck, J. (1994). Wissenserwerb mit darstellenden Bildern. In B. Weidenmann (Hrsg.), *Wissenserwerb mit Bildern* (S. 59-94). Bern: Huber.
- Peirce, C. S. (1906). Prolegomena to an apology for pragmatism. *Monist*, 16, 492-546.
- Perrig, W. & Kintsch, W. (1985). Propositional and situational representation of text. *Journal of Memory and Language*, 24, 503-518.
- Pettersson, R. (1988). Interpretation of image content. *Educational Communication Technology Journal*, 36, 45-55.
- Pinker, S. (1984). Visual cognition: An introduction. *Cognition*, 18, 1-63.
- Plass, J., Chun, D. M., Mayer, R. E. & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology*, 90, 25-36.
- Powell, C. S. (1995). Trends in der Geophysik: Innenansicht der Erde. In P. Giese (Hrsg.), *Geodynamik und Plattentektonik* (S. 10-22). Heidelberg: Spektrum Akademischer Verlag.

- Pressley, M., Levin, J. R. & Delaney, H. D. (1981). The mnemonic keyword method. *Review of Educational Research*, 52, 61-91.
- Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, 80, 1-24.
- Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, 88, 16-45.
- Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, 25, 157-238.
- Raaijmakers, J. G. W. & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, 88, 93-134.
- Rasch, B., Friese, M., Hofmann, W. & Naumann, E. (2010). *Quantitative Methoden Band 2: Einführung in die Statistik für Psychologen und Sozialwissenschaftler* (3., erweiterte Aufl.). Berlin: Springer.
- Rasch, T. (2006). *Verstehen abstrakter Sachverhalte*. Berlin: Wissenschaftlicher Verlag.
- Ratcliff, R. & McKoon, G. (1981). Does activation really spread? *Psychological Review*, 88, 454-462.
- Rayner, K. (1998). Eye movement in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.
- Rayner, K. & Pollatsek, A. (1992). Eye movement and scene perception. *Canadian Journal of Psychology*, 46, 342-376.
- Reber, A. S. (1967). Implicit learning. *Journal of Verbal Learning and Verbal Behaviour*, 6, 855-863.
- Recht, D. R. & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, 80, 16-20.
- Reigeluth, C. M., Merrill, M. D., Wilson, B. G. & Spiller, R. T. (1980). The elaboration theory of instruction: A model for sequencing and synthesizing instruction. *Instructional Science*, 9, 195-219.
- Renkl, A., Mandl, H. & Gruber, H. (1996). Inert knowledge: Analyses and remedies. *Educational Psychologist*, 31, 115-121.
- Rheinberg, F., Vollmeyer, R. & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung Aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica*, 47, 57-66.
- Richardson, A. (1977). Verbalizer-visualizer: A cognitive style dimension. *Journal of Mental Imagery*, 1, 109-126.
- Rickheit, G. & Sichelschmidt, L. (1999). Mental models: Some answers, some questions, some suggestions. In G. Rickheit & C. Habel (Eds.), *Mental models in discourse processing and reasoning* (p. 9-40). Amsterdam: Elsevier Science B.V.
- Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, 86, 638-641.

- Rost, J. (2004). *Lehrbuch Testtheorie-Testkonstruktion* (2., vollständig überarbeitete und erweiterte Aufl.). Verlag Hans Huber.
- Rumelhart, D. E. & Norman, D. A. (1978). Accretion, tuning and restructuring: Three modes of learning. In J. W. Cotton & R. L. Klatzky (Eds.), *Semantic factor in cognition* (p. 37-53). Hillsdale, N.J: Lawrence Erlbaum.
- Rumelhart, D. E. & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. Spiro & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (p. 99-135). Hillsdale, N.J: Lawrence Erlbaum.
- Sadoski, M. (1981). Right forest, wrong tree? a critique of carey, harste, and smith's research. *Reading Research Quarterly*, 16, 600-603.
- Sadoski, M., Paivio, A. & Goetz, E. T. (1991). A critique of schema theory in reading and a dual coding alternative. *Reading Research Quarterly*, 26, 463-484.
- Salomon, G. (1984). Television is easy and print is tough: The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, 76, 647-658.
- Sanchez, C. A. & Branaghan, R. J. (2009). The interaction of map resolution and spatial abilities on map learning. *International Journal of Human-Computer Studies*, 67, 475-481.
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13(3), 501-518.
- Schneider, D. W. & Dixon, P. (2009). Visuospatial cues reinstating models in working memory during interrupted reading. *Canadian Journal of Experimental Psychology*, 63, 161-172.
- Schneider, W., Körkel, J. & Weinert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison on high- and low-aptitude children. *Journal of Educational Psychology*, 81, 306-312.
- Schnotz, W. (1994a). *Aufbau von Wissenstrukturen. Untersuchungen zur Kohärenzbildung bei Wissenserwerb mit Texten*. Weinheim: Beltz PVU.
- Schnotz, W. (1994b). Wissenserwerb mit logischen Bildern. In B. Weidenmann (Hrsg.), *Wissenserwerb mit Bildern* (S. 95-147). Bern: Huber.
- Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, 14(1), 101-120.
- Schnotz, W. (2003). Informationsintegration mit Sprache und Bild. In G. Rickheit & T. Herrmann (Hrsg.), *Psycholinguistik. Ein internationales Handbuch* (S. 577-587). Berlin: Walter de Gruyter.
- Schnotz, W. (2005). A integrated model of text and picture comprehension. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (p. 49-69). New York: Cambridge University Press.

- Schnotz, W. (2006). *Pädagogische Psychologie Workbook*. Weinheim: Beltz.
- Schnotz, W. (2010). Reanalyzing the expertise reversal effect. *Instructional Science*, 38, 315-323.
- Schnotz, W. & Bannert, M. (1999). Einflüsse der Visualisierungsform auf die Konstruktion mentaler Modelle beim Text- und Bildverstehen. *Zeitschrift für experimentelle Psychologie*, 46, 217-236.
- Schnotz, W. & Bannert, M. (2003). Construction and inference in learning from multiple representations. *Learning and Instruction*, 13, 141-156.
- Schnotz, W. & Dutke, S. (2004). Kognitionspsychologische Grundlagen der Lesekompetenz: Mehrebenenverarbeitung anhand multipler Informationsquellen. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung der Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000* (S. 61-99). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schnotz, W., Fries, S. & Horz, H. (2009). Motivational aspects of cognitive load theory. In M. Wosnitza, S. A. Karabenick, A. Efklides & P. Nenniner (Eds.), *Contemporary motivation research* (p. 69-96). Göttingen: Hogrefe.
- Schnotz, W. & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychological Review*, 19, 469-508.
- Schnotz, W. & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology: Research and Development*, 53(3), 47-58.
- Schnotz, W. & Rasch, T. (2008). Functions of animation in comprehension and learning. In R. Lowe & W. Schnotz (Eds.), *Learning with animation* (p. 92-113). New York: Cambridge University Press.
- Schwartz, N. H. & Kulhavy, R. W. (1981). Map features and the recall of discourse. *Contemporary Educational Psychology*, 6, 151-158.
- Scriven, M. (1972). The concept of comprehension. In R. O. Freedle & J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge* (p. 31-39). Washington, D.C.: Winston.
- Shah, P., Mayer, R. E. & Hegarty, M. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology*, 91, 690-702.
- Shepard, R. N. & Chipman, S. (1970). Second-order isomorphism of internal representations: Shapes of states. *Cognitive Psychology*, 1(1), 1-17.
- Shepard, R. N. & Feng, C. (1972). A chronometric study of mental paper folding. *Cognitive Psychology*, 3, 228-243.
- Shepard, R. N. & Metzler, J. (1971). Mental rotation of three dimensional objects. *Science*, 171, 701-703.

- Shiffrin, R. M. & Schneider, W. (1977). Controlled and automatic human information processing: II. perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127-190.
- Simon, H. A. (1974). How big is a chunk? *Science*, 183, 482-488.
- Sims, V. K. & Mayer, R. E. (2002). Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology*, 16, 97-115.
- Spada, H., Ernst, A. M. & Ketterer, W. (1998). Lernen. In H. Spada (Hrsg.), *Allgemeine Psychologie* (2 Aufl., S. 323-372). Hans Huber: Bern.
- Spiro, R. J. (1980). Accomodative reconstruction in prose recall. *Journal of Verbal Learning and Verbal Behaviour*, 19, 84-95.
- Staresina, B. P. & Davachi, L. (2006). Differential encoding mechanisms for subsequent associative recognition and free recall. *The Journal of Neuroscience*, 26, 9162-9172.
- Steiner, G. (2006). Lernen und Wissenserwerb. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (5 Aufl., S. 137-202). Beltz: Weinheim.
- Strahan, R. F. (1980). More on averaging judges' ratings: Determining the most reliable composite. *Journal of Consulting and Clinical Psychology*, 48, 587-589.
- Stüwe, K. (2000). *Einführung in die Geodynamik der Lithosphäre*. Berlin: Springer.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257-285.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In *The cambridge handbook of multimrdia learning* (p. 19-30). New York: Cambridge University Press.
- Sweller, J. & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction*, 8, 351-362.
- Sweller, J. & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, 12, 185-233.
- Sweller, J. & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology*, 4, 434-458.
- Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. W. C. (1998). Cognitve architecture and insturctional design. *Educational Psychology Review*, 10, 251-296.
- Tergan, S.-O. (1989). Psychologische Grundlagen der Erfassung individueller Wissenspraesentationen. Teil I: Grundlagen der Wissensmodellierung. *Sprache & Kognition*, 8(3), 152-165.
- Tversky, B. (1973). Encoding processes in recognition and recall. *Cognitive Psychology*, 5, 275-287.
- Ullman, S. (1984). Visual routines. *Cognition*, 18, 97-159.

- Underwood, G. & Everatt, J. (1992). The role of eye movements in reading: Some limitations of the eye mind assumption. In E. Chekaluk & K. R. Llewellyn (Eds.), *The role of eye movements in perceptual processes* (p. 111-169). Amsterdam: North-Holland.
- van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. San Diego, CA: Academic Press.
- van Merriënboer, J., Schuurman, J., Crook, M. de & Paas, F. (2002). Redirecting learners' attention during training: effects on cognitive load, transfer test performance and training efficiency. *Learning and Instruction*, 12, 11-37.
- van Merriënboer, J. J. G. & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17, 147-177.
- Verdi, M. P., Johnson, J. T., Stock, W. A., Kulhavy, R. W. & Whitman-Ahern, P. (1997). Organized spatial displays and texts: Effects of presentation order and display type on learning outcomes. *The Journal of Experimental Education*, 65, 303-317.
- Verdi, M. P. & Kulhavy, R. W. (2002). Learning with maps and texts: An overview. *Educational Psychology Review*, 14, 27-46.
- Vicente, K. J. & Wang, J. H. (1998). An ecological theory of expertise effects in memory recall. *Psychological Review*, 105, 33-57.
- Ward, M. & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, 7, 1-39.
- Waskan, J. A. (2003). Intrinsic cognitive models. *Cognitive Science*, 27, 259-283.
- Weidenmann, B. (1988). *Psychische Prozesse beim Verstehen von Bildern*. Bern: Huber.
- Weidenmann, B. (1994a). Codes of instructional pictures. In W. Schnotz & R. W. Kulhavy (Eds.), *Comprehension of graphics* (p. 29-42). Amsterdam: Elsevier Science B.V.
- Weidenmann, B. (1994b). Informierende Bilder. In B. Weidenmann (Hrsg.), *Wissenserwerb mit Bildern* (S. 9-58). Bern: Huber.
- Weidenmann, B. (2002a). Abbilder in Multimediaanwendungen. In L. J. Issing & P. Klimsa (Hrsg.), *Informationen und Lernen mit Multimedia und Internet* (S. 83-96). Weinheim: Beltz.
- Weidenmann, B. (2002b). Multicodierung und Multimodalität. In L. J. Issing & P. Klimsa (Hrsg.), *Informationen und Lernen mit Multimedia und Internet* (S. 45-62). Weinheim: Beltz.
- Weidenmann, B. (2006). Lernen mit Medien. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (Vol. 5, S. 423-476). Weinheim: Beltz.
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt, II. *Psychologische Forschung*, 4, 301-350.
- Winn, W. D. (1990). A theoretical framework for research on learning from graphics. *International Journal of Educational Research*, 14, 553-564.

- Winn, W. D. (1994). Contributions of perceptual and cognitive processes to the comprehension of graphics. In W. Schnotz & R. W. Kulhavy (Eds.), *Comprehension of graphics* (p. 3-27). Amsterdam: North Holland.
- Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist*, 24, 345-376.
- Wolfe, M. B. W., Schreiner, M. E., Rehder, B., Laham, D., Foltz, P. W., Kintsch, W. et al. (1998). Learning from text: Matching readers and texts by latent semantic analysis. *Discourse Processes*, 25, 309-336.
- Wygotski, L. (1964). *Denken und Sprechen*. Berlin: Akademie Verlag.
- Yuan, K., Steedle, J., Shavelson, R., Alonzo, A. & Oppizzo, M. (2006). Working memory, fluid intelligence and science learning. *Educational Research Review*, 1, 83-98.
- Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 44, p. 35-62). New York: Academic Press.
- Zwaan, R. A. & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162-185.

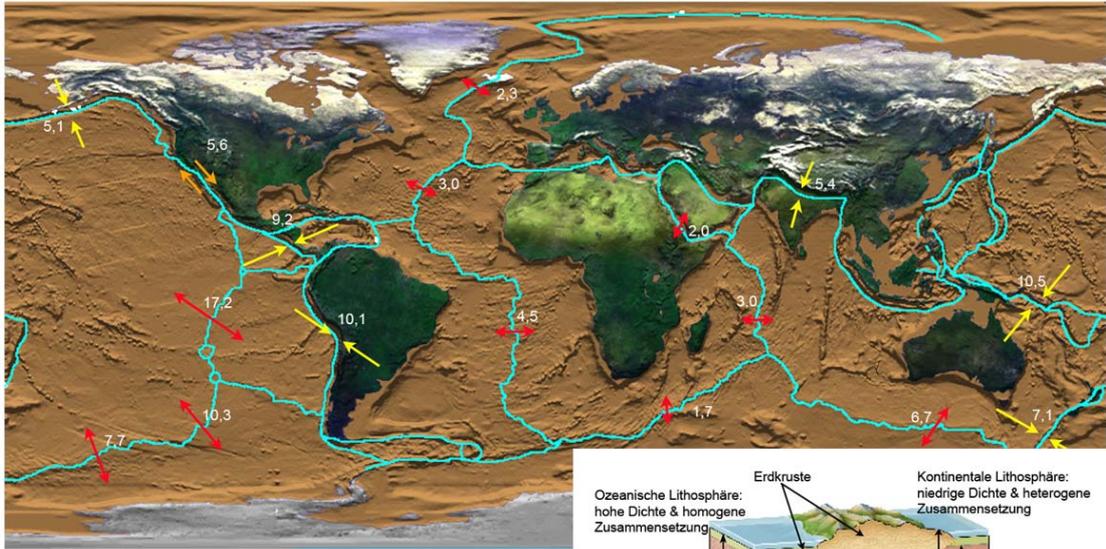
Teil IV

Anhang

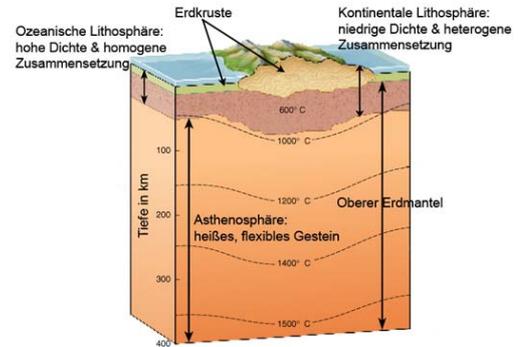
A Lernmaterial

A.1 Kongruente Bilder & Texte

Geodynamik und Plattentektonik



Zeichenerklärung



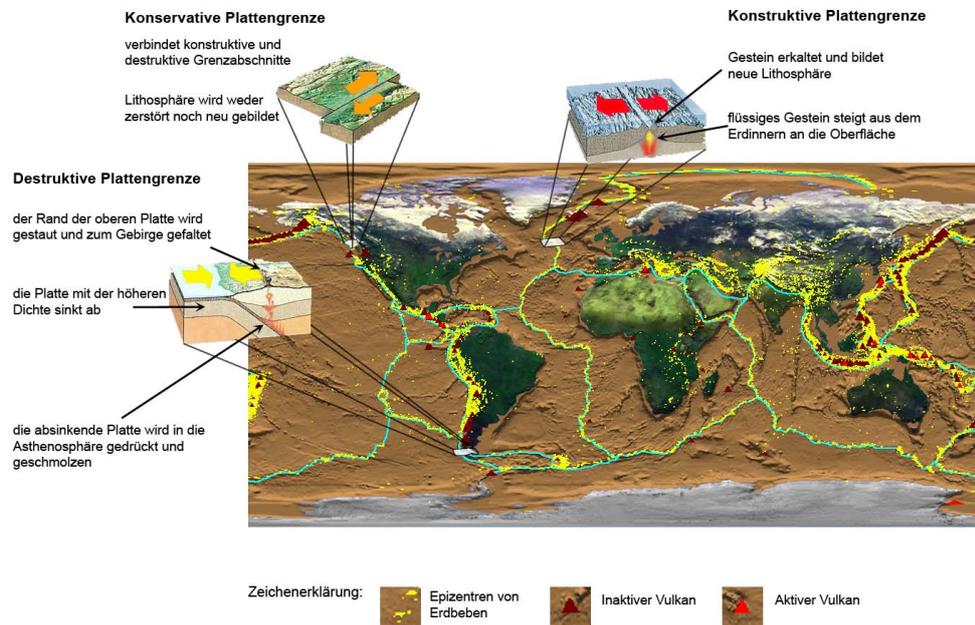
Geodynamik und Plattentektonik

Die Geodynamik befasst sich mit großräumigen Bewegungsvorgängen im Erdinnern und auf der Erdoberfläche. Zugleich erforscht sie die Prozesse, Antriebsmechanismen und Kräfte, mit denen die einzelnen Phänomene in Zusammenhang stehen. Die bekannteste Theorie der Geodynamik ist die Plattentektonik, mit deren Hilfe geologische Phänomene an der Erdoberfläche und im Erdmantel beschrieben werden.

Im Zentrum der Plattentektonik steht die Annahme, dass die Oberfläche der Erde von einer Schicht aus festem Gestein, der Lithosphäre, umhüllt ist. Die Lithosphäre umfasst die Erdkruste und Teile des Erdmantels. Sie bildet aber keine geschlossene Schicht, sondern besteht aus sieben großen und zahlreichen kleinen voneinander getrennten Platten. Neben Größe und Form unterscheiden sich diese Platten vor allem darin, ob sie unter einem Kontinent oder einem Ozean liegen. So sind Kontinentalplatten durchschnittlich dicker, in ihrer Zusammensetzung weit weniger homogen als ozeanische Platten und haben eine deutlich geringere physikalische Dichte ($2,7 \text{ g/m}^3 < 3,2 \text{ g/m}^3$).

Satellitenmessungen haben ergeben, dass sich diese kilometerdicken, gewaltigen Gesteinsplatten um mehrere cm im Jahr bewegen. Je nach Abschnitt, Plattengrenze und den anzutreffenden geologischen Bedingungen werden mit Hilfe von GPS Positionsveränderungen von jährlich circa 2 bis 17 cm gemessen. Nach der Theorie der Plattentektonik können sich die gewaltigen Lithosphärenplatten bewegen, weil sie auf einer Schicht aus sehr heißem und damit im geologischen Sinne „weichen“ Gestein liegen. Diese als Asthenosphäre bezeichnete Gesteinsschicht enthält im Vergleich zu der kalten Lithosphäre geringe Anteile von Schmelzmaterial (ca 3%). Die hohe Temperatur und das geschmolzene Gestein machen die Asthenosphäre flexibel und verformbar. Stark vereinfacht könnte man sagen, dass die Lithosphärenplatten auf der Asthenosphäre schwimmen.

Plattengrenzen



Plattengrenzen und ihre geologische Bedeutung

Den Plattengrenzen kommt im Rahmen der Geodynamik besondere Bedeutung zu, da die meisten geologischen Phänomene an Plattengrenzen auftreten. Man unterscheidet drei unterschiedliche Grenztypen:

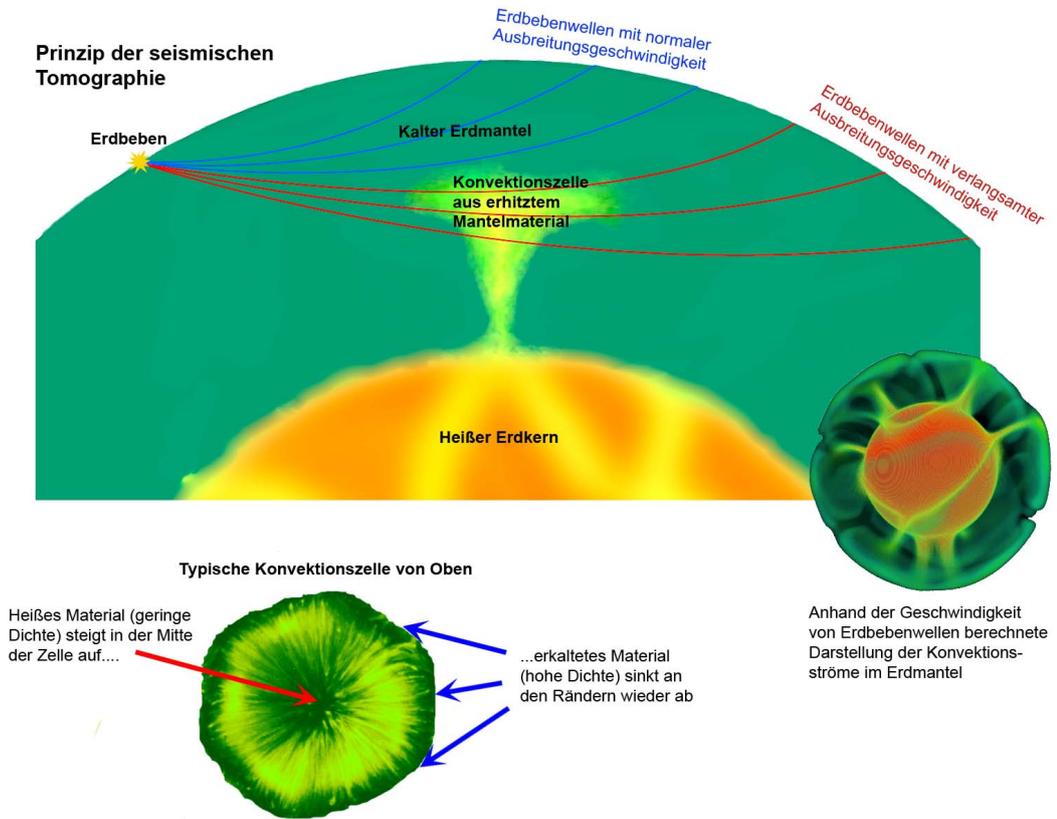
An konstruktiven Plattengrenzen driften zwei Platten auseinander, so dass ein Riss in der Kruste entsteht. Durch diesen Riss steigt geschmolzenes Gestein (Magma) aus dem Erdinneren an die Oberfläche, erkaltet und lagert sich an den Rändern ab. Auf diese Weise entsteht neue Lithosphäre. Die meisten konstruktiven Plattengrenzen lassen sich in der Tiefsee finden und werden aufgrund der Formen des abgelagerten Gesteins auch als Mittelozeanische Rücken bezeichnet.

An destruktiven Plattengrenzen bewegen sich zwei Platten aufeinander zu. Die Platte mit der höheren Dichte sinkt ab und wird von der leichteren Platte, die sich über die Platte mit der hohen Dichte schiebt, in Richtung Mantel gedrückt (subduziert). Die subduzierte Platte wird in den heißen sub-lithosphärischen Schichten allmählich abgeschmolzen und wieder in den Mantel integriert. Der Rand der oberen Platte wird hingegen gestaucht und gefaltet, so dass häufig Gebirge (z.B. die Anden) oder Inseln (z.B. Island) entstehen.

An konservativen Plattengrenzen gleiten zwei Platten aneinander vorbei, ohne dass Lithosphäre zerstört oder neu gebildet wird. Diese häufig auch als Transformstörung bezeichneten Plattenränder sind Verwerfungen oder Verschiebungen innerhalb konstruktiver oder destruktiver Grenzverläufe. Das prominenteste Beispiel für eine konservative Plattengrenze ist der St. Andreasgraben in Kalifornien.

Die Relativbewegung zwischen den Platten verläuft aber weder spannungsfrei noch kontinuierlich. Wenn sich die Platten verhaken, entstehen gewaltige Spannungen innerhalb des Gesteins, die sich ruckartig an den Plattenrändern entladen. Trägt man die Verteilung der Erdbeben-Epizentren der letzten zehn Jahre auf einer Weltkarte ab, so werden eindrucksvoll die Plattengrenzen nachgezeichnet.

Messungen im Erdinneren



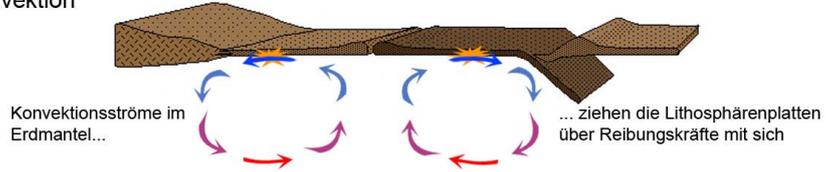
Messungen im Erdinneren

Um die Ursachen für verschiedene geologische Phänomene zu verstehen, ist es nötig, einen Blick in das Innere der Erde zu werfen. Aus diesem Grund bedienen sich die Geologen der seismischen Tomographie. Mit dieser Methode wird der Verlauf einer großen Anzahl von seismischen Wellen, wie sie bei Erdbeben entstehen, im dreidimensionalen Raum des Erdmantels erfasst und gemessen. Durch den Vergleich der Laufzeiten kann geschlossen werden, ob die Wellen in dem betrachteten Bereich die erwartete Geschwindigkeit aufweisen, beschleunigt oder verzögert werden. Ist das Gestein heißer, werden die seismischen Wellen verzögert, ist es kühler, werden sie beschleunigt. Die auf diese Weise erzeugten dreidimensionalen Wärmebilder der Erde zeigen, dass es im Erdmantel offensichtlich Konvektionsströmungen gibt.

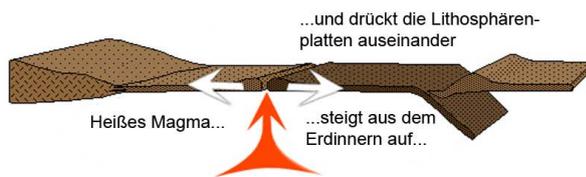
Geologen erklären sich die thermischen Strömungen damit, dass Gesteinsmaterial in der Nähe des extrem heißen Erdkerns erhitzt wird, dabei bei konstanter Masse an Volumen zunimmt und so seine Dichte verringert. Wie bei einer kochenden Flüssigkeit die einzelnen heißen Teilchen aufsteigen, steigt auch das heiße Gestein nach oben, da es im Vergleich zur Umgebung mit hoher Dichte einen sehr viel stärkeren Auftrieb besitzt. An der kalten Oberfläche kühlt das Gestein dann wieder ab und sinkt als erkaltetes Material in die tieferen Schichten des Erdmantels zurück. Blickt man von oben auf eine Konvektionsströmung, so bilden sich bei sehr vielen Stoffen hoch strukturierte und symmetrische Formen aus, die wie ein Bienenwabenmuster aussehen. Auch wenn die Konvektionszellen im Erdmantel keine symmetrische Form annehmen, so steigt das heiße Material in der Mitte der Zelle auf und sinkt, nachdem es abgekühlt ist, an den Rändern wieder ab.

Mechanismen der Plattenbewegung

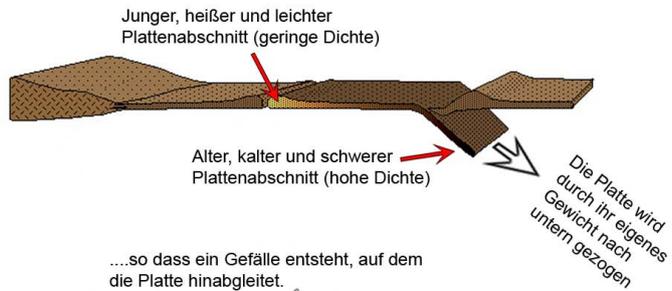
Mantelkonvektion



Rückendruck



Plattenzug



Hangabtrieb



Mechanismen der Plattenbewegung

Eines der zentralen Anliegen der Geodynamik ist es, die Ursache für das Driften der Platten zu erklären. Einige Wissenschaftler sehen einen direkten Zusammenhang mit dem Konvektionsstrom im Erdmantel. So wird angenommen, dass das heiße Gesteinsmaterial, während es an der Lithosphärenplatte entlang fließt, diese über Reibungskräfte regelrecht mit sich zieht, bevor es genug Wärme abgegeben hat und wieder in die tieferen Schichten des Mantels zurück sinkt.

Auch ein zweiter Erklärungsansatz geht zumindest indirekt auf die Konvektionsströme im Erdmantel zurück. An konstruktiven Plattengrenzen bringen aufwärtsgerichtete Ströme geschmolzenes Gesteinsmaterial an die Erdoberfläche. Das aufsteigende Magma erzeugt dabei so lange einen seitlich auf die Lithosphäre übertragenen Druck bis die noch junge und dünne Gesteinsschicht ruckartig aufbricht und das Magma austreten kann. Die Platten werden an konstruktiven Grenzen durch aufsteigendes Mantelmaterial quasi auseinandergedrückt.

Die meisten Anhänger hat allerdings eine Theorie, die Gravitationskräfte als den entscheidenden Antriebsmechanismus der Plattenbewegung sieht. Denn während sich die ozeanische Lithosphärenplatte von den Mittelozeanischen Rücken entfernt und dabei auskühlt, nehmen ihre Mächtigkeit und Dichte stetig zu. Aus diesem Grund sinkt bei einer Kollision zweier ozeanischer Platten die ältere der beiden aufgrund ihrer höheren Dichte in den Mantel ein. Hat das Absinken dann begonnen, wird die Platte durch ihr eigenes Gewicht tiefer gezogen.

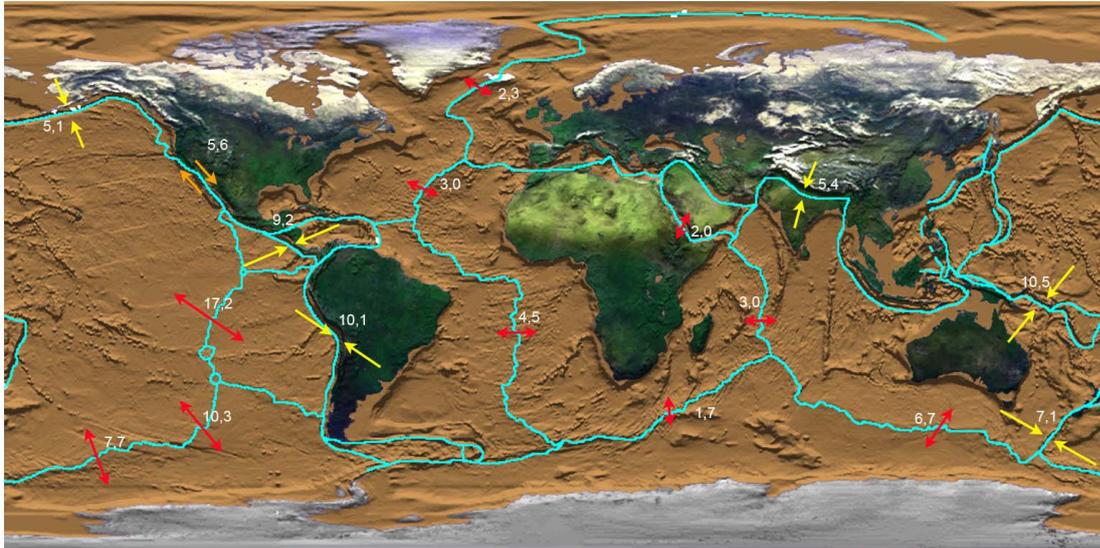
Andere Wissenschaftler hingegen glauben, dass der Aufstrom heißen Materials aus dem Mantel die noch junge Lithosphäre an den konstruktiven Plattenrändern anhebt. Auf diese Weise entsteht ein Gefälle, auf dem die Lithosphärenplatte hinabgleitet.

Aktuelle Modelle postulieren das Zusammenwirken von mehreren Kräften, deren Bedeutung in Abhängigkeit des jeweiligen Modells stark variiert.

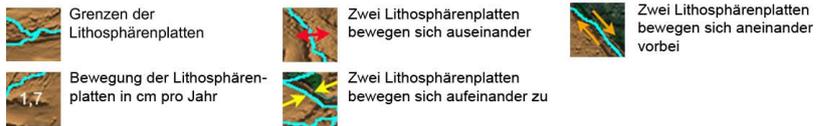
A Lernmaterial

A.2 Komplementäre Bilder & Texte

Geodynamik und Plattentektonik



Zeichenerklärung



Geodynamik und Plattentektonik

Die Geodynamik befasst sich mit großräumigen Bewegungsvorgängen im Erdinnern und auf der Erdoberfläche. Zugleich erforscht sie die Prozesse, Antriebsmechanismen und Kräfte, mit denen die einzelnen Phänomene in Zusammenhang stehen. Die bekannteste Theorie der Geodynamik ist die Plattentektonik, mit deren Hilfe geologische Phänomene an der Erdoberfläche und im Erdmantel beschrieben werden.

Im Zentrum der Plattentektonik steht die Annahme, dass die Oberfläche der Erde von einer Schicht aus festem Gestein, der Lithosphäre, umhüllt ist. Die Lithosphäre umfasst die Erdkruste und Teile des Erdmantels. Neben Größe und Form unterscheiden sich diese Platten vor allem darin, ob sie unter einem Kontinent oder einem Ozean liegen. So sind Kontinentalplatten durchschnittlich dicker, in ihrer Zusammensetzung weit weniger homogen als ozeanische Platten und haben eine deutlich geringere physikalische Dichte ($2,7 \text{ g/m}^3 < 3,2 \text{ g/m}^3$).

Nach der Theorie der Plattentektonik liegen die gewaltigen Lithosphärenplatten auf einer Schicht aus sehr heißem und damit im geologischen Sinne „weichen“ Gestein. Diese als Asthenosphäre bezeichnete Gesteinsschicht enthält im Vergleich zu der kalten Lithosphäre geringe Anteile von Schmelzmaterial (ca 3%). Die hohe Temperatur (1000 - 1500 °C) und das geschmolzene Gestein machen die Asthenosphäre flexibel und verformbar. Stark vereinfacht könnte man sagen, dass die Lithosphärenplatten auf der Asthenosphäre schwimmen.

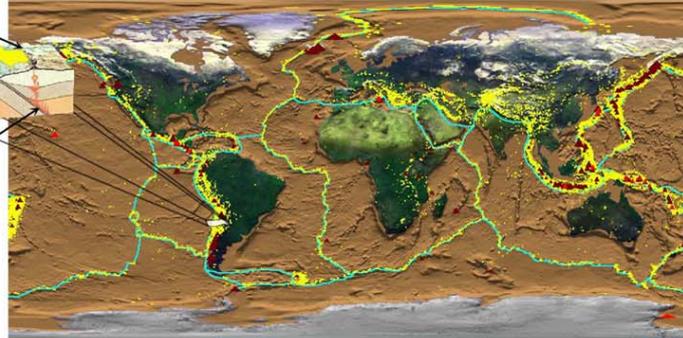
Plattengrenzen

Beispiel: Destruktive Plattengrenze

der Rand der oberen Platte wird gestaut

die Platte mit der höheren Dichte sinkt ab

die absinkende Platte wird in die Asthenosphäre gedrückt und geschmolzen



Zeichenerklärung:



Epizentren von Erdbeben



Inaktiver Vulkan



Aktiver Vulkan

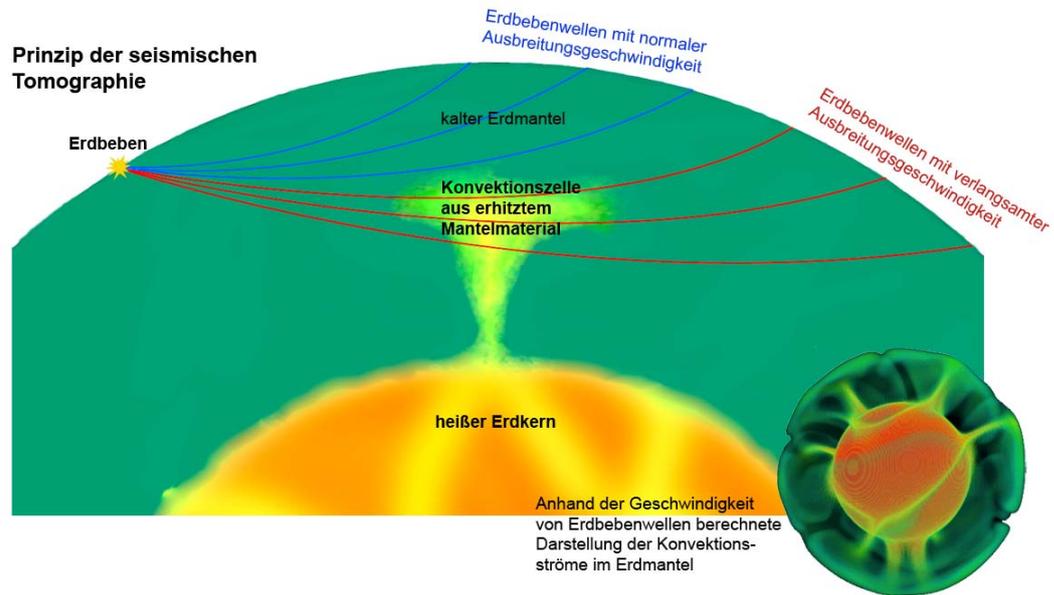
Plattengrenzen und ihre geologische Bedeutung

Den Plattengrenzen kommt im Rahmen der Geodynamik besondere Bedeutung zu, da die meisten geologischen Phänomene an Plattengrenzen auftreten. Man unterscheidet folgende Grenztypen:

An konstruktiven Plattengrenzen driften zwei Platten auseinander, so dass ein Riss in der Kruste entsteht. Durch diesen Riss steigt geschmolzenes Gestein (Magma) aus dem Erdinneren an die Oberfläche, erkaltet und lagert sich an den Rändern ab. Auf diese Weise entsteht neue Lithosphäre. Die meisten konstruktiven Plattengrenzen lassen sich in der Tiefsee finden und werden aufgrund der Formen des abgelagerten Gesteins auch als Mittelozeanische Rücken bezeichnet.

An konservativen Plattengrenzen gleiten zwei Platten aneinander vorbei, ohne dass Lithosphäre zerstört oder neu gebildet wird. Diese häufig auch als Transformstörung bezeichneten Plattenränder sind Verwerfungen oder Verschiebungen innerhalb konstruktiver oder destruktiver Grenzverläufe. Das prominenteste Beispiel für eine konservative Plattengrenze ist der St. Andreasgraben in Kalifornien.

Messungen im Erdinneren



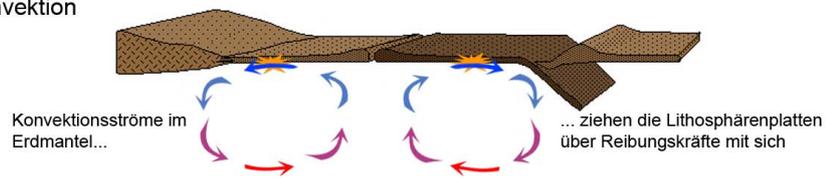
Messungen im Erdinneren

Um die Ursachen für verschiedene geologische Phänomene zu verstehen, ist es nötig, einen Blick in das Innere der Erde zu werfen. Mittels komplexer Messverfahren konnte so festgestellt werden, dass es im Erdmantel Konvektionsströmungen gibt.

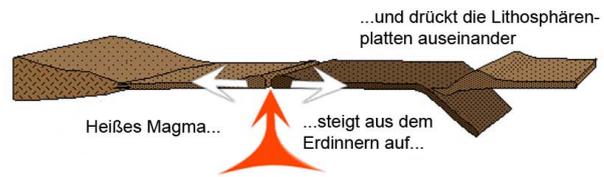
Geologen erklären sich diese thermischen Strömungen damit, dass Gesteinsmaterial in der Nähe des extrem heißen Erdkerns erhitzt wird, dabei bei konstanter Masse an Volumen zunimmt und so seine Dichte verringert. Wie bei einer kochenden Flüssigkeit die einzelnen heißen Teilchen aufsteigen, steigt auch das heiße Gestein nach oben, da es im Vergleich zur Umgebung mit hoher Dichte einen sehr viel stärkeren Auftrieb besitzt. An der kalten Oberfläche kühlt das Gestein dann wieder ab und sinkt als erkaltetes Material in die tieferen Schichten des Erdmantels zurück. Blickt man von oben auf eine Konvektionsströmung, so bilden sich bei sehr vielen Stoffen hoch strukturierte und symmetrische Formen aus, die wie ein Bienenwabenmuster aussehen. Auch wenn die Konvektionszellen im Erdmantel keine symmetrische Form annehmen, so steigt das heiße Material in der Mitte der Zelle auf und sinkt, nachdem es abgekühlt ist, an den Rändern wieder ab.

Mechanismen der Plattenbewegung

Mantelkonvektion



Rückendruck



Mechanismen der Plattenbewegung

Eines der zentralen Anliegen der Geodynamik ist es, die Ursache für das Driften der Platten zu erklären. Die meisten Anhänger hat eine Theorie, die Gravitationskräfte als den entscheidenden Antriebsmechanismus der Plattenbewegung sieht. Denn während sich die ozeanische Lithosphärenplatte von der konstruktiven Plattengrenze entfernt und dabei auskühlt, nehmen ihre Mächtigkeit und Dichte immer weiter zu. Aus diesem Grund sinkt bei einer Kollision zweier ozeanischer Platten die ältere der beiden aufgrund ihrer höheren Dichte in den Mantel ein. Hat das Absinken dann begonnen, wird die Platte durch ihr eigenes Gewicht immer tiefer gezogen.

Andere Wissenschaftler hingegen glauben, dass der Aufstrom heißen Materials aus dem Mantel die noch junge Lithosphäre an den konstruktiven Plattenrändern anhebt. Auf diese Weise entsteht ein Gefälle, auf dem die Lithosphärenplatte hinabgleitet.

Aktuelle Modelle postulieren das Zusammenwirken von mehreren Kräften, deren Bedeutung in Abhängigkeit des jeweiligen Modells stark variiert.

B Instrumente

B.1 Wissenstest

Wissenstest

Es folgen einige Fragen zu den Texten und Bildern, die Sie gerade am Computer bearbeitet haben. Bei Fragen mit vorgegebenen Antwortalternativen ist immer genau eine Antwort richtig. Markieren Sie die Ihrer Meinung nach richtige Antwortalternative mit einem Kreuz in dem entsprechenden Kästchen. Dabei bitten wir Sie, nicht zu raten. Wissen Sie die Antwort wirklich nicht, dann lassen Sie bitte die Frage aus. Bei offenen Fragen reichen zur Beantwortung Stichworte.

1. Wie nennt man die bekannteste Theorie der Geodynamik?

2. Welche der folgenden Aussagen trifft zu?
 - Die Erdkruste ist Teil der Lithosphäre
 - Die Erdkruste ist Teil des oberen Erdmantels
 - Die Erdkruste ist durchschnittlich 80 km dick

3. Wie hoch ist der Anteil (in %) geschmolzenen Gesteins in der Asthenosphäre?

4. Wie viele große Lithosphärenplatten gibt es?

5. Warum können sich die gewaltigen Gesteinplatten bewegen?
 - weil die darunter liegende Schicht aus flüssigem Gestein besteht
 - weil die darunter liegende Schicht heiß ist
 - weil die darunter liegende Schicht weich und verformbar ist

6. Nennen sie zwei Eigenschaften, in denen sich die kontinentale und die ozeanische Lithosphäre unterscheiden?

7. Um wie viele cm bewegen sich die Gesteinsplatten jährlich?
- 1,7 bis 19,4 cm
 - 4 ,1 bis 17,2 cm
 - 1,7 bis 17,2 cm
8. Welche Temperaturen herrschen in der Asthenosphäre?
- zwischen 1000 und 2000° C
 - zwischen 1000° und 1500° C
 - zwischen 600° und 1500°C
9. An welchem Kontinent gleiten zwei Platten aneinander vorbei?
- Afrika
 - Nordamerika
 - Südamerika
10. Warum sind in der Geodynamik die Plattengrenzen besonders wichtig?
11. An welcher der drei Plattengrenzen entsteht neue Lithosphäre?
- Konservative Plattengrenze
 - Destruktive Plattengrenze
 - Konstruktive Plattengrenze
12. Wie werden konservative Plattengrenzen noch bezeichnet?
13. Wo findet man die meisten konstruktiven Plattengrenzen?
14. Auf welche Weise wird an destruktiven Plattengrenzen Lithosphäre zerstört?

B Instrumente

15. An welcher Plattengrenze gibt es mehr Erdbeben?

- Konstruktive Plattengrenze
- Destruktive Plattengrenze

16. Gibt es Vulkane nur an Plattengrenzen?

- ja
- nein

17. Wie bezeichnet man die Methode, mit der Messungen im Erdinneren gemacht werden können?

18. Warum ist es gerade für die Geologie wichtig, Messungen im Erdinneren vorzunehmen?

19. Wenn Erdbebenwellen durch heißes Gestein wandern werden sie....

- langsamer
- schneller
- Erdbebenwellen haben immer eine konstante Geschwindigkeit

20. Nennen sie die Energiequelle, welche die Konvektionsströme im Erdmantel antreibt?

21. Bilden Konvektionszellen im Erdmantel symmetrische Formen aus?

22. Wo steigen bei einer Konvektionszelle die heißen Teilchen auf?

- am Rand
- in der Mitte
- am Rand und in der Mitte

23. Welche Aussage trifft zu. Heißes Gestein...

- beschleunigt seismische Wellen
- dehnt sich aus
- nimmt an Masse zu

24. Wie viel Ecken hat eine typische Konvektionszelle?

- 5
- 6
- 8

25. Wie viele verschiedene Erklärungsansätze für das Driften der Platten werden in Text und Bild genannt?

26. Wenn bei der Kollision zweier ozeanischer Platten, die eine absinkt, so geschieht das, weil die subduzierte Platte....

- dicker...
- größer...
- älter ist

27. Warum sinkt eine subduzierte Platte immer weiter ab?

28. Durch welche Kräfte wirkt die Konvektionsströmung im Erdmantel direkt auf die Plattenbewegung ein?

29. An welchen Plattengrenzen wirkt der Mechanismus des Plattendrucks?

- Konservative Plattengrenze
- Destruktive Plattengrenze
- Konstruktive Plattengrenze

30. Bei welchen beiden Erklärungsansätzen sehen Forscher die Schwerkraft als entscheidende Kraft hinter der Plattenbewegung? (**2 Antworten!!!**)

- Rückendruck
- Konvektionsströmung im Erdmantel
- Plattenzug
- Hangabtrieb

31. Kreuzen Sie bitte die Kräfte der Plattenbewegung an, die unwirksam werden, wenn der Erdkern erkaltet (**mehrere Antworten möglich!!!**)

- Rückendruck
- Konvektionsströmung im Erdmantel
- Plattenzug
- Hangabtrieb

32. Wie Sie gelernt haben, treten Erdbeben häufig dann auf, wenn sich ineinander verhakte Lithosphärenplatten ruckartig voneinander lösen. An welchen Plattengrenzen sind daher die Erdbeben mit der geringsten Stärke zu beobachten?

- Konservative Plattengrenze
- Destruktive Plattengrenze
- Konstruktive Plattengrenze

33. Nehmen Sie an, eine Erdbebenwelle bewegt sich von der äußeren Kruste exakt auf den Mittelpunkt der Erde zu
- Sie wird in Abhängigkeit der Konvektionsströme tendenziell langsamer
 - Sie bewegt sich mit konstanter Geschwindigkeit auf den Erdkern zu
 - Sie wird in Abhängigkeit der Konvektionsströme tendenziell schneller
34. Erdbeben sind vor allem dann gefährlich, wenn sie sich nahe der Erdoberfläche und bewohnter Gebiete ereignen. Welche der drei Plattengrenzen macht dieser Umstand besonders gefährlich?
- Konservative Plattengrenzen
 - Destruktive Plattengrenzen
 - Konstruktive Plattengrenzen
35. Nehmen wir an, die Konvektionsströmung ist der einzige Mechanismus, der für die Bewegung der Lithosphärenplatten verantwortlich ist. Was würde an einer konstruktiven Plattengrenze geschehen, wenn sich die Strömungsrichtung der Konvektionszelle umkehrt?
- die Platten würden sich langsamer voneinander weg bewegen
 - die Platten würden sich schneller voneinander weg bewegen
 - die Platten würden sich gar nicht mehr bewegen
 - die Platten würden sich aufeinander zu bewegen und kollidieren
36. Stellen Sie sich vor, in 10 Millionen Jahren wollen Sie mit einer heutzutage üblichen Verkehrsmaschine von Frankfurt nach New York fliegen. Wie lange dauert der Flug?
- länger
 - kürzer
 - gleich lang

37. Jemand vertritt die These, dass die Ozeane kontinuierlich größer werden.

- die These trifft zu, da der Meeresspiegel ansteigt
- die These trifft zu, weil die ozeanischen Platten schneller wachsen, als dass sie abgebaut werden
- die These trifft nicht zu, da kontinentale Platten von ozeanischen Platten zwar gestaut, aber nicht subduziert werden
- die These trifft nicht zu, da ozeanische Platten genauso schnell wachsen, wie sie abgebaut werden

38. Mitten im Ozean entsteht in der Nähe einer Plattengrenze eine Inselgruppe.

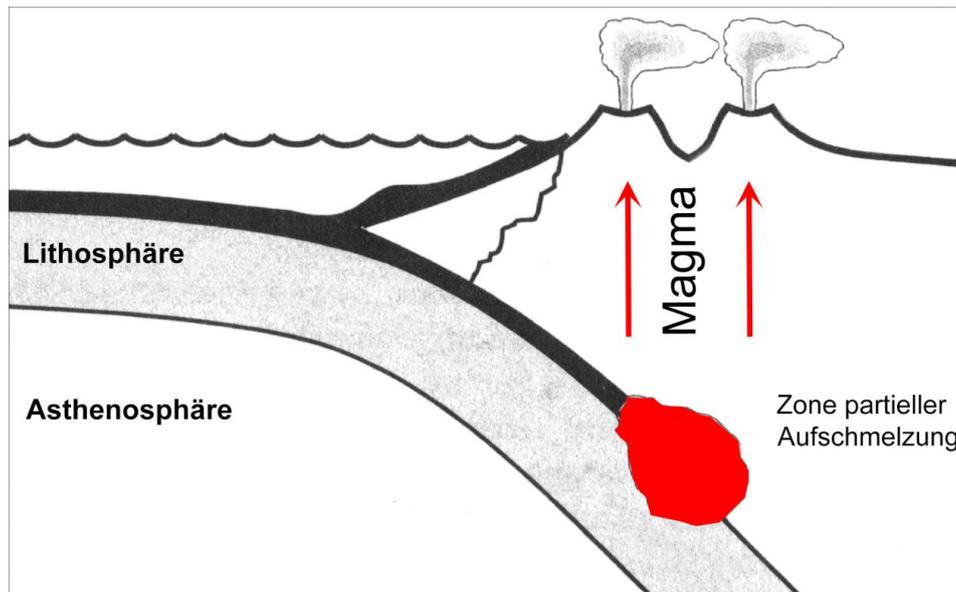
Um was für eine Plattengrenze muss es sich dabei handeln?

- Konservative Plattengrenze
- Destruktive Plattengrenze
- Konstruktive Plattengrenze

39. Nehmen wir an, eine ozeanische und eine kontinentale Lithosphärenplatte kollidieren miteinander. Die kontinentale Platte ist älter als die ozeanische Platte. Was geschieht?

- die beiden Platten stauen sich gegenseitig und es entsteht ein Küstengebirge
- die kontinentale Platte wird subduziert
- die ozeanische Platte wird subduziert

40. Wie man an in dem Bild erkennen kann, entstehen an destruktiven Plattengrenzen Vulkane, da Teile der subduzierten Platte abgeschmolzen werden und durch Erdspalten als Magma nach oben wandern.



Erklären Sie bitte mit wenigen Stichworten, warum das Magma durch Erdspalten an die Oberfläche steigt und dort als Lava austritt.

Itemkennwerte für die finale Version des Wissenstests

Item	Skala	p	r _{it}	κ	S _{ei}
01	freies Erinnern	.57	.397	.978	.401
02	Wiederkennen	.71	.207	-	.227
03	freies Erinnern	.43	.342	1.000	.345
04	freies Erinnern	.32	.311	.917	.333
05**	Inferenz-MC	.35	-.016	-	-
06	freies Erinnern	.53	.405	.980	-
07	Wiederkennen	.69	.163	-	.176
08**	Weiderkennen	.47	-.038	-	-
09	Wiederkennen	.65	.184	-	.193
10	Inferenz offen	.32	.294	.883	.315
11	Wiederkennen	.75	.186	-	.215
12	freies Erinnern	.02	.133	.954	.475
13	Inferenz offenen	.34	.189	.932	.198
14	freies Erinnern	.42	.207	.980	.209
15	Wiederkennen	.64	.177	-	.184
16	Wiedererkennen	.57	.142	-	.143
17	freies Erinnern	.09	.157	.957	.274
18	Inferenz offen	.38	.263	.602	.259
19*	Wiederkennen	.77	.046	-	.055
20	Freies Erinnern	.51	.190	1.000	.190
21	Freies Erinnern	.34	.289	.945	.305
22	Wiederkennen	.75	.290	-	.335
23	Inferenz-MC	.65	.332	-	.350
24	Wiederkennen	.60	.232	-	.237
25***	freies Erinnern				
26	Wiedererkennen	.58	.288	-	.293
27	freies Erinnern	.62	.262	.771	.270
28	freies Erinnern	.21	.240	.981	.294
29	Wiederkennen	.25	.168	-	.194
30	Inferenz-MC	.26	.156	-	.178
31	Inferenz-MC	.11	.139	-	.224
32	Inferenz-MC	.32	.284	-	.300
33	Inferenz-MC	.69	.366	-	.389
34	Inferenz-MC	.29	.289	-	.320
35	Inferenz-MC	.58	.263	-	.265
36	Inferenz-MC	.49	.215	-	.217
37*	Inferenz-MC	.40	.086	-	.087
38	Inferenz-MC	.40	.142	-	.147
39	Inferenz-MC	.24	.245	-	.286
40	Inferenz offen	.07	.078	.791	.158

Anmerkungen: p=Schwierigkeit; r_{it}=Trennschärfe; κ = Übereinstimmungsmaß bei der Auswertung; S_{ei} = Selektionkriterium; *= Item wurde entfernt, da Sel < .1; ** = Item wurde wegen negativer Trennschärfe entfernt; *** Item wurde aufgrund eines konzeptionellen Fehlers entfernt

B.2 Weitere Instrumente

B Instrumente

Skalen zur Erfassung der mentalen Belastung (cognitive load) und der empfundenen Schwierigkeit ...

nach der Lernphase:

Wie stark haben Sie sich beim Lernen der Texte und Bilder angestrengt?

sehr wenig sehr stark

Wie schwierig fanden Sie es, die Textinformationen zu lernen?

sehr leicht sehr schwierig

Wie schwierig fanden Sie es, die Bildinformationen zu lernen?

sehr leicht sehr schwierig

nach dem Wissenstest:

Wie stark haben Sie sich bei der Bearbeitung der Fragen angestrengt?

sehr wenig sehr stark

Wie schwer fanden Sie es die Fragen des Wissenstests zu beantworten?

sehr leicht sehr schwierig

Vorwissenstest

Fragen zum Thema Geologie

1. Die Lithosphäre ist eine Schicht aus festem Gestein, die aus der Erdkruste und Teilen des oberen Erdmantels besteht.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
2. Es gibt 9 große und etliche kleine Kontinentalplatten	Richtig Falsch	<input type="radio"/> <input type="radio"/>
3. Konvektionsströme sind zirkuläre Strömungen, die kalte Materie nach oben befördern und heiße Materie absinken lassen.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
4. Als konservativ bezeichnet man Plattengrenzen, die aneinander vorbeigleiten.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
5. An destruktiven Grenzen bewegen sich zwei Kontinentalplatten aufeinander zu.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
6. Kontinentalplatten bewegen sich mit einer Geschwindigkeit von bis zu 2,5 m im Jahr	Richtig Falsch	<input type="radio"/> <input type="radio"/>
7. Geodynamik ist eine wissenschaftliche Disziplin, die sich mit geologischen Phänomenen, wie Vulkanismus und Erdbeben beschäftigt.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
8. Die Kontinentalplatten liegen auf einer Schicht aus flüssigem Gestein.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
9. Das Abtauchen einer lithosphärischen Platte in die Asthenosphäre bezeichnet man als Subduktion.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
10. An der Stelle, an der eine Lithosphärenplatte unter eine andere Platte abtaucht, entsteht ein Tiefseegebirge.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
11. Erdbebenwellen breiten sich mit konstanter Geschwindigkeit im Erdmantel aus.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
12. Konvektionszellen nehmen häufig die Form von Bienenwaben an.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
13. Die modernsten Theorien gehen davon aus, dass die Konvektionsströme im Erdmantel für das Driften der Platten verantwortlich sind.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
14. Der Motor der Konvektionsströme im Erdmantel ist der heiße Erdkern.	Richtig Falsch	<input type="radio"/> <input type="radio"/>
15. Die meisten Erdbeben gibt es an Plattengrenzen	Richtig Falsch	<input type="radio"/> <input type="radio"/>

B Instrumente

Adaptierte Version des Fragebogens zur Erfassung der aktuellen Motivation (FAM, Rheinberg, Vollmeyer & Burns, 2001)

Fragen zur aktuellen Motivation:

1. Das Lernen mit den Texten und Bildern hat mir Spaß gemacht.
trifft nicht zu trifft zu

2. Ich bin mir sicher, die Zusammenhänge verstanden zu haben.
trifft nicht zu trifft zu

3. Ich bin fest entschlossen, mich bei dem Lerntest voll anzustrengen.
trifft nicht zu trifft zu

4. Ich fand das Lernen anstrengend.
trifft nicht zu trifft zu

5. Ich habe keinerlei Schwierigkeiten, meine Gedanken beisammen zu halten.
trifft nicht zu trifft zu

6. Der kommende Lerntest ist eine richtige Herausforderung für mich.
trifft nicht zu trifft zu

7. Ich glaube, ich werde schlecht bei dem kommenden Lerntest abschneiden.
trifft nicht zu trifft zu

8. Das Lernthema war interessant.
trifft nicht zu trifft zu

9. Eigentlich möchte ich am liebsten aufhören.
trifft nicht zu trifft zu

10. Als ich die Texte und Bilder bearbeitet habe, merkte ich gar nicht, wie die Zeit vergeht.
trifft nicht zu trifft zu

Erfassung der Lernstils: Verbalizer-Visualizer-Questionnaire nach Kirby, Moore & Schofield (1988). Die Items der Visualizerskala sind durch *Kursivschrift* hervorgehoben.

Präferenz für verbale oder visuelle Sprachverarbeitung

Im folgenden sollen Sie nun einen Fragebogen ausfüllen, in dem es darum geht, wie Sie mit Sprache und Bildern umgehen. Bitte kreuzen Sie an, ob die Aussagen für Sie zutreffen oder nicht.

1	Ich habe Spaß an Arbeiten, bei denen man Wörter benutzen muss.	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht
2.	Ich lerne gerne neue Wörter.	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
3.	Es fällt mir leicht Synonyme für Wörter zu finden.	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
4	Ich lese sehr langsam.	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
5	Ich lese lieber eine Anleitung, als dass mir jemand zeigt, wie eine Sache funktioniert.	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
6	Ich kann überdurchschnittlich gut mit Worten umgehen.	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
7	Ich verwende wenig Zeit, um meinen Wortschatz aufzubessern.	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
8	<i>Ich finde Zeitungsartikel mit Graphiken gut.</i>	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
9	Ich mag keine Wörterspiele, wie z. B. Kreuzworträtsel.	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu

B Instrumente

10	Es fällt mir schwer, mir Liedertexte einzuprägen.	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
11	<i>Bilder oder Diagramme finde ich beim Lesen sehr hilfreich.</i>	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
12	<i>Es fällt mir schwer, mir einen Platz vorzustellen, an dem ich erst selten war.</i>	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
13	<i>Um etwas zu erklären benutze ich selten Diagramme.</i>	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
14	<i>Ich mag keine Karten oder Diagramme in Büchern.</i>	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
15	<i>Wenn ich ein Buch lese, in dem Karten sind, schaue ich mir diese öfter an.</i>	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
16	<i>Dem alten Sprichwort „Ein Bild sagt mehr als tausend Worte“ stimme ich voll zu.</i>	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
17	Ich mag es nicht, Wörter im Lexikon nachzuschlagen.	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
18	<i>Um sich in einer neuen Stadt zurechtzufinden, finde ich eine Karte sehr sinnvoll</i>	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
19	<i>Ich glaube nicht, dass jemand in Bildern denken kann.</i>	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu
20	<i>Ich mochte noch nie Puzzlespiele.</i>	<input type="radio"/> Trifft zu <input type="radio"/> Trifft nicht zu

Ergänzende Fragen:

Fragen zu den Lerngewohnheiten (prae)

Wie häufig nutzen Sie informative Bilder in der Schule bzw. im Studium (schematische Darstellungen, Diagramme, etc.)?

nie sehr häufig

Wie häufig nutzen Sie beim Lesen von Zeitschriften- bzw. Zeitungsartikeln als zusätzliche Informationsquelle Bilder?

nie sehr häufig

Wie leicht bzw. schwer fällt es Ihnen in der Regel Bilder zu verstehen und sich zu merken?

sehr schwer sehr leicht

Fragen zum Lernen während der Präsentationsphase (post)

Haben Sie allgemein eher versucht, sich Details zu merken oder eher die wesentlichen Zusammenhänge grob zu verstehen?

nur Details nur Zusammenhänge

Haben Sie sich die Inhalte mehr „sprachlich“ gemerkt (Zahlenwerte, Eigenheiten,...) oder haben Sie versucht, sich ein „Bild“ des dargestellten Sachverhaltes zu konstruieren und zu merken?

nur sprachlich nur bildlich

Beim Lernen, haben Sie sich stärker auf den Text oder stärker auf die Bilder konzentriert?

nur auf den Text nur auf die Bilder

War es eher positiv oder eher negativ, die Texte und Bilder an einem Computer und nicht aus einem Buch zu lernen?

sehr positiv sehr negativ

