

**Interferieren (geschriebene) räumliche Textinhalte
mit der Bildverarbeitung im visuell-räumlichen
Arbeitsgedächtnis?
Testung einer erweiterten kognitiven Theorie
multimedialen Lernens**

Dissertation

der Fakultät für Informations- und Kognitionswissenschaften
der Eberhard-Karls-Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Dipl.-Psych. Anne Julia Schüler
aus Solingen

Tübingen
2010

Tag der mündlichen Qualifikation:

02.12.2010

Dekan:

Prof. Dr. Oliver Kohlbacher

1. Berichterstatter:

Prof. Dr. Katharina Scheiter

2. Berichterstatter:

Prof. Dr. Dr. Friedrich Hesse

3. Berichterstatter:

Prof. Dr. Ralf Rummer

An dieser Stelle möchte ich allen danken, die mich darin unterstützt haben, dieses Projekt zu beginnen, durchzuführen und – zu guter Letzt – auch zu vollenden.

Mein Dank gilt insbesondere...

Prof. Dr. Katharina Scheiter

Prof. Dr. Peter Gerjets

Prof. Dr. Dr. Friedrich Hesse

Prof. Dr. Ralf Rummer

der NG Wissenserwerb mit Multimedia

der AG Wissenserwerb mit Hypermedia

den Hiwis meiner Projekte

Lino und meiner Familie

Vielen Dank!

Inhaltsverzeichnis

1	<u>Einleitung</u>	1
<u>I. Theoretischer Hintergrund</u>.....		4
2	<u>Lernen mit Multimedia</u>	4
2.1	Lernen mit Multimedia: Definition und empirische Evidenz	4
2.2	Theoretische Grundlagen für das Lernen mit Multimedia	6
2.2.1	Theoretische Grundlagen der Cognitive Theory of Multimedia Learning	6
2.2.2	Die Annahmen der Cognitive Theory of Multimedia Learning.....	13
2.3	Die Rolle des Arbeitsgedächtnisses beim Lernen mit Multimedia	17
2.3.1	Untersuchungsmethoden der Arbeitsgedächtnisforschung.....	17
2.3.2	Die Verarbeitung multimedialen Materials im Arbeitsgedächtnis.....	25
2.4	Zusammenfassung und Implikationen für die vorliegende Arbeit	31
3	<u>Das Modalitätsprinzip beim Lernen mit Multimedia</u>	33
3.1	Definition und empirische Evidenz	33
3.2	Theoretische Erklärungen des Modalitätseffekts	35
3.2.1	Fehlende räumliche und zeitliche Kontiguität bei geschriebenem Text	35
3.2.2	Interferenzen bei geschriebenem Text und zusätzliche kognitive Ressourcen bei gesprochenem Text	36
3.3	Beurteilung der Erklärungen des Modalitätseffekts	40
3.3.1	Plausibilität	40
3.3.2	Geltungsbereich	44
3.4	Zusammenfassung und Implikationen für die vorliegende Arbeit	46
4	<u>Herleitung einer weiteren Erklärung des Modalitätseffekts vor dem Hintergrund der VSSP-Struktur</u>	48
4.1	Die VSSP-basierte Erklärung des Modalitätseffekts	48
4.2	Empirische Evidenz: Informationsverarbeitung innerhalb des VSSP	49
4.2.1	Die Struktur des VSSP.....	49
4.2.2	Die Kontrolle von Blickbewegungen im VSSP.....	50
4.2.3	Die Verarbeitung verbaler Informationen im VSSP	52
4.3	Empirische Evidenz: Modalitätseffekte in Abhängigkeit des Textinhalts	58
4.3.1	Analyse der Textinhalte in Untersuchungen zum Modalitätsprinzip beim Lernen mit Multimedia.....	58
4.3.2	Empirische Evidenz aus monomedialen Untersuchungen	62
4.3.3	Empirische Evidenz aus multimedialen Untersuchungen.....	67
4.4	Zusammenfassung und Implikationen für die vorliegende Arbeit	68
5	<u>Erweiterungen der Cognitive Theory of Multimedia Learning</u>	70
5.1	Der strukturelle Aufbau der ECTML	70
5.2	Ableitungen aus der ECTML	71
5.2.1	Annahmen bezüglich des Einflusses der Textmodalität beim Lernen mit Multimedia	71

5.2.2	Annahmen bezüglich des Zusammenspiels von Textmodalität und Textinhalt.....	73
5.2.3	Annahmen bezüglich des Einflusses der Textinhalte beim Lernen mit Multimedia ..	74
5.3	Zusammenfassung und Implikationen für die vorliegende Arbeit.....	78

6	<u>Ableitung der Hypothesen.....</u>	<u>80</u>
---	--------------------------------------	-----------

II. Empirischer Teil..... **86**

7	<u>Gestaltung der Lernumgebungen</u>	<u>86</u>
---	--	-----------

8	<u>Die Operationalisierung der Variablen</u>	<u>91</u>
---	--	-----------

8.1	Operationalisierung der unabhängigen Variablen	91
8.1.1	Textmodalität	91
8.1.2	Textinhalt	92
8.1.3	Arbeitsgedächtniskapazität	95
8.1.4	Zweitaufgabe	95
8.2	Operationalisierung der abhängigen Variablen.....	96
8.2.1	Erinnerungsleistung	96
8.2.2	Bewertung der Lernphase	104
8.3	Personenbezogene Variablen	105

9	<u>Experiment 1</u>	<u>106</u>
---	---------------------------	------------

9.1	Methode	106
9.1.1	Versuchsdesign	106
9.1.2	Versuchspersonen	107
9.1.3	Geräte und Hilfsmittel	107
9.1.4	Versuchsablauf.....	108
9.2	Ergebnisse.....	109
9.2.1	Reliabilitätsanalysen	109
9.2.2	Datenexploration und Überprüfung inferenzstatistischer Analysevoraussetzungen.	110
9.2.3	Beschreibung der Stichprobe	111
9.2.4	Deskriptive Daten	113
9.2.5	Prüfung der Textmodalität-Hypothese.....	115
9.2.6	Prüfung der Textinhalt-Hypothese.....	119
9.2.7	Prüfung der Arbeitsgedächtnis-Hypothese	121
9.3	Zusammenfassung und Diskussion	123

10	<u>Experiment 2</u>	<u>128</u>
----	---------------------------	------------

10.1	Erweiterung der Hypothesen.....	128
10.2	Methode	133
10.2.1	Versuchsdesign	133
10.2.2	Versuchspersonen	136
10.2.3	Geräte und Hilfsmittel	136
10.2.4	Versuchsablauf.....	136
10.3	Ergebnisse.....	137
10.3.1	Reliabilitätsanalysen	138
10.3.2	Datenexploration und Überprüfung inferenzstatistischer Analysevoraussetzungen.	141

10.3.3	Beschreibung der Stichprobe	141
10.3.4	Deskriptive Daten	144
10.3.5	Baseline der Zweitaufgabe: Vergleichbarkeit zwischen den Gruppen und Einfluss in der Lernphase.....	146
10.3.6	Prüfung der Textmodalität-Hypothese.....	147
10.3.7	Prüfung der Textinhalt-Hypothese.....	151
10.3.8	Prüfung der Arbeitsgedächtnis-Hypothese	157
10.3.9	Prüfung der Zweitaufgaben-Hypothese	159
10.4	Zusammenfassung und Diskussion	162

11 Experiment 3 168

11.1	Erweiterung der Hypothesen	168
11.2	Methode	170
11.2.1	Versuchsdesign	170
11.2.2	Versuchspersonen	172
11.2.3	Geräte und Hilfsmittel	173
11.2.4	Versuchsablauf.....	173
11.3	Ergebnisse.....	174
11.3.1	Reliabilitätsanalysen	175
11.3.2	Datenexploration und Überprüfung inferenzstatistischer Analysevoraussetzungen.....	176
11.3.3	Beschreibung der Stichprobe	177
11.3.4	Deskriptive Daten	179
11.3.5	Prüfung der Textmodalität-Hypothese.....	181
11.3.6	Prüfung der Textinhalt-Hypothese.....	186
11.3.7	Prüfung der Arbeitsgedächtnis-Hypothese	190
11.3.8	Prüfung der Blickbewegungs-Hypothese.....	192
11.4	Zusammenfassung und Diskussion	196

12 Abschließende Zusammenfassung und Diskussion 201

12.1	Welchen Einfluss nimmt die Textmodalität beim Lernen mit Multimedia?	205
12.2	Welchen Einfluss nimmt der Textinhalt beim Lernen mit Multimedia?.....	210
12.3	Welchen Einfluss nimmt das Arbeitsgedächtnis beim Lernen mit Multimedia?.....	213
12.4	Stärken und Schwächen der empirischen Studien.....	215

13 Fazit und Ausblick 219

14 Literaturverzeichnis..... 223

Anhang 240

Abbildungsverzeichnis

Abbildung 1. Das Modell des Arbeitsgedächtnisses (a) in seiner ursprünglichen Konzeption und (b) nach seiner Erweiterung.....	11
Abbildung 2. Die <i>Cognitive Theory of Multimedia Learning</i>	14
Abbildung 3. Schematische Abbildung (a) des Corsi Blocks und (b) zweier Matrizen unterschiedlicher Komplexität des VPT.....	23
Abbildung 4. Die Ergebnisse der Untersuchung von Gyselinck et al.....	28
Abbildung 5. Die Verarbeitung (A) gesprochener Wörter und Bilder und (B) geschriebener Wörter und Bilder nach der CTML.....	38
Abbildung 6. Die Verarbeitung geschriebener Wörter und Bilder nach (A) den Annahmen der CTML von Mayer (2009) und (B) dem Arbeitsgedächtnismodell von Baddeley (1992).....	42
Abbildung 7. Lokalisierung der Ursachen des Modalitätseffekts im Verarbeitungsprozess nach der CTML und der zeitlichen und räumlichen Kontiguitätsklärung.....	44
Abbildung 8. Die von Mayer und Moreno (1998; Moreno & Mayer, 1999) verwendeten Textinhalte in Studien zum Modalitätseffekt.....	59
Abbildung 9. Der strukturelle Aufbau der ECTML.....	71
Abbildung 10. Die angenommene Belastung des Arbeitsgedächtnisses bei der Darbietung verschiedener Textmodalitäten (unabhängig vom Textinhalt).....	71
Abbildung 11. Die angenommene Belastung des Arbeitsgedächtnisses bei der Darbietung verschiedener Textmodalitäten und Textinhalte.....	74
Abbildung 12. Die angenommene Belastung des Arbeitsgedächtnisses bei der Verarbeitung unterschiedlicher Textinhalte (A-C).....	75
Abbildung 13. Darstellung des erwarteten Ergebnismusters für die Textmodalität-Hypothese nach der ECTML.....	82
Abbildung 14. Darstellung des erwarteten Ergebnismusters für die Textinhalt-Hypothese nach der ECTML.....	83
Abbildung 15. Beispielhafte Abbildung eines in der Lernumgebung dargebotenen Fisches und einer Pflanze.....	87
Abbildung 16. Das Verhältnis von Text- und Bildinformation.....	89
Abbildung 17. Die Bildschirmdarstellung mit gesprochenem und geschriebenem Text.....	92
Abbildung 18. Verbale und bildhafte Verifikationsitems.....	98
Abbildung 19. Darstellung des Zusammenhangs des VPT Maßes mit der Erinnerung an visuelle Bildinhalte in Abhängigkeit des Textinhalts und der Textmodalität.....	117
Abbildung 20. Die angenommene Belastung des Arbeitsgedächtnisses bei der Darbietung verschiedener Textmodalitäten und Textinhalte zusammen mit einer räumlichen Zweitaufgabe (ZA).....	129
Abbildung 21. Darstellung der erwarteten Ergebnismuster für die Textmodalität-Hypothese nach der CTML in Abhängigkeit verschiedener AVn.....	130
Abbildung 22. Darstellung der erwarteten Ergebnismuster für die Textmodalität-Hypothese nach der ECTML in Abhängigkeit verschiedener AVn.....	131
Abbildung 23. Darstellung des erwarteten Ergebnismusters für die Textinhalt-Hypothese.....	132

Abbildung 24. Darstellung des Zusammenhangs des Corsi Block Maßes mit der in die Lernphase investierten Mühe in Abhängigkeit der Zweitaufgabe und des Textinhalts.....	156
Abbildung 25. Darstellung des erwarteten Ergebnismusters für die Textinhalt-Hypothese.....	169
Abbildung 26. AOIs in der Lernphase.....	172

Tabellenverzeichnis

Tabelle 1 Die sieben Kernprinzipien der CTML und ihre Annahmen	17
Tabelle 2 Übersicht über empirische Studien zur Involviertheit des Arbeitsgedächtnisses beim Lernen mit Multimedia.	30
Tabelle 3 Räumliche (links) und nicht-räumliche (rechts) Sätze der Brooks-Aufgabe	54
Tabelle 4 Übersicht über die verwendeten Textinhalte in empirischen Studien zum Modalitätseffekt (ME)	61
Tabelle 5 Gegenüberstellung der Annahmen der CTML und ECTML	85
Tabelle 6 Darstellung des Aufbaus eines Texts für die Bedingungen mit visuellem Textinhalt in den ersten beiden Experimenten.....	88
Tabelle 7 Gegenüberstellung von visuellen und räumlichen Textinhalten.....	94
Tabelle 8 Indizes und ihre Bedeutung.	97
Tabelle 9 Instruktionen zur Erfassung der bildbezogenen Erinnerungsleistung (Verifikationsitems).....	99
Tabelle 10 Instruktionen zur Erfassung der textbezogenen Erinnerungsleistung (Verifikationsitems).....	101
Tabelle 11 Fragebogen zur Bewertung der Lernphase	105
Tabelle 12 Darstellung des Versuchsplans mit den Faktoren Textmodalität und Textinhalt (Experiment 1).	107
Tabelle 13 Cronbach´s α der einzelnen Indizes vor und nach Itemausschluss.....	110
Tabelle 14 Variablen mit Abweichungen von der Normalverteilung sowie ungleichen Fehlervarianzen.....	111
Tabelle 15 Mittelwerte und Standardabweichungen (in Klammern) der intervallskalierten Variablen in Abhängigkeit des Textinhalts und der Textmodalität.....	111
Tabelle 16 Die Verteilung der nominalen Daten in Abhängigkeit des Textinhalts und der Textmodalität.....	112
Tabelle 17 Adjustierte Mittelwerte und Standardfehler (in Klammern) der Verifikationsitems in Abhängigkeit des Textinhalts und der Textmodalität.....	114
Tabelle 18 Adjustierte Mittelwerte und Standardfehler (in Klammern) der Bewertungsitems in Abhängigkeit des Textinhalts und der Textmodalität.....	114
Tabelle 19 Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der MANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die Verifikationsitems.....	116
Tabelle 20 Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der MANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die Bewertungsitems.....	118
Tabelle 21 Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der ANCOVAs für den Faktor Textmodalität (TM) für die Bewertungsitems.	118
Tabelle 22 Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der MANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) für die Verifikationsitems.	120
Tabelle 23 Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der MANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) für die Bewertungsitems	121
Tabelle 24 Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der MANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und Digit Span für die Verifikationsitems.....	122

Tabelle 25 Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der MANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und Digit Span für die Bewertungsitems.....	122
Tabelle 26 Darstellung des dreifaktoriellen Versuchsplans mit den interindividuell variierten Faktoren Textinhalt und Zweitaufgabe und dem intraindividuell variierten Faktor Textmodalität.	134
Tabelle 27 Cronbach's α der einzelnen Indizes der Verifikationsitems vor und nach Itemausschluss.	139
Tabelle 28 Cronbach's α der einzelnen Indizes der offenen Fragen vor und nach Itemausschluss.	140
Tabelle 29 Variablen mit Missing Data, Abweichungen von der Normalverteilung sowie ungleichen Fehlervarianzen.	141
Tabelle 30 Mittelwerte und Standardabweichungen (in Klammern) der intervallskalierten personenbezogenen Daten in Abhängigkeit des Textinhalts und der Zweitaufgabe (ZA).....	142
Tabelle 31 Die Verteilung der nominalen personenbezogenen Daten in Abhängigkeit des Textinhalts und der Zweitaufgabe (ZA).....	142
Tabelle 32 Mittelwerte und Standardabweichungen (in Klammern) der intervallskalierten Kennwerte der Baseline in Abhängigkeit des Textinhalts.....	143
Tabelle 33 Adjustierte Mittelwerte und Standardfehler (in Klammern) der Verifikationsitems in Abhängigkeit des Textinhalts, der Zweitaufgabe (ZA) und der Textmodalität.....	144
Tabelle 34 Adjustierte Mittelwerte und Standardfehler (in Klammern) der offenen Fragen in Prozent in Abhängigkeit des Textinhalts, der Zweitaufgabe (ZA) und der Textmodalität.....	145
Tabelle 35 Adjustierte Mittelwerte und Standardfehler (in Klammern) der Leistung in der Zweitaufgabe in Abhängigkeit des Textinhalts und der Textmodalität.....	145
Tabelle 36 Adjustierte Mittelwerte und Standardfehler (in Klammern) der Bewertungsitems in Abhängigkeit des Textinhalts, der Zweitaufgabe (ZA) und der Textmodalität.....	146
Tabelle 37 Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Verifikationsitems.	148
Tabelle 38 Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die offenen Fragen.....	149
Tabelle 39 Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Leistung in der Zweitaufgabe.	150
Tabelle 40 Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Bewertungsitems.....	150
Tabelle 41 Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Verifikationsitems.....	152

Tabelle 42 Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die offenen Fragen.	154
Tabelle 43 Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Leistung in der Zweitaufgabe.	154
Tabelle 44 Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Bewertungsitems.	155
Tabelle 45 Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Verifikationsitems.	157
Tabelle 46 Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Erinnerung der offenen Fragen.	158
Tabelle 47 Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Leistung in der Zweitaufgabe.	158
Tabelle 48 Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Bewertungsitems.	159
Tabelle 49 Zur Prüfung der Zweitaufgaben-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Zweitaufgabe (ZA) und Corsi Block (Corsi) für die Verifikationsitems.	160
Tabelle 50 Zur Prüfung der Zweitaufgaben-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Zweitaufgabe (ZA) und Corsi Block (Corsi) für die offenen Fragen.	161
Tabelle 51 Zur Prüfung der Zweitaufgaben-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Zweitaufgabe (ZA) und Corsi Block (Corsi) für die Bewertungsitems.	161
Tabelle 52 Cronbach's α der einzelnen Indizes vor und nach Itemausschluss.	175
Tabelle 53 Cronbach's α der einzelnen Indizes der offenen Fragen vor und nach Itemausschluss.	176
Tabelle 54 Variablen mit Missing Data, Abweichungen von der Normalverteilung sowie ungleichen Fehlervarianzen.	177
Tabelle 55 Mittelwerte und Standardabweichungen (in Klammern) der personenbezogenen Daten in Abhängigkeit des Textinhalts und der Textmodalität.	177
Tabelle 56 Die Verteilung der nominalen personenbezogenen Daten in Abhängigkeit des Textinhalts und der Textmodalität.	178
Tabelle 57 Die Verteilung der nominalen personenbezogenen Daten in Abhängigkeit des Textinhalts und der Textmodalität.	179
Tabelle 58 Adjustierte Mittelwerte und Standardfehler (in Klammern) der offenen Fragen in Prozent in Abhängigkeit des Textinhalts und der Textmodalität.	180
Tabelle 59 Adjustierte Mittelwerte und Standardfehler (in Klammern) der Blickverhaltensindizes in Abhängigkeit des Textinhalts und der Textmodalität.	180
Tabelle 60 Adjustierte Mittelwerte und Standardfehler (in Klammern) der Bewertungsitems in Abhängigkeit des Textinhalts und der Textmodalität.	181

Tabelle 61 Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die Verifikationsitems.....	182
Tabelle 62 Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die offenen Fragen.	183
Tabelle 63 Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die Blickverhaltensindizes.	184
Tabelle 64 Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die Bewertungsitems.....	185
Tabelle 65 Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) für die Verifikationsitems.	187
Tabelle 66 Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) bei der Erinnerung der offenen Fragen.	188
Tabelle 67 Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) für die Blickverhaltensindizes.	189
Tabelle 68 Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) für die Bewertungsitems.	189
Tabelle 69 Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Verifikationsitems	190
Tabelle 70 Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Erinnerung der offenen Fragen.....	191
Tabelle 71 Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Blickverhaltensindizes.....	191
Tabelle 72 Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Bewertungsitems	192
Tabelle 73 Zur Prüfung der Blickbewegungs-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Anzahl an Sakkaden (AS), Textinhalt (TI) und Corsi Block (Corsi) für die Verifikationsitems.	193
Tabelle 74 Zur Prüfung der Blickbewegungs-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Anzahl an Sakkaden (AS), Textinhalt (TI) und Corsi Block (Corsi) für die offenen Fragen.....	194
Tabelle 75 Zur Prüfung der Blickbewegungs-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Anzahl an Sakkaden (AS), Textinhalt (TI) und Corsi Block (Corsi) für die Bewertungsitems	195
Tabelle 76 Zusammenfassung der empirischen Befunde hinsichtlich unterschiedlicher Vorhersagen der CTML und ECTML.	204
Tabelle 77 Zusammenfassung der empirischen Befunde hinsichtlich Vorhersagen der CTML und ECTML.....	205

1 Einleitung

„Die Darbietung von Texten zusammen mit Bildern führt zu besseren Lernergebnissen als die Darbietung von Texten ohne Bildern“ (Mayer, 2009) – dieses sogenannte *Multimediaprinzip* bildet die Grundlage einer Vielzahl an Forschungsarbeiten in der empirischen Lehr-Lernforschung, die in den letzten Jahrzehnten durchgeführt wurden. Dabei stand im Vordergrund vieler Arbeiten die Frage nach der „optimalen“ Darbietungsweise von Text und Bild, d. h. der dem Lernen förderlichsten Darbietungsweise. Diese Entwicklung hängt sicherlich nicht zuletzt mit den technischen Fortschritten der letzten Jahre zusammen, die eine Vielzahl an unterschiedlichen Text-Bilddarbietungen ermöglichen. Man denke hier z. B. an die computerbasierte Darbietung dynamischer Bilder zusammen mit gesprochenen Texten im Gegensatz zur zuvor üblichen Darbietung statischer Bilder zusammen mit geschriebenen Texten in Büchern.

Die theoretische Basis vieler dieser Arbeiten bildet die *Cognitive Theory of Multimedia Learning* (CTML; Mayer, 2009). Diese Theorie hat den Anspruch kognitive Grundlagentheorien auf das Lernen mit Text und Bild anzuwenden. Insbesondere das Arbeitsgedächtnismodell nach Baddeley (1992) nimmt eine zentrale Rolle in der CTML ein.

Problematischerweise erfolgte die empirische Prüfung der theoretischen Annahmen dabei fast ausschließlich indirekt, indem die aus den Modellen vorhergesagten Effekte empirisch nachgewiesen wurden, ohne jedoch Evidenz für den zugrundeliegenden Entstehungsmechanismus zu liefern. Nicht-replizierte Effekte wurden oftmals mit einem schlechten Design der Materialien erklärt, ohne zu prüfen, ob nicht auf theoretischer Ebene eine Fehlkonzeption vorliegt. Clark und Feldon (2005) merken hierzu an: “Even when easily available research findings contradict widely shared beliefs about benefits, it is tempting to ignore the research by assuming, without careful analysis, that the multimedia instruction has been poorly designed“ (Clark & Feldon, S. 97). Dies führt dazu, dass die CTML seit den 90er Jahren in fast unveränderter Form existiert, obwohl es immer mehr empirische Evidenz gibt, die gegen ihre Gültigkeit spricht.

In der vorliegenden Arbeit wird aufgezeigt, dass die theoretischen Annahmen der CTML teilweise fehlerhaft sind. Dies manifestiert sich insbesondere bei der

Erklärung der CTML für den sogenannten Modalitätseffekt, dem empirischen Befund, dass die Darbietung geschriebener Texte zusammen mit Bildern zu schlechterer Leistung führt als die Darbietung gesprochener Texte zusammen mit Bildern. Dieser Befund geht nach der CTML auf Interferenzen im Arbeitsgedächtnis zwischen den Repräsentationen, sogenannten *Abbildern*, des geschriebenen Texts und des Bilds zurück, was theoretisch unplausibel ist. Eine weitere Erklärung für den Modalitätseffekt, welche von fehlender räumlicher Kontiguität bei geschriebenem Text und Bild ausgeht, ist aus empirischer Sicht unplausibel, weil der angenommene Mechanismus extensiver visueller Suchprozesse bei geschriebenem Text und Bild empirisch nicht nachgewiesen wurde (z. B. Hegarty, 1992). Eine dritte Erklärung für den Modalitätseffekt geht von fehlender zeitlicher Kontiguität bei geschriebenem Text und Bild aus. Sie sollte allerdings vor allem dann Gültigkeit besitzen, wenn Text und Bild integriert werden müssen, was sich insbesondere in der Transferleistung widerspiegeln sollte. Für die Erinnerung von Faktenwissen sollte eine Integration von Text und Bild hingegen nicht erforderlich sein, so dass hier die fehlende zeitliche Kontiguität bei geschriebenem Text und Bild für die Erklärung des Modalitätseffekts keine bzw. nur eine untergeordnete Rolle spielen sollte.

Da der Modalitätseffekt aber auch dann nachgewiesen wurde, wenn das Erinnern von text- oder bildbezogenen Fakten erforderlich war, scheint es einen weiteren Mechanismus für den Modalitätseffekt zu geben. In der vorliegenden Arbeit wird entsprechend eine alternative Erklärung für den Modalitätseffekt aufgezeigt, die bisher in der (Multimedia)literatur nicht existierte. Diese Erklärung basiert auf neueren Befunden hinsichtlich der Struktur des visuell-räumlichen Notizblocks (*Visuo-Spatial Sketchpad; VSSP*), einer Subkomponente des Arbeitsgedächtnisses. Dazu werden die oberflächlichen Annahmen der CTML hinsichtlich der Struktur des Arbeitsgedächtnisses und insbesondere des VSSP spezifiziert, so dass ein erweitertes Modell der CTML, die *Extendend Cognitive Theory of Multimedia Learning* (ECTML), resultiert.

Wie im Theorieteil aufgezeigt wird, lässt sich aus der ECTML – im Einklang mit der Arbeitsgedächtnisliteratur – die Vorhersage ableiten, dass der Modalitätseffekt hinsichtlich der Erinnerung von Faktenwissen nur bei der Darbietung räumlicher Textinhalte, d. h. bei der Beschreibung räumlicher Konfigurationen, auftritt. Zudem sollten räumliche Textinhalte auf das Lernen mit Multimedia nach der ECTML generell einen negativen Einfluss haben. Für diese

Annahmen gibt es auch erste empirische Hinweise, die im Theorieteil aufgeführt werden.

Im empirischen Teil der Arbeit werden drei Experimente beschrieben, mit deren Hilfe die aus der ECTML ableitbaren Hypothesen empirisch geprüft wurden. In den Experimenten werden durch die Erfassung der Arbeitsgedächtniskapazität der Lerner sowie prozessnaher Maße wie Blickbewegungen und der Performanz in einer Zweitaufgabe die Arbeitsgedächtnisprozesse fokussiert.

Ziel dieser Forschungsarbeit ist es neue Wege der Multimediaforschung aufzuzeigen, damit in Zukunft nicht nur die empirische Untersuchung der Effekte selber, sondern auch die Mechanismen für diese Effekte stärker in den Vordergrund rücken.

I. Theoretischer Hintergrund

2 Lernen mit Multimedia

2.1 Lernen mit Multimedia: Definition und empirische Evidenz

Für den Terminus „Multimedia“ gibt es bis heute keine einheitliche Definition. Im umgangssprachlichen Gebrauch ist der Begriff „Multimedia“ oftmals stark mit technischen Aspekten, speziell dem Einsatz von Computern, assoziiert (vgl. z. B. Clark & Feldon, 2005). Die wortwörtliche Bedeutung des Begriffs Multimedia impliziert hingegen nur den Einsatz „vieler vermittelnder Elemente“ (multus [lat.]: viel, zahlreich; Medium: vermittelndes Element [Wermke, Kunkel-Razum & Scholze-Stubenrecht, 2004, S. 643]).

Auch in der Fachliteratur gibt es unterschiedlich differenzierte Definitionen des Begriffs Multimedia. Während Weidenmann (1995, 1997) drei Dimensionen unterscheidet, auf denen der Begriff definiert werden kann, nämlich das Medium an sich, z. B. Fernseher oder Videorekorder, die Art der Kodalität, z. B. Wörter oder Bilder, und die angesprochene Sinnesmodalität, z. B. Auge oder Ohr, wird der Begriff in der Multimedialiteratur in Anlehnung an Mayer (2009, 2005b) meist weitaus weniger differenziert wie folgt definiert: „I define *multimedia* as presenting both words (such as spoken text or printed text) and pictures (such as illustrations, photos, animation, or video)” (Mayer, 2005b, S. 2; Hervorhebung im Original). Damit ist das entscheidende Kriterium für multimediale Lernumgebungen die gleichzeitige Darbietung von Wörtern und Bildern, also nach Weidenmann (1997) die Kodalität. Sinnesmodalität und Medium werden hingegen nur als Spezifika der Wörter oder Bilder aufgeführt: So können Wörter gesprochen oder geschrieben dargeboten werden und sprechen dadurch eine bestimmte Sinnesmodalität an, während das Medium nur indirekt als Träger verschiedener Bildformate eine Rolle spielt – so würde man eine Illustration eher mit einem Buch verbinden, während Animationen sehr stark mit der Verwendung digitaler Medien verknüpft sind. Im Vordergrund steht also nach Mayer die Darbietung von Texten und Bildern – also die Instruktion an sich. Da auch in der vorliegenden Arbeit die Instruktion mit Hilfe von Text und Bild im Vordergrund steht und die theoretischen Annahmen Mayers sowie

die darauf aufbauenden Befunde eine wichtige Rolle einnehmen, wird dieser Arbeit die Definition Mayers von Multimedia zugrunde gelegt.

Eine grundlegende Annahme bezüglich des Lernens mit Multimedia trifft Mayer (2005b, 2009) mit seinem *Multimediaprinzip*: “People can learn more deeply from words and pictures than from words alone” (Mayer, 2005b, S. 1). Mit anderen Worten: Multimediale Lernumgebungen, die nach Mayers Definition von Multimedia durch die gleichzeitige Darbietung von Bildern und Wörtern charakterisiert sind, führen zu höherem Lernerfolg als monomediale Lernumgebungen, in denen nur Wörter dargeboten werden. Bezüglich der Erinnerungsleistung wurde dieses Multimediaprinzip schon in den 70er und 80er Jahren in zahlreichen Studien untersucht und bestätigt. So kamen Levie und Lentz (1982) nach einer Übersicht über 46 Studien zu dem Schluss, dass in 98% der Studien die Darbietung von Text und Bild der Darbietung von Text ohne Bild bezüglich der Erinnerungsleistung für illustrierte Textinformation überlegen war. In keiner der Studien war hingegen die Bedingung Text *ohne* Bild überlegen. Ähnliche Befunde werden auch in einer Metaanalyse von Levin, Anglin und Carney (1987) berichtet: Bilder, die nicht nur dekorativen Zwecken dienen, hatten einen positiven Effekt auf das Erinnern der Inhalte. Weitere Überblicksarbeiten, die die Überlegenheit von Text mit Bild über Text ohne Bild bestätigen, wurden z. B. von Levin und Lesgold (1978), Anglin, Towers und Levie (1996) sowie Carney und Levin (2002) publiziert. Studien in neuerer Zeit untersuchten nicht nur die Gültigkeit des Multimediaprinzips hinsichtlich der Erinnerungsleistung, sondern auch hinsichtlich der Transferleistung. So konnte vor allem Mayer zeigen, dass auch bei Transferleistungen die Darbietung von Text und Bild der alleinigen Textdarbietung überlegen war (für eine Übersicht vgl. Mayer, 2009). Man kann somit festhalten, dass das Multimediaprinzip empirisch sowohl bezüglich der Erinnerungsleistung als auch bezüglich der Transferleistung sehr gut abgesichert ist: In einer Vielzahl von Studien war die Präsentation von Text mit Bild der Präsentation von Text ohne Bild überlegen.

2.2 Theoretische Grundlagen für das Lernen mit Multimedia

Die theoretische Grundlage für das Lernen mit Texten und Bildern bildet vor allem die *Cognitive Theory of Multimedia Learning* (CTML; Mayer, 2009)¹. Die CTML lehnt sich stark an kognitionspsychologische Vorhersagen bezüglich der Verarbeitung von Informationen im kognitiven System an. Zentral ist dabei insbesondere das Arbeitsgedächtnismodell von Baddeley in seiner ursprünglichen Konzeption (z. B. 1992, 1999). Weitere Grundlagen bilden die Theorie der dualen Kodierung von Paivio (1990) sowie das Drei-Speicher-Modell von Atkinson und Shiffrin (1968). Diese theoretischen Grundlagen werden zunächst kurz skizziert (Abschnitt 2.2.1), bevor die CTML (Abschnitt 2.2.2) vorgestellt wird.

2.2.1 Theoretische Grundlagen der Cognitive Theory of Multimedia Learning

Nach der CTML findet die Verarbeitung von Texten und Bildern in verschiedenen Gedächtnissystemen, nämlich dem sensorischen Gedächtnis, dem Arbeitsgedächtnis und dem Langzeitgedächtnis, statt. Insbesondere dem Arbeitsgedächtnis kommt dabei eine zentrale Rolle zu, denn hier werden die externen Repräsentationen, also Texte und Bilder, modalitätsspezifisch bzw. in einem späteren Stadium kodialitätsspezifisch weiterverarbeitet. Im Folgenden werden die theoretischen Grundlagen für diese Annahmen kurz vorgestellt, nämlich das Drei-Speicher-Modell (Atkinson & Shiffrin, 1968), das Arbeitsgedächtnismodell (Baddeley, 1992) sowie die Theorie der dualen Kodierung (Paivio, 1990).

Das Drei-Speicher-Modell von Atkinson und Shiffrin (1968)

Die CTML bettet die Prozesse bei der Text-Bildverarbeitung in das von Atkinson und Shiffrin (1968) postulierte Drei-Speicher-Modell ein. Demnach besteht das menschliche Gedächtnis aus drei Komponenten, nämlich dem sensorischen Speicher (*Sensory Register*), dem Kurzzeitspeicher (*Short-Term Store*) und dem Langzeitspeicher (*Long-Term Store*). Wird ein Reiz präsentiert, so wird er zunächst von dem sensorischen Speicher der entsprechenden Sinnesmodalität registriert. In

¹ Auf eine Beschreibung des *Integrated Model of Text and Picture Comprehension* von Schnotz (2005) und der *Cognitive Load Theory* (Sweller, 1999) wird im Rahmen dieser Arbeit verzichtet, da beide Theorien Annahmen machen, die in hohem Maße redundant zu den Annahmen der CTML sind und sie dementsprechend die gleichen Vorhersagen machen.

der visuellen Modalität herrscht beispielsweise für einige Millisekunden ein exaktes Abbild des visuellen Reizes vor, das aber verfällt, wenn es nicht in den Kurzzeitspeicher übertragen wird. Die Übertragung vom sensorischen Speicher in den Kurzzeitspeicher findet durch eine gezielte Suche nach relevanten Informationen bzw. Aufmerksamkeitsausrichtung statt. Auch im Kurzzeitspeicher verfällt die Information nach einiger Zeit, aber diese Zeitspanne bis zum Verfall ist länger als die im sensorischen Gedächtnis. Um Informationen längerfristig zu speichern, muss die Information in den Langzeitspeicher übertragen werden. Die Übertragung vom Kurzzeit- in den Langzeitspeicher hängt vor allem von der Verweildauer der Information im Kurzzeitspeicher ab: Je länger eine Information im Kurzzeitspeicher gehalten wird, desto größer ist die Wahrscheinlichkeit, dass sie übertragen wird. Aber auch ein umgekehrter Transfer findet statt: Für alle bewussten Prozesse, wie Denken oder Problemlösen, muss die Information in den Kurzzeitspeicher übertragen werden.

Baddeley (Baddeley & Hirsch, 1974) spezifizierte die Annahmen bezüglich des Kurzzeitgedächtnisses einige Jahre später in seinem Arbeitsgedächtnismodell, das im Folgenden beschrieben wird.

Das Arbeitsgedächtnismodell nach Baddeley (1992)

Unter dem Begriff Arbeitsgedächtnis (*Working Memory*) wird ein System verstanden, das es erlaubt, verschiedene Informationseinheiten zur gleichen Zeit im Gedächtnis zu halten, sie miteinander zu verknüpfen (Baddeley, 1999) und sie kurzfristig zu speichern (Kintsch, Healy, Hegarty, Pennington & Salthouse, 1999). Während Atkinson und Shiffrin (1968) den Kurzzeitspeicher als eine einheitliche Struktur zur kurzfristigen Speicherung von Informationen konzeptualisierten, geht Baddeley (1992) von mehreren miteinander agierenden Subsystemen aus, in denen Informationen nicht nur gespeichert, sondern auch verändert werden können. Die CTML bezieht sich auf die ursprüngliche Konzeption des Baddeley-Modells, wie sie bis 2000 existierte (z. B. Baddeley, 1986, 1992). Nach dieser Konzeption besteht das Arbeitsgedächtnis aus drei Komponenten, nämlich einer zentralen Exekutive (*Central Executive*) und zwei Subsystemen (*Slave Systems*), der phonologischen Schleife (*Phonological Loop*, PL), in der sprachbasierte Informationen verarbeitet werden, und dem VSSP, in dem visuell-räumliche Informationen verarbeitet werden. Mit Hilfe dieser beiden Subsysteme werden Informationen eingelesen und

verarbeitet. Nach Baddeley ist die Verarbeitungskapazität der einzelnen Komponenten jeweils beschränkt. Daher kann die Verarbeitung von Informationen in der gleichen Komponente zu Interferenzen führen. Werden die Informationen hingegen in zwei verschiedenen Komponenten verarbeitet, so stehen mehr kognitive Ressourcen zur Verfügung und es treten keine Interferenzen bei der Verarbeitung auf.

Bezüglich der PL wurden bereits in der ursprünglichen Arbeitsgedächtniskonzeption differenzierte Aussagen getroffen. Demnach ist die PL für die Speicherung verbaler Informationen über kurze Zeit zuständig (Baddeley, Gathercole & Papagno, 1998). Sie besteht aus einer passiven Speicherkomponente (*Phonological Store*) und einem aktiven Aufrechterhaltungsmechanismus (*Rehearsal Process* oder *Inner Speech*). Im phonologischen Speicher (*Phonological Store*) können Informationen circa 1-2 Sekunden aktiviert bleiben, bevor sie zerfallen. Mit Hilfe des artikulatorischen Aufrechterhaltungsmechanismus (*Rehearsal Process*) können die Informationen jedoch durch Reartikulation aufgefrischt und im phonologischen Speicher aktiv gehalten werden. Die phonologische Information wird also, sobald sie zu zerfallen droht, durch die subvokale Reartikulation zurück in den phonologischen Speicher eingelesen, wodurch sich auch die Bezeichnung des Systems als phonologische *Schleife* erklärt (Baddeley, 2006). Allerdings ist die Kapazität dieses Aufrechterhaltungsmechanismus beschränkt, da die Artikulation in Echtzeit stattfindet: Wenn die Anzahl an zu artikulierenden Items ansteigt, ist irgendwann ein Punkt erreicht, an dem das erste Item bereits zerfallen ist, bevor es erneut artikuliert werden kann (Baddeley, 2003). Dieser Effekt sollte eben dann eintreten, wenn für den artikulatorischen Aufrechterhaltungsmechanismus des Materials mehr als jene 1-2 Sekunden notwendig sind, die das Material auch passiv im phonologischen Speicher gehalten werden kann. Neben dem aktiven Halten von Informationen in der PL hat der artikulatorische Aufrechterhaltungsmechanismus eine weitere Funktion, nämlich das Einlesen von visuellem, ursprünglich nicht phonologisch dargebotenem, Material in die PL: Hiermit sind vor allem visuell dargebotene Wörter, also geschriebener Text, gemeint, aber auch benennbare Bilder – z. B. könnte das Bild eines Apfels mit dem verbalen, d. h. phonologischem, Label „Apfel“ versehen werden. Dieses Einlesen visuellen Materials in die PL geschieht durch die subvokale Artikulation der jeweiligen Wörter (Baddeley, 1992). Während gesprochene Sprache also direkten Zugang zum phonologischen Speicher hat, muss

geschriebene Sprache zunächst durch die artikulatorische Subvokalisation in eine phonologische Form transformiert werden, um so Zugang zum phonologischen Speicher zu erhalten (Logie, 1995). Die Annahme einer PL und ihrer Subkomponenten wird durch verschiedene experimentelle Befunde gestützt [vgl. z. B. phonologischer Ähnlichkeitseffekt, Baddeley (1966); Wortlängeneffekt, Baddeley, Thomson und Buchanan (1975); artikulatorischer Suppressionseffekt, Murray (1967); Effekt irrelevanter Sprache, Colle und Welsh (1976)]. Die CTML übernimmt diese spezifischen Annahmen zur Struktur der PL jedoch nicht, sondern betrachtet die PL als einheitliche Komponente.

Weniger differenziert waren in älteren Versionen des Baddeley-Modells zunächst die Annahmen hinsichtlich des VSSP. Es wurde jedoch bereits eine Unterteilung des VSSP in eine visuelle (vVSSP) und eine räumliche Komponente (rVSSP) vorgeschlagen (Baddeley, 1992). Evidenz für diese Unterteilung kam aus neuropsychologischen Studien, die zeigten, dass einige Patienten Schwierigkeiten hatten, visuelle Merkmale wie z. B. Formen oder Farben bestimmter Objekte zu erinnern, während sie keine Schwierigkeiten hatten, räumliche Aufgaben zu lösen, wie z. B. Wege zu beschreiben oder Objekte zu lokalisieren (Farah, Hammond, Levine & Calvanio, 1988). Diese Zweiteilung des VSSP wurde bis heute beibehalten, allerdings wurden die Mechanismen im VSSP noch weiter spezifiziert. So wird im rVSSP der Steuermechanismus für Bewegungen wie z. B. Armbewegungen, Fingerbewegungen oder Blickbewegungen lokalisiert (vgl. Abschnitt 4.2.2). Darüber hinaus weiß man, dass nicht nur bildhafte Informationen, sondern auch verbale Informationen, wenn sie visuelle oder räumliche Merkmale beschreiben, im VSSP verarbeitet werden (vgl. Abschnitt 4.2.3). Es sei darauf verwiesen, dass diese Annahmen bezüglich des VSSP nicht Bestandteil der CTML sind. Stattdessen wird der VSSP als eine einheitliche Komponente betrachtet (vgl. Abschnitt 2.2.2).

Auch die Funktionen der zentralen Exekutive waren in älteren Versionen des Baddeley-Modells noch wenig geklärt und wurden von Baddeley (1992, S. 557) wie folgt festgelegt: „One proposed role for the central executive is that of coordinating information from two or more slave systems“. Die zentrale Exekutive hatte somit vor allem koordinatorische Aufgaben innerhalb des Arbeitsgedächtnissystems. In neuerer Zeit wurden die exekutiven Prozesse weiter spezifiziert (vgl. z. B. Baddeley, 2007; Miyake, Friedman, Emerson, Witzki & Howerter, 2000; Smith & Jonides, 1999).

Smith und Jonides (1999) unterscheiden dabei die folgenden fünf zentralen Prozesse: (1) *Attention and Inhibition*: Dies impliziert die Fokussierung der Aufmerksamkeit auf relevante Informationen und die Abschirmung gegenüber irrelevanten Informationen. (2) *Task Management*: Hierunter wird die Planung notwendiger Prozesse zur Lösung komplexer Aufgaben sowie der Wechsel des Aufmerksamkeitsfokus auf die jeweils zur Lösung relevante Information verstanden. (3) *Planning*: Hierunter fällt die Planung spezifischer Unterziele sowie ihres Ablaufs zum Erreichen eines übergeordneten Ziels. (4) *Monitoring*: Dies impliziert die Aktualisierung der Inhalte im Arbeitgedächtnis bzw. in den Subkomponenten sowie die Kontrolle der aktuell stattfindenden Prozesse und die Entscheidung, welcher Prozess als nächstes notwendig ist, um das Ziel zu erreichen. (5) *Coding*: Hierunter verstehen Smith und Jonides die Kodierung der Repräsentationen im Arbeitsgedächtnis nach Zeit und Ort ihres Erscheinens.

Während Baddeley in älteren Konzeptionen annahm, die zentrale Exekutive verbinde auch Informationen aus dem Langzeitgedächtnis und dem Arbeitsgedächtnis miteinander, wurde diese Ansicht in neueren Entwicklungen des Modells revidiert. Stattdessen wurde eine weitere Komponente in das Modell aufgenommen: der episodische Buffer (*Episodic Buffer*; z. B. Baddeley, 2000, 2003), in dem die Verbindung der Informationen aus den unterschiedlichen Gedächtnissystemen stattfindet. In Abbildung 1 sind das ursprüngliche Modell (Abschnitt a), auf das sich die CTML bezieht, sowie das erweiterte Modell (Abschnitt b) dargestellt.

Der episodische Buffer bildet demnach die Schnittstelle zwischen den zwei Subsystemen, der zentralen Exekutive und dem Langzeitgedächtnis (Baddeley, 2007; vgl. auch Abbildung 1). Er dient zur Verknüpfung von wahrnehmungsbasierter Information mit Information aus den Subsystemen und Informationen aus dem Langzeitgedächtnis. Die Informationen werden dabei zu episodischen Einheiten verknüpft und sind im episodischen Buffer der bewussten Wahrnehmung zugänglich. Der episodische Buffer wurde von Baddeley (2000) eingeführt, da mit seiner Hilfe das sogenannte *binding problem*, also das Problem der Integration verschiedener Informationen, gelöst werden konnte. Baddeley ging davon aus, dass der episodische Buffer in seiner Kapazität begrenzt ist, da zur Integration der Information auf eine Vielzahl an Codes gleichzeitig zugegriffen werden muss. Zudem nahm er an, dass der episodische Buffer über die zentrale Exekutive kontrolliert wird, welche den

Inhalt des Buffers durch Aufmerksamkeitsprozesse determiniert. Durch die Lenkung der Aufmerksamkeit auf bestimmte Informationen werden diese in den Buffer übernommen. Diese Ansicht revidierte Baddeley jedoch in einem neueren Modell (2007, vgl. Abbildung 1, Abschnitt b). Grund für diese Modifikation waren empirische Befunde, die zeigten, dass die Verknüpfung von unterschiedlichen Informationen, z. B. gesprochener Wörter und visueller Formen, unabhängig von der zentralen Exekutive erfolgte (vgl. z. B. Allen, Hitch & Baddeley, 2009). Dieser Befund legt nahe, dass der episodische Buffer direkten Zugang zu den Informationen des VSSP und der PL hat und die zentrale Exekutive nicht, wie zuvor angenommen, in diesen Prozess involviert sein muss. Wie bereits erwähnt, bezieht sich die CTML auf ältere Konzeptionen des Baddeley-Modells (z. B. 1992, 1999). Daher spielt die Komponente des episodischen Buffers in der CTML keine Rolle.

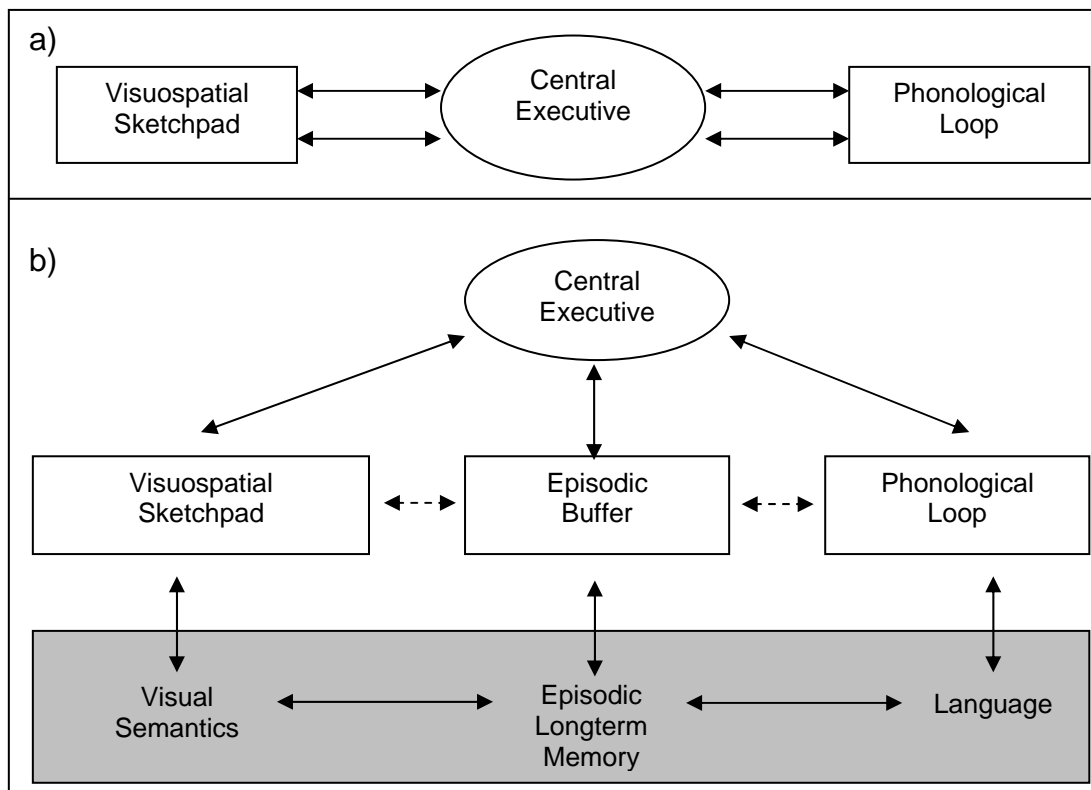


Abbildung 1. Das Modell des Arbeitsgedächtnisses (a) in seiner ursprünglichen Konzeption und (b) nach seiner Erweiterung (nach Baddeley, 2000, S. 418; Baddeley, 2007, S. 147).

Die Theorie der dualen Kodierung nach Paivio (1990)

Eine weitere theoretische Grundlage der CTML bildet Paivios Theorie der dualen Kodierung (*Dual Coding Theory*, 1990), die sich vor allem mit der Repräsentation

von Inhalten im Langzeitgedächtnis beschäftigt, aber auch einige Annahmen zur Verarbeitung von Informationen macht. „The most general assumption in *Dual Coding* theory is that there are ... separate subsystems, one specialized for the representation and processing of information concerning nonverbal objects and events, the other specialized for dealing with language” (Paivio, 1990, S.53). Nach der Theorie der dualen Kodierung besteht das kognitive System also aus einem verbalen System, in dem verbales, sprachbezogenes Material verarbeitet und repräsentiert wird, und einem nonverbalen System, in dem nonverbales, bildliches Material verarbeitet und repräsentiert wird. Paivio unterscheidet drei Arten der kognitiven Verarbeitung: den Aufbau von repräsentativen Verknüpfungen (*Representational Connections*), von assoziativen Strukturen (*Associative Structures*) und von referentiellen Verknüpfungen (*Referential Connections*). Unter repräsentativen Verknüpfungen versteht Paivio die Aktivierung interner verbaler Repräsentationen (*Logogens*) durch externe verbale Stimuli und die Aktivierung interner nonverbaler Repräsentationen (*Imagens*) durch externe nonverbale Stimuli. Assoziative Strukturen sind Aktivierungen innerhalb desselben Systems, die jeweils durch *Imagens* oder *Logogens* ausgelöst werden können: So kann ein *Logogen*, wie z. B. „Rose“, dazu führen, dass man andere *Logogens*, wie z. B. „Blume“ oder „Liebe“, mit ihm assoziiert. Unter referentiellen Verknüpfungen versteht Paivio die Aktivierung des verbalen Systems durch das nonverbale System und vice versa. Demnach kann beispielsweise das Bild einer Rose dazu führen, dass im verbalen System der entsprechende Begriff „Rose“ aktiviert wird. Durch die referentiellen Verknüpfungen können also Bilder verbal benannt und Wörter können bildlich vorgestellt werden. Diese doppelte Kodierung erhöht nach Paivio die Wahrscheinlichkeit des Abrufs der Information, da eine doppelte Gedächtnisspur vorliegt. Evidenz für diese Annahme kommt aus Studien, die zeigen, dass konkrete Nomen (wie z. B. Katze) einfacher erinnert werden als abstrakte Nomen (wie z. B. Tugend), was darauf zurückgeführt wird, dass konkrete Nomen sowohl verbal als auch bildhaft enkodiert werden, während abstrakte Nomen nur verbal enkodiert werden (Paivio, 1965). Nach der Theorie der dualen Kodierung kann der Multimediaeffekt, also die Überlegenheit der Darbietung von Texten zusammen mit Bildern im Vergleich zur reinen Textdarbietung, somit durch die Tatsache erklärt werden, dass die Darbietung von Text mit Bild die Wahrscheinlichkeit einer dualen Kodierung im Gegensatz zur Darbietung von Text ohne Bild erhöht.

Die CTML verbindet die Annahmen der skizzierten Grundlagentheorien miteinander, um daraus Vorhersagen bezüglich der Verarbeitung von Texten und Bildern bzw. Vorhersagen über die kognitive Beanspruchung bei der Informationsverarbeitung abzuleiten, wie im Folgenden aufgezeigt wird.

2.2.2 Die Annahmen der Cognitive Theory of Multimedia Learning

Mayer (z. B. 2005a, 2009) verbindet in der CTML die Annahmen des Drei-Speicher-Modells (Atkinson & Shiffrin, 1968), des Arbeitsgedächtnismodells in seiner ursprünglichen Konzeption (z. B. Baddeley, 1992) sowie der *Theorie der dualen Kodierung* (Paivio, 1986) miteinander. So nimmt die CTML (Mayer, 2009) eine dem Drei-Speicher-Modell entsprechende Dreiteilung des kognitiven Systems an. Diese Dreiteilung wird in Abbildung 2 durch die Unterscheidung in *Sensory Memory*, *Working Memory* und *Long-Term Memory* verdeutlicht.

Die Basis der CTML bilden drei grundlegende Annahmen bezüglich der menschlichen Informationsverarbeitung:

(1) Die Informationsverarbeitung basiert auf zwei Kanälen (*Dual Channel Assumption*): Nach der *Dual Channel Assumption* findet die Informationsverarbeitung in einem auditiv-verbale Kanal (in Abbildung 2 oben) und einem visuell-bildlichen Kanal (in Abbildung 2 unten) statt. Die Differenzierung der beiden Kanäle erfolgt dabei zunächst nach der Sinnesmodalität (auditiv vs. visuell), die durch das zu verarbeitende Material beansprucht wird. Die theoretische Grundlage für diese Differenzierung bildet die Unterteilung des Arbeitsgedächtnisses in die PL und den VSSP nach Baddeley (1992). In einem späteren Stadium der Verarbeitung werden die beiden Kanäle dann in Anlehnung an Paivio's *Theorie der dualen Kodierung* (1990) nach der Kodalität (verbal vs. bildlich) des zu verarbeitenden Materials differenziert. Im auditiv-verbale Kanal wird somit Material verarbeitet, das auditiv dargeboten wird (z. B. Geräusche) und/oder verbaler Natur ist (z. B. Text). Im visuell-bildlichen Kanal wird Material verarbeitet, das visuell dargeboten wird (z. B. Schrift) und/oder bildhafter Natur ist (z. B. Bilder). Es sei bereits hier angemerkt, dass die Informationen zwischen den Kanälen wechseln können: So wird geschriebener Text nach Mayer zunächst im visuellen Teil des *visuell-bildlichen* Kanals verarbeitet, aber dann in den verbale Teil des *auditiv-verbale* Kanals überführt.

(2) Jeder der beiden Kanäle ist in seiner Kapazität begrenzt (*Limited Capacity Assumption*): In Anlehnung an Baddeley (1992) ist die Menge an Information, die jeweils zu einem Zeitpunkt in einem der beiden Kanäle verarbeitet werden kann, begrenzt. Deshalb sollten Informationen so dargeboten werden, dass die Verarbeitung mit Hilfe beider Kanäle stattfinden kann, denn dadurch kann nach Mayer (2005a) zum Einen die zur Verfügung stehende kognitive Kapazität erhöht werden, zum Anderen werden kognitive Engpässe vermieden. Die *Limited Capacity Assumption* ähnelt stark den im Rahmen der *Cognitive Load Theory* (Sweller, 1999; Sweller, van Merriënboer & Paas, 1998) postulierten Annahmen hinsichtlich der kognitiven Belastung. Nach der *Cognitive Load Theory* sollte die dem Lernen abträgliche kognitive Belastung (*Extraneous Cognitive Load*) möglichst gering gehalten werden, damit der Lerner die kognitive Kapazität des Arbeitsgedächtnisses für lernrelevante Prozesse (*Germane Cognitive Load*) nutzen kann.

(3) Lernen setzt die aktive Durchführung kognitiver Prozesse durch den Lerner voraus (*Active Processing Assumption*): Nach der *Active Processing Assumption* ist der Mensch durch die Ausführung bestimmter relevanter kognitiver Prozesse aktiv am Lernprozess beteiligt. Nur durch diese aktive Beteiligung am Lernprozess kann die dargebotene Information angemessen verarbeitet werden und ein kohärentes mentales Modell aufgebaut werden. Die für das Lernen relevanten kognitiven Prozesse sind nach Mayer (2005a): (a) die *Selektion* relevanter Wörter oder relevanter Bilder, (b) die *Organisation* der ausgewählten Wörter oder Bilder in ein verbales bzw. bildhaftes Modell (c) die *Integration* des verbalen Modells, des bildhaften Modells und des Vorwissens in ein kohärentes mentales Modell.

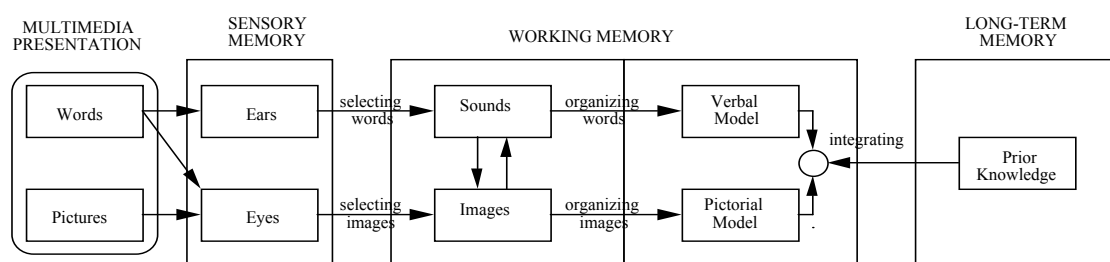


Abbildung 2. Die *Cognitive Theory of Multimedia Learning* (aus Mayer, 2005a, S. 37)

Wenn Lernern multimediales Material, also Texte und Bilder, dargeboten werden, laufen nach der CTML die folgenden Prozesse bei der Informationsverarbeitung ab:

Die Wörter (visuell oder auditiv dargeboten) und Bilder (visuell dargeboten) passieren zunächst das sensorische Gedächtnis. In diesem werden die exakten visuellen oder auditiven Abbilder aller einkommenden Informationen für sehr kurze Zeit festgehalten. Die beiden Kanäle des Verarbeitungssystems haben hier ihren Ursprung: Auditive Information wird über die Ohren in den auditiv-verbale Kanal aufgenommen, visuelle Information wird über die Augen in den visuell-bildlichen Kanal aufgenommen. Durch Aufmerksamkeitszuwendung auf bestimmte Aspekte des im sensorischen Gedächtnis befindlichen Materials findet die *Selektion* relevanter Wörter und Bilder statt. Diese selektierten Aspekte werden dann in das Arbeitsgedächtnis überführt. Im Arbeitsgedächtnis wird die ausgewählte auditive Information zunächst als *Sound*, die ausgewählte visuelle Information als *Image* repräsentiert. Die Pfeile zwischen den *Sounds* und *Images* in Abbildung 2 deuten an, dass der Lerner die Information in den jeweils anderen Kanal übertragen kann. Dies ist vor allem für die Verarbeitung des visuellen verbalen Materials, also geschriebenen Texts, von Bedeutung: Aufgrund der Trennung der Kanäle zunächst nach Modalität (visuell vs. auditiv) und anschließend nach Kodalität (verbal vs. bildlich) ergibt sich für die Verarbeitung visuellen verbalen Materials, dass die Verarbeitung zunächst im visuell-bildlichen Kanal, dann aber im auditiv-verbale Kanal stattfindet. Der Übergang der schriftlich dargebotenen Wörter vom visuell-bildlichen in den auditiv-verbale Kanal erfolgt nach Mayer dadurch, dass der Lerner zu den visuellen Abbildern der Wörter das auditive Äquivalent generiert, so dass die Wörter in den auditiv-verbale Kanal überwechseln.

Im nächsten Schritt findet dann die *Organisation* der Bilder und Wörter statt, indem der Lerner nun die *Sounds* und *Images* in eine kohärente mentale Repräsentation überführt. Dabei bildet der Lerner jeweils innerhalb der *Sounds* und *Images* Verbindungen zwischen den einzelnen Elementen, so dass die *Sounds* in einem verbalen Modell², die *Images* in einem bildhaften, nonverbalen Modell resultieren. Die Unterscheidung der Kanäle erfolgt in dieser Phase der Verarbeitung nach der Kodalität (verbal vs. bildlich) und nicht mehr nach der Modalität. Im letzten Schritt werden das verbale Modell, das bildhafte Modell und Wissen aus dem Langzeitgedächtnis integriert. Korrespondierende Elemente werden dabei aufeinander bezogen und es resultiert eine kohärente mentale Repräsentation, die auch als mentales Modell (*mental model*) bezeichnet wird. Dieses mentale Modell

² Nonverbale auditive Stimuli, wie z. B. non-verbale Musik, werden bei Mayer (2009, 2005a) kaum berücksichtigt.

enthält nach Mayer (2009) die wichtigsten Elemente des Lerninhalts und sollte insbesondere für die Transferleistung von Bedeutung sein, da z. B. Ursache-Wirkungs-Beziehungen aus ihm abgeleitet werden können. Für die reine Erinnerungsleistung sollte das mentale Modell hingegen keine Vorteile bringen, da hierfür die isolierte Text- bzw. Bildinformation dienen kann. Bezüglich der Modalität oder Kodalität des mentalen Modells werden keine Annahmen getroffen.

Nach der CTML erklärt sich der Vorteil der Darbietung von Texten zusammen mit Bildern im Vergleich zur Darbietung von Texten ohne Bildern durch die folgenden zwei Punkte: Bei der Text-Bilddarbietung werden beide Kanäle genutzt, so dass hier die Wahrscheinlichkeit der Bildung eines verbalen *und* eines piktorialen Modells und somit der dualen Kodierung (vgl. Paivio, 1990) gegenüber der reinen Textdarbietung erhöht ist, da die gleichen Informationen sowohl bildhaft als auch verbal dargestellt werden können. Diese duale Kodierung sollte insbesondere die Erinnerungsleistung fördern. Des Weiteren ist bei monomedialen Darbietungen die Wahrscheinlichkeit einer simultanen Aktivierung von verbalem und piktorialem Modell geringer – diese ist aber Voraussetzung, um Verbindungen zwischen den beiden Modellen aufzubauen und ein integriertes mentales Modell zu bilden, was wiederum der Transferleistung zugute kommen sollte: „The multimedia presentation allows learners to hold corresponding verbal and pictorial representations in working memory at the same time, thus increasing the chances that learners will be able to build mental connections between them.“ (Mayer, 2001, S. 69)

Im Rahmen der CTML wurden verschiedene Prinzipien postuliert, wie Text und Bild dargeboten werden sollten, um dem Lernen zuträglich zu sein. Neben dem zentralen Multimediaprinzip werden sechs weitere Kernprinzipien formuliert, die in Tabelle 1 dargestellt sind (für eine Übersicht über weitere Prinzipien: vgl. Mayer, 2005, 2009). Für die vorliegende Arbeit ist insbesondere das Modalitätsprinzip (vgl. Abschnitt 3) von Bedeutung.

Tabelle 1

Die sieben Kernprinzipien der CTML und ihre Annahmen

Name	Annahme des Prinzips
Multimediaprinzip	Die gemeinsame Darbietung von Texten und Bildern führt zu besseren Lernergebnissen als die alleinige Textdarbietung.
Modalitätsprinzip	Bilder sollten mit gesprochenen Texten und nicht mit geschriebenen Texten dargeboten werden.
Räumliches Kontiguitätsprinzip	Korrespondierende Elemente von geschriebenen Texten und Bildern sollten räumlich nah beieinander anstatt räumlich getrennt dargeboten werden.
Zeitliches Kontiguitätsprinzip	Korrespondierende Elemente von gesprochenen Texten und Bildern sollten zeitlich nah, d. h. simultan, anstatt zeitlich getrennt, d. h. sequentiell, dargeboten werden.
Kohärenzprinzip	Für das Lernziel irrelevantes Material (Bilder, Texte, Geräusche, Musik) sollte aus der Darbietung ausgeschlossen werden.
Redundanzprinzip	Bilder sollten nur zusammen mit gesprochenen Texten und nicht mit gesprochenen <i>und</i> geschriebenen Texten dargeboten werden.
Prinzip individueller Unterschiede	Designeffekte sind für Lerner mit niedrigem Vorwissen und hohen räumlichen Fertigkeiten größer als für Lerner mit hohem Vorwissen und niedrigen räumlichen Fertigkeiten.

2.3 Die Rolle des Arbeitsgedächtnisses beim Lernen mit Multimedia

Wie aufgezeigt, geht die CTML von einer Involviertheit des Arbeitsgedächtnisses (Baddeley, 1992) beim Lernen aus. Im vorliegenden Abschnitt werden daher empirische Studien vorgestellt, die diese a priori Annahme der Theorie prüfen. Dazu ist es notwendig, zunächst zwei zentrale Forschungsparadigmen einzuführen, die in der Arbeitsgedächtnisforschung genutzt werden, und deren Kenntnis zum Verständnis der darauf folgenden Beschreibungen der empirischen Studien notwendig ist.

2.3.1 Untersuchungsmethoden der Arbeitsgedächtnisforschung

Zur Untersuchung der Involviertheit der einzelnen Komponenten des Arbeitsgedächtnisses in Informationsverarbeitungsprozesse werden insbesondere zwei zentrale Forschungsparadigmen eingesetzt: das sogenannte Zweitaufgabenparadigma (*Dual Task Paradigm*) sowie Methoden zur Messung der Kapazität der einzelnen Arbeitsgedächtniskomponenten.

Bei der Methode des Zweitaufgabenparadigmas handelt es sich um eine Methode, die eine experimentelle Manipulation erforderlich macht, weil die Lerner während der Lernphase zusätzlich eine Zweitaufgabe durchführen müssen. Bei der Messung der individuellen Arbeitsgedächtniskapazität handelt es sich um eine Methode, bei der aufgrund der unterschiedlichen Leistung von Lernern mit unterschiedlicher Arbeitsgedächtniskapazität auf die Involviertheit der Komponenten beim Verarbeitungsprozess zurückgeschlossen wird.

Das Zweitaufgabenparadigma

Beim Zweitaufgabenparadigma müssen die Probanden simultan zur Hauptaufgabe eine zweite Aufgabe bearbeiten. Die Logik des Zweitaufgabenparadigmas ist die folgende: Werden zur Lösung der beiden Aufgaben die gleichen kognitiven Ressourcen benötigt, dann sollte es in der entsprechenden Arbeitsgedächtniskomponente zu Interferenzen kommen, was dann wiederum die Leistung im Vergleich zu einer Bedingung, in der die beiden Aufgaben einzeln durchgeführt werden, senken sollte. Liegt hingegen keine Interferenz vor, sinkt die Leistung also bei simultaner Durchführung der beiden Aufgaben nicht, so impliziert dieses, dass unterschiedliche Arbeitsgedächtniskomponenten in die Verarbeitung involviert sind. In Experimenten, die das Zweitaufgabenparadigma nutzen, werden also verschiedene Hauptaufgaben mit verschiedenen Zweitaufgaben kombiniert, so dass Interferenzen eindeutig auf eine gemeinsame Ressourcennutzung zurückgeführt werden können (Pickering, 2001).

Voraussetzung für die Anwendbarkeit des Zweitaufgabenparadigmas ist, dass es Zweitaufgaben gibt, von denen man weiß, welche kognitiven Ressourcen sie beanspruchen, d. h. also in welchem Arbeitsgedächtnissystem sie verarbeitet werden. Kombiniert man solche Zweitaufgaben mit einer Hauptaufgabe, wie z. B. dem Lösen einer arithmetischen Gleichung, kann man aufgrund der vorhandenen Interferenzen schließen, wo die Hauptaufgabe, also in diesem Fall die arithmetische Gleichung, verarbeitet wird (Andrade, 2001). Im Laufe der Zeit haben sich zur Untersuchung der Arbeitsgedächtnissysteme jeweils spezifische Aufgaben als geeignet erwiesen: So wird zur Untersuchung der PL meist die artikulatorische Suppression, zur Untersuchung des rVSSP räumliches Tastendrücker (*Spatial Tapping*) und zur Untersuchung des vVSSP die Darbietung irrelevanter visueller Stimuli (*Irrelevant [visual] Noise*) genutzt. Zur Untersuchung der zentralen Exekutiv werden Aufgaben

verwendet, bei denen zufällige Sequenzen generiert werden müssen. Bezüglich der Untersuchung des episodischen Buffers gibt es bisher noch keine etablierte Zweitaufgabe. Die Zweitaufgaben zur Untersuchung der Subsysteme sowie der zentralen Exekutive werden im Folgenden kurz beschrieben.

(1) *Artikulatorische Suppression*. Die artikulatorische Suppression wird zur Belastung der PL eingesetzt. Dabei müssen die Probanden während der Ausführung der Hauptaufgabe kontinuierlich Silben, kurze Wörter oder Zahlen artikulieren, wie z. B. „ba be bi bo bu ba be bi bo bu“ etc. (vgl. z. B. Gyselinck, Jamet & Dubois, 2008) oder „eins zwei drei vier fünf sechs eins zwei“ etc. (vgl. z. B. Baddeley, Eldridge & Lewis, 1981). Durch die artikulatorische Suppression wird der *Rehearsal Process* bzw. der Einleseprozess geschriebenen verbalen Materials in die PL behindert (Gathercole & Baddeley, 1993). Kommt es also zu Interferenzen zwischen der artikulatorischen Suppressionsaufgabe und einer Hauptaufgabe, so kann man schließen, dass die PL an der Hauptaufgabe beteiligt ist.

(2) *Spatial Tapping*. Das räumliche Tastendrücken wird zur Belastung des rVSSP verwendet (vgl. Farmer, Berman & Fletcher, 1986). Dabei müssen die Probanden während der Ausführung der Hauptaufgabe meist im Sekundentakt Tasten auf einer Tastatur, die nicht in ihrem Blickfeld liegt, drücken. Die Erklärung für den Effekt des räumlichen Tastendrückens wird in den durch das Drücken induzierten Bewegungen gesehen: Da Bewegungen durch den rVSSP gesteuert werden, führt die Ausführung von Bewegungen wie dem Tastendrücken zu einer Belastung des rVSSP. Diese durch die Bewegung induzierte Belastung führt wiederum zu Interferenzen mit der Verarbeitung von Informationen, die ebenfalls im rVSSP verarbeitet werden. Entsprechend interferieren nicht nur das räumliche Tastendrücken, sondern auch andere Bewegungen, wie z. B. Armbewegungen oder Blickbewegungen, mit der Verarbeitung räumlicher Informationen (z. B. Postle, Idzikowski, Della Sala, Logie & Baddeley, 2006; Quinn & Ralstone, 1986). In einer Reihe von Studien konnte die Interferenz des *Spatial Tapping* mit räumlichen Arbeitsgedächtnisaufgaben gezeigt werden (vgl. z. B. Darling, Della Sala & Logie, 2007; Farmer et al., 1986), während keine Interferenzen mit visuellen Arbeitsgedächtnisaufgaben auftraten (Logie & Marchetti, 1991). Interferiert das räumliche Tastendrücken also mit einer Hauptaufgabe, so kann man schließen, dass für diese Hauptaufgabe der rVSSP benötigt wird.

(3) *Irrelevant Noise*. Zur Untersuchung des vVSSP wird die Darbietung irrelevanter visueller Stimuli während der Durchführung der Hauptaufgabe verwendet. Es wird angenommen, dass solche irrelevanten visuellen Stimuli direkten Zugang zum vVSSP erlangen, ohne dass ihnen Aufmerksamkeit geschenkt werden muss (vgl. z. B. Logie, 1986; McConnell & Quinn, 2000). McConnell und Quinn verwendeten eine Technik, die sie als *Dynamic Visual Noise* bezeichneten. Dabei werden auf einem Computerbildschirm Pixel zufällig dargeboten, welche kontinuierlich mit einer Rate von 291 Hz von schwarz zu weiß und vice versa wechseln. Die Autoren konnten zeigen, dass diese Technik den Vorteil einer visuellen Memorierungsstrategie aufhebt. Mit ansteigender Pixelzahl, Dichte und Größe des Feldes nimmt dieser Interferenzeffekt zu. Ein statisches Bild mit Pixeln zeigte diesen Effekt hingegen nicht (vgl. McConnell & Quinn, 2000, Experiment 1). Problematisch ist, dass diese visuelle Zweitaufgabe anscheinend nur mit bestimmten visuellen Aufgaben interferiert (vgl. z. B. Andrade, Kemps, Werniers, May & Szmalec, 2002) und dass die praktische Umsetzung teilweise problematisch sein kann. Daher wird auch häufig auf ein sogenanntes *Pre-load Design* (z. B. Cocchini, Logie, Della Sala, MacPherson & Baddeley, 2002) zurückgegriffen, wenn der vVSSP belastet werden soll.

Die Logik des *Pre-load Design*s besteht darin, vor der Hauptaufgabe eine Gedächtnisaufgabe darzubieten, die ein bestimmtes Arbeitsgedächtnissystem belastet. Die Testung der Gedächtnisaufgabe erfolgt erst nach der Hauptaufgabe. Da die Inhalte der Gedächtnisaufgabe über die Ausführung der Hauptaufgabe hinweg im Arbeitsgedächtnis aktiv gehalten werden müssen, wird davon ausgegangen, dass während der Ausführung der Hauptaufgabe ein spezifischer Teil des Arbeitsgedächtnisses belastet ist. Findet man, dass die Leistung in der Hauptaufgabe im Vergleich zu einer Bedingung ohne Gedächtnisaufgabe sinkt, so kann man schließen, dass die Hauptaufgabe im gleichen Arbeitsgedächtnissystem verarbeitet wird wie die Gedächtnisaufgabe. Zur Belastung des vVSSP werden den Probanden z. B. visuelle Muster dargeboten, die sie über die Lernphase hinweg erinnern müssen (vgl. Cocchini et al., 2002).

(4) *Zufällige Sequenzgenerierung*. Zur Untersuchung der zentralen Exekutive werden Aufgaben verwendet, bei denen die Probanden zufällige Sequenzen generieren sollen, z. B. durch Buchstabennennungen oder Tastendrücken (z. B. Baddeley, Emslie, Kolodny & Duncan, 1998). Durch die Instruktion zur Zufälligkeit

muss der Proband stereotypische und überlernte Sequenzen, wie z. B. 123 oder ABC, explizit meiden und seine Aufmerksamkeit auf die Produktion einer zufälligen Abfolge richten. Dies sollte die Prozesse der zentralen Exekutiv beanspruchen. Da die Aufgaben neben der zentralen Exekutiv auch das jeweilige Subsystem beanspruchen – z. B. Buchstabennennung die PL – werden meist mehrere Aufgaben, die unterschiedliche Subsysteme belasten, miteinander kombiniert. So kann man zwischen solchen Interferenzen, die auf die spezifischen Subsysteme zurückgehen, und solchen, die auf die zentrale Exekutiv zurückgehen, differenzieren.

An dieser Stelle sei erwähnt, dass man in der Literatur eine Reihe an Untersuchungen findet, die sich das Zweitaufgabenparadigma zu Nutze machen, um auf objektive Weise die *generelle* kognitive Belastung zu messen (z. B. Brünen, Steinbacher, Plass & Leutner, 2002). Das Ziel solcher Untersuchungen besteht nicht darin herauszufinden, welche Subkomponenten in die Verarbeitung involviert sind, sondern die Gesamtbelastung des Systems zu messen. Hierzu werden in der Regel Entscheidungsaufgaben verwendet, bei denen die Probanden ihre Entscheidung so schnell wie möglich fällen sollen, während sie gleichzeitig eine Hauptaufgabe ausführen. Während die Methoden der artikulatorischen Suppression, des räumlichen Tastendrückens, des *Irrelevant Noise* oder der zufälligen Sequenzgenerierung also die Frage beantworten sollen, *wo* Informationen verarbeitet werden, soll mit Hilfe der Entscheidungsaufgaben die Frage beantwortet werden, wie viel kognitive Kapazität generell für eine Aufgabe verwendet wird.

Methoden zur Messung der individuellen Arbeitsgedächtniskapazität

Eine weitere Möglichkeit auf die Rolle der verschiedenen Arbeitsgedächtniskomponenten bei der Verarbeitung von Informationen zurückzuschließen, besteht in der Messung der Kapazität der einzelnen Arbeitsgedächtniskomponenten. Wenn die Kapazität der Komponenten unterschiedlich stark mit der Leistung korreliert ist, kann rückgeschlossen werden, welche Komponenten in die Verarbeitung bestimmter Materialien involviert sind.

Zur Messung der Kapazität der PL und des VSSP haben sich insbesondere drei Aufgaben etabliert, die im Folgenden kurz vorgestellt werden. Gemein ist diesen Aufgaben, dass jeweils relativ einfache, wenig sinnhafte Informationen erinnert werden müssen, wie z. B. Zahlenreihen oder bestimmte räumliche Sequenzen. Charakteristischerweise beginnen die Testungen mit sehr kurzen Sequenzen, die

dann im Verlaufe des Tests immer weiter zunehmen, bis der Proband sie nicht mehr korrekt erinnern kann. Die Sequenzlänge, die der Proband in circa 60%³ der Fälle korrekt erinnert, entspricht der Kapazität des jeweils gemessenen Subsystems (Pickering, 2006).

(1) *Digit Span Task*. Der Zahlenspanntest (*Digit Span Task*) ist nach Pickering (2006) die am häufigsten genutzte Methode zur Messung der Kapazität der PL. Den Probanden werden hierbei Sequenzen der Ziffern null bis neun auditiv dargeboten, die sie im Anschluss erinnern sollen. Die Länge der Zahlensequenzen nimmt im Verlauf des Tests zu. Die Ziffern sind zufällig angeordnet und werden mit gleichmäßiger Stimmlage und Geschwindigkeit präsentiert (meist eine Ziffer pro Sekunde) um zu gewährleisten, dass die Zahlen einzeln erinnert werden und *Chunking*-Prozesse keine Rolle spielen. Der Ablauf des Zahlenspanntests sieht wie folgt aus: Es werden jeweils fünf Zahlensequenzen gleicher Länge dargeboten (im Folgenden als *Set* bezeichnet). Das erste Set besteht dabei aus fünf Sequenzen à drei Ziffern, das zweite Set aus fünf Sequenzen à vier Ziffern etc. Nach Darbietung einer Zahlensequenz muss diese durch Nachsprechen wiederholt werden. Es müssen mindestens drei Sequenzen der fünf Sequenzen innerhalb eines Sets korrekt wiedergegeben werden, um zum nächsten Set überzugehen, dessen Sequenzen jeweils eine Ziffer länger sind als die des vorherigen Sets. Die Länge der Zahlensequenzen innerhalb des letzten korrekt erinnerten Sets entspricht der phonologischen Gedächtnisspanne. Bei der Erinnerung an die Zahlensequenzen sollten die *Rehearsal* Prozesse der PL eine entscheidende Rolle spielen – je mehr Informationen man subvokal artikulieren kann, desto höher die Spanne der PL.

(2) *Corsi Block Test*. Zur Untersuchung der Kapazität des rVSSP dient der Corsi Block Test von Milner (1971), der auch als Block-Bord Test (Schellig & Hättig, 1993) bezeichnet wird. Bei diesem Test müssen die Probanden räumliche Sequenzen erinnern, die der Versuchsleiter ihnen vorgibt, indem er bestimmte Blöcke, die auf einem Holzbrett befestigt sind, sequentiell antippt. Zwischen Proband und Versuchsleiter liegt dabei ein Brett, auf dem neun Blöcke in unregelmäßiger Anordnung befestigt sind. Auf der dem Versuchsleiter zugewandten Seite des Corsi Blocks sind Nummern vermerkt (vgl. Abbildung 3, Teil a).

Der Versuchsleiter tippt im Sekundentakt eine mithilfe der Blocknummern festgelegte Sequenz von Blöcken mit dem Finger an, der Proband muss diese

³ Dieses Kriterium variiert je nach Test.

Sequenz dann nachtippen. Die Sequenzlänge nimmt dabei im Verlauf des Experiments zu. Es werden jeweils drei Sequenzen mit gleicher Blockanzahl pro Set dargeboten. Erinnerung der Proband zwei dieser drei Sequenzen eines Sets korrekt, wird zum nächsten Set, dessen Sequenzen einen Block länger sind, übergegangen. Der Test startet mit einer Sequenzlänge von drei Blöcken und er wird abgebrochen, sobald der Proband zwei Fehler innerhalb eines Sets macht. Die Länge der Sequenzen des letzten korrekt erinnerten Sets entspricht dann der räumlichen Gedächtnisspanne des Probanden. Da die Probanden räumliche Sequenzen erinnern müssen, wird angenommen, dass mithilfe des Corsi Block Tests die Kapazität des rVSSP gemessen wird. Ein Vorteil des Corsi Block Tests ist, dass verbale Enkodierungsstrategien, die häufig auch bei visuell-räumlichen Stimuli genutzt werden, nicht zum Tragen kommen. Die Kapazität der PL sollte somit keine Rolle für die Leistung im Corsi Block Test spielen (vgl. Pickering, 2006).

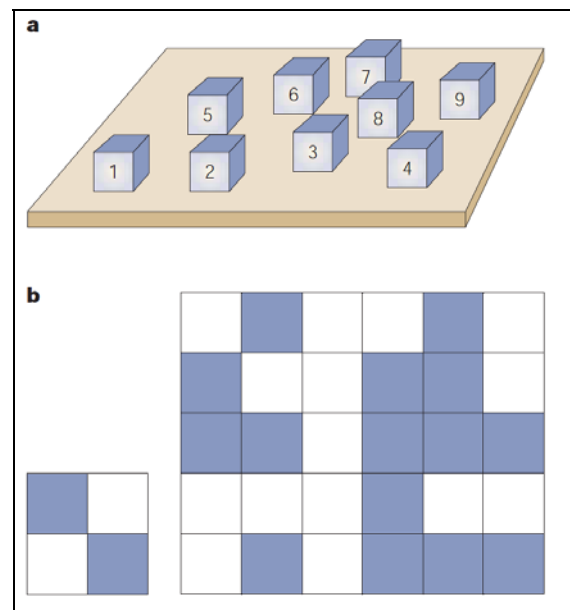


Abbildung 3. Schematische Abbildung (a) des Corsi Blocks und (b) zweier Matrizen unterschiedlicher Komplexität des VPT (aus Baddeley, 2003, S. 833).

(3) *Visual Pattern Test* (VPT). Der VPT von Della Sala, Gray, Baddeley und Wilson (1997) dient der Bestimmung der Kapazität des rVSSP und erfordert das Erinnern abstrakter Muster, die in Form von zweidimensionalen Matrizen dargeboten werden und schwierig verbal zu enkodieren sind. Die Muster ergeben sich jeweils durch die Einfärbung der Hälfte der Zellen der dargebotenen Matrix (vgl. Abbildung 3, Teil b). Der Test beginnt mit einer 2×2 Matrix. Nach einer dreisekündigen Darbietung der

Matrix durch den Versuchsleiter wird die jeweilige Matrix entfernt. Auf einem Antwortblatt wird dem Probanden dann eine unausgefüllte Matrix vorgelegt, die die gleiche Anzahl an Zellen und den gleichen Aufbau hat. Der Proband kennzeichnet auf dem Antwortblatt die Zellen, die seiner Meinung nach bei der Originalmatrix ausgefüllt waren. Wenn die Anzahl der erinnerten Zellen nicht der Anzahl der ausgefüllten Zellen beim Stimulusmaterial entspricht, wird der Proband entsprechend darauf hingewiesen. Er kann dann Korrekturen vornehmen. Nachdem der Proband eine Matrix erinnert hat, wird ihm erneut die Ausgangsmatrix gezeigt und er bekommt Feedback über seine Leistung. Pro Set werden drei Matrizen gleicher Größe gezeigt. Werden diese drei Items korrekt beantwortet, so wird zum nächsten Set übergegangen. Hier kommen pro Matrix zwei Zellen im Vergleich zum vorherigen Set hinzu. Die Testung wird solange fortgesetzt, bis der Proband eine der Matrizen eines Sets nicht mehr korrekt erinnert. Die Anzahl der ausgefüllten Zellen der zuletzt korrekt erinnerten Matrix entspricht der visuellen Gedächtnisspanne. Nach Della Sala, Gray, Baddeley, Allamano und Wilson (1999) misst der VPT die Kapazität des vVSSP, während der Corsi Block Test die des rVSSP misst.

Problematisch ist allerdings, dass beim VPT die Lokalisationen der gefärbten Zellen innerhalb der Matrizen erinnert werden müssen. Dies legt nahe, dass auch hier Räumlichkeit eine Rolle spielen könnte (vgl. Mammarella, Pazzaglia & Cornoldi, 2006). Leider gibt es bis heute kein alternatives Verfahren zum VPT, so dass diese Ungewissheit in Kauf genommen werden muss. Es wurden allerdings zwischen der Leistung im VPT und Corsi Block nur geringe Korrelationen beobachtet (Logie & Pearson, 1997), was nach Meinung vieler Autoren dafür spricht, dass sie in voneinander unabhängigen Systemen des VSSP verarbeitet werden. Zudem zeigten Della Sala et al. (1999), dass die Corsi Block Aufgabe stärker durch räumliche als durch visuelle Interferenzaufgaben gestört wird, während der VPT stärker durch visuelle als durch räumliche Interferenzaufgaben gestört wird. Dies spricht dafür, dass man mit den beiden Aufgaben tatsächlich die Kapazität der beiden Subsysteme des VSSP misst.

Nach dieser Vorstellung der Methoden zur Untersuchung des Arbeitsgedächtnisses werden im Folgenden nun Studien aufgeführt, die die Verarbeitung multimedialen Materials im Arbeitsgedächtnis mit Hilfe der beschriebenen Methoden untersuchten.

2.3.2 Die Verarbeitung multimedialen Materials im Arbeitsgedächtnis

Das Arbeitsgedächtnismodell nach Baddeley (1992) bildet eine zentrale Grundlage der CTML. Es stellt sich allerdings die Frage, inwiefern die Befunde der Arbeitsgedächtnisforschung, in der häufig sehr basales Material verwendet wird, auf die Verarbeitung multimedialen Materials überhaupt übertragbar sind. So wurden den Probanden in Untersuchungen der Arbeitsgedächtnisforschung häufig nur einzelne Wörter, Buchstaben oder einfache Matrizen dargeboten. Bei der Darbietung von multimedialem Material im Sinne der Multimediaforschung geht es aber um die Darbietung von größeren Sinneinheiten, d. h. hier werden ganze Textabschnitte zusammen mit – im Vergleich zu den in der Grundlagenforschung verwendeten Matrizen – komplexen Bildern dargeboten. Daher ist die Frage berechtigt, inwiefern das Arbeitsgedächtnis an der Verarbeitung von multimedialem Material überhaupt beteiligt ist. Es sei angemerkt, dass in den folgenden Studien keine Unterteilung des VSSP in eine visuelle und räumliche Komponente vorgenommen wurde, sondern dass von einer Gesamtbelastung des VSSP durch die Durchführung visuell-räumlicher Zweitaufgaben ausgegangen wurde.

Empirische Überprüfung mit Hilfe des Zweitaufgabenparadigmas

Eine erste Untersuchung zur Involviertheit des Arbeitsgedächtnisses bei der Verarbeitung multimedialen Materials mit Hilfe des Zweitaufgabenparadigmas wurde von Kruley, Sciana und Glenberg (1994) durchgeführt. Sie untersuchten, inwiefern der VSSP in die Verarbeitung von Bildern und inwiefern die PL in die Verarbeitung von gesprochenen Texten involviert ist. Dazu wählten sie ein *Pre-load* Design (vgl. Abschnitt 2.3.1), d. h. die Probanden mussten sich über die experimentelle Lernphase hinweg entweder visuell-räumliche Matrizen zur Belastung des VSSP (Experiment 1) oder eine Anzahl an Zahlen zur Belastung der PL (Experiment 4) merken. In der Lernphase wurden den Lernern entweder gesprochene Texte ohne Bilder oder gesprochene Texte zusammen mit statischen Bildern zur Funktionsweise eines Vulkans präsentiert. Nach der Lernphase wurde ihnen erneut eine Matrix oder zwei Zahlen dargeboten, die sie mit den vor der Lernphase dargebotenen Matrizen bzw. Zahlen abgleichen mussten. In einer Kontrollbedingung wurden den Probanden vor der Lernphase leere Matrizen (Experiment 1) oder keine Informationen (Experiment 4) präsentiert; nach der experimentellen Lernphase mussten die Probanden der Kontrollbedingung eine

andere Entscheidung bezüglich der Matrizen bzw. der Zahlen treffen – sie mussten die Matrizen und Zahlen aber nicht über die Lernphase hinweg im Gedächtnis behalten. Somit waren der VSSP und die PL in der Kontrollbedingung nicht über die Lernphase hinweg belastet. Die Annahme der Autoren war, dass die visuell-räumliche Zweitaufgabe mit der Bildverarbeitung interferieren sollte, wenn auch das Bild im VSSP verarbeitet wird. Die Autoren fanden eine solche Interferenz bezüglich der Leistung in der Zweitaufgabe: Wenn die Probanden über die Lernphase hinweg die Matrizen erinnern mussten, sank in der Bedingung mit Bild – im Gegensatz zur Bedingung ohne Bild – die Wiedererkennungsleistung für die visuell-räumlichen Matrizen. Dies spricht für eine Interferenz der visuell-räumlichen Zweitaufgabe mit der Bildverarbeitung im VSSP. Mit der Zweitaufgabe zur Belastung der PL (Experiment 4) wurde dagegen keine Interaktion beobachtet. Stattdessen sank die Leistung mit verbaler Zweitaufgabe sowohl in der Bedingung mit als auch ohne Bild gleichstark ab. Dies kann auf eine Interferenz der Textverarbeitung und der verbalen Zweitaufgabe in der PL zurückgeführt werden. Bezüglich der Leistung in der Hauptaufgabe fanden die Autoren allerdings keine Interferenzen. Sie erklären dies damit, dass die Lerner die Hauptanstrengung in die Hauptaufgabe investierten, so dass hier keine Unterschiede zum Vorschein kamen. Leider berichten die Autoren nicht die konkrete Instruktion unter der die Teilnehmer die Aufgaben bearbeiteten und die eventuell Aufschluss darüber hätte geben können, inwiefern das Hauptaugenmerk auf die Bearbeitung der Hauptaufgabe gelenkt wurde.

Eine ähnliche Untersuchung – allerdings mit geschriebenen Texten – wurde von Gyselinck, Ehrlich, Cornoldi, De Beni und Dubois (2000) durchgeführt. Die Autoren boten ihren Probanden in der Lernphase verschiedene physikalische Inhalte entweder nur mit geschriebenen Texten oder mit geschriebenen Texten zusammen mit statischen Bildern dar. Genau wie Kruley et al. (1994) wählten auch sie ein *Pre-load* Design, bei dem entweder der VSSP oder die PL während der Lernphase belastet wurde. Sie konnten die Befunde von Kruley et al. (1994) jedoch nicht bestätigen. Stattdessen fanden sie keine spezifischen Interferenzen zwischen den Zweitaufgaben und der Verarbeitung der Bilder und Texte. Problematisch an dieser Untersuchung war jedoch die gewählte Kontrollbedingung: In dieser mussten die Probanden über die Lernphase hinweg sowohl Wörter als auch deren Orte innerhalb einer Matrix erinnern. Diese Bedingung sollte jedoch sowohl die PL als auch den VSSP belasten. Insofern ist es nicht erstaunlich, dass zwischen der

Kontrollbedingung und den Experimentalbedingungen keine Unterschiede nachweisbar waren.

In einer zweiten Untersuchung wechselten die Autoren (Gyselinck, Cornoldi, Dubois, De Beni & Ehrlich, 2002) daher die verwendete Zweitaufgabe. Statt eines *Pre-load* Designs wurden zwei simultane Zweitaufgaben gewählt, nämlich räumliches Tastendrücken und artikulatorische Suppression. Wie bereits in der im Jahr 2000 durchgeführten Untersuchung (Gyselinck et al., 2000) boten Gyselinck und Kollegen den Lernern Inhalte physikalischer Natur dar und variierten zum Einen, ob die geschriebenen Texte mit oder ohne statische Bilder dargeboten wurden, und zum Anderen ob eine Zweitaufgabe zur Belastung des VSSP, zur Belastung der PL oder keine Zweitaufgabe durchgeführt werden musste. Die Ergebnisse sind in Abbildung 4 dargestellt und sind im Einklang mit den Ergebnissen Kruleys et al. (1994): In der Kontrollbedingung, in der keine Zweitaufgabe durchgeführt wurde, trat der Multimediaeffekt auf, d. h. Probanden mit Bild zeigten ein besseres Verständnis der Inhalte als Probanden ohne Bild (vgl. Abbildung 4 links). Dieser Multimediaeffekt wurde in der Bedingung mit räumlicher Zweitaufgabe aufgehoben, da die Leistung der Lerner mit Bild im Vergleich zur Kontrollgruppe stark absank, während die Leistung der Lerner ohne Bild konstant blieb (vgl. Abbildung 4, Mitte). Die räumliche Zweitaufgabe interferierte also nur mit dem Bild, nicht mit dem dargebotenen geschriebenen Text. Für die verbale Zweitaufgabe wurde hingegen eine Interferenz in beiden Bedingungen nachgewiesen (vgl. Abbildung 4 rechts), was auf die Interferenz zwischen verbaler Zweitaufgabe und der Textverarbeitung in der PL zurückgeht. Da die Textverarbeitung sowohl in der Text ohne Bild als auch in der Text mit Bild Bedingung notwendig war, führte die verbale Zweitaufgabe auch in den beiden Bedingungen zu Interferenzen. Erwartungsgemäß interferierte die verbale Zweitaufgabe nicht mit der Bildverarbeitung – vielmehr blieb der Multimediaeffekt auch innerhalb der Gruppe mit verbaler Zweitaufgabe bestehen (vgl. Abbildung 4, rechts).

Nam und Pujari (2005) untersuchten, inwiefern der VSSP nicht nur in die Verarbeitung statischer, sondern auch in die Verarbeitung dynamischer Bilder involviert ist. Sie boten den Lernern geschriebene Texte dar, die entweder alleine, mit statischen Bildern oder mit dynamischen Bildern präsentiert wurden. Inhaltlich ging es um verschiedene technische Domänen, z. B. die Funktionsweise eines Kühlschranks oder eines Dieselmotors. Die Lerner mussten während der Lernphase

entweder eine räumliche Tappingaufgabe zur Belastung des VSSP oder keine Zweitaufgabe durchführen. Die Autoren konnten zeigen, dass auch bei der Verarbeitung dynamischer Bilder der VSSP involviert ist. Besonders bei Personen mit hoher räumlicher Kapazität führte die räumliche Zweitaufgabe zu einem starken Leistungsabfall im Vergleich zur Kontrollgruppe. Die Autoren erklären diesen Befund dahingehend, dass vor allem Lerner mit hoher räumlicher Kapazität von den dargebotenen Bildern und Animationen profitierten, so dass hier die Effekte einer Zweitaufgabe auf die Verarbeitung der Bilder besonders deutlich werden.

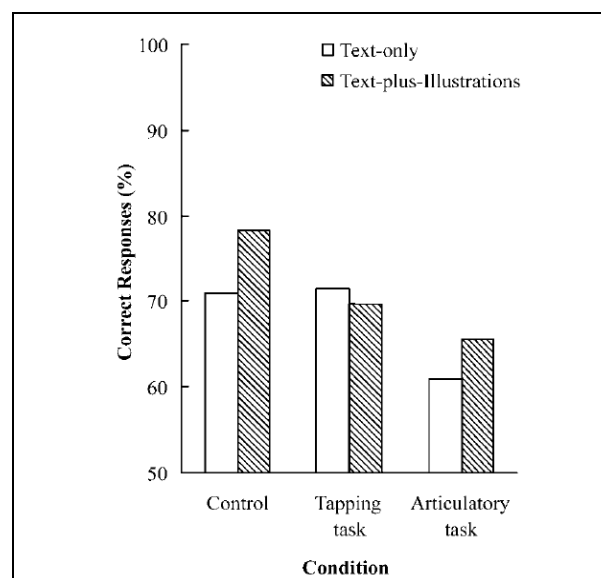


Abbildung 4. Die Ergebnisse der Untersuchung von Gyselinck et al. (2002, S. 672).

Eine weitere Untersuchung zur Rolle des Arbeitsgedächtnisses beim Lernen mit Multimedia, diesmal allerdings bezüglich prozeduraler Inhalte und nicht kausaler Inhalte wie in den zuvor aufgeführten Studien, wurde von Brunyé, Taylor, Rapp und Spiro (2006) durchgeführt. Neben der Rolle des VSSP und der PL beim Lernen mit Multimedia untersuchten Brunyé et al. (2006) auch die Rolle der zentralen Exekutive bei der Verarbeitung dieser Lernmaterialien. Die Probanden mussten daher während der Lernphase in verschiedenen Bedingungen – neben dem räumlichen Tastendrücken und der artikulatorischen Suppression – zwei *Random Generation* Aufgaben zur Belastung der zentralen Exekutive durchführen, d. h. sie mussten entweder zufällig Tasten drücken oder zufällig Silben generieren. Zudem variierten die Autoren, ob den Probanden nur geschriebene Texte, nur Bilder oder geschriebene Texte zusammen mit Bildern dargeboten wurden. Die Untersuchung bestätigte erneut, dass räumliche Zweitaufgaben mit der Bildverarbeitung interferieren,

während verbale Zweitaufgaben mit der Textverarbeitung interferieren. Bilder werden somit im VSSP verarbeitet, während Texte in der PL verarbeitet werden. Bei der Darbietung von Text *und* Bild kam es zudem zu stärkeren Interferenzen mit beiden *Random Generation* Aufgaben als bei der Darbietung von nur Text oder nur Bild. Die Autoren schließen daraus, dass diese stärkere Interferenz bei der Text-Bild Darbietung auf die Beanspruchung der zentralen Exekutive rückzuführen ist, die somit bei der Text-Bild Bearbeitung ebenfalls eine entscheidende Rolle zu spielen scheint. Nach Meinung der Autoren besteht die Rolle der zentralen Exekutive vor allem in der Steuerung der Aufmerksamkeit auf Text und Bild sowie der Integration von Text und Bild.

Die beschriebenen Studien belegen, dass das Arbeitsgedächtnis auf spezifische Weise in die Verarbeitung multimedialen Materials involviert ist: Verbale Informationen werden in der PL verarbeitet, bildhafte Informationen werden im VSSP verarbeitet. Tabelle 2 fasst die Studien sowie deren Befunde zusammen.

Tabelle 2

Übersicht über empirische Studien zur Involviertheit des Arbeitsgedächtnisses beim Lernen mit Multimedia

Autoren	Design	Art der Zweitaufgabe	Lernmaterial	Haupteffekte	Interaktionen
Kruley et al. (1994) Exp. 1	ZA (VSSP vs. Kontrolle) × Bilddarbietung (mit vs. ohne)	VSSP: Matrizen über Lernphase hinweg erinnern Kontrolle: Matrizen nicht über Lernphase hinweg erinnern	gesprochene Texte kausal (Vulkan)	L-HA: mit Bild > ohne Bild mit ZA < Kontrolle	L-HA: keine Interaktion L-ZA: ZA interferiert mit Bildverarbeitung
Kruley et al. (1994) Exp. 2	ZA (PL vs. Kontrolle) × Bilddarbietung (mit vs. ohne)	PL: Zahlen über Lernphase hinweg erinnern Kontrolle: Zahlen nicht über Lernphase hinweg erinnern	gesprochene Texte kausal (Vulkan)	L-HA & L-ZA: mit Bild > ohne Bild mit ZA < Kontrolle	L-HA: keine Interaktion L-ZA: keine Interaktion
Gyselinck et al. (2000)	ZA (VSSP vs. PL vs. Kontrolle) × Bilddarbietung (mit vs. ohne)	VSSP: Matrizen erinnern PL: Phantasiewörter erinnern Kontrolle: Phantasiewörter in Matrix erinnern	geschriebene Texte kausal (physikalische Domänen)	L-HA: mit Bild > ohne Bild mit ZA = Kontrolle	L-HA: keine Interaktion
Gyselinck et al. (2002) Exp. 1	ZA (rVSSP vs. PL vs. Kontrolle) × Bilddarbietung (mit vs. ohne)	rVSSP: räumliches Klopfen PL: artikulatorische Suppression Kontrolle: keine ZA	geschriebene Texte kausal (physikalische Domänen)	L-HA: mit Bild = ohne Bild mit ZA < Kontrolle	L-HA: Kontrolle: mit Bild > ohne Bild ZA rVSSP: Interferenz mit Bildverarbeitung ZA PL: Interferenz mit Textverarbeitung
Gyselinck et al. (2002) Exp. 2	ZA (rVSSP vs. PL vs. Kontrolle) × Darbietung (Text vs. Bild)	rVSSP: räumliches Klopfen PL: artikulatorische Suppression Kontrolle: keine ZA	geschriebene Texte kausal (physikalische Domänen)	L-HA: Text = Bild mit ZA < Kontrolle	L-HA: Kontrolle: Text = Bild ZA rVSSP: Interferenz mit Bildverarbeitung ZA PL: Interferenz mit Textverarbeitung
Nam & Pujari (2005)	ZA (rVSSP vs. Kontrolle) × Bilddarbietung (statisch vs. dynamisch vs. ohne)	rVSSP: räumliches Klopfen Kontrolle: keine ZA	geschriebene Texte kausal (technische Domänen)	L-HA: mit Bild > ohne Bild mit ZA < Kontrolle	L-HA: Kontrolle: mit Bild > ohne Bild ZA rVSSP: Interferenz mit Bildverarbeitung
Brunyé et al. (2006)	ZA (rVSSP vs. PL vs. zentrale Exekutive vs. Kontrolle) × Darbietung (nur Text vs. nur Bild vs. Text und Bild)	rVSSP: räumliches Klopfen PL: artikulatorische Suppression Zentrale Exekutive: Zufallsgenerierung von Abfolgen Kontrolle: keine ZA	geschriebene Texte prozedural (Aufbauen von Figuren)	L-HA: Text + Bild > nur Bild > nur Text ohne ZA > mit ZA	L-HA: Kontrolle: Text + Bild > nur Bild > nur Text ZA rVSSP: Interferenz mit Bildverarbeitung ZA PL: Interferenz mit Textverarbeitung ZA zentrale Exekutive: Interferenz mit Multimediarbeitung

Anmerkung. ZA: Zweitaufgabe, L-HA: Leistung in der Hauptaufgabe, L-ZA: Leistung in der Zweitaufgabe

Empirische Überprüfung durch Messung der individuellen Arbeitsgedächtniskapazität

Eine Alternative zur Anwendung des Zweitaufgabenparadigmas stellt die Messung der individuellen Arbeitsgedächtniskapazität und ihre Relation zum Lernergebnis in Abhängigkeit der experimentellen Bedingung dar. Allerdings gibt es im Bereich der Multimediaforschung nur wenige Studien, in denen die Arbeitsgedächtnismaße der PL, des vVSSP und des rVSSP eingesetzt wurden.

Eine Ausnahme bilden die Studien von Gyselinck und Kollegen (Gyselinck et al., 2000; Gyselinck et al., 2002). Neben der Anwendung des Zweitaufgabenparadigmas (siehe oben) erfassten Gyselinck et al. auch die Kapazität des VSSP mit Hilfe des Corsi Blocks. Die Lerner wurden aufgrund ihrer Leistung im Corsi Block Test in zwei Gruppen unterteilt, nämlich Lerner mit hoher und Lerner mit niedriger VSSP Kapazität. Hinsichtlich der PL Kapazität waren die beiden Gruppen vergleichbar. Die Ergebnisse zeigten, dass Lerner mit hoher VSSP Kapazität von der Präsentation der Bilder profitierten, während Lerner mit niedriger VSSP Kapazität keine Lernvorteile bei der Darbietung von Bildern zeigten. Dies impliziert, dass das VSSP bei der Verarbeitung von Bildern involviert ist. In einem zweiten Experiment untersuchten Gyselinck et al. (2000) die Rolle der PL Kapazität beim Lernen mit Multimedia. Erneut wurden die Lerner in zwei Gruppen aufgeteilt, nämlich Lerner mit hoher versus niedriger PL Kapazität. Die VSSP Kapazität der beiden Gruppen war dabei vergleichbar. Die Ergebnisse zeigten, dass Lerner mit hoher PL Kapazität mehr von Textdarbietungen als von Bilddarbietungen profitierten, während Lerner mit niedriger PL Kapazität sowohl bei Text- als auch Bilddarbietung vergleichbare Leistungen zeigten. Diese Ergebnisse legen nahe, dass die PL in die Verarbeitung von Texten nicht aber von Bildern involviert ist. Ein Nachteil der aufgeführten Studien von Gyselinck et al. ist jedoch, dass die Gruppen jeweils nur aus wenigen Versuchspersonen (11-15) bestanden, da nur Lerner mit extremen Werten in die Analysen eingingen.

2.4 Zusammenfassung und Implikationen für die vorliegende Arbeit

Im vorliegenden Kapitel wurde zunächst der Begriff Multimedia definiert: Im Rahmen der vorliegenden Arbeit wird darunter in Anlehnung an Mayer (2005b,

2009) das Lernen mit Texten und Bildern verstanden. Zudem wurde aufgezeigt, dass die empirische Evidenz für den Multimediaeffekt sowohl bezüglich der Erinnerungsleistung als auch der Transferleistung hoch ist: Inhalte werden somit besser erinnert und transferiert, wenn den Lernern Texte und Bilder dargeboten werden anstatt Texte ohne Bilder.

Insbesondere eine Theorie hat sich hinsichtlich des Lernens mit Multimedia etabliert, die CTML. Sie verbindet kognitive Grundlagentheorien wie das Arbeitsgedächtnismodell nach Baddeley (1992), die Theorie der dualen Kodierung nach Paivio (1990) und das Drei-Speicher-Modell nach Atkinson und Shiffrin (1968) miteinander, um das Lernen mit Multimedia zu erklären. Dabei resultiert der Multimediaeffekt nach der CTML zum Einen aus der dualen Kodierung der Inhalte und zum Anderen aus der simultanen Aktivierung eines verbalen und bildhaften Modells, was die mentale Modellbildung unterstützen sollte.

Insbesondere ältere Konzeptionen des Arbeitsgedächtnisses nach Baddeley (z. B., 1992, 1999) nehmen eine zentrale Rolle in der CTML ein. Dass diese a priori Annahme auch empirisch begründbar ist, zeigen Studien, die mit Hilfe des Zweitaufgabenparadigmas bzw. der Messung der individuellen Arbeitsgedächtniskapazität eine Involviertheit des VSSP bei der Verarbeitung von Bildern und der PL bei der Verarbeitung verbaler Informationen nachweisen konnten (z. B. Brunyé et al., 2006; Gyselinck et al., 2002; Kruley et al., 1994). Im folgenden Abschnitt wird eines der im Rahmen der CTML formulierten Prinzipien näher betrachtet – das sogenannte Modalitätsprinzip. Dieses Prinzip ist in der vorliegenden Arbeit vor allem deshalb von besonderem Interesse, da sich an ihm die Annahmen bezüglich der kognitiven Verarbeitungsmechanismen besonders gut exemplifizieren lassen.

3 Das Modalitätsprinzip beim Lernen mit Multimedia

Auf Grundlage der in den Abschnitten 2.2.2 beschriebenen CTML formulierte Mayer verschiedene Prinzipien, wie (multimediales) Lernmaterial gestaltet werden sollte, um das Lernen zu erleichtern. Eines dieser Prinzipien ist das Modalitätsprinzip (Mayer, 2009). In Abschnitt 3.1 erfolgen zunächst eine Definition des Modalitätsprinzips und die Vorstellung entsprechender empirischer Befunde. In Abschnitt 3.2 werden dann zwei gebräuchliche Erklärungen für das Modalitätsprinzip vorgestellt, nämlich zum Einen die Annahme fehlender räumlicher und zeitlicher Kontiguität bei geschriebenem Text sowie zum Anderen die Annahme von Interferenzen bei geschriebenem Text und zusätzlichen kognitiven Ressourcen bei gesprochenem Text. Im Anschluss (Abschnitt 3.3) erfolgt dann eine Beurteilung der Erklärungen hinsichtlich ihrer theoretischen Plausibilität sowie ihres Geltungsbereichs.

3.1 Definition und empirische Evidenz

Nach dem sogenannten Modalitätsprinzip sollten bei einer multimedialen Darbietung die Texte gesprochen und nicht geschrieben dargeboten werden. Mayer (2005c) definiert das Modalitätsprinzip wie folgt: „People learn more deeply from multimedia messages when the words are presented as spoken rather than printed text” (S. 178). Das Modalitätsprinzip beruht auf dem Modalitätseffekt, dem empirischen Befund, dass die Kombination von Bild und gesprochenem Text zu besserer Leistung führt als die Kombination von Bild und geschriebenem Text.

Der dem Modalitätsprinzip zugrundeliegende Modalitätseffekt wurde in einer Reihe von Studien bestätigt, wobei der Versuchsaufbau typischerweise wie folgt aussah: Den Lernern wurden Visualisierungen in Form von statischen oder dynamischen Bildern vorgelegt. Diese Visualisierungen wurden je nach experimenteller Bedingung entweder durch gesprochene oder geschriebene Texte, die sich im Wortlaut entsprachen, begleitet. Meistens wurde die Geschwindigkeit der Präsentation durch die Geschwindigkeit der gesprochenen Darbietung bestimmt. Im Anschluss wurden verschiedene abhängige Variablen (AVn) erhoben, anhand derer die Lerner dann verglichen wurden.

Die Resultate solcher Studien deuten auf eine starke Evidenz zu Gunsten des Modalitätsprinzips hin. So führte die Darbietung von Bild und gesprochenem Text zu höheren Werten im nahen und/oder weiten Transfer (z. B. Craig, Gholson & Driscoll, 2002; Mayer & Moreno, 1998; Moreno & Mayer, 1999; Schmidt-Weigand, Kohnert & Glowalla, 2009; Tindall-Ford, Chandler & Sweller, 1997), zu größerem Wissenszuwachs (Brünken, Plass & Leutner, 2004), zu besseren Erinnerungsleistungen (z. B. Chung, 2007; Craig et al., 2002; Mayer & Moreno, 1998; Moreno & Mayer, 1999, 2002; Schmidt-Weigand et al., 2009), zu kürzeren Lösungszeiten (Jeung, Chandler & Sweller, 1997; Mousavi, Low & Sweller, 1995, Rinck & Glowalla, 1996) und zu geringerer kognitiver Belastung, sowohl in subjektiven (Tindall-Ford et al., 1997; van Gerven, 2002) als auch in objektiven Maßen (Brünken et al., 2002). Diese Ergebnisse zeigen, dass die Darbietung von gesprochenem Text und Bild nicht nur zu besserer Transferleistung führt, wie es insbesondere Mayer (2009) postuliert, sondern dass auch die Erinnerungsleistung von der Darbietung gesprochenen Texts mit Bild profitiert.

Ginns (2005) kommt in einer Metaanalyse 43 unabhängiger Effekte aus Studien zum Modalitätseffekt zu dem folgenden Schluss: „Across a broad range of instructional materials, age groups and outcomes, students who learned from instructional materials using graphics with spoken text outperformed those who learned from a graphics with printed text“ (S. 326). Die durchschnittliche gewichtete Effektgröße⁴ von 39 Studien mit interindividuellem Design betrug $d = 0.72$, was auf einen mittleren bis hohen Effekt hinweist (Bortz & Döring, 2002).

Der Modalitätseffekt wird dabei durch zwei Faktoren moderiert: Zum Einen tritt der Modalitätseffekt insbesondere dann auf, wenn die Präsentation systemgesteuert ist. Bei selbstgesteuerter Präsentation konnte kein oder nur ein abgeschwächter Modalitätseffekt nachgewiesen werden (Atkinson, 2002; Tabbers, 2002). Dies ist vermutlich darauf zurückzuführen, dass durch Selbststeuerung der Materialdarbietung die dem Modalitätseffekt zugrundeliegenden Prozesse (vgl. Abschnitt 3.2 und Abschnitt 4.1) aufgehoben werden können, weil der Lerner das Material adaptiver verarbeiten kann. Zum Anderen kann das Vorwissen der Lerner einen Einfluss auf den Modalitätseffekt nehmen: So konnte der Modalitätseffekt mit zunehmendem Vorwissen der Lerner nicht mehr nachgewiesen werden (Kalyuga,

⁴ Die Effektgröße berechnete sich wie folgt: Mittelwerte der Bedingungen mit gesprochenem Text minus Mittelwerte der Bedingungen mit geschriebenem Text relativiert an der gemeinsamen Standardabweichung.

Chandler & Sweller, 2000). Dies ist vermutlich darauf zurückzuführen, dass Lerner mit hohem Vorwissen nicht auf Text *und* Bild angewiesen sind, sondern bereits eine der beiden Informationsquellen ausreicht, um ihr Vorwissen zu aktivieren. Daher sollten auch hier die dem Modalitätseffekt zugrundeliegenden Prozesse (vgl. Abschnitt 3.2 und Abschnitt 4.1) aufgehoben werden.

3.2 Theoretische Erklärungen des Modalitätseffekts

Insbesondere zwei Erklärungen für den Modalitätseffekt sind in der Literatur verbreitet: Erstens, die Annahme fehlender räumlicher und zeitlicher Kontiguität bei geschriebenem Text. Zweitens, die Annahme von Interferenzen zwischen den Abbildern des geschriebenen Texts und den Abbildern der Bilder im Arbeitsgedächtnis bzw. die Bereitstellung zusätzlicher Ressourcen bei der Darbietung von gesprochenem Text. Im Folgenden werden diese beiden Erklärungen für den Modalitätseffekt vorgestellt.

3.2.1 *Fehlende räumliche und zeitliche Kontiguität bei geschriebenem Text*

Eine in der Multimediaforschung verbreitete Erklärung für den Modalitätseffekt ist die Annahme fehlender räumlicher und zeitlicher Kontiguität bei der Darbietung geschriebenen Texts im Vergleich zur Darbietung gesprochenen Texts.

Nach der räumlichen Kontiguitätserklärung des Modalitätseffekts gibt es auf Wahrnehmungsebene (in Abbildung 2 *Sensory Memory*) bei der Darbietung von geschriebenem Text und Bild einen Engpass bei der Verarbeitung, denn es müssen Blickwechsel zwischen geschriebenem Text und Bild vollzogen werden. Dies ist eine typische *Split Attention* Situation, d. h. der Lerner muss seine visuelle Aufmerksamkeit zwischen den beiden Informationsquellen aufteilen und visuelle Suchprozesse durchführen. Bei gesprochenem Text und Bild ist ein entsprechendes Ausmaß an visuellen Suchprozessen hingegen nicht notwendig, weil der Lerner gleichzeitig das Bild betrachten und den Text hören kann. Nach der räumlichen Kontiguitätserklärung geht der Modalitätseffekt demnach darauf zurück, dass bei gesprochener Textdarbietung kaum visuelle Suchprozesse von Nöten sind, um Text und Bild aufeinander zu beziehen, was im Vergleich zu geschriebener Textdarbietung in besserer Leistung resultieren sollte.

Nach der zeitlichen Kontiguitätserklärung des Modalitätseffekts wird die mentale Integration von Text und Bild (vgl. Abbildung 2 *Integration*) erleichtert, wenn Text und Bild simultan wahrgenommen und verarbeitet werden können. Da bei der Darbietung gesprochenen Texts der Lerner simultan den Text hören und das Bild betrachten kann, sollten beide Informationen zeitgleich in das Arbeitsgedächtnis eingelesen werden und dort simultan aktiviert sein, was die Integration von Text und Bild erleichtern sollte. Bei der Darbietung von geschriebenem Text und Bild kommt es hingegen zu einer Zeitverzögerung zwischen der Verarbeitung des Texts und des Bilds, da immer nur eine der beiden Informationsquellen im Fokus der visuellen Aufmerksamkeit stehen kann. Daher können die beiden Informationsquellen nicht zeitgleich in das Arbeitsgedächtnis eingelesen werden, was wiederum die Wahrscheinlichkeit der simultanen Aktivierung im Arbeitsgedächtnis sowie der Integration der beiden Informationsquellen senkt. Stattdessen ist es bei geschriebenem Text und Bild notwendig, eine der beiden Informationsquellen durch *Rehearsal* Prozesse aufrecht zu erhalten, wodurch zusätzliche Gedächtnisanforderungen entstehen. Somit kann der Modalitätseffekt gemäß der zeitlichen Kontiguitätsannahme darauf zurückgeführt werden, dass bei gesprochenem Text und Bild nicht aber geschriebenem Text und Bild die synchrone Verarbeitung bildhafter und verbaler Informationen möglich ist, wodurch die mentale Modellbildung bei gesprochener Textdarbietung erleichtert wird.

Neben diesen beiden Kontiguitätserklärungen gibt es eine weitere verbreitete Erklärung für den Modalitätseffekt, die insbesondere von Mayer (2009) präferiert wird und die im Folgenden vorgestellt wird.

3.2.2 *Interferenzen bei geschriebenem Text und zusätzliche kognitive Ressourcen bei gesprochenem Text*

Nach Mayer (2009) erklärt sich der Modalitätseffekt durch eine Bereitstellung zusätzlicher Arbeitsgedächtnisressourcen bei der Darbietung von gesprochenem Text mit Bild bzw. durch Interferenzen im Arbeitsgedächtnis bei der Darbietung von geschriebenem Text mit Bild. Entscheidend für diese theoretische Erklärung ist die Annahme der CTML, dass die Verarbeitung im Arbeitsgedächtnis zunächst getrennt nach der Darbietungsmodalität erfolgt. So hält Mayer (2005a, S. 34) fest: “According to the sensory-modality approach, one channel processes visually represented material and the other channel processes auditorily represented material. This

conceptualization is most consistent with Baddeley's (1992) distinction between the visuo-spatial sketchpad and the articulatory (or phonological) loop."

Diese Annahme der CTML ist in Abbildung 5 verdeutlicht. Demnach werden im Anfangsstadium der Verarbeitung im Arbeitsgedächtnis Abbilder des geschriebenen Texts sowie Abbilder der bildhaften Information im visuell-bildhaften Kanal verarbeitet (vgl. Abbildung 5, Abschnitt B, *Images*). Aufgrund der begrenzten Kapazität des Kanals kann es zu Interferenzen zwischen den Abbildern des geschriebenen Texts und der Bilder kommen und die für das Lernen relevanten Prozesse können nicht mehr ausgeführt werden. Gesprochene Texte werden hingegen nach Mayer (2009) von Anfang an im auditiv-verbale Kanal verarbeitet (vgl. Abbildung 5, Abschnitt A). Daher kommt es nicht zu Interferenzen mit der Bildverarbeitung. Stattdessen stehen den Lernern sogar zusätzlich die kognitiven Ressourcen des auditiv-verbale Kanals zur Verfügung, um das Multimediamaterial zu verarbeiten. Mayer (2009) schreibt diesbezüglich:

(...) both pictures and printed words must enter the learner's information processing through the eyes and initially be represented as images in working memory – thus both compete for resources within the visual channel. (...) By contrast, the most efficient way to present verbal material is through the verbal channel – that is, as spoken text only – because in this way it does not compete with pictures for cognitive resources in the visual channel. (Mayer, 2009, S. 206 f.)

Interessanterweise siedelt Mayer (2009) den Interferenzprozess zwischen geschriebenem Text und Bild vor der eigentlichen Integration des bildhaften und des verbalen Modells an. Dies kann erklären, warum der Modalitätseffekt nicht nur für die Transferleistung, für die ein integriertes mentales Modell Voraussetzung ist, sondern auch für die Erinnerungsleistung, für die bildhafte und verbale Informationen isoliert erinnert werden können, beobachtet wurde.

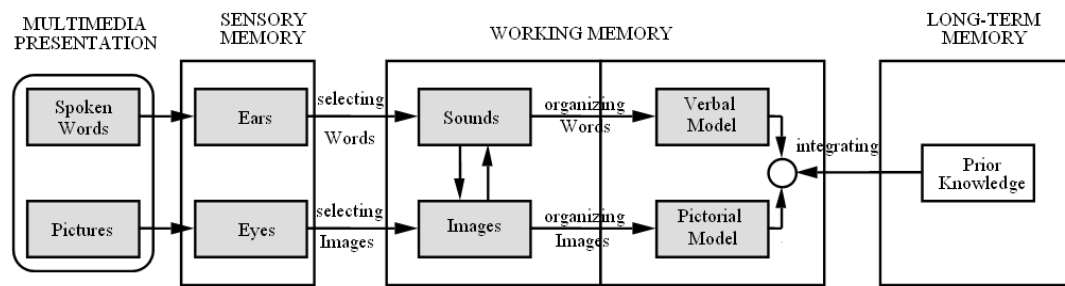
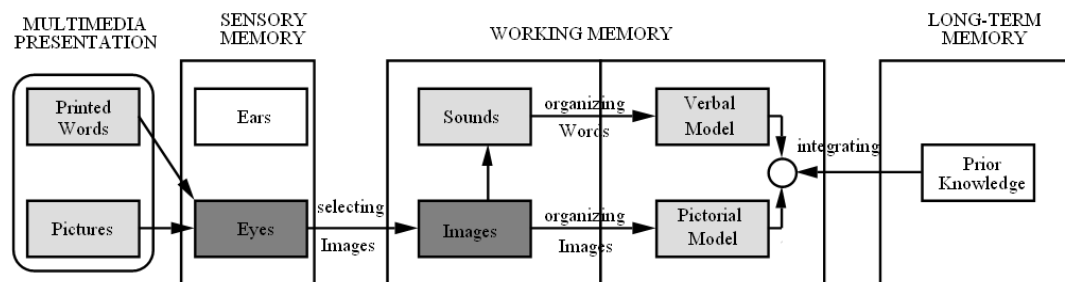
A. Gesprochene Wörter und Bilder:B. Geschriebene Wörter und Bilder:

Abbildung 5. Die Verarbeitung (A) gesprochener Wörter und Bilder und (B) geschriebener Wörter und Bilder nach der CTML (modifiziert nach Mayer, 2005a). Dunklere Einfärbungen entsprechen stärkeren Beanspruchungen.

Die in Abschnitt 3.2.1 aufgeführte Erklärung des Modalitätseffekts durch fehlende zeitliche und räumliche Kontiguität bei geschriebenem Text wird von Mayer (Moreno & Mayer, 1999) explizit zu Gunsten der hier aufgeführten Erklärung abgelehnt. Als empirischen Beleg gegen die zeitliche und räumliche Kontiguitätsklärungen und für die Annahme zusätzlicher zur Verfügung stehender Kapazität bei gesprochenem Text bzw. Interferenzen bei geschriebenem Text wird häufig eine Studie von Moreno und Mayer (1999) angeführt. Hier konnten die Autoren zum Einen zeigen, dass der Modalitätseffekt auch bei sequentieller Darbietung von Text und Bild auftritt. Dies widerspricht der zeitlichen Kontiguitätsklärung, da bei sequentieller Darbietung von gesprochenem Text und Bild keine zeitliche Kontiguität von Text und Bild mehr gegeben ist. Somit ist die zeitliche Kontiguität bei sequentieller Darbietung von gesprochenem Text und Bild mit der sequentiellen Darbietung von geschriebenem Text und Bild vergleichbar. Wird also ein Modalitätseffekt auch bei sequentieller Darbietung gefunden, so kann dieser zumindest nicht auf das Ausmaß der zeitlichen Kontiguität rückgeführt werden, da diese bei der Darbietung geschriebenen und gesprochenen Texts nun

gleich ist. Da in der Untersuchung von Moreno und Mayer (1999) auch bei sequentieller Darbietung ein Modalitätseffekt beobachtet wurde, gehen die Autoren davon aus, dass dieser auf Kapazitätsunterschiede bei der Verarbeitung geschriebenen vs. gesprochenen Texts zurückgehen muss. Es gibt allerdings auch eine weitere Erklärungsmöglichkeit für diesen Befund, die auf die Eigenschaften des in der Studie verwendeten Materials zurückgeht (Rummer, Schweppe, Scheiter & Gerjets, 2008). Demnach kann der von Moreno und Mayer (1999) nachgewiesene Modalitätseffekt bei sequentieller Darbietung auf einen sogenannten *Recency*-Effekt für akustisch dargebotenes Material (vgl. Penney, 1989) zurückgeführt werden: Nach dem *Recency*-Effekt bleiben gesprochene Wörter für kurze Zeit aktiviert, wenn nicht ein neues gesprochenes Wort oder ein anderer akustischer Reiz folgt. Aufgrund dieses akustischen Nachhalls können gesprochene Sätze, bei denen das letzte Wort eine zentrale Rolle spielt, besser rekonstruiert werden als geschriebene Sätze, bei denen es keinen entsprechenden Mechanismus gibt. Dieser *Recency*-Effekt für akustisch dargebotene Information sollte sich aber nur dann auf die Leistung auswirken, wenn Ein-Satz-Material dargeboten wird, denn hier sollte der *Recency*-Effekt die Rekonstruktion des entsprechenden Satzes unterstützen. Bei längeren Texten sollte die Verfügbarkeit eines akustischen Nachhalls des letzten Worts hingegen kaum ins Gewicht fallen, da nur die Rekonstruktion des letzten Satzes, nicht aber des gesamten Texts unterstützt wird. Somit sollte ein Vorteil gesprochenen Texts im Sinne dieser Annahmen auf einzelne Sätze beschränkt sein, nicht aber bei der Darbietung längerer Textabschnitte auftreten. Interessanterweise boten Moreno und Mayer (1999) tatsächlich sehr kurze Sätze dar (vgl. hierzu auch Rummer, Fürstenberg & Schweppe, 2008). Entsprechend gibt es Untersuchungen, die bei vergleichbarer geteilter Aufmerksamkeit zwischen Bedingungen mit geschriebenem und gesprochenem Text und *längerer* Textdarbietung keinen Modalitätseffekt mehr fanden (z. B. Baggett & Ehrenfeucht, 1983; Schüler, Scheiter, Gerjets & Rummer, 2008; Tiene, 2000).

In einem weiteren Experiment (Moreno & Mayer, 1999) konnten die Autoren zudem zeigen, dass der Modalitätseffekt von dem Ausmaß der räumlichen Kontiguität des verwendeten Materials unabhängig zu sein scheint, indem sie eine gesprochene Textdarbietung mit einer entweder räumlich integrierten oder einer räumlich getrennten Präsentation von geschriebenem Text und Bild verglichen. Es zeigten sich ein Modalitätseffekt sowie ein Vorteil der integrierten gegenüber der

nicht-integrierten Text-Bilddarbietung. Dies interpretierten Moreno und Mayer (1999) als Evidenz für die von ihnen präferierte Erklärung des Modalitätseffekts. Problematischerweise wurde der für diese Interpretation zentrale statistische Vergleich zwischen integrierter geschriebener und gesprochener Text-Bilddarbietung aber nicht berichtet. Die absoluten Werte deuten allerdings darauf hin, dass beim Vergleich dieser beiden kritischen Bedingungen kein Modalitätseffekt bzw. nur ein abgeschwächter Modalitätseffekt nachweisbar wäre, wie man es erwarten würde, wenn das Ausmaß an räumlicher Kontiguität bei geschriebenem Text die Ursache für den Modalitätseffekt ist.

Somit sind die von Moreno und Mayer (1999) aufgeführten Befunde gegen die Erklärung des Modalitätseffekts durch fehlende räumliche und zeitliche Kontiguität bei geschriebenem Text als unzureichend zu bewerten. Es kann demnach nicht ausgeschlossen werden, dass die fehlende Kontiguität bei geschriebenem Text den Modalitätseffekt (mit) verursacht.

Für eine Einschätzung der Rolle der beschriebenen Mechanismen bei der Entstehung des Modalitätseffekts erfolgt im nächsten Abschnitt eine Beurteilung der Erklärungen hinsichtlich ihrer Plausibilität und ihres Geltungsbereichs.

3.3 Beurteilung der Erklärungen des Modalitätseffekts

3.3.1 *Plausibilität*

Es erfolgt zunächst eine Einschätzung der Plausibilität der Erklärung des Modalitätseffekts durch Interferenzen bei geschriebenem Text und zusätzliche kognitive Ressourcen bei gesprochenem Text, anschließend dann eine Einschätzung der Plausibilität der beiden Kontiguitätserklärungen.

Interferenzen bei geschriebenem Text und zusätzliche kognitive Ressourcen bei gesprochenem Text

Eine entscheidende Annahme der von Mayer (2009) präferierten Erklärung des Modalitätseffekts ist, dass die Verarbeitung im Arbeitsgedächtnis zunächst getrennt nach der Darbietungsmodalität erfolgt. Demnach werden Abbilder des geschriebenen Texts zunächst im visuell-bildhaften Kanal verarbeitet, während gesprochener Text von Anfang an im auditiv-verbale Kanal verarbeitet wird.

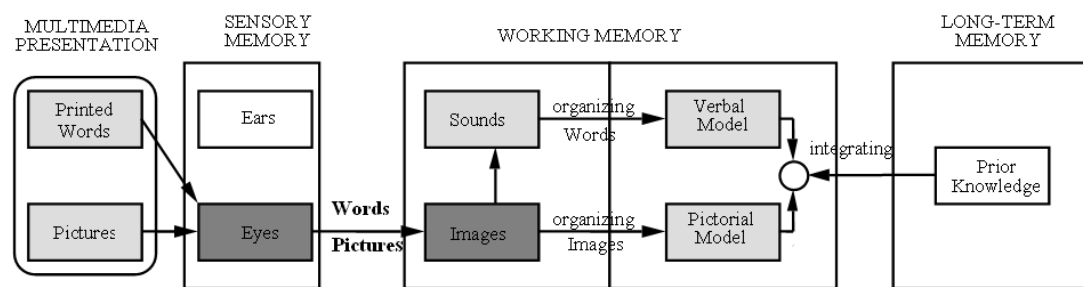
Diese modalitätsspezifische Trennung der beiden Kanäle ist jedoch aus theoretischer Sicht problematisch und wurde z. B. von Tabbers (2002) und Rummer, Schweppe et al. (2008) kritisiert: Denn nach Baddeley (1992) unterscheiden sich die PL und der VSSP nicht hinsichtlich der in ihnen verarbeiteten Modalität, sondern hinsichtlich der in ihnen verarbeiteten *Kodalität*. Demnach wird *verbales*, sprachbasiertes Material in der PL verarbeitet, während *bildhaftes* Material im VSSP verarbeitet wird⁵.

Diese fehlerhafte Interpretation der Verarbeitung innerhalb der beiden Subkomponenten führt dazu, dass die CTML einen kognitiven Engpass für die Darbietung geschriebener Texte mit Bildern im Arbeitsgedächtnis vorhersagt, den es den Annahmen Baddeleys (1992) zufolge gar nicht gibt. Stattdessen wird nach Baddeley sowohl gesprochener als auch geschriebener Text von Anfang an in der PL verarbeitet, da es sich bei beiden Texten um *verbales, sprachbasiertes* Material handelt. Zwar geht Baddeley davon aus, dass geschriebener Text zunächst durch den Prozess der *Inner Speech* in die PL übertragen wird, aber – und das ist der entscheidende Kritikpunkt – Baddeley geht nicht davon aus, dass Abbilder des geschriebenen Texts im VSSP verarbeitet werden, so dass es zu einem Konflikt um die kognitiven Ressourcen bei der Verarbeitung geschriebenen Texts und bildhafter Information kommen könnte. Stattdessen wird nach Baddeley verbale Information unabhängig von der Textmodalität in der PL verarbeitet, auch wenn geschriebener Text zunächst über *Inner Speech* eingelesen werden muss. Die in der CTML getroffene Differenzierung der beiden Kanäle nach der sensorischen Verarbeitungsmodalität im Anfangsstadium der Verarbeitung im Arbeitsgedächtnis ist somit zumindest mit den Annahmen Baddeleys nicht begründbar. Vielmehr müsste in Bezug auf Baddeley auch bereits hier die Unterscheidung der beiden Kanäle anhand der Kodalität des Materials getroffen werden. Abbildung 6 stellt nochmals die Verarbeitung geschriebenen Texts mit Bild nach der CTML (Mayer, 2009) sowie zudem die Verarbeitung von geschriebenem Text mit Bild nach Baddeley (1992) dar.

⁵ Eine Ausnahme bildet verbales Material, welches räumliche oder visuelle Aspekte beschreibt. Dieses wird, wie neuere Befunde zeigen, ebenfalls im VSSP verarbeitet (vgl. Abschnitt 4.3). Allerdings ist dieses unabhängig von der Textmodalität. Neuere Befunde zeigen zudem, dass Blickbewegungen ebenfalls das VSSP belasten können (vgl. Abschnitt 4.2). Daher kann bei geschriebenem Text von einer höheren Belastung des VSSP ausgegangen werden, da mehr Blickbewegungen zur Rezeption notwendig sind. Dies rechtfertigt allerdings ebenfalls nicht die modalitätsspezifische Trennung der Kanäle, da die Texte unabhängig von ihrer Modalität von Anfang an in der PL verarbeitet werden und sowohl Blickbewegungen bei geschriebenem als auch gesprochenem Text den VSSP belasten.

Empirisch werden die Annahmen zur Verarbeitung von geschriebenem Text und Bild im Sinne Baddeleys (1992, vgl. Abbildung 6, B) auch durch die in Abschnitt 2.3.2 aufgeführten Studien unterstützt, die die Verarbeitung multimedialen Materials mit Hilfe des Zweitaufgabenparadigmas untersuchten. So zeigte sich in keiner der Studien, die geschriebene Texte verwendeten, eine spezifische Interferenz zwischen der Textverarbeitung und der räumlichen Zweitaufgabe (vgl. Brunyé et al., 2006; Gyselinck et al., 2002; Nam & Pujari, 2005). Dies zeigt, dass geschriebener Text den rVSSP zumindest nicht so stark belastet, dass es zu Interferenzen kommt. Eine weitere Untersuchung von Gyselinck et al. (2000) zeigte, dass die Verarbeitung geschriebenen Texts auch nicht mit der Erinnerung visueller Matrizen interferiert. Da diese vermutlich eher den vVSSP belasten, scheint geschriebener Text ebenfalls nicht den vVSSP stark zu belasten.

A. Geschriebene Wörter und Bilder nach Mayer (2009)



B. Geschriebene Wörter und Bilder nach Baddeley (1992)

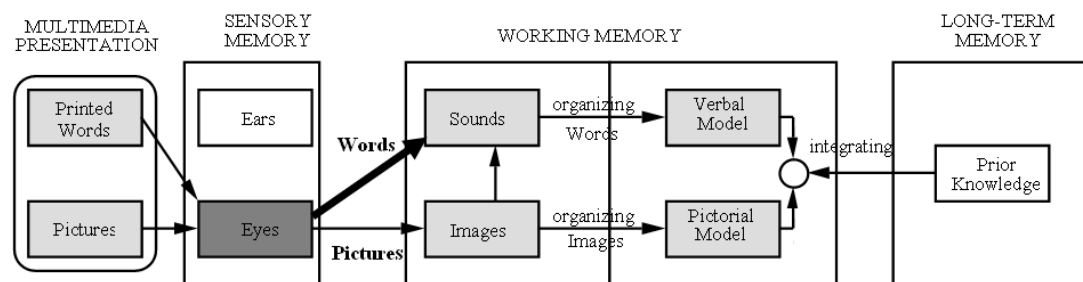


Abbildung 6. Die Verarbeitung geschriebener Wörter und Bilder nach (A) den Annahmen der CTML von Mayer (2009) und (B) dem Arbeitsgedächtnismodell von Baddeley (1992) (nach Rummer, Schweppe et al., 2008).

Gyselinck et al. (2008) führten eine weitere Studie durch, in der es explizit um die Frage ging, welche Arbeitsgedächtnisstrukturen an der Verarbeitung geschriebener und gesprochener Texte beteiligt sind. Den Lernern wurden Bilder mit entweder gesprochenen oder geschriebenen Texten dargeboten. Während der Lernphase

führten sie eine artikulatorische Suppressionsaufgabe, räumliches Tastendrücken oder keine Zweitaufgabe durch. Wenn geschriebener Text im VSSP verarbeitet wird, so sollte bei geschriebenem Text eine Interferenz mit dem räumlichen Tastendrücken auftreten – bei gesprochenem Text hingegen nicht. Die Ergebnisse bestätigten diese Annahme der CTML nicht: Die räumliche Zweitaufgabe führte sowohl mit gesprochenem Text und Bild als auch mit geschriebenem Text und Bild zu gleichem Absinken der Leistung – dieses Absinken der Leistung kann durch die Interferenz mit der Bildverarbeitung erklärt werden. Die artikulatorische Suppressionsaufgabe führte ebenfalls zu gleichem Absinken der Leistung bei geschriebenem und gesprochenem Text, was darauf zurückgeführt werden kann, dass beide Texte – unabhängig von ihrer Darbietungsmodalität – in der PL verarbeitet wurden.

Es kann somit festgehalten werden, dass die von der CTML vorhergesagte Verarbeitung von Texten in Abhängigkeit ihrer Textmodalität im VSSP nicht bestätigt werden kann – dagegen sprechen zum Einen explizit die Befunde der Studie Gyselincks et al. (2008), aber auch indirekt der Befund, dass visuell-räumliche Zweitaufgaben in keiner der in Tabelle 2 aufgeführten Studien mit geschriebenem Text interferierten. Somit ist die Erklärung des Modalitätseffekts im Sinne der CTML nach Mayer (2009) theoretisch nicht plausibel und somit als nicht valide zu beurteilen.

Fehlende räumliche und zeitliche Kontiguität bei geschriebenem Text

Hinsichtlich der Annahmen der Kontiguitätsklärungen wird insbesondere die Plausibilität der räumlichen Kontiguitätsannahme durch empirische Befunde in Frage gestellt. Nach der räumlichen Kontiguitätsannahme führen visuelle Suchprozesse zwischen geschriebenem Text und Bild zu einer schlechteren Leistung im Vergleich zu gesprochenem Text und Bild, da sie zu einer erhöhten Belastung führen. Problematischerweise konnte die a priori Annahme extensiver visueller Suchprozesse zwischen geschriebenem Text und Bild jedoch empirisch nicht bestätigt werden. Stattdessen zeigte sich, dass Lerner mit geschriebenem Text und Bild meist zunächst den gesamten Text lesen und erst am Ende das Bild betrachten (vgl. z. B. Hegarty, 1992; Schmidt-Weigand et al., 2009, 2010), wobei sie dem Text mehr Aufmerksamkeit widmen als dem Bild. Ein hohes Ausmaß an visuellen Suchprozessen zwischen geschriebenem Text und Bild wurde jedoch nicht beobachtet. Somit erscheint die Annahme, dass visuelle Suchprozesse bei

geschriebenem Text und Bild die Ursache für den Modalitätseffekt sind, eher unwahrscheinlich.

Die Annahmen der zeitlichen Kontiguitätserklärung sind hingegen sowohl aus theoretischer als auch empirischer Sicht plausibel und im Einklang mit den Annahmen der Arbeitsgedächtnisforschung. Allerdings ist hinsichtlich der zeitlichen Kontiguitätserklärung der Geltungsbereich eingeschränkt, wie im folgenden Abschnitt aufgezeigt wird.

3.3.2 Geltungsbereich

Wie aufgeführt, machen die drei Erklärungsansätze für den Modalitätseffekt unterschiedliche Annahmen darüber, wo im Verarbeitungsprozess die Ursache für den Modalitätseffekt lokalisiert ist (vgl. Abbildung 7). Während die CTML und die räumliche Kontiguitätserklärung die Engpässe der Verarbeitung geschriebenen Texts mit Bild relativ früh ansiedeln, nämlich vor der Bildung eines integrierten mentalen Modells, erklärt die zeitliche Kontiguitätserklärung den Modalitätseffekt durch eine Erschwerung der integrierten mentalen Modellbildung bei geschriebenem Text und Bild.

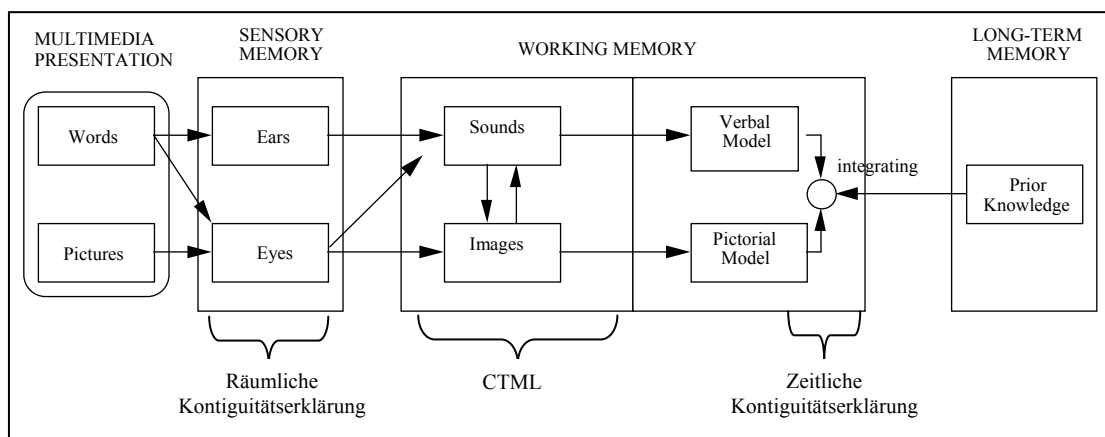


Abbildung 7. Lokalisierung der Ursachen des Modalitätseffekts im Verarbeitungsprozess nach der CTML und der zeitlichen und räumlichen Kontiguitätserklärung.

Eine weitere grundlegende Annahme ist, dass die Bildung eines integrierten mentalen Modells Voraussetzung für die Transferleistung aber nicht für die Erinnerungsleistung von Lernern ist. Somit sollte die Transferleistung dann erschwert werden, wenn Prozesse, die vor oder während der mentalen Modellbildung stattfinden, gestört werden. Die Erinnerungsleistung sollte hingegen nur durch die

Störung von Prozessen erschwert werden, die *vor* der mentalen Modellbildung lokalisiert sind. Werden also Prozesse gestört, die nur die mentale Modellbildung betreffen, sollte die Erinnerungsleistung der Lerner dadurch nicht beeinträchtigt werden. Aufgrund der unterschiedlichen Verortung der Ursachen des Modalitätseffekts im Verarbeitungsprozess sowie die aufgeführten Annahmen hinsichtlich der Erinnerungs- und Transferleistung resultieren unterschiedliche Geltungsbereiche für die drei Erklärungen, wie im Folgenden aufgezeigt wird.

Interferenzen bei geschriebenem Text und zusätzliche kognitive Ressourcen bei gesprochenem Text

Der Geltungsbereich der CTML erstreckt sich sowohl auf die Erinnerungs- als auch die Transferleistung. Dies ist darin begründet, dass die angenommenen Prozesse im Arbeitsgedächtnis, die zu einem Modalitätseffekt führen, der mentalen Modellbildung (vgl. Abbildung 7) vorgeschaltet sind. Demnach sollte ein Modalitätseffekt auch dann beobachtbar sein, wenn die Erinnerung von Faktenwissen abgefragt wird. Dies wird u. a. dadurch belegt, dass wie in Abschnitt 3.1 aufgezeigt, der Modalitätseffekt auch für die Erinnerungsleistung nachgewiesen wurde und nicht nur für die Transferleistung. Der Geltungsbereich der im Rahmen der CTML postulierten Erklärung ist somit nicht eingeschränkt. Problematisch ist allerdings, dass wie in Abschnitt 3.3.1 aufgeführt, die theoretischen Annahmen der Erklärung nach der CTML unplausibel sind.

Fehlende räumliche und zeitliche Kontiguität bei geschriebenem Text

Der Geltungsbereich der räumlichen Kontiguitätserklärung erstreckt sich ebenfalls sowohl auf die Erinnerungs- als auch die Transferleistung, da der angenommene Engpass im visuellen System durch visuelle Suchprozesse der Integration von verbalem und bildhaften Modell mit Vorwissen vorgeschaltet ist (vgl. Abbildung 7). Problematisch ist hier allerdings, dass die angenommenen visuellen Suchprozesse empirisch nicht nachgewiesen werden konnten, so dass die Erklärung wenig plausibel erscheint. Nicht ausgeschlossen werden kann hingegen, dass Lerner mit geschriebenem Text aufgrund der Vernachlässigung des Bildes die Bildinformationen nur oberflächlich verarbeiten. Dies sollte im Vergleich zu Lernern mit gesprochenem Text und Bild in schlechterer Leistung hinsichtlich der Erinnerung des Bildes sowie schlechterer Transferleistung resultieren. Dies konnte in einer

Untersuchung von Schüler et al. (2008) gezeigt werden. Daher sollte der Geltungsbereich der räumlichen Kontiguitätserklärung auf die Erinnerung von bildbezogenem Faktenwissen sowie die Transferleistung beschränkt sein, aber nicht für die textbezogene Erinnerung gelten.

Hinsichtlich des Geltungsbereichs der zeitlichen Kontiguitätserklärung stellt sich die Frage, inwiefern diese nur dann Gültigkeit besitzt, wenn die Bildung eines integrierten mentalen Modells für die Beantwortung der AVn erforderlich ist. Wie in Abschnitt 3.2.1 aufgeführt, wird nach Mayer (2009) durch gegebene zeitliche Kontiguität die Bildung eines integrierten mentalen Modells gefördert. Entsprechend bezieht sich der Effekt der zeitlichen Kontiguität insbesondere auf den Transfer des Gelernten auf neue Situationen, aber nicht auf die Erinnerung von Faktenwissen (für eine Übersicht vgl. Ginns, 2006). Der Geltungsbereich der zeitlichen Kontiguitätserklärung für den Modalitätseffekt ist somit insbesondere für die Transferleistung der Lerner von Bedeutung.

Zusammenfassend kann man aufgrund dieser Analyse festhalten, dass die fehlende zeitliche Kontiguität bei geschriebenen Texten den Modalitätseffekt insbesondere dann erklären kann, wenn die Transferleistung erfasst wird. Die fehlende räumliche Kontiguität sollte den Modalitätseffekt hingegen insbesondere bei bildbezogener Erinnerungsleistung oder Transferleistung erklären können. Da der Modalitätseffekt aber nicht nur hinsichtlich der bildbezogenen Erinnerungsleistung nachgewiesen wurde, deutet dies darauf hin, dass es neben dem Mechanismus fehlender räumlicher und zeitlicher Kontiguität bei geschriebenem Text und Bild einen weiteren Mechanismus gibt, der den Modalitätseffekt erklären kann. Diese dritte Erklärung für den Modalitätseffekt wird in Abschnitt 4 vor dem Hintergrund der VSSP-Struktur abgeleitet.

3.4 Zusammenfassung und Implikationen für die vorliegende Arbeit

Im vorliegenden Abschnitt wurde das Modalitätsprinzip beim Lernen mit Multimedia vorgestellt. Diesem Prinzip zufolge sollten Bilder zusammen mit gesprochenen Texten und nicht zusammen mit geschriebenen Texten präsentiert werden. Die empirische Evidenz für das Modalitätsprinzip ist umfassend. Insbesondere zwei Erklärungen für den Modalitätseffekt sind in der Literatur verbreitet: Die Annahme

fehlender zeitlicher und räumlicher Kontiguität bei geschriebenem Text und die von Mayer (2009) präferierte Annahme spezifischer Interferenzen zwischen geschriebenem Text und Bild bzw. zusätzlicher kognitiver Ressourcen bei gesprochenem Text. Der Geltungsbereich letzterer Erklärung sollte dabei unabhängig davon sein, ob die Erinnerungsleistung oder die Transferleistung erfasst wird, weil die Prozesse der mentalen Modellbildung vorgeschaltet sind. Aus theoretischer Sicht problematisch ist jedoch die im Rahmen der CTML postulierte Annahme, der Modalitätseffekt gehe auf mehr zur Verfügung stehende Ressourcen bei der Darbietung von gesprochenem Text mit Bild bzw. auf generelle Interferenzen zwischen den Abbildern geschriebener Texte und Bilder im VSSP zurück, da die Verarbeitung im Arbeitsgedächtnis modalitätsspezifisch erfolge. Diese Annahme ist nicht mit der Konzeption des Arbeitsgedächtnisses von Baddeley (1992) vereinbar.

Die Annahmen der räumlichen Kontiguitätserklärung sind zwar theoretisch plausibel, konnten aber empirisch nicht bestätigt werden, da Blickbewegungsstudien zeigen, dass Lerner keine extensiven visuellen Suchprozesse durchführen, sondern stattdessen ihre Aufmerksamkeit dem Text widmen und erst danach das Bild, oftmals sehr kurz, betrachten. Somit sollte die räumliche Kontiguitätserklärung nur dann den Modalitätseffekt erklären können, wenn die bildbezogene Erinnerungsleistung oder Transfer erfasst wird. Die zeitliche Kontiguitätserklärung ist theoretisch ebenfalls plausibel, aber auch ihr Geltungsbereich ist eingeschränkt, da sie nur zum Tragen kommen sollte, wenn die Transferleistung erfasst wird.

Aufgrund der mangelnden Plausibilität der Erklärungen bzw. des eingeschränkten Geltungsbereichs der Erklärungen wird im Folgenden nun eine weitere Erklärung für den Modalitätseffekt vorgestellt. Diese besitzt sowohl für die Erinnerungsleistung als auch die Transferleistung Gültigkeit, weil sie den Prozessen der mentalen Modellbildung vorgeschaltet ist. Sie beruht ebenfalls auf Annahmen zu der Struktur des Arbeitsgedächtnisses, insbesondere des VSSP, befindet sich aber im Einklang mit Befunden aus der Arbeitsgedächtnisforschung.

4 Herleitung einer weiteren Erklärung des Modalitätseffekts vor dem Hintergrund der VSSP-Struktur

4.1 Die VSSP-basierte Erklärung des Modalitätseffekts

Die dritte Erklärung für den Modalitätseffekt leitet sich ebenfalls aus der Struktur des Arbeitsgedächtnisses ab. Ihre Basis bilden Befunde hinsichtlich der Informationsverarbeitung im VSSP, die bisher in der Multimedialiteratur nicht zur Kenntnis genommen wurden. Im Gegensatz zu der zeitlichen Kontiguitätsklärung kann sie den Modalitätseffekt auch für die Erinnerung von Faktenwissen erklären, da die angenommenen Prozesse der mentalen Modellbildung vorgeschaltet sind, und im Gegensatz zur räumlichen Kontiguitätsklärung ist ihr Geltungsbereich nicht auf die bildbezogene Erinnerungsleistung beschränkt. Die Erklärung des Modalitätseffekts vor dem Hintergrund der VSSP-Struktur geht davon aus, dass Modalitätseffekte in Abhängigkeit der dargebotenen Textinhalte auftreten können. Demnach sollten geschriebene Texte, die *räumliche* Beschreibungen beinhalten, zu generell schlechterer Erinnerungs- und Transferleistung führen als gesprochene Texte, die *räumliche* Beschreibungen beinhalten. Bei Texten ohne räumliche Inhalte sollten hingegen nur dann Modalitätseffekte zu beobachten sein, wenn speziell die bildbezogene Erinnerungsleistung oder die Transferleistung erfasst werden, denn unter diesen Umständen führen die fehlende zeitliche und räumliche Kontiguität bei geschriebenem Text zu einem Modalitätseffekt.

Die Annahme, dass die Textinhalte den Modalitätseffekt moderieren können, ist darin begründet, dass sowohl räumliche Textinhalte als auch die Kontrolle von Blickbewegungen den rVSSP belasten. Werden also räumliche Textinhalte dargeboten, kommt es zu einer erhöhten Belastung des rVSSP. Werden die räumlichen Textinhalte zusätzlich in geschriebener Form präsentiert, ist ein im Vergleich zur gesprochenen Textdarbietung höheres Ausmaß an Blickbewegungen erforderlich, da Lesen mehr Blickbewegungen erforderlich macht als Hören. Diese zusätzliche Belastung des rVSSP sollte zu einer *Overload* Situation bei der Darbietung räumlicher Textinhalte in geschriebener Form führen, d. h. es sollte ein

Modalitätseffekt resultieren. Bei der Darbietung von Texten ohne räumliche Inhalte sollte hingegen die Belastung des rVSSP bei geschriebener und gesprochener Darbietung jeweils moderat sein, so dass hier das höhere Ausmaß an Blickbewegungen bei geschriebenem Text nicht in einem Modalitätseffekt resultiert. Im Folgenden wird zum Einen die entsprechende empirische Evidenz für die angenommenen Informationsverarbeitungsprozesse im VSSP aufgeführt (vgl. Abschnitt 4.2). Zum Anderen werden empirische Befunde präsentiert, die zeigen, dass der Modalitätseffekt tatsächlich insbesondere dann auftritt, wenn räumliche Textinhalte dargeboten werden (vgl. Abschnitt 4.3).

4.2 Empirische Evidenz: Informationsverarbeitung innerhalb des VSSP

Im folgenden Abschnitt wird empirische Evidenz für die grundlegenden Annahmen der VSSP-basierten Erklärung des Modalitätseffekts berichtet, nämlich erstens, dass der VSSP aus einem visuellen und einem räumlichen Subsystem besteht (vgl. Abschnitt 4.2.1), zweitens, dass im rVSSP die Kontrolle von (Blick)Bewegungen stattfindet (vgl. Abschnitt 4.2.2) und drittens, dass verbal dargebotene Informationen unter bestimmten Umständen ebenfalls im VSSP verarbeitet werden können, nämlich dann, wenn sie Informationen räumlicher oder visueller Natur beinhalten (vgl. Abschnitt 4.2.3).

4.2.1 Die Struktur des VSSP

Die Annahme getrennter visueller und räumlicher Komponenten des VSSP, wie sie bereits von Baddeley im Jahre 1992 vorgeschlagen wurde, ist inzwischen sowohl aus kognitionspsychologischer als auch neuropsychologischer Sicht gut bestätigt. Während im vVSSP visuelle Informationen gespeichert werden, wie Farbe, Form oder Oberflächenmerkmale, werden im rVSSP räumliche Informationen, wie z. B. Bewegungssequenzen oder die Orte bestimmter Objekte, verarbeitet (Logie, 1995). Im Folgenden wird ein kurzer Überblick über die entsprechende empirische Evidenz gegeben. .

Kognitionspsychologische Untersuchungen nutzten vor allem das in Abschnitt 2.3.1 beschriebene Zweitaufgabenparadigma, um die Trennung des VSSP in eine visuelle und räumliche Komponente zu überprüfen. So zeigten z. B. Logie und

Marchetti (1991), dass die Erinnerung an räumliche Muster durch eine räumliche Zweitaufgabe (Armbewegungen während des Erinnerungsintervalls) gestört wurde, während die Erinnerung an visuelle Informationen (Farbe und Form) durch diese räumliche Zweitaufgabe nicht gestört wurde. Andererseits wurde die Erinnerung an visuelle Informationen durch eine visuelle Zweitaufgabe im Erinnerungsintervall gestört, während die Erinnerung an räumliche Muster durch eine solche visuelle Aufgabe nicht gestört wurde. Weitere Evidenz für eine Trennung des VSSP in eine visuelle und eine räumliche Komponente unter zur Zuhilfenahme des Zweitaufgabenparadigmas kommt z. B. aus kognitionspsychologischen Studien von Tresch, Sinnamon und Seamon (1993), Kessels, Postma und De Haan (1999), Quinn und McConnell (1996), Hecker und Mapperson (1997) sowie Hamilton, Coates und Heffernan (2003).

Auch neuropsychologische Studien bestätigen die Trennung des VSSP in eine visuelle und eine räumliche Komponente. So zeigten Studien, die mit bildgebenden Verfahren arbeiteten, dass bei der räumlichen Verarbeitung vor allem dorsale Areale des Gehirns aktiviert sind, während bei der Verarbeitung visueller Informationen eher ventrale Gebiete des Gehirns aktiviert sind (z. B. Courtney, Ungerleider, Keil & Haxby, 1996; Owen, Milner, Petrides & Evans, 1996; Smith et al., 1995). Ein Problem bei der Interpretation bildgebender Verfahren ist jedoch, dass die Trennung zwischen Wahrnehmungsprozessen und Verarbeitungsprozessen schwierig ist. Dies ist insofern problematisch, als dass man nur Letztere als Indikatoren für Prozesse im Arbeitsgedächtnis interpretieren kann. Weitere neuropsychologische Studien zeigten, dass Patienten mit Ausfällen in räumlichen Aufgaben durchschnittliche Leistungen in visuellen Aufgaben erbrachten und vice versa (vgl. z. B. Darling, Della Sala, Logie & Cantagallo, 2006). Diese Befunde bestätigen somit die Trennung des VSSP in eine visuelle und eine räumliche Komponente.

Im Folgenden werden zwei Funktionen des VSSP, die in der visuellen und/oder räumlichen Komponente lokalisiert werden, genauer betrachtet: die Kontrolle von Blickbewegungen (Abschnitt 4.2.2) sowie die Verarbeitung verbal dargebotener visueller und räumlicher Informationen (Abschnitt 4.2.3).

4.2.2 *Die Kontrolle von Blickbewegungen im VSSP*

Die Kontrolle von Bewegungen wird speziell im rVSSP lokalisiert. In der vorliegenden Arbeit sind dabei insbesondere Befunde hinsichtlich *Blickbewegungen*

von Interesse, da Blickbewegungen mit gesprochenen und geschriebenen Texten unterschiedlich stark assoziiert sind. So sollte das Lesen geschriebener Texte zu mehr Blickbewegungen führen als das Hören gesprochener Texte. Daher sind insbesondere Blickbewegungen für die Erklärung des Modalitätseffekts basierend auf der Struktur des VSSP von Bedeutung.

Die empirische Evidenz für die Annahme, dass Blickbewegungen im rVSSP kontrolliert werden, kommt aus Studien, die spezifische Interferenzen zwischen im rVSSP angesiedelten Aufgaben und der Ausführung von Blickbewegungen nachweisen konnten. So zeigten Postle et al. (2006, Experiment 4), dass Sakkaden, die die Probanden zwischen der Darbietung eines vorgegebenen Items (Target) und eines Testitems (Probe) ausführen sollten, speziell mit einer räumlichen Aufgabe, bei der die Probanden den Ort des Targets wiedererkennen mussten, nicht aber mit einer visuellen Aufgabe, bei der die Probanden die Form des Targets wiedererkennen mussten, interferierten.

Auch Woodman und Luck (2004) zeigten, dass das Behalten räumlicher Information mit Blickbewegungen – in diesem Falle induziert durch eine Suchaufgabe – interferierte: Mit steigender Komplexität der Suchaufgabe – und somit zunehmenden Blickbewegungen – sank die Leistung in der räumlichen Aufgabe stärker ab. Bezüglich visueller Aufgaben (Woodman & Luck, 2004) sowie verbaler, nicht-räumlicher Aufgaben (Lawrence, Myerson, Oonk & Abrams, 2001) konnte eine Interferenz mit Blickbewegungen dagegen nicht nachgewiesen werden

Postle et al. (2006) zeigten, dass es nicht die Veränderungen des visuellen Feldes sind, die mit der räumlichen Information interferieren, sondern tatsächlich das Ausführen der Blickbewegungen an sich: Die Veränderung des Hintergrundes, während die Probanden einen Punkt fixierten, also keine Blickbewegungen durchführten, nahm keinen Einfluss auf die Erinnerung räumlicher Information. Ähnliche Ergebnisse fanden Pearson und Sharaie (2003): Auch hier nahm die Veränderung des visuellen Feldes während ein Punkt fixiert wurde keinen Einfluss auf das räumliche Gedächtnis. Zudem führte eine Bedingung, in der Sakkaden mit geschlossenen Augen durchgeführt wurden, ebenfalls zu einer Interferenz mit den räumlichen Inhalten. Entscheidend für eine Interferenz zwischen Blickbewegungen und räumlichen Inhalten ist des Weiteren, dass die Blickbewegungen aktiv durch den Probanden durchgeführt werden und nicht von außen induziert werden (vgl. Postle et al., 2006).

Diese Befunde zeigen, dass es zwischen der Verarbeitung räumlicher Informationen und der aktiven, kontrollierten Ausführung von Blickbewegungen eine Interferenz gibt. Da solche Interferenzen bei der Verarbeitung visueller Informationen bzw. nicht visuell-räumlicher Informationen nicht auftraten, kann man schließen, dass die Interferenz im rVSSP lokalisiert ist. Hier werden somit nicht nur räumliche Informationen verarbeitet, sondern auch Blickbewegungen kontrolliert.

Neben diesen Befunden zur Kontrolle von Blickbewegungen im rVSSP gibt es in neuerer Zeit auch Befunde zur Verarbeitung verbaler Informationen im VSSP. Diese werden im Folgenden vorgestellt.

4.2.3 Die Verarbeitung verbaler Informationen im VSSP

In Abschnitt 2.2.1 wurde bereits aufgezeigt, dass die PL für die Verarbeitung verbaler Informationen verantwortlich ist, während der VSSP für die Verarbeitung visueller und räumlicher Informationen zuständig ist. Im Folgenden wird ein Überblick über Studien gegeben, die zeigen, dass auch *verbale* Informationen unter bestimmten Umständen im VSSP verarbeitet werden, nämlich dann, wenn sie räumliche oder visuelle Aspekte beschreiben⁶. Unter räumlichen Textinhalten werden dabei Informationen verstanden, die räumliche Beziehungen zwischen Objekten beschreiben. Unter visuellen Textinhalten werden hingegen Informationen verstanden, die visuelle Aspekte in den Vordergrund rücken, wie z. B. Farben und Formen. Unter abstrakten Textinhalten werden Informationen verstanden, die weder räumlicher noch visueller Natur sind. Logie und Pearson (1997, S. 242) definieren visuelle und räumliche Information wie folgt: „(...) visual information such as colour and shape, (...) spatial information such as object location and movement sequences.”

Es ist dabei wichtig, die Verarbeitung verbaler Information im VSSP von der Verarbeitung verbaler Information in der PL abzugrenzen. Das Entscheidungskriterium, ob die verbale Information im VSSP verarbeitet wird, ist *semantischer* Natur: Nur wenn visuelle oder räumliche Inhalte vermittelt werden, ist der VSSP involviert. In der PL werden hingegen jegliche verbalen Informationen verarbeitet – unabhängig von ihrer semantischen Bedeutung. Diese Unterscheidung ist wichtig, da demzufolge die PL bei der Darbietung verbaler Informationen *immer*

⁶ Man beachte, dass das Kriterium zur Verarbeitung verbaler Informationen im VSSP der Textinhalt ist und nicht die Textmodalität, wie es die CTML annimmt.

in die Verarbeitung involviert ist, der VSSP hingegen nur, wenn Inhalte mit einem bestimmten semantischen Gehalt dargeboten werden. Im Folgenden werden Befunde zur Verarbeitung verbaler Informationen im VSSP aufgeführt.

Die Verarbeitung räumlicher Textinhalte im VSSP

Bezüglich der Verarbeitung räumlicher Textinhalte im rVSSP gibt es eine Reihe an Untersuchungen. Leider erfolgt die Trennung zwischen visuellen und räumlichen Textinhalten in diesen Studien oftmals nicht explizit, stattdessen wird der VSSP als ein Gesamtsystem betrachtet. Aufgrund der spezifischen Interferenzen mit räumlichen Zweitaufgaben (meist *Spatial Tapping*) bzw. aufgrund der Materialbeschaffenheit ist es aber möglich, die im vorliegenden Abschnitt zitierten Studien dem räumlichen Subsystem zuzuordnen.

In einigen Studien wurde die nach Brooks (1967, 1970) benannte Brooks-Aufgabe verwendet, um die Verarbeitung räumlicher verbaler Informationen versus nicht-räumlicher verbaler Informationen im VSSP zu untersuchen. Bei dieser Aufgabe werden den Probanden (fünf oder acht) Sätze dargeboten, die sie nach der Darbietung allesamt erinnern sollen (vgl. Tabelle 3). Es gibt zwei verschiedene Versionen der Sätze: Die erste Version der Sätze enthält räumliche Informationen (wie z. B. links, rechts, unten, oben usw.), während bei der zweiten Version die räumlichen Wörter durch nicht-räumliche Adjektive (wie gut, schlecht, langsam, schnell) ersetzt werden. Zudem wird eine Matrix dargeboten, in der eine Zelle als Startpunkt markiert ist. Bei der Darbietung der räumlichen Sätze kann sich der Proband einen Weg durch die Matrix vorstellen, während dies bei nicht-räumlichen Sätzen nicht möglich ist, da sie keinen Bezug zur Matrix aufweisen. In Untersuchungen, die die Brooks-Aufgabe nutzten (Baddeley, Grant, Wight & Thomson, 1973; Pazzaglia & Cornoldi, 1999), zeigten sich Interferenzen zwischen räumlichen Zweitaufgaben und der Verarbeitung der Sätze mit räumlichen Inhalten (Tabelle 3, linke Spalte), während solche Interferenzen bei der Verarbeitung nicht-räumlicher Sätze (Tabelle 3, rechte Spalte) nicht nachweisbar war. Dieses Befundmuster spricht dafür, dass bei der Verarbeitung der verbal dargebotenen räumlichen Information der rVSSP involviert ist, während bei der Verarbeitung der nicht-räumlichen Information der rVSSP keine Rolle spielt.

Tabelle 3

Räumliche (links) und nicht-räumliche (rechts) Sätze der Brooks-Aufgabe

Räumliche Inhalte	Nicht-räumliche Inhalte
In the starting square put a 1.	In the starting square put a 1.
In the next square to the <i>right</i> put a 2.	In the next square to the <i>quick</i> put a 2.
In the next square <i>up</i> put a 3.	In the next square to the <i>good</i> put a 3.
In the next square to the <i>right</i> put a 4.	In the next square to the <i>quick</i> put a 4.
In the next square <i>down</i> put a 5.	In the next squarer to the <i>bad</i> pt a5.
In the next square <i>down</i> put a 6.	In the next square to the <i>bad</i> put a 6.
In the next square to the <i>left</i> put a 7.	In the next square to the <i>slow</i> put a 7.
In the next square <i>down</i> put an 8.	In the next square to the <i>bad</i> put an 8.

Untersuchungen von Pazzaglia, De Beni und Meneghetti (2007) und De Beni, Pazzaglia, Gyselinck und Meneghetti (2005) zeigten, dass nicht nur bei der Darbietung einzelner Sätze, sondern auch bei längeren Texten die räumlichen Informationen im rVSSP verarbeitet werden: Pazzaglia et al. (2007) boten ihren Probanden räumliche Inhalte oder nicht visuell-räumliche Inhalte dar. Räumliche Texte bestanden erneut aus einer Situationsbeschreibung, z. B.: „On your right, on the corner of the holiday farm, there is the entrance gate...Once you have reached the restaurant, turn left and, leaving the restaurant behind, you will soon pass a little bridge crossing a small lake” (Pazzaglia et al., 2007, S. 493). Bei nicht visuell-räumlichen Inhalten wurde die Produktion von Wein beschrieben, z. B.: „Before bottling, crystallisation takes place by bringing the wine to sub-zero temperatures, about -5°C. This procedure lasts 2 days and allows the excess tartar to deposit so it can be eliminated later” (Pazzaglia et al., 2007, S. 493). Beide Texte waren mit circa 265 Wörtern relativ lang. Die Texte wurden entweder mit einer verbalen Zweitaufgabe (artikulatorische Suppression) oder einer räumlichen Zweitaufgabe (räumliches Tastendrücken) gepaart oder – in der Kontrollbedingung – ohne Zweitaufgabe ausgeführt. Die Ergebnisse bestätigen die Annahme, dass der rVSSP in die Verarbeitung verbal dargebotener räumlicher Informationen involviert ist: das räumliche Tastendrücken interferierte stärker mit Texten mit räumlichen Inhalten als mit Texten ohne räumliche Inhalte. Die Ergebnisse zeigten außerdem, dass die PL in die Verarbeitung beider Informationsarten involviert war. Die Ergebnisse von Pazzaglia et al. (2007) bestätigen somit, dass bei der Verarbeitung verbaler Informationen die PL generell beteiligt ist, während der rVSSP bei der Verarbeitung von Texten dann eine Rolle spielt, wenn diese räumliche Informationen beinhalten. Es gibt eine Vielzahl weiterer Studien, die mit Hilfe des Zweitaufgabenparadigmas zeigten, dass räumliche verbale Informationen im rVSSP verarbeitet werden (vgl.

z. B. Deyzac, Logie & Denis, 2006; Noordzij, van der Lubbe, Neggers & Postma, 2004)

Neben den Belegen aus kognitionspsychologischen Untersuchungen, dass der VSSP in die Verarbeitung verbal dargebotener räumlicher Informationen involviert ist, gibt es auch Belege aus neuropsychologischen Studien. In einer Untersuchung, die mit bildgebenden Verfahren arbeitete (Wallentin, Weed, Østergaard, Mouridsen & Roepstorff, 2008), wurden den Probanden Sätze vorgegeben, die sowohl eine räumliche als auch eine nicht-räumliche Konfiguration beschrieben. Bei Fragen bezüglich der räumlichen Konfiguration waren solche Teile des Gehirns involviert, die auch bei der Beantwortung von Fragen bezüglich räumlicher *Bild*aspekte involviert sind. Bei Sätzen mit nicht-räumlichen Aspekten war dies nicht der Fall. Phillips, Jarrold, Baddeley, Grant und Karmiloff-Smith (2004) untersuchten Patienten, die an dem sogenannten Williams-Syndrom erkrankt waren. Das Williams-Syndrom zeichnet sich unter anderem dadurch aus, dass die betroffenen Personen relativ gute Sprachfähigkeiten besitzen, aber Defizite in visuellen und räumlichen Fähigkeiten aufweisen. Phillips et al. (2004) untersuchten, inwiefern sich diese Defizite auf das Verständnis von verbalen Informationen auswirkten. Zur Messung des Verständnisses sollten die Probanden auf Bilder zeigen, die einem vom Versuchsleiter vorgelesenen Satz am besten entsprachen. Es zeigte sich, dass die Patienten Schwierigkeiten hatten, Sätze zu verstehen, die räumliche Präpositionen – wie über, unter, in, auf, hinter – enthielten. Bei Sätzen ohne räumliche Inhalte hatten sie hingegen keine Verständnisschwierigkeiten. Dies spricht dafür, dass für das Verständnis verbal dargebotener räumlicher Inhalte visuell-räumliche Fähigkeiten notwendig sind.

Zusammengenommen sprechen diese Befunde dafür, dass verbale räumliche Information ähnlich wie bildhafte räumliche Information im rVSSP verarbeitet wird. Analog zu diesen Befunden sollte sich folglich auch zeigen, dass verbal dargebotene visuelle Inhalte vVSSP verarbeitet werden. Im folgenden Abschnitt werden die entsprechenden Befunde aufgeführt.

Die Verarbeitung visueller Textinhalte im VSSP

Bezüglich der Frage, ob visuelle verbale Informationen – also Informationen über Farben und Formen – im vVSSP verarbeitet werden, gibt es leider nur sehr wenige Studien. Deyzac et al. (2006, S. 218) bemerken hierzu treffend an:

It is no longer sufficient simply to postulate that spatial text or discourse calls for undifferentiated “visuospatial capacities”. What is required is an approach based on the current distinctions made by the working memory model, in particular, the distinction between the visual and spatial components.

Pazzaglia und Cornoldi (1999) untersuchten die Interferenz von räumlichen und visuellen Textinformationen mit räumlichen und visuellen Zweitaufgaben. Sie boten ihren Probanden jeweils Beschreibungen derselben Umgebung dar, wobei entweder räumliche oder visuelle Aspekte der Umgebung für die Beschreibung hervorgehoben wurden. Eine räumliche Beschreibung enthielt beispielsweise die folgenden Informationen:

You are in front of the only entrance to the zoo. Enter the zoo through the entrance which is situated at the centre of one of its four sides. After having passed the entrance you can see in front of you a bar. Turn left and go straight ahead until you arrive in a park. (S. 41)

Eine visuelle Beschreibung bestand hingegen aus der der Beschreibung visueller Merkmale der Situation:

The zoo is situated in a green area, rich in low bushes and tall old trees. In the zoo there is only one entrance, with a monumental gate which has a wide arch. The arch has wonderful red roses, and the gate is made of wrought iron. Inside there is a kiosk with elegant little tables and a green tent providing protection against the strong summer sun. (S.41)

Die Befunde der Studie von Pazzaglia und Cornoldi (1999) sprechen für eine Verarbeitung der räumlichen Informationen im rVSSP: So interferierte eine räumliche Zweitaufgabe stärker mit der räumlichen Textinformation als mit der visuellen Textinformation. Allerdings konnte ein selektiver Einfluss der visuellen Zweitaufgabe auf die visuelle Beschreibung nicht nachgewiesen werden: Die räumliche Zweitaufgabe und die visuelle Zweitaufgabe hatten den gleichen Einfluss auf die Verarbeitung der visuellen Textinformation, was gegen eine Verarbeitung der visuellen Informationen speziell im vVSSP spricht. Stattdessen deutet dieses Ergebnis darauf hin, dass der Text mit visuellen Inhalten sowohl im vVSSP als auch im rVSSP verarbeitet wurde. Problematisch ist allerdings in dieser Studie die Operationalisierung der visuellen Textinhalte: Farben und Formen – von denen man weiß, dass sie im vVSSP verarbeitet werden (vgl. z. B. Farah et al., 1988) – wurden nur wenig genannt, stattdessen wurde der Fokus eher auf eine generelle

Beschreibung des Zoos gelegt, die auch räumliche Signalwörter enthielt (z. B. situated in, inside).

In einer Untersuchung von Deyzack et al. (2006) wurden den Probanden zwei verschiedene Texte dargeboten: In einem Text wurde eine Umgebung mit verschiedenen Sehenswürdigkeiten aus der Vogelperspektive beschrieben, während in einem zweiten Text ein Spaziergang durch die Umgebung mit den gleichen Sehenswürdigkeiten beschrieben wurde. Die Annahme war, dass bei der Spaziergangsbeschreibung eher die visuellen Merkmale der Sehenswürdigkeiten, wie Form, Farbe und Details, enkodiert werden, da man sich hier die Sehenswürdigkeit aus einer frontalen Perspektive vorstellt. Dies sollte bei der Beschreibung aus der Vogelperspektive nicht der Fall sein. Bezüglich einer visuellen Interferenzaufgabe, bei der während der Verarbeitung der Beschreibungen irrelevantes visuelles Material dargeboten wurde (*Dynamic Visual Noise*), konnten die Autoren ihre Annahmen bestätigen: Die Erinnerung an räumliche Informationen (Bewegung/Ort) sowie an die Sehenswürdigkeiten, wenn sie aus der Vogelperspektive beschrieben wurden, wurde durch die visuelle Interferenzaufgabe nicht gestört. Eine Interferenz trat hingegen bei der Spaziergangsbeschreibung auf, bei der nach Meinung der Autoren die visuellen Merkmale der beschriebenen Sehenswürdigkeiten enkodiert wurden. Dieses Ergebnis weist darauf hin, dass auch visuelle Informationen spezifisch im VSSP verarbeitet werden, nämlich im vVSSP.

Ein Problem der beiden beschriebenen Studien besteht allerdings darin, dass die visuellen Informationen im Rahmen von räumlichen Beschreibungen operationalisiert wurden. Dadurch wird eine klare Trennung zwischen visueller und räumlicher verbaler Information schwierig und somit auch der Rückschluss erschwert, wo visuelle und räumliche verbale Informationen im VSSP verarbeitet werden.

Die Ursache für die Verarbeitung verbaler Informationen im VSSP ist bis heute ungeklärt. So gehen einige Forscher davon aus, dass der VSSP bei der Verarbeitung semantischer Informationen aufgrund mentaler Vorstellungsprozesse involviert ist, andere Forscher hingegen vermuten, dass im VSSP tatsächlich Sprachverarbeitungsprozesse stattfinden (siehe Barsalou, 2008; Fischer & Zwaan, 2008; Hauk, Johnsrude & Pulvermüller, 2004; Pulvermüller, 2005). An dieser Stelle wird auf eine weitere Ausführung der einzelnen Erklärungsansätze verzichtet, da die

Ursache für die Verarbeitung verbaler Informationen im VSSP für die vorliegende Arbeit keine Rolle spielt.

Zusammenfassend kann man festhalten, dass die empirische Evidenz die Annahmen der VSSP-basierten Erklärung des Modalitätseffekts bestätigt: Sowohl Blickbewegungen als auch räumliche verbale Inhalte führen zu einer Belastung des rVSSP. Somit sollte die Darbietung geschriebener räumlicher Textinhalte den rVSSP im Gegensatz zur Darbietung gesprochener räumlicher Textinhalte stark belasten. Im nächsten Abschnitt wird empirische Evidenz aufgeführt, die zeigt, dass diese erhöhte Belastung des rVSSP bei geschriebenen räumlichen Textinhalten tatsächlich in schlechterer Leistung resultiert als die Darbietung gesprochener räumlicher Textinhalte. Zunächst erfolgt eine Analyse der Textinhalte in Untersuchungen zum Modalitätsprinzip beim Lernen mit Multimedia, um die Annahme zu prüfen, dass auch hier speziell räumliche Textinhalte verwendet wurden. Allerdings können die Modalitätseffekte hier auch teilweise auf die fehlende räumliche oder zeitliche Kontiguitätsklärung zurückgehen, und zwar dann, wenn sie nur hinsichtlich der Transferleistung oder der bildbezogenen Erinnerungsleistung nachgewiesen wurden. Daran anschließend folgen empirische Befunde aus monomedialen Studien und multimedialen Studien, die experimentell zeigen konnten, dass Modalitätseffekte speziell in Abhängigkeit des Textinhalts auftreten, auch wenn die Erinnerung von Fakten und keine Transferleistung erfasst wurden. Es sei darauf verwiesen, dass die aufgeführten Studien nicht explizit im Kontext der Arbeitsgedächtnisforschung angesiedelt sind. Ihre Ergebnisse lassen jedoch Rückschlüsse darüber zu, wie Textinhalte und Textmodalität im Arbeitsgedächtnis miteinander interagieren.

4.3 Empirische Evidenz: Modalitätseffekte in Abhängigkeit des Textinhalts

4.3.1 *Analyse der Textinhalte in Untersuchungen zum Modalitätsprinzip beim Lernen mit Multimedia*

Die im vorherigen Abschnitt aufgeführten Befunde aus der Arbeitsgedächtnisforschung deuten darauf hin, dass der Textinhalt tatsächlich eine entscheidende Rolle für das Auftreten eines Modalitätseffekts spielen kann. Eine Frage, die sich in diesem Zusammenhang stellt, ist allerdings, inwiefern in klassischen Untersuchungen der Multimedialiteratur tatsächlich räumliche Inhalte

verwendet wurden, so dass die dort gefundenen Modalitätseffekte, neben dem Mechanismus der fehlenden räumlichen und zeitlichen Kontiguität bei geschriebenen Texten, auf die Textinhalte zurückgeführt werden können.

Exemplarisch wird im Folgenden der Fokus auf die von Richard Mayer verwendeten Materialien gelegt. In seinen Untersuchungen zum Modalitätseffekt verwendete er z. B. Materialien, in denen die Entstehung von Blitzen (Mayer & Moreno, 1998; Moreno & Mayer, 1999) oder die Funktionsweise von Autobremsten (Mayer & Moreno, 1998) beschrieben wurden. In Abbildung 8 sind Ausschnitte der Textinhalte dargestellt. Hervorgehoben sind solche Wörter oder Inhalte, die räumliche Aspekte beschreiben. Es wird deutlich, dass die Texte einen hohen Anteil räumlicher Beschreibungen enthielten. Somit können die Modalitätseffekte neben dem Mechanismus fehlender Kontiguität bei geschriebenem Text und Bild auch auf eine höhere Belastung des rVSSP bei der Darbietung der Inhalte in geschriebener Form zurückgehen.

A) Blitzenstehung (aus Mayer, 2009, S. 36f)

Cool moist air moves **over** a warmer surface and becomes heated. Warmed moist air **near** the earth's surface **rises** rapidly. As the air in this **updraft** cools, water vapour condenses into water droplets and forms a cloud. The cloud's top **extends above** the freezing level, so the **upper** portion of the cloud is composed of tiny ice crystals. Eventually, the water droplets and ice crystals become too large to be suspended by the **updrafts**. As raindrops and ice crystals **fall through** the cloud, they drag some of the air in the cloud **downward**, producing **downdrafts**. When **downdrafts** strike the ground, they spread out **in all directions**, producing the gusts of cool wind people feel just before the start of the rain. Within the cloud, the **rising and falling** air currents cause electrical charges to build. The charge results from the **collision** of the cloud's **rising** water droplets against heavier, **falling** pieces of ice. The negatively charged particles **fall** to the bottom of the cloud, and most of the positively charged particles **rise** to the top. A stepped leader of negative charges moves **downward** in a series of steps. It **nears** the ground. A positively charged leader travels **up** from such objects as trees and buildings. The two leaders generally meet about 165-feet **above** the ground. Negatively charged particles then **rush from the cloud to the ground** along the path created by the leaders. It is not very bright. As the leader stroke **nears** the ground, it induces an opposite charge, so positively charged particles from the ground rush **upward** along the same path. This **upward** motion of the current is the **return stroke**. It produces the bright light that people notice as a flash of lightning.

B) Funktionsweise von Autobremsten (aus Mayer, 2009, S. 42)

When the driver steps **on** the car's brake pedal, a piston moves **forward inside** the master cylinder. The piston forces brake fluid **out** of the master cylinder and **through** the tubes to the wheel cylinders. In the wheel cylinders, the **increase** in fluid pressure makes a set of smaller pistons move. These smaller pistons activate the brake shoes. When the brake shoes press **against** the drum, both the drum and the wheel stop or slow down.

Abbildung 8. Die von Mayer und Moreno (1998; Moreno & Mayer, 1999) verwendeten Textinhalte in Studien zum Modalitätseffekt.

Tabelle 4 gibt eine Übersicht über die Inhalte anderer Studien zum Modalitätseffekt sowie, wenn möglich, ein Beispiel für den verwendeten Textinhalt. Zudem wird aufgezeigt, ob der Modalitätseffekt auch hinsichtlich der Erinnerung von

Faktenwissen bestätigt wurde. Es wird deutlich, dass auch in vielen anderen Studien zum Modalitätseffekt die verwendeten Texte einen hohen Anteil räumlicher Information enthielten. Zudem konnte der Modalitätseffekt in einem Großteil der Studien auch für die Erinnerungsleistung bestätigt werden, was die Annahme belegt, dass es neben der zeitlichen Kontiguität einen weiteren Mechanismus gibt, der an der Entstehung des Modalitätseffekts beteiligt ist. Leider wird in den empirischen Studien nicht zwischen text- und bildbezogener Erinnerung unterschieden, so dass die Ergebnisse unter Umständen auch auf die fehlende räumliche Kontiguität zurückgeführt werden können. Interessanterweise zeigte sich jedoch in der Untersuchung von Tabbers, in der Texte mit geringer visuell-räumlicher Information verwendet wurden, auch für die Erinnerungsleistung kein Modalitätseffekt. Dies unterstützt die Annahme, dass neben der räumlichen Kontiguität auch die verwendeten Textinhalte eine Rolle bei der Entstehung des Modalitätseffekts spielen.

Im Folgenden wird Evidenz für die Annahme eines Modalitätseffekts in Abhängigkeit des Textinhalts aus monomedialen als auch multimedialen empirischen Studien berichtet. Dabei wird insbesondere zwischen räumlichen, visuellen und nicht visuell-räumlichen (abstrakten) Textinhalten unterschieden. Leider gibt es keine genauere Festlegung, welche Textinhalte als räumlich und welche Textinhalte als visuell zu beurteilen sind. Diese Entscheidung liegt in der Hand des jeweiligen Forschers. Somit bleibt für eine Konkretisierung der Begriffe „visuell“ und „räumlich“ nur die konkrete Betrachtung des in den Studien verwendeten verbalen Materials. Daher werden im Folgenden, wenn möglich, Materialbeispiele dargeboten.

Tabelle 4

Übersicht über die verwendeten Textinhalte in empirischen Studien zum Modalitätseffekt (ME)

Inhaltsdomäne	Autoren	Beispielsatz	Ergebnis
Autobremsen	Mayer & Moreno (1998)	When the driver steps on the car's brake pedal, a piston moves forward inside the master cylinder	ME Erinnerung
Biologie	Harkskamp, Mayer & Suhre (2007)	A behavioral system is a group of related actions mostly with a clear aim. In this example: sitting and watching the prey, creeping up on the prey, preparing to jump staying low, jumping on the prey and carrying the prey away.	ME Transfer
Blitzentstehung	Moreno & Mayer (1999)	As raindrops and ice crystals fall through the cloud, they drag some of the air in the cloud downward, producing downdrafts.	ME Erinnerung
Blutkreislauf	Brünken, Seufert, Zander (2005) Brünken et al. (2004)	Um seine vielfältigen Funktionen erfüllen zu können, muss das Blut alle Zellen des Körpers erreichen können: es benötigt ein eigenes leistungsfähiges Transportwegesystem: den Blutkreislauf mit seinen mehrere Tausend Kilometer langen Blutgefäßbahnen.	ME Erinnerung
Botanik	Moreno, Mayer, Spiers, Lester (2001) Moreno & Mayer (2002)	A short stem here in this shade is dangerous for the plant, because its leaves won't get any sunlight. The stem should be long enough to put the leaves in the sun.	ME Erinnerung
Elektrische Schaltkreise	Tindall-Ford et al. (1997)	Make sure the appliance's switch is "on". Place the earth lead on the active pin of the appliance's plug. Place the other lead on the frame of the appliance. Press the test	ME Transfer
Elektronische Motoren	Mayer, Dow & Mayer (2003)	When the motor is switched on, electrons flow from the negative terminal of the battery through the yellow wire and through the red wire to the positive terminal of the battery.	ME Transfer
Energie und Konstruktion	Segers, Verhoeven & Hulstijn-Hendrikse (2008)	In some constructions, concrete is poured over a metal framework, thus creating reinforced concrete.	ME Erinnerung
Geometrie	Jeung et al. (1997)	When two parallel lines intersect with a third line, four pairs of corresponding angles are equal.	ME Erinnerung
Inselbeschreibung	Tiene (2000)	The only river can be found on the west coast of the island.	ME Erinnerung
Instruktionsdesign	Tabbers (2002)	The casetype we start our training with is an experiment with just one factor. This is a simple experiment with regard to the data analysis.	Kein ME
Schnittgeschwindigkeiten bei verschiedenen Materialien	Kalyuga et al. (2000)	At the right upper corner of the diagram, select the diagonal line that corresponds to the lowest available cutting speed within the suggested range for bronze.	ME Erinnerung
Stadtbeschreibung	Brünken et al. (2002)	Vor dem Palazzo Vecchio stehen zwei monumentale Marmorskulpturen.	ME Dual Tasks
Temperaturgraphen	Leahy, Chandler & Sweller (2003)	Keine Angaben	ME Transfer

4.3.2 *Empirische Evidenz aus monomedialen Untersuchungen*

Eine Reihe von Studien untersuchte den Zusammenhang zwischen Textinhalt und Textmodalität in monomedialen Untersuchungen, d. h. den Lernern wurden keine Bilder, sondern ausschließlich Texte dargeboten. Man kann dabei unterscheiden, ob die verwendeten Textinhalte visuelle *und* räumliche Aspekte beschrieben, ob sie ausschließlich räumliche Aspekte beschrieben oder ob sie ausschließlich visuelle Aspekte beschrieben. Als Kontrollbedingung dienten in der Mehrzahl der Studien Textinhalte, die nicht visuell-räumlicher Natur waren. Im Folgenden werden zunächst Studien dargestellt, die keine explizite Trennung zwischen visuellen und räumlichen Inhalten machten, um anschließend Studien vorzustellen, die eine solche explizite Trennung vornahmen.

Eddy und Glass (1981) boten ihren Probanden entweder Sätze dar, die visuell-räumliche Informationen enthielten (z. B. „Der Davidstern hat sechs Ecken“) oder Sätze, die nach Einschätzung der Autoren keine visuell-räumliche Information enthielten (z. B. „Die Woche hat sieben Tage“) Zudem variierten sie die Textmodalität, mit der die Sätze dargeboten wurden. Aufgabe der Probanden war es die Sätze zu verifizieren. Die Ergebnisse zeigten, dass Probanden mit geschriebenen Texten länger brauchten, um Sätze mit visuell-räumlichen Inhalten zu verifizieren als Sätze mit nicht visuell-räumlichen Inhalten. Das gleiche Ergebnismuster wurde für die Beurteilung der Sinnhaftigkeit der Sätze beobachtet. Wurde der Text gesprochen dargeboten, gab es hingegen keine Unterschiede bei der Verifikation oder der Beurteilung der Sätze mit hohem und niedrigem visuell-räumlichem Informationsgehalt.

Leider berichten die Autoren hinsichtlich der Textmodalität nur einen allgemeinen Haupteffekt zugunsten gesprochener Textdarbietung, aber nicht, ob dieser Unterschied nur bei Sätzen mit hohen visuell-räumlichen Inhalten statistisch bedeutsam war – was man aufgrund der gefundenen Interaktion zwischen Textinhalt und Textmodalität sowie der numerischen Werte vermuten könnte. Die Ergebnisse zeigen jedoch, dass bei geschriebenen Texten ein Einfluss des Textinhalts besteht, nämlich in der Form, dass die Verarbeitung geschriebener Texte mit visuell-räumlichen Inhalten interferieren kann, während dieses Zusammenspiel bei gesprochenen Texten nicht besteht.

Die Interferenz zwischen Textmodalität und Textinhalt konnte auch für die Erinnerungsleistung nachgewiesen werden (Unnava, Agarwal & Haugtvedt, 1996): Lerner mit geschriebenem Text erinnerten weniger Informationen bei Texten mit hohem visuell-räumlichen Inhalt als bei Texten mit niedrigem visuell-räumlichen Inhalt. Bei gesprochenem Text war das Muster genau umgekehrt: Hier wurde mehr Information erinnert, wenn die Texte einen hohen Anteil visuell-räumlicher Information hatten. Bei der Darbietung von Texten mit hohem visuell-räumlichen Inhalt wurde zudem ein Modalitätseffekt gefunden, d. h. gesprochener Text führte hier zu besserer Leistung, während dieser Modalitätseffekt bei niedrigem visuell-räumlichen Inhalten nicht nachweisbar war.

Ein Nachteil der zitierten Studien ist, dass die Interferenz zwischen Darbietungsmodalität und Textinhalt jedoch sowohl auf räumliche als auch visuelle Textinhalte zurückgehen kann. Für eine Erklärung des zugrunde liegenden Mechanismus ist es jedoch hilfreich, die Textinhalte nochmals genauer zu spezifizieren. Daher werden im Folgenden Studien vorgestellt, die spezifisch den Einfluss der Textmodalität bei räumlicher Information untersuchen bzw. aufgrund ihrer Materialauswahl dieser Kategorie zugeordnet werden können.

Erste Studien zum Einfluss der Textmodalität auf die Verarbeitung räumlicher Informationen wurden von Brooks (1967, 1970) durchgeführt. Er untersuchte mit der Brooks-Aufgabe (vgl. Abschnitt 4.2.3), inwiefern die Textmodalität (gesprochen vs. geschrieben/gesprochen) und die Textinhalte (räumlich vs. nicht-räumlich) miteinander interagieren. Es zeigte sich eine Interaktion zwischen der Modalität der Textdarbietung und der Erinnerung an Sätze mit räumlicher Information bzw. an Sätze ohne räumliche Information (Brooks, 1967, Experiment 1): Während bei Sätzen mit räumlicher Information die gesprochene Darbietung im Vergleich zur geschriebenen/gesprochenen Darbietung in besserer Erinnerungsleistung resultierte, führte bei Sätzen mit nicht-räumlicher Information die Darbietung von geschriebenen/gesprochenen Texten im Vergleich zu gesprochenen Texten zu besserer Erinnerungsleistung. Brooks erklärt diese Interaktion damit, dass geschriebene Texte mit der Verarbeitung der räumlichen Informationen interferieren, während dies bei gesprochenen Texten nicht der Fall ist. Problematisch ist allerdings, dass Brooks keine Bedingung mit ausschließlich geschriebenem Text darbot. Es könnte sich somit auch um einen Redundanzeffekt handeln, d. h. die redundante Darbietung des Texts in sowohl gesprochener als auch geschriebener Form führte

zusammen mit der Verarbeitung räumlicher Informationen zu einer kognitiven Überlastung und somit zu schlechterer Leistung dieser Gruppe.

Eine weitere Untersuchung zum Einfluss der Textmodalität auf die Verarbeitung räumlicher Inhalte wurde von Beech (1977) durchgeführt: Den Probanden wurden Beschreibungen räumlicher Konfigurationen von Objekten entweder gesprochen oder geschrieben dargeboten. Eine solche Konfiguration lautete beispielweise: „Base pot, left comb, left nail, up tape“ (S. 952). Die Probanden sollten die beschriebenen Konfigurationen mental visualisieren und ein Zeichen geben, sobald sie das gesamte Bild visualisiert hatten. Gemessen wurden die Zeiten, die die Probanden benötigten, um das gesamte Bild zu visualisieren. Die Textmodalität nahm Einfluss auf die Verarbeitung der räumlichen Informationen: Bei geschriebenem Text stieg die benötigte Zeit für die Visualisierung exponentiell mit der Anzahl der zu visualisierenden Objekte an. Dies war bei gesprochenem Text nicht der Fall. Zudem brauchten Leser generell mehr Zeit als Hörer für die Visualisierung der Inhalte, allerdings wird nicht berichtet, ob dieser Unterschied signifikant ist. Problematisch ist allerdings, dass es keine Kontrollbedingung mit nicht-räumlichen Texten gab. Es kann daher nicht ausgeschlossen werden, dass auch bei nicht-räumlichen Sätzen das Lesen mit der Vorstellungsbildung interferiert hätte.

Hörnig und Kollegen (Hörnig, Claus & Eyferth, 2000; Hörnig, Eyferth & Claus, 1999) verglichen den Einfluss der Textmodalität auf die Verarbeitung räumlicher Situationsbeschreibungen über zwei Experimente hinweg. Die räumlichen Beschreibungen lauteten z. B.: „An der Wand links von Torsten steht der Kühlschrank. Auf dem Kühlschrank steht die Schüssel“ (Hörnig et al., 1999, S. 145). Bei geschriebenem Text war die Lesezeit selbstgesteuert, gleiches versuchten die Autoren bei der Darbietung gesprochener Texte zu erreichen, indem die Lerner einzelne Textpassagen aussuchen und erneut hören konnten. Bei der Darbietung geschriebener Texte machten die Lerner mehr Fehler in einer anschließenden Überprüfung des räumlichen Modells als Lerner mit gesprochenem Text. Zudem konnten die Autoren zeigen, dass Leser im Gegensatz zu Hörern ihr mentales Modell nicht aktualisierten, wenn eine neue Betrachtungsperspektive der Situation eingeführt wurde. Problematisch ist allerdings, dass der Vergleich der Textmodalitäten nur indirekt über den Vergleich zweier voneinander unabhängiger Experimente erfolgte.

Einen direkten Vergleich der Darbietungsmodalität nahmen Kürschner, Seufert, Hauck, Schnotz und Eid (2006, Experiment 1) vor: Sie variierten, inwiefern das

Ausmaß an räumlicher Information die Überlegenheit von gesprochenem Text moderiert. Hierfür boten sie den Probanden drei verschiedene Arten an Materialien dar: (a) Material, dessen räumlicher Informationsanteil als hoch einzuschätzen war, (b) Material, das räumliche Informationen und nicht-räumliche Informationen enthielt und (c) Material, das keine räumliche Information enthielt. Bei Material, das einen hohen Anteil räumlicher Informationen hatte, wurde – ähnlich der Brooks-Aufgabe – der Verlauf einer Linie in einer Matrix beschrieben. Beispielsweise lautete der Text: „Sie befinden sich am Startpunkt. Der Verlauf der Linie ist zweimal hoch, zweimal rechts, zweimal hoch, viermal links, zweimal runter, einmal rechts und fünfmal hoch“ (Kürschner, Seufert et al., 2006, S. 121). Die Probanden wurden instruiert, sich den Verlauf der Linie vorzustellen und ihn anschließend in eine leere Matrix einzuzeichnen. Es zeigte sich ein Modalitätseffekt: Lerner mit gesprochenem Text schnitten besser ab als Lerner mit geschriebenem Text. Beim zweiten Materialtyp – also Material sowohl mit als auch ohne räumlichen Informationsgehalt – sollten sich die Probanden merken, welche Objekte in einem Regal standen (nicht-räumliche Information) und wo sie in diesem Regal standen (räumliche Information). Beispielsweise lautete der Text: „Oben links im Regal steht eine Vase, rechts daneben ein Putzeimer, daneben ein Stapel alter Zeitungen.“ (S. 122). Auch bei diesem Materialtyp trat ein Modalitätseffekt auf, allerdings war die Effektstärke hier niedriger. Bei Materialien ohne räumlichen Charakter sollten sich die Probanden Objekte einer Einkaufsliste merken. Der Text lautete beispielsweise: „Hans kaufte eine Jacke, zwei Äpfel, einen Hammer, drei Brötchen, vier Liter Orangensaft, zwei Ohrringe, drei Schachteln Nägel und einen Fußball.“ (S. 122). Bei dieser Aufgabe wurde kein Unterschied mehr zwischen Lernern mit geschriebenem und gesprochenem Text beobachtet. Diese Ergebnisse deuten erneut darauf hin, dass die Verarbeitung räumlicher Textinformation mit geschriebenem Text interferieren kann, während die Verarbeitung nicht-räumlicher Textinformation nicht zu einer solchen Interferenz führte.

Den bisher aufgeführten Studien ist gemein, dass sie als Kontrollbedingung Texte ohne visuell-räumliche Inhalte verwendeten. Interessant ist jedoch auch der Vergleich von räumlichen und visuellen Textinhalten hinsichtlich ihres Zusammenspiels mit der Textmodalität, weil wie in Abschnitt 4.2.3 aufgeführt, visuelle und räumliche Textinhalte in unterschiedlichen Systemen des VSSP verarbeitet werden. Unter der Annahme, dass speziell räumliche Textinhalte den

rVSSP belasten und zu einem Modalitätseffekt führen, sollte man bei der Darbietung visueller Textinhalte, die den rVSSP nicht belasten, keinen Modalitätseffekt erwarten.

Leider gibt es zu dieser Fragestellung kaum empirische Befunde. Eine Ausnahme bildet ein Experiment von Glass, Millen, Beck und Eddy (1985, Experiment 3). Die Autoren boten den Probanden Sätze mit entweder visuellem Inhalt (Farbe / Form und Größe), mit räumlichem Inhalt (Orte / Bewegungsrichtungen) oder ohne visuell-räumlichen Inhalt dar. Aufgabe der Probanden war es, diese Sätze zu verifizieren. Ein Satz mit visuellen Farbinformationen lautete z. B.: „The spots on a giraffe are brown/yellow“, ein Satz mit visuellen Inhalten bezüglich der Form lautete z. B.: „The Statue of Liberty holds a torch in her right/left hand“ und ein Satz mit räumlichen Inhalten lautete z. B.: „To turn on a light you move the switch up/down.“ (Glass et al., 1985, S. 456). Ein Satz ohne visuell-räumliche Information war beispielsweise: „Salt is used more/less often than pepper“ (Glass et al., S. 446). Dabei indiziert die Auswahl an zwei Wörtern in den dargestellten Sätzen jeweils die richtig/falsch Version der Sätze. Die Sätze wurden entweder gesprochen oder geschrieben dargeboten. Die Ergebnisse zeigten, dass das Lesen im Vergleich zum Hören mit der Verifikation von Sätzen mit räumlichen Inhalten sowie von Sätzen mit visuellen Informationen über Form und Größe interferierte, während dies bei der Verifikation von Sätzen mit visuellen Inhalten über Farbe sowie von Sätzen mit nicht visuell-räumlicher Information nicht der Fall war. Hinsichtlich der Operationalisierung der Sätze mit visuellen Inhalten bezüglich Form und Größe, muss jedoch kritisch angemerkt werden, dass die dort enthaltene Information auch als räumlich eingestuft werden kann (vgl. Beispiel: left/right). Insofern kann der hier gefundene Modalitätseffekt auch auf Interferenzen mit der räumlichen Information zurückgehen.

Diese Übersicht über Befunde zum Lernen mit monomedialen Materialien, d. h. ausschließlich Texten, zeigt, dass geschriebene Textdarbietungen insbesondere mit der Verarbeitung räumlicher Inhalte zu interferieren scheinen, nicht jedoch mit der Verarbeitung nicht-räumlicher Inhalte bzw. visueller Inhalte. Da in den zitierten Studien eine Interaktion zwischen Textinhalt und Textmodalität beobachtet wurde, können die Effekte tatsächlich auf die verwendeten Textinhalte zurückgeführt werden, nicht aber auf die fehlende zeitliche oder räumliche Kontiguität bei geschriebenem Text, nach der ein genereller Modalitätseffekt unabhängig vom

Textinhalt erwartet wird. Es folgt nun ein Überblick über Studien, die das Zusammenspiel von Textinhalt und Textmodalität auch bei multimedialen Darbietungen untersuchten.

4.3.3 *Empirische Evidenz aus multimedialen Untersuchungen*

Während es bezüglich des Zusammenspiels von Textmodalität und Textinhalt bei reiner Textdarbietung eine Reihe an Studien gibt, ist dies bezüglich der Darbietung multimedialen Materials nicht der Fall.

Eine Untersuchung in diesem Zusammenhang wurde von Langston, Kramer und Glenberg (1998) durchgeführt. Die Autoren boten den Probanden Sätze dar, die eine räumliche Anordnung beschrieben, z. B. den Aufbau eines Aquariums. In jedem der Texte wurden insgesamt sechs Objekte zueinander in Beziehung gesetzt. Ein Text lautete beispielsweise:

Sam was setting up his fish aquarium decorations. Sam put the castle down first. Next, he put the plastic diver right of the castle. Then Sam put the big rock under the plastic diver. Then Sam put the seaweed left of the big rock. Next, he put the treasure chest left of the seaweed. Finally, Sam put the sunken boat over the treasure chest. When Sam put the fish in, it was scared by all of the stuff. (Langston et al., 1998, S. 249)

Neben der Textmodalität variierten die Autoren, ob den Probanden zusätzlich ein Bild dargeboten wurde, auf dem die ersten drei Objekte und ihre räumlichen Verhältnisse zueinander dargestellt waren. Bei einem anschließenden Verifikationstest bezüglich der dargebotenen Bilder zeigte sich eine Überlegenheit von Probanden mit gesprochener Textdarbietung. Diese spricht dafür, dass auch bei der Darbietung von Bildern eine Interferenz zwischen räumlichen Textinhalten und dem Leseprozess vorliegen kann. Problematisch ist allerdings, dass es keine Bedingung mit nicht-räumlichen Textinhalten gab. Insofern könnte der Modalitätseffekt hier auch auf die fehlende räumliche Kontiguität bei geschriebenem Text und Bild zurückgehen, auch da die Verifikationsitem sich auf die bildbezogene Erinnerung bezogen (vgl. Abschnitt 3.2.1).

Eine Untersuchung von Kürschner et al. (2007) untersuchte explizit den Einfluss der Textmodalität auf die Verarbeitung von Texten und Bildern in Abhängigkeit des Textinhalts. Den Lernern wurde ein relativ langer Text (1240 Wörter) über den Aufbau einer fiktiven Stadt (visuell-räumliche Information) und

den dort stattfindenden Bürgermeisterwahlen (abstrakte Information) dargeboten. Beide Inhaltsarten wurden von Bildern begleitet: Bilder, die visuell-räumliche Informationen begleiteten, stellten entweder die im Text genannten Objekte oder ihre räumliche Anordnung dar, während abstrakte Informationen durch Diagramme begleitet wurden. Es zeigte sich ein Modalitätseffekt, also eine Überlegenheit der gesprochenen Textdarbietung, bezüglich der Erinnerung an visuell-räumliche Informationen (gemessen über Lokalisationsaufgaben) und bezüglich der Erinnerung von Oberflächeninformation (gemessen über Wiedererkennung von verbalen oder bildlichen Informationen). Bezüglich der abstrakten Informationen konnte kein Modalitätseffekt beobachtet werden. Die Autoren erklärten sich dieses Ergebnis anhand der unterschiedlichen Visualisierungen, die dargeboten wurden. Allerdings ist es aufgrund der bisher aufgeführten Befunde auch plausibel anzunehmen, dass die Textinhalte den Modalitätseffekt moderierten. Demzufolge würde man die spezifische Interferenz zwischen Textinhalt und Textmodalität auch bei gleichbleibenden Visualisierungen erwarten.

In einer weiteren Untersuchung variierten Kürschner und Schnotz (2007) neben der Textmodalität, ob Bilder dargeboten wurden oder nicht. Sie verwendeten das gleiche Material wie in der zuvor berichteten Studie von Kürschner et al. (2007), nur dass zwei Gruppen Text (gesprochen oder geschrieben) ohne Bilder dargeboten bekamen. Der Vorteil gesprochener Sprache zeigte sich erneut nur für visuell-räumliche Informationen. Dieser Modalitätseffekt war unabhängig von der Darbietung der Bilder, d. h. nicht nur in den Gruppen ohne Bilder, sondern auch in den Gruppen mit Bildern war gesprochener Text überlegen. Dies spricht dafür, dass die Befunde aus Studien ohne Bilddarbietung (vgl. Abschnitt 4.3.2) auch auf die kombinierte Darbietung von Texten und Bildern übertragbar sind.

4.4 Zusammenfassung und Implikationen für die vorliegende Arbeit

Neben der fehlenden zeitlichen und räumlichen Kontiguität bei geschriebenen Texten scheint es einen weiteren Wirkmechanismus zu geben, auf den der Modalitätseffekt zurückgeführt werden kann. Diese alternative theoretische Erklärung für Modalitätseffekte basiert auf neueren Befunden bezüglich der VSSP-Struktur. Demnach besteht der VSSP aus einer visuellen und einer räumlichen Komponente, in

denen visuelle und räumliche Informationen verarbeitet werden. Dabei ist der VSSP nicht nur in die Verarbeitung visueller und räumlicher *bildhafter* Informationen involviert, sondern auch in die Verarbeitung räumlicher oder visueller *verbaler* Informationen. Des Weiteren kontrolliert der rVSSP Blickbewegungen, wie sie z. B. beim Lesen erforderlich sind. Aufgrund dieser Merkmale der Informationsverarbeitung im rVSSP kann man annehmen, dass räumliche Textinhalte in geschriebener Form den rVSSP stark belasten, da sowohl die räumlichen Inhalte als auch die Blickbewegungen hier verarbeitet bzw. kontrolliert werden. Da bei gesprochener Darbietung räumlicher Textinhalte weniger Blickbewegungen von Nöten sind, sollte die Belastung des rVSSP geringer ausfallen. Diese Belastungsunterschiede des rVSSP induziert über das Ausmaß an Blickbewegungen resultieren dann bei der Darbietung räumlicher Textinhalte in einem Modalitätseffekt, wie er auch empirisch nachgewiesen werden konnte.

So zeigten sich sowohl in Studien, in denen monomediales Material verwendet wurde, als auch in Studien, in denen multimediales Material verwendet wurde, Modalitätseffekte insbesondere dann, wenn die Textinhalte räumlicher Natur waren. Bei der Darbietung nicht visuell-räumlicher Textinhalte bzw. visueller Textinhalte gab es hingegen keine Überlegenheit gesprochener Textdarbietungen.

Diese dritte Erklärung für den Modalitätseffekt ist den Prozessen der mentalen Modellbildung vorgeschaltet. Somit wird ein Modalitätseffekt in Abhängigkeit des Textinhalts nicht nur für die Transferleistung sondern auch für die Erinnerungsleistung erwartet.

Im folgenden Abschnitt werden die Annahmen zur Informationsverarbeitung im VSSP in die CTML integriert, so dass eine *Extended Cognitive Theory of Multimedia Learning* (ECTML) resultiert. Anschließend wird aufgezeigt, welche weiteren Vorhersagen neben der Interaktion von Textinhalt und Textmodalität aus der ECTML resultieren.

5 Erweiterungen der Cognitive Theory of Multimedia Learning

Wie aufgeführt haben die Annahmen hinsichtlich der Informationsverarbeitung im VSSP Auswirkungen auf die von der CTML postulierten Verarbeitungsmechanismen von Texten und Bildern im Arbeitsgedächtnis. Dies ist aber nicht nur für die Erklärung des Modalitätseffekts von Bedeutung, sondern hat auch Konsequenzen für die generelle Text-Bildverarbeitung. Im vorliegenden Abschnitt werden deshalb die Befunde bezüglich des VSSP in die CTML integriert, so dass die ECTML resultiert.

5.1 Der strukturelle Aufbau der ECTML

Der strukturelle Aufbau der ECTML ist in Abbildung 9 dargestellt. Dabei wurden die folgenden Erweiterungen auf Grundlage der in Abschnitt 4 zitierten Befunde bezüglich des VSSP vorgenommen: Der VSSP wird in eine visuelle und eine räumliche Komponente unterteilt. Innerhalb des vVSSP werden visuelle Bildmerkmale (*visual images*) sowie visuelle verbale Inhalte (*visual content*) bezüglich Farbe und Form verarbeitet. Innerhalb des rVSSP werden räumliche Bildmerkmale (*spatial images*) sowie räumliche verbale Inhalte (*spatial content*) bezüglich Positionen und Relationen verarbeitet. Zudem findet im rVSSP die Kontrolle von Blickbewegungen (*eye movements*) statt. Blickbewegungen können dabei in multimedialen Lernumgebungen durch die Betrachtung von Bildern, durch Lesen oder durch Blickwechsel zwischen geschriebenem Text und Bild induziert werden. Des Weiteren wurden die Annahmen zum Verarbeitungsweg geschriebenen Texts korrigiert: Die Übertragung des geschriebenen Texts in die PL erfolgt nun nicht mehr über den VSSP wie in der CTML angenommen, sondern geschriebener Text erhält direkten Zugang zur PL über *Inner Speech*.

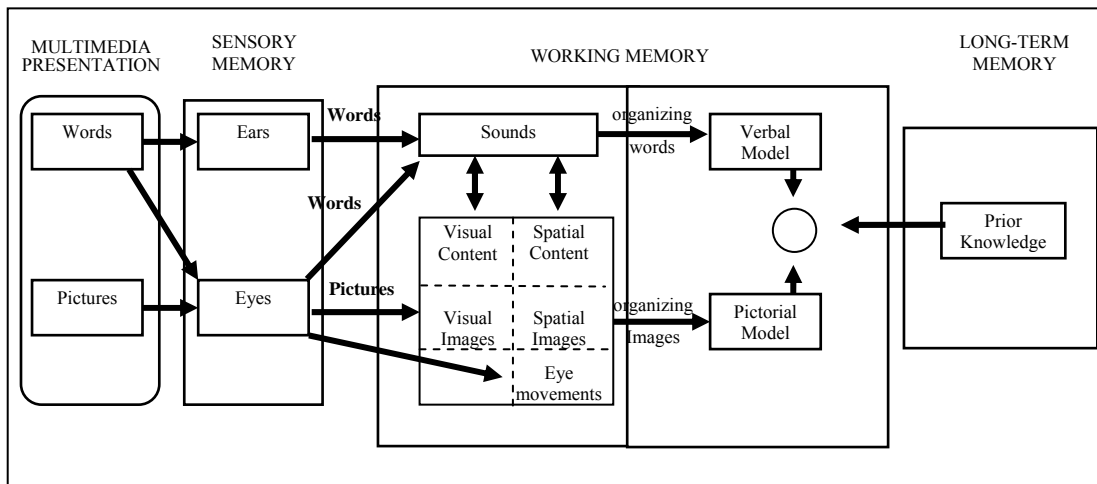


Abbildung 9. Der strukturelle Aufbau der ECTML.

Im Folgenden werden nun Ableitungen aus der ECTML aufgeführt.

5.2 Ableitungen aus der ECTML

5.2.1 Annahmen bezüglich des Einflusses der Textmodalität beim Lernen mit Multimedia

Nach der ECTML würde man bei der Darbietung geschriebener und gesprochener Texte zusammen mit Bildern die in Abbildung 10 dargestellten Belastungen erwarten.

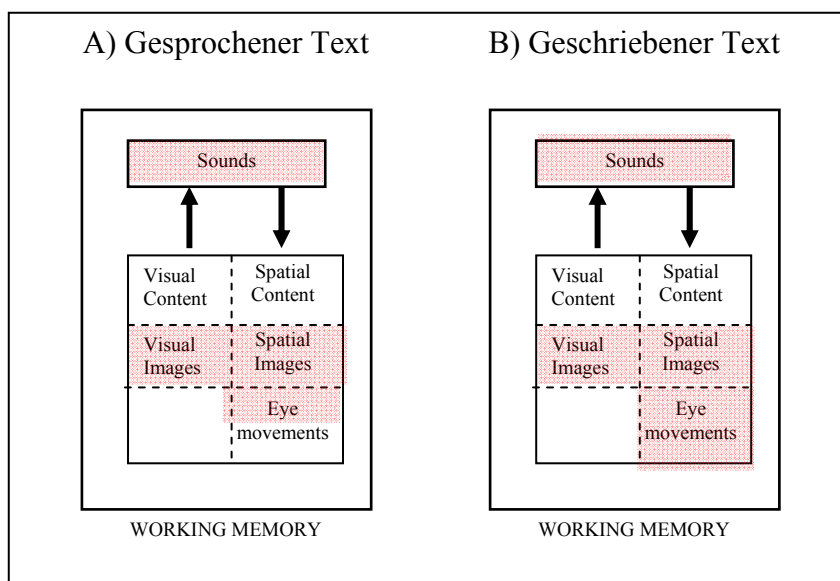


Abbildung 10. Die angenommene Belastung des Arbeitsgedächtnisses bei der Darbietung verschiedener Textmodalitäten (unabhängig vom Textinhalt). Färbungen indizieren Beanspruchungen der Komponenten.

Abbildung 10 verdeutlicht nochmals, dass ein Belastungsunterschied bei der Darbietung gesprochener und geschriebener Texte im Hinblick auf den rVSSP zu erwarten ist, da bei geschriebenem Text erhöhte Anforderungen bezogen auf die Kontrolle der Blickbewegungen resultieren: So sind Blickbewegungen zum Lesen des Texts, zur Betrachtung des Bildes und zum Wechsel zwischen Text und Bild notwendig. Bei gesprochenen Texten werden Blickbewegungen hingegen nur durch die Bildbetrachtung evoziert.

Es sei an dieser Stelle darauf verwiesen, dass diese zusätzliche Belastung des rVSSP bei geschriebenem Text auf den ersten Blick den in Abschnitt 3.3.1 kritisierten Annahmen der CTML einer modalitätsspezifischen Trennung der Kanäle gleicht. Die entscheidenden Unterschiede bestehen jedoch darin, dass (a) eine Belastung durch Augenbewegungen sowohl für gesprochenen als auch für geschriebenen Text angenommen wird und nur das Ausmaß der Belastung variiert, (b) diese Mehrbelastung bei geschriebenem Text nicht darauf zurückgeht, dass Abbilder des Texts im VSSP verarbeitet werden, was den Annahmen des Arbeitgedächtnismodells widerspricht, (c) nur eine Belastung des rVSSP und nicht des gesamten VSSP erwartet wird, (d) der Text unabhängig von der Textmodalität in der PL verarbeitet wird und (e) davon ausgegangen wird, dass diese Mehrbelastung des rVSSP durch die Blickbewegungen bei geschriebenem Text allein nicht ausreicht, um einen Modalitätseffekt zu erzeugen. Erst wenn das rVSSP zusätzlich durch räumliche Textinhalte belastet wird, sollte das höhere Ausmaß an Blickbewegungen bei geschriebenem Text die Leistung verschlechtern (vgl. Abschnitt 5.2.2).

Für letzteren Punkt sprechen zum Einen die in Abschnitt 2.3.2 und 3.3.1 aufgeführten Befunde, dass die räumliche Zweitaufgaben nicht mit dem Lesen geschriebener Texte interferierte (Gyselinck et al., 2002; Gyselinck et al., 2008; Brunyé et al., 2006). Dies spricht für eine eher geringe Belastung des rVSSP beim Lesen geschriebener Texte. Zum Anderen sprechen die in Abschnitt 4.3 aufgeführten Befunde zur Moderation des Modalitätseffekts durch den Textinhalt (z. B. Kürschner & Schnotz, 2007; Kürschner et al., 2007) ebenfalls dafür, dass die generelle Belastung durch Blickbewegungen beim Lesen eher gering ist. Denn würden die durch Blickbewegungen induzierten Belastungsunterschiede an sich ausreichen, um einen Modalitätseffekt zu erzeugen, dürfte der Textinhalt den Modalitätseffekt nicht moderieren, sondern man würde eine generelle Überlegenheit gesprochener Textdarbietung erwarten. Daher wird in der vorliegenden Arbeit ein Zusammenspiel

zwischen Textmodalität und Textinhalt erwartet, wie es bereits in Abschnitt 4.3 beschrieben wurde. Da, wie aufgeführt, auch ein Einfluss der räumlichen Kontiguität auf die bildbezogene Erinnerungsleistung nicht ausgeschlossen werden kann, ist es ein Ziel der vorliegenden Arbeit, explizit zwischen bild- und textbezogener Erinnerung zu trennen. Ein genereller Modalitätseffekt für die bildbezogene Erinnerung unabhängig vom Textinhalt würde für einen Einfluss der räumlichen Kontiguität sprechen.

5.2.2 *Annahmen bezüglich des Zusammenspiels von Textmodalität und Textinhalt*

Da im VSSP sowohl bestimmte Textinhalte verarbeitet werden als auch die Kontrolle von Blickbewegungen stattfindet, sollte es nach der ECTML eine Interaktion zwischen Textinhalt und Textmodalität geben, da die Textmodalität das Ausmaß an Blickbewegungen beeinflusst.

Abbildung 11 verdeutlicht die Annahmen der ECTML: Die Darbietung räumlicher Inhalte führt generell zu einer hohen Belastung des rVSSP (vgl. C und F in Abbildung 11). Bei geschriebener Textdarbietung kommt zusätzlich ein höheres Ausmaß an Blickbewegungen hinzu, dass durch den Leseprozess induziert wird. Demnach sollte eine sehr hohe Belastung des rVSSP bei der Darbietung geschriebener Texte mit räumlichen Inhalten vorliegen (vgl. F in Abbildung 11), die in einem *Overload* resultiert. Bei der Darbietung räumlicher Inhalte in gesprochener Form (vgl. C) sollte die Belastung des rVSSP zwar ebenfalls hoch sein, aber – da ein geringeres Ausmaß an Blickbewegungen notwendig ist als bei geschriebenem Text – sollte hier die Gefahr einer Überlastung nicht bestehen. Die *Overload*-Situation bei geschriebener Darbietung räumlicher Textinhalte sollte zu schlechterer Leistung im Vergleich zur Darbietung gesprochener räumlicher Textinhalte führen, d. h. es wird ein Modalitätseffekt erwartet. In allen anderen Bedingungen sollte die Belastung des rVSSP hingegen vergleichsweise geringer ausfallen, da keine räumlichen Textinhalte verarbeitet werden müssen (vgl. Abschnitt A-B, D-E). Die Gefahr eines *Overloads* besteht somit nicht. Daher werden hier keine Unterschiede zwischen gesprochener und geschriebener Textdarbietung erwartet, wenn die Erinnerungsleistung erfasst wird. Die kritische Bedingung für die Entstehung eines Modalitätseffekts ist somit die Darbietung räumlicher Textinhalte. Empirisch konnte dieses wie aufgeführt

sowohl in monomedialen als auch multimedialen Studien nachgewiesen werden (vgl. Abschnitt 4.3).

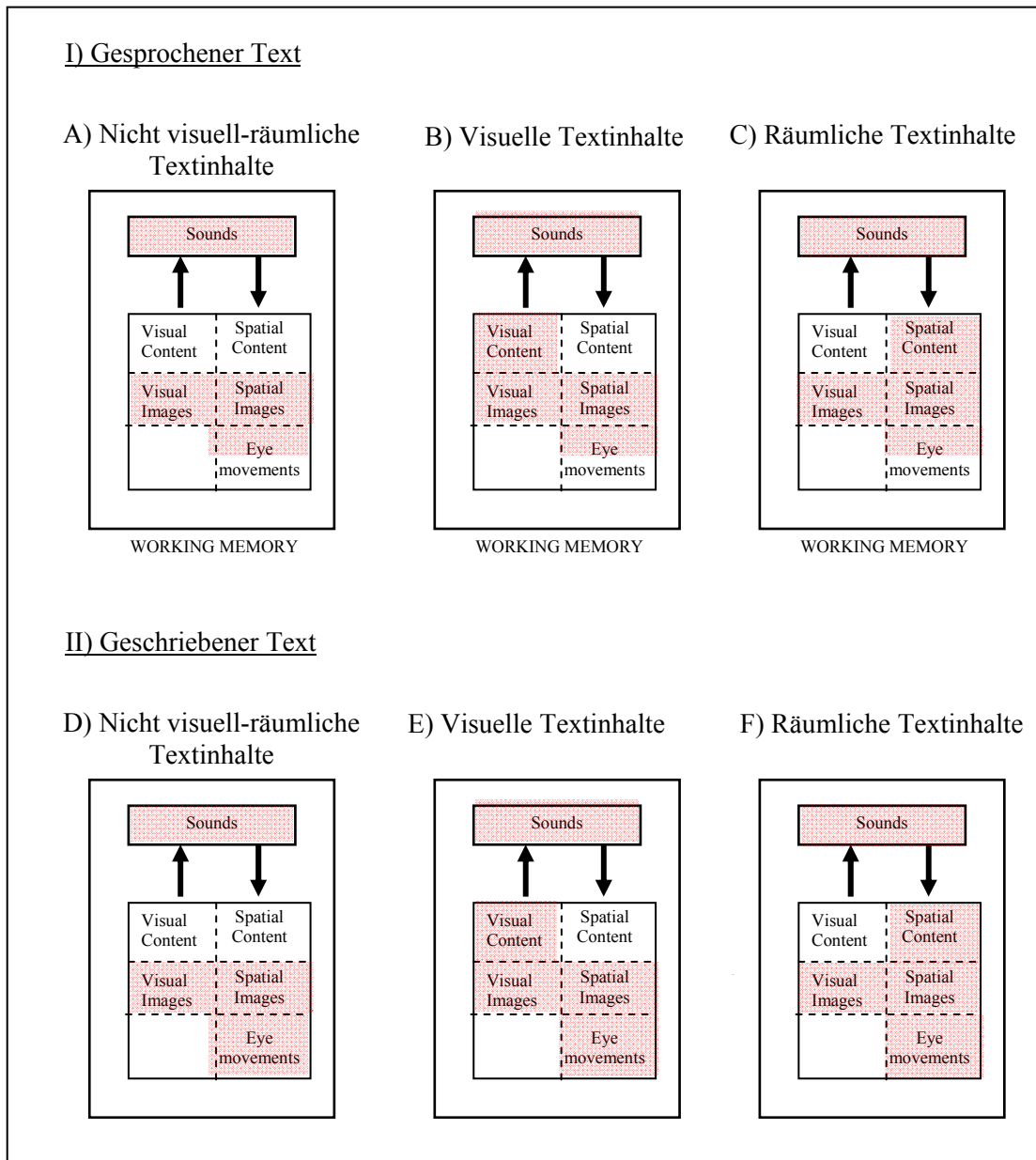


Abbildung 11. Die angenommene Belastung des Arbeitsgedächtnisses bei der Darbietung verschiedener Textmodalitäten und Textinhalte. Färbungen indizieren Beanspruchungen der Komponenten.

5.2.3 Annahmen bezüglich des Einflusses der Textinhalte beim Lernen mit Multimedia

Neben den Vorhersagen hinsichtlich der Textmodalität können mit Hilfe der ECTML auch Vorhersagen hinsichtlich des Einflusses der Textinhalte getroffen werden. Denn

geht man davon aus, dass visuelle und räumliche Textinhalte eine zusätzliche Belastung des VSSP darstellen, so kann man für nicht visuell-räumliche, visuelle und räumliche Textinhalte die in Abbildung 12 dargestellten Belastungen ableiten.

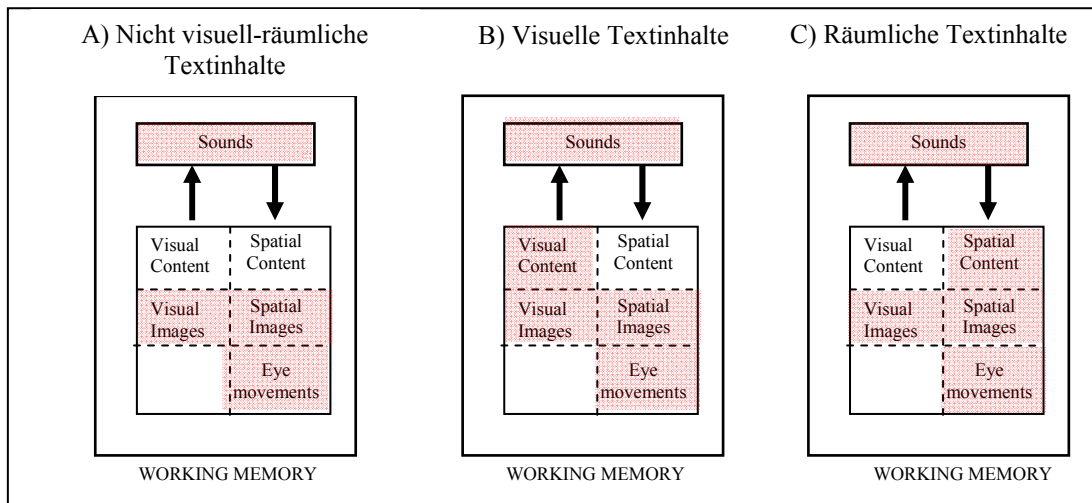


Abbildung 12. Die angenommene Belastung des Arbeitsgedächtnisses bei der Verarbeitung unterschiedlicher Textinhalte (A-C). Färbungen indizieren Beanspruchungen der Komponenten.

Es wird deutlich, dass sich die unterschiedlichen Textinhalte in unterschiedlichen Belastungen des VSSP niederschlagen. Während bei der Darbietung nicht visuell-räumlicher Textinhalte die Belastung des VSSP aus Blickbewegungen und der Verarbeitung des Bildes resultiert, wird bei der Darbietung visueller Textinhalte der vVSSP zusätzlich belastet, während bei der Darbietung räumlicher Textinhalte der rVSSP zusätzlich belastet wird.

Abbildung 12 zeigt, dass bei der Darbietung räumlicher Textinhalte zusammen mit Bildern eine starke Belastung des rVSSP nicht aber des vVSSP resultieren sollte, da sowohl räumliche Textinhalte, räumliche Bildinhalte als auch die Kontrolle der Blickbewegungen den rVSSP belasten. Der vVSSP wird hingegen bei der Darbietung räumlicher Textinhalte nur durch visuelle Bildinhalte belastet (vgl. Abbildung 12C). Bei der Darbietung visueller Textinhalte zusammen mit Bildern sollten hingegen die Belastungen des vVSSP und des rVSSP vergleichbar sein, da im vVSSP visuelle Bildinhalte und visuelle Textinhalte verarbeitet werden, während der rVSSP weiterhin durch räumliche Bildinhalte und die Blickbewegungen belastet wird (vgl. Abbildung 12B).

Welche Vorhersage kann man aus diesen Annahmen ableiten? Folgt man den angenommenen Prozessen in Abbildung 12, so erwartet man aufgrund der hohen Belastung des rVSSP bei der Darbietung räumlicher Textinhalte zusammen mit Bild ein Absinken der Leistung im Vergleich zur Darbietung visueller Textinhalte oder nicht visuell-räumlicher Textinhalte.

Dieser Annahme widersprechen auf den ersten Blick Befunde, die zeigten, dass die Darbietung von Bildern (in diesem Falle Landkarten) einen positiven Einfluss bei der Verarbeitung räumlicher Textinformation hatte (z. B. Ferguson & Hegarty, 1994; Pazzaglia, 2008): Probanden mit Landkarten schnitten in verschiedenen Performanzaufgaben besser ab als Probanden, denen nur der Text dargeboten wurde. Aber auch wenn einige Studien eine Überlegenheit von Text-Bilddarbietungen gegenüber reinen Textdarbietungen für die Darbietung räumlicher Textinhalte bestätigen, so belegt dies noch nicht, dass die Verarbeitung räumlicher Textinhalte und Bilder nicht miteinander interferiert. Denn es kann durchaus sein, dass es die *visuellen* Aspekte des Bildes sind, die zu einem Vorteil der Text-Bild Darbietung gegenüber einer reinen Textdarbietung führen.

Im Rahmen von Forschung zur sogenannten *Conjoint-Retention* Hypothese (z. B. Kulhavy, Lee & Caterino, 1985) wurden bezüglich der Frage, ob die visuellen oder die räumlichen Merkmale der Bilder das Lernen erleichtern, Untersuchungen durchgeführt. Die *Conjoint-Retention* Hypothese macht dabei die folgende Annahme:

According to CR [*Conjoint-Retention*], information that is referenced in both the text and the accompanying map is encoded both verbally and spatially, whereas information that is referenced in a text but not in a map is encoded only verbally. Due to this dual coding advantage, map information is “conjointly retained” and thus is more likely to be retrieved because the spatial representation can be accessed when an attempt to retrieve the verbal representation fails. (Griffin & Robinson, 2000, S. 125f.)

Nach dieser Definition handelt es sich bei der *Conjoint-Retention* Hypothese um eine Eingrenzung des Multimediaprinzips auf das Lernen räumlicher Informationen, die im Text und im Bild dargeboten werden. Die *Conjoint-Retention* Hypothese geht explizit davon aus, dass es die räumliche Information in der Landkarte ist, die den Lernprozess unterstützt. Diese Annahme widerspricht den Ableitungen aus der ECTML, nach denen die Verarbeitung räumlicher Bildinhalte und räumlicher

Textinhalte interferieren sollte. Um die Annahme der Unterstützung der Lernprozesse durch die räumlichen Bildinformationen zu überprüfen, boten Griffin und Robinson (2005) Lernern schriftliche Beschreibungen einer Stadt dar. Beispielsweise lautete ein Auszug aus dem verwendeten Text:

There are many interesting places and landmarks within Rome. The city itself is located next to the Tiber River where the bones of prehistoric animals have been found. Then there is the Appian Way which was one of the first paved roads to run out of Rome. Up on Mount Palatine are vast gardens that cover the ruins below them. (Griffin & Robinson, 2005, S. 36)

Dieser Text wurde entweder durch räumliche Abbildungen der Stadt oder durch Listen ergänzt. Dabei gab es zwei Arten räumlicher Abbildungen: Mimetische Abbildungen, in denen die Sehenswürdigkeiten visuell abgebildet waren, und symbolische Abbildungen, in denen die Sehenswürdigkeiten nur symbolisch als Quadrate abgebildet waren. Entsprechend gab es zwei Arten an Listen: Solche, in denen die Sehenswürdigkeiten (unter- und nebeneinander) mimetisch abgebildet waren und solche, in denen sie nur symbolisch als Quadrate abgebildet waren. Die Autoren konnten eine Überlegenheit der räumlichen Abbildungen nicht bestätigen, was einem Befund der Autoren aus dem Jahre 2000 entspricht (Griffin & Robinson, 2000). Stattdessen fanden sie bezüglich der Erinnerung von Namen der Sehenswürdigkeiten sogar eine Überlegenheit der Listen. Interessanterweise fanden sie teilweise auch eine Überlegenheit der mimetischen Abbildungen gegenüber den symbolischen Quadraten. Dies spricht nach Meinung der Autoren dafür, dass es vor allem mimetische Aspekte in Abbildungen sind, die den Lernprozess unterstützen. Dass die räumlichen Abbildungen das Lernen nicht unterstützten, sondern sogar teilweise zu schlechteren Ergebnissen führten, unterstützt die aufgestellte Annahme, dass räumliche Textinhalte, räumliche Bildinhalte und Blickbewegungen zu einer zusätzlichen Belastung des VSSP führen sollten.

Ein weiterer Hinweis auf eine Interferenz zwischen räumlichen Textinhalten und Bildverarbeitung kommt aus einer Untersuchung von Scheiter und Schmidt-Weigand (2008). Die Autoren boten den Lernern zwei Texte dar, deren Ausmaß an räumlicher Information variierte (hoch vs. niedrig). Zudem variierten sie, ob die Texte mit oder ohne dynamische Bilder dargeboten wurden. Die Ergebnisse zeigten, dass der Vorteil des Bildes sowohl hinsichtlich der verbalen als auch der bildhaften Erinnerungsleistung aufgehoben wurde, wenn der Text in hohem Ausmaß räumliche

Informationen enthielt. Dieses Ergebnis deutet zum Einen auf eine Interferenz zwischen räumlichen Textinhalten und der Bildverarbeitung hin. Zum Anderen zeigt der Befund, dass die Prozesse wie angenommen der Bildung eines integrierten mentalen Modells vorgeschaltet sind, da die Effekte hinsichtlich der Erinnerungsleistung beobachtet wurden.

5.3 Zusammenfassung und Implikationen für die vorliegende Arbeit

Im vorliegenden Abschnitt wurden die Befunde hinsichtlich der Informationsverarbeitung im VSSP in die CTML integriert, so dass die ECTML resultierte. Mit Hilfe der ECTML lassen sich nicht nur die Befunde hinsichtlich der Moderation des Modalitätseffekts durch den Textinhalt erklären, sondern es lassen sich auch generelle Annahmen zum Einfluss der Textmodalität sowie des Textinhalts ableiten: Auf Grundlage der Verarbeitung im VSSP wird kein genereller Modalitätseffekt erwartet. Allerdings können bei der Erfassung von Transferwissen und bildbezogener Erinnerungsleistung die fehlende zeitliche und räumliche Kontiguität bei geschriebenem Text einen Modalitätseffekt erzeugen. Nach der vVSSP-basierten Erklärung für den Modalitätseffekt sollte sich ein Einfluss der Textmodalität hingegen nur im Zusammenspiel mit den Textinhalten zeigen, da es dann bei der Darbietung geschriebener räumlicher Textinhalte zu einem *Overload* im frühen Verarbeitungsstadium des Arbeitsgedächtnis kommen sollte. Des Weiteren sollte die Darbietung räumlicher Textinhalte zusammen mit Bildern ebenfalls zu einer erhöhten Belastung des rVSSP bereits im frühen Verarbeitungsstadium führen und in schlechterer Leistung resultieren als die Darbietung nicht visuell-räumlicher Textinhalte bzw. visueller Textinhalte.

Für diese Annahmen gibt es auch erste empirische Evidenz in der Literatur: So fiel die Verarbeitung geschriebener Texte mit räumlichen Inhalten sowohl in mono- als auch in multimedialen Lernumgebungen im Vergleich zu gesprochenen Texten schlechter aus (z. B. Beech, 1977; Glass et al., 1985; Kürschner et al., 2007; Kürschner, Seufert et al., 2006) und räumliche Textinhalte interferierten mit der Verarbeitung von Bildern (z. B. Griffin & Robinson, 2000; Scheiter & Schmidt-Weigand, 2008).

Im empirischen Teil der vorliegenden Arbeit wurden die unterschiedlichen Annahmen der ECTML und der CTML geprüft. Zwar unterstützen die in den Abschnitten 4.3 und 5.2.3 aufgeführten Studien die Annahmen der ECTML, keine der Studien variierte jedoch systematisch alle genannten Einflussfaktoren (räumliche vs. visuelle vs. abstrakte Textinhalte sowie gesprochene vs. geschriebene Textdarbietung) bei gleichbleibender Inhaltsdomäne.

6 Ableitung der Hypothesen

Ziel der vorliegenden Arbeit ist die Prüfung der Grundannahmen der ECTML im Vergleich zu den Annahmen der CTML hinsichtlich der Rolle des VSSP beim Lernen multimedial dargebotener Inhalte. Nach der CTML werden im VSSP geschriebene Texte und Bilder verarbeitet. Es wird keine spezifische Unterteilung des VSSP vorgenommen. Nach der ECTML wird der VSSP in eine visuelle und eine räumliche Komponente unterteilt. Innerhalb des vVSSP werden bildhafte und verbale Informationen bezüglich visueller Merkmale (Farbe und Form) verarbeitet. Innerhalb des rVSSP werden bildhafte und verbale Informationen bezüglich räumlicher Merkmale (Positionen, Relationen) verarbeitet. Zudem findet im rVSSP die Kontrolle von Blickbewegungen statt. Blickbewegungen können dabei durch die Betrachtung von Bildern, durch Lesen oder durch Blickwechsel zwischen geschriebenem Text und Bild induziert werden. Die Prüfung der Vorhersagen der beiden Modelle erfolgt in der vorliegenden Arbeit zum Einen durch gezielte experimentelle Variation. Zum Anderen werden die Kapazitäten des vVSSP und rVSSP gemessen und in Beziehung zur Leistung innerhalb der verschiedenen experimentellen Gruppen gesetzt. Des Weiteren wurden für die Prüfung der Vorhersagen die text- und die bildbezogene Erinnerungsleistung erfasst, um zu zeigen, dass die angenommenen Prozesse tatsächlich der mentalen Modellbildung vorgeschaltet sind. Da keine Transferleistung erfasst wurde, sollte die fehlende zeitliche Kontiguität bei geschriebenem Text in der vorliegenden Arbeit keinen Einfluss auf die Ergebnisse nehmen. Hinsichtlich der bildbezogenen Erinnerungsleistung konnte ein Einfluss der fehlenden räumlichen Kontiguität bei geschriebenem Text nicht ausgeschlossen werden, allerdings sollte dieser Einfluss unabhängig vom Textinhalt erfolgen, so dass die jeweiligen Ergebnisse eindeutig auf die VSSP-basierte bzw. räumliche Kontiguitätserklärung rückgeführt werden konnten.

Die erste Hypothese bezog sich auf den Einfluss der Textmodalität beim Lernen mit Multimedia und wird deshalb im Folgenden als *Textmodalität-Hypothese* bezeichnet. Sie macht die folgenden Annahmen:

Unter Gültigkeit der CTML wird ein genereller Haupteffekt der Textmodalität für alle AVn erwartet, da geschriebene Texte im VSSP mit der Bildverarbeitung interferieren. Statistisch sollte sich diese Hypothese in einem

Haupteffekt der Textmodalität ausdrücken. Des Weiteren wird eine Moderation des Modalitätseffekts durch die Kapazität des vVSSP und rVSSP erwartet: Für Lerner mit höherer Kapazität des VSSP sollte die Interferenz zwischen geschriebenem Text und Bildverarbeitung geringer ausfallen, so dass auch der Modalitätseffekt abgeschwächt wird. Statistisch sollte sich diese Interferenz in einer Interaktion zwischen Textmodalität und vVSSP-Maß bzw. rVSSP-Maß ausdrücken.

Unter Gültigkeit der ECTML ist der Modalitätseffekt im rVSSP lokalisiert. Er wird nur dann erwartet, wenn der rVSSP stark belastet ist. Dies ist der Fall, wenn räumliche Textinhalte dargeboten werden und zusätzliche Blickbewegungen durch geschriebene Textdarbietung notwendig sind. Unter dieser Bedingung wird für die Abfrage räumlicher Informationen ein Modalitätseffekt erwartet, für die Abfrage nicht-räumlicher Informationen sollte es hingegen keinen Modalitätseffekt geben, da nicht-räumliche Informationen nicht im rVSSP verarbeitet werden und ihre Verarbeitung somit nicht beeinträchtigt werden sollte. Somit wird hinsichtlich der Abfrage nicht-räumlicher Inhalte kein Zusammenspiel von Textmodalität und Textinhalt erwartet. Statistisch sollte sich diese Hypothese in einer Interaktion der Faktoren Textinhalt und Textmodalität ausdrücken, wenn räumliche Inhalte erinnert werden müssen. In Abbildung 13 ist die Textmodalität-Hypothese nach der ECTML veranschaulicht. Im Hinblick auf die Kapazitäten des VSSP wird erwartet, dass für Lerner mit hohen Kapazitäten des rVSSP die Interferenz zwischen räumlichen Textinhalten und Blickbewegungen geringer ausfällt, so dass der Modalitätseffekt abgeschwächt wird. Statistisch sollte sich diese Hypothese in einer Interaktion zwischen Textinhalt, Textmodalität und rVSSP ausdrücken. Für die Kapazität des vVSSP wird hingegen kein Einfluss erwartet, da hier keine Interferenzen zwischen Textinhalt und Textmodalität erwartet werden.

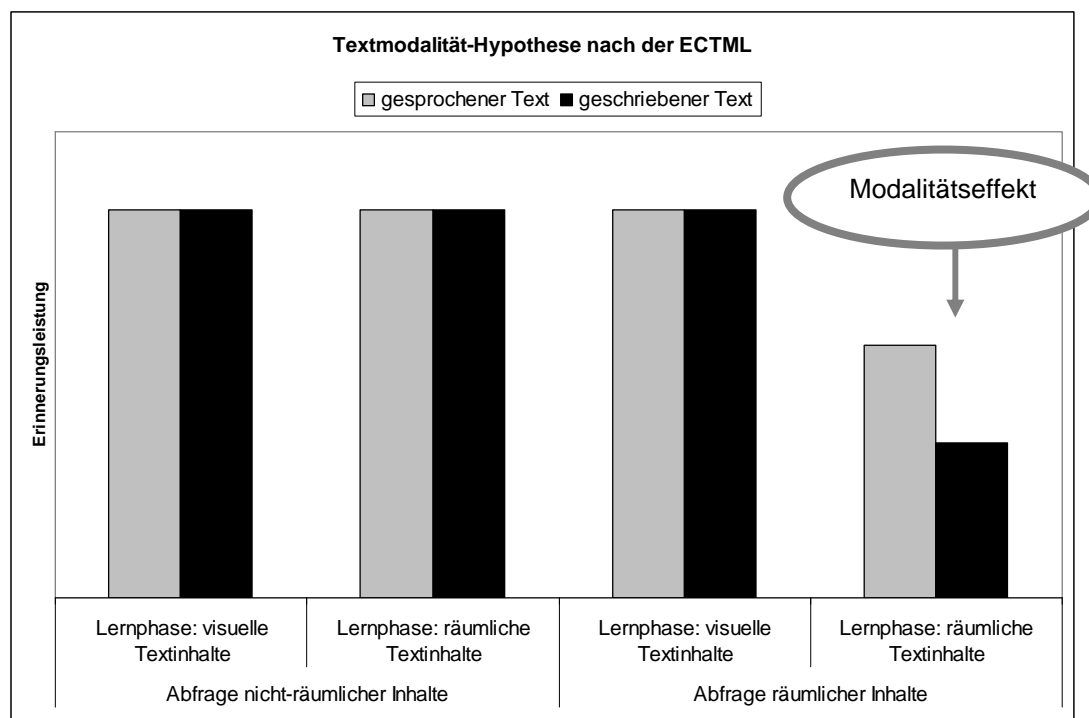


Abbildung 13. Darstellung des erwarteten Ergebnismusters für die Textmodalität-Hypothese nach der ECTML.

Die zweite Hypothese bezog sich auf den Einfluss der Textinhalte beim Lernen mit Multimedia. Sie wird deshalb im Folgenden als *Textinhalt-Hypothese* bezeichnet. Sie macht die folgenden Annahmen:

Unter Gültigkeit der CTML wird kein Effekt des Textinhalts erwartet, da keine spezifische Verarbeitung von Textinhalten angenommen wird. Statistisch sollte sich diese Hypothese in einem Nulleffekt für den Textinhalt ausdrücken.

Unter Gültigkeit der ECTML wird angenommen, dass es bei der Darbietung räumlicher Textinhalte in einer multimedialen Lernumgebung zu einer erhöhten Belastung des rVSSP kommt, da hier neben der Verarbeitung der räumlichen Text- und Bildinhalte auch die Kontrolle der Augenbewegungen stattfindet. Diese erhöhte Belastung sollte sich im Vergleich zur Darbietung visueller Textinhalte in einer schlechteren Erinnerung räumlicher Informationen ausdrücken. Bei der Erinnerung nicht-räumlicher Inhalte sollte es hingegen keine Unterschiede zwischen Lernen mit visuellen und räumlichen Textinhalten geben, da diese Inhalte nicht im rVSSP verarbeitet werden. Statistisch sollte sich die Textinhalt-Hypothese somit in einem Haupteffekt des Textinhalts ausdrücken, wenn räumliche Inhalte erinnert werden müssen. Bei der Abfrage

nicht-räumlicher Inhalte wird hingegen kein statistisch bedeutsamer Unterschied zwischen Lernern mit visuellen und räumlichen Inhalten erwartet. Diese Hypothese ist in Abbildung 14 dargestellt. Im Hinblick auf die Kapazität des VSSP wurde erwartet, dass bei der Darbietung räumlicher Textinhalte für Lerner mit hohen Kapazitäten des rVSSP die Interferenz zwischen Textinhalt und Bildverarbeitung geringer sein sollte, während bei der Darbietung visueller Textinhalte kein Einfluss der rVSSP erwartet wird. Statistisch sollte sich dieses in einer Interaktion zwischen Textinhalt und rVSSP-Maß ausdrücken. Bei der Darbietung visueller Textinhalte sollte hingegen ein Einfluss des vVSSP zu beobachten sein, was bei der Darbietung räumlicher Textinhalte nicht der Fall sein sollte. Statistisch sollte sich dieses in einer Interaktion zwischen Textinhalt und vVSSP-Maß ausdrücken.

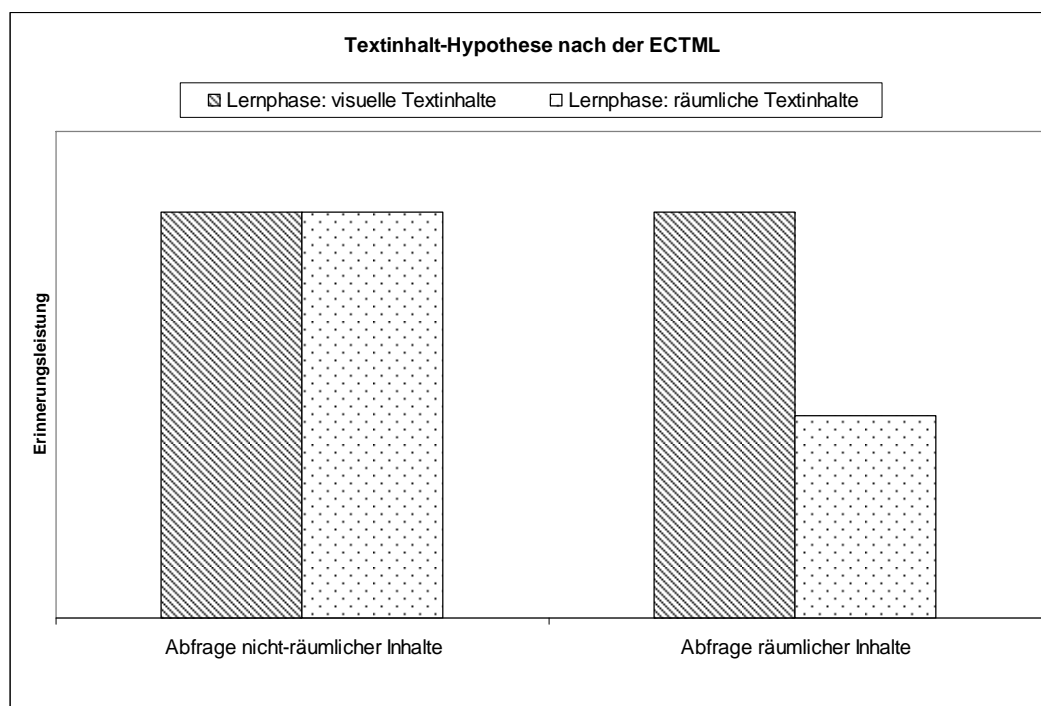


Abbildung 14. Darstellung des erwarteten Ergebnismusters für die Textinhalt-Hypothese nach der ECTML.

Wie in Abschnitt 2.3 aufgeführt, gibt es bisher nur wenige empirische Studien (eine Ausnahme bilden Gyselinck et al., 2000; Gyselinck et al., 2002) zum Zusammenhang der Arbeitsgedächtniskapazität und des Lernerfolgs. Daher wurde in der vorliegenden Forschungsarbeit des Weiteren untersucht, inwiefern die Arbeitsgedächtniskapazität generell einen Einfluss auf die Lernleistung nimmt. Nach

der *Arbeitsgedächtnis-Hypothese* wurde erwartet, dass hohe Kapazitäten der Arbeitsgedächtniskomponenten (PL, vVSSP, rVSSP) zu besserer Leistung führen sollten. Statistisch sollte sich ein Einfluss der Arbeitsgedächtniskapazität in einem Haupteffekt des vVSSP-Maß, des rVSSP-Maß und des PL-Maß ausdrücken. Diese Annahme war sowohl im Einklang mit der CTML als auch der ECTML.

Die Annahme, dass der VSSP in die Verarbeitung von Bildern (vgl. Abschnitt 2.3.2) und unter bestimmten Umständen auch verbaler Informationen involviert ist (vgl. Abschnitt 4.2.3) wurde in Experiment 2 mit Hilfe einer räumlichen Zweitaufgabe geprüft. Die *Zweitaufgaben-Hypothese* machte die folgenden Annahmen:

Unter Gültigkeit der CTML wird ein Effekt der Zweitaufgabe auf die Erinnerung bildhafter Informationen erwartet, da die Bilder nach der CTML im VSSP verarbeitet werden. Es wird kein Effekt der Zweitaufgabe auf die Erinnerung der verbalen Informationen erwartet.

Unter Gültigkeit der ECTML wird ein Effekt der Zweitaufgabe auf die Erinnerung räumlicher bildhafter Informationen erwartet, da sowohl räumliche Bildinhalte als auch die räumliche Zweitaufgabe speziell den rVSSP belasten sollten. Des Weiteren wird ein Absinken der Leistung bei der Erinnerung räumlicher verbaler Inhalte erwartet, da räumliche verbale Inhalte ebenfalls mit der räumlichen Zweitaufgabe im rVSSP interferieren sollten.

In Tabelle 5 sind die Annahmen der CTML und ECTML hinsichtlich der Hypothesen nochmals zusammengefasst. Die Prüfung der Hypothesen erfolgte mit Hilfe dreier Experimente, die im Folgenden beschrieben werden. Da allen drei Experimenten eine ähnliche multimediale Lernumgebung zugrunde lag und die Operationalisierung der Variablen äquivalent in den Experimenten erfolgte, werden diese beiden Punkte zunächst vorgestellt, um anschließend die durchgeführten Experimente zu beschreiben.

Tabelle 5

Gegenüberstellung der Annahmen der CTML und ECTML

	CTML	ECTML
Textmodalität-Hypothese	Alle Abfragen (AVn): Haupteffekt Textmodalität Textmodalität × rVSSP Textmodalität × vVSSP	Räumliche Abfrage (AVn) Textmodalität × Textinhalt Textmodalität × Textinhalt × rVSSP Textmodalität × Textinhalt × vVSSP Nicht-räumliche Abfragen (AVn): —
Textinhalt-Hypothese	—	Räumliche Abfragen (AVn): Haupteffekt Textinhalt Textinhalt × rVSSP Textinhalt × vVSSP Nicht-räumliche Abfragen (AVn): —
Arbeitsgedächtnis-Hypothese	Alle Abfragen (AVn): Haupteffekt PL Haupteffekt rVSSP Haupteffekt vVSSP	Alle Abfragen (AVn): Haupteffekt PL Haupteffekt rVSSP Haupteffekt vVSSP
Zweitaufgaben-Hypothese	Bildbezogene Abfragen (AVn) Haupteffekt Zweitaufgabe Nicht-bildbezogene Abfragen (AVn): —	Räumliche Abfragen (AVn): Haupteffekt Zweitaufgabe Nicht-räumliche Abfragen (AVn): —

II. Empirischer Teil

Im vorliegenden zweiten Teilabschnitt wird das empirische Vorgehen vorgestellt. Zunächst erfolgt in Abschnitt 7 eine Vorstellung der Gestaltung der Lernumgebung sowie in Abschnitt 8 die Beschreibung der Operationalisierung der Variablen. Im Anschluss erfolgt dann eine Beschreibung der jeweiligen Experimente im Hinblick auf die konkret verwendete Methode, die Ergebnisse sowie eine Diskussion der Ergebnisse.

7 Gestaltung der Lernumgebungen

In jedem der drei Experimente wurden Lernumgebungen verwendet, deren Gestaltung vergleichbar war. Bei den verwendeten Inhaltsdomänen handelte es sich jeweils um Domänen zur Vermittlung konzeptueller Inhalte. Diese Auswahl war zum Einen darin begründet, dass hier eine sinnvolle Variation der räumlichen und visuellen Inhalte vorgenommen werden konnte (s. u.). Zudem war eine differenzierte Erfassung der Erinnerungsleistung getrennt nach textbezogener und bildbezogener sowie visueller, räumlicher und abstrakter Information möglich. Die vermittelten Inhalte waren dabei artifizieller Natur, um das Vorwissen der Lerner möglichst niedrig zu halten, da wie in Abschnitt 3.1 aufgezeigt, bei hohem Vorwissen keine Modalitätseffekte beobachtet werden (Kalyuga et al., 2000). Die Experimente unterschieden sich hinsichtlich der konkreten Inhaltsbereiche: Während in den ersten beiden Experimenten Inhalte bezüglich Fischarten vermittelt wurden, wurden im dritten Experiment Informationen über Pflanzenarten vermittelt. Die Auswahl der beiden Inhaltsbereiche war dadurch begründet, dass sowohl bei Fischen als auch bei Pflanzen visuelle Aspekte (wie z. B. Farbe und Form bestimmter Merkmale), räumliche Aspekte (wie z. B. die Konfiguration bestimmter Merkmale) und abstrakte Aspekte (wie z. B. Lebensraum) sinnvoll vermittelt werden konnten. Abgesehen von dieser inhaltlichen Differenzierung war die Konzeption der Lernumgebungen identisch.

Die Lernumgebung war computerbasiert und bestand jeweils aus einer Vorphase, einer Lernphase und einer Testphase.

In der *Vorphase* wurden die Versuchspersonen über den Versuchsablauf instruiert (vgl. Anhang A), es wurden die personenbezogenen Daten erfasst (vgl. Abschnitt 8.3) und die Versuchspersonen erlernten wichtige Begriffe, deren Kenntnis Voraussetzung für das Verständnis in der Lernphase war. Hierfür wurde ihnen eine schematische Abbildung des Lernobjekts dargeboten, dessen einzelne Merkmale beschriftet waren (vgl. Anhang B). Diese Begrifflichkeiten mussten erlernt werden und nach dem Erlernen schriftlich wiedergegeben werden. Bei Fehlern wurden die Probanden aufgefordert die korrekten Benennungen erneut zu lernen.

In der *Lernphase* wurden den Versuchspersonen sechs verschiedene, artifizielle Fischarten (Experiment 1 und 2) bzw. Pflanzenarten (Experiment 3) präsentiert. Dabei wurden pro Lernobjekt zwei wesentliche Gestaltungsmerkmale verwendet, nämlich (a) ein statisches Bild, auf dem das Lernobjekt abgebildet war, und (b) eine verbale Beschreibung der spezifischen Charakteristika des Lernobjekts. Jedes Lernobjekt wurde auf einer eigenen Seite präsentiert.

Die sechs statischen Bilder wurden jeweils mit Hilfe des Programms Macromedia Flash Professional 8 erstellt. Auf den Bildern war jeweils die linke Körperhälfte des Fisches (vgl. Abbildung 15, links) bzw. die Pflanze aus der Frontalsicht (vgl. Abbildung 15, rechts) auf weißem Hintergrund abgebildet. Die dargebotenen Bilder waren in allen Versuchsgruppen gleich, d. h. alle Lerner bekamen unabhängig von den Textinhalten und der Textmodalität die gleichen Bilder dargeboten. Abbildung 15 zeigt zwei Screenshots der Lernphase, jeweils aus der Bedingung mit geschriebener Textdarbietung.

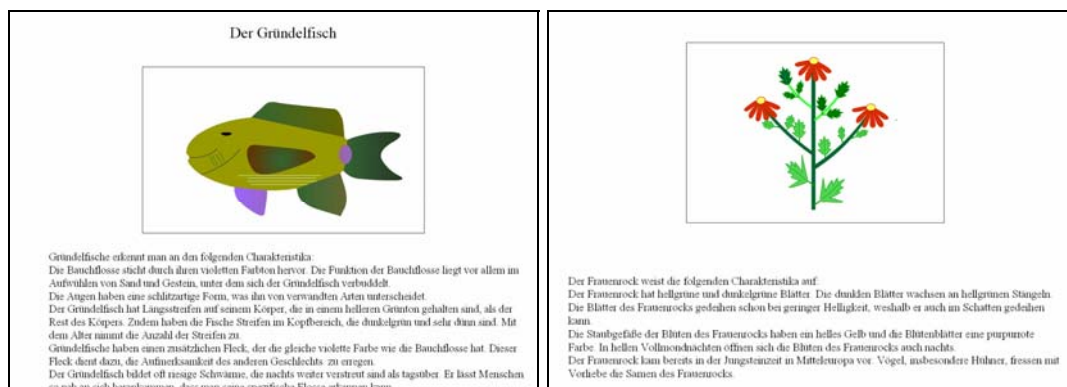


Abbildung 15. Beispielhafte Abbildung eines in der Lernumgebung dargebotenen Fisches (links, Experiment 1 und Experiment 2) und einer Pflanze (rechts, Experiment 3).

Die verbalen Erläuterungen unterschieden sich im Gegensatz zu den Bildern je nach Bedingung, da bezüglich der Textmodalität und der Textinhalte die experimentelle Manipulation erfolgte (vgl. Abschnitt 8.1.2). Abgesehen von diesen experimentellen Unterschieden wurde darauf geachtet, die Texte sowohl bezüglich ihres Aufbaus als auch bezüglich ihrer Länge in allen Bedingungen möglichst gleich zu gestalten. So bestanden alle Texte zum Einen aus der Beschreibung spezifischer, auch auf dem Bild abgebildeter Merkmale der Lernobjekte – im Folgenden als Charakteristika bezeichnet (vgl. Tabelle 6; Abschnitt II und IV) – und zum Anderen aus zusätzlichen abstrakten und nicht auf das Bild bezogenen Informationen (vgl. Tabelle 6, Abschnitt III, V und VI). Während bezüglich der Charakteristika die experimentelle Variation des Textinhalts erfolgte (vgl. Abschnitt 8.1.2), waren die abstrakten Sachverhalte in allen Bedingungen gleich. Sie spezifizierten zum Einen die Charakteristika, wobei darauf geachtet wurde, dass sie im gleichen Ausmaß mit den visuellen und räumlichen Textinhalten zusammenhingen. Zum Anderen handelte es sich um allgemeine Informationen wie z. B. den Lebensraum des Lernobjekts. Tabelle 6 verdeutlicht den Aufbau der Texte exemplarisch anhand eines in den ersten beiden Experimenten verwendeten Texts.

Tabelle 6

Darstellung des Aufbaus eines Texts mit visuellem Textinhalt (Experiment 1 & 2)

	Textbaustein	Erklärung
I	Typisch für den Ferdefisch sind die folgenden Charakteristika:	Einleitender Satz
II	Die Afterflosse hat die gleiche hellbraune Farbe wie die Rückenflossen. Die Rückenflossen zeichnen sich jeweils durch eine, sich aus gut erkennbaren dunkleren Punkten ergebende, Musterung aus.	Beschreibung der ersten zwei (visuellen) Charakteristika
III	Rücken- und Afterflossen werden vom Ferdefisch zur Verteidigung eingesetzt.	Abstrakter Sachverhalt bezüglich der ersten beiden (visuellen) Charakteristika
IV	Die Brustflosse hat eine rundliche Form.	Beschreibung des dritten (visuellen) Charakteristikums
V	Der Ferdefisch nutzt die Brustflosse, um sich damit im Grund einzugraben.	Abstrakter Sachverhalt bezüglich des dritten Merkmals.
VI	Der Ferdefisch lebt auf lockerem Untergrund, sandig oder schlammig, gelegentlich auch zwischen Pflanzen oder in der Nähe von Felsen. Man trifft ihn oft in sehr flachem Wasser an, wo er auf dem Grund liegt oder sich teilweise eingräbt.	Abstrakte Information über den Lebensraum des Tieres.

In Abbildung 16 ist das Verhältnis von Text und Bild dargestellt. Es wird deutlich, dass es sowohl Informationen gab, die in Text und Bild genannt wurden, als auch Informationen, die nur im Bild oder nur im Text genannt wurden. Dabei konnte die folgende Zuordnung getroffen werden: Abstrakte Informationen wurden ausschließlich im Text genannt, Informationen über Charakteristika der Lernobjekte wurden in Text und Bild dargeboten und die reine Bildinformation wurde ausschließlich im Bild gezeigt, aber nicht im Text erwähnt. Es wurde davon ausgegangen, dass sich die erwartete Interferenz zwischen Textinhalt und Textmodalität speziell hinsichtlich der Erinnerung der reinen räumlichen Bildinformation sowie hinsichtlich der Erinnerung der Charakteristika zeigen sollte. Hinsichtlich der Erinnerung der abstrakten Textinformationen wurde keine Interaktion erwartet, da hier keine räumliche Textinformation verarbeitet werden musste.

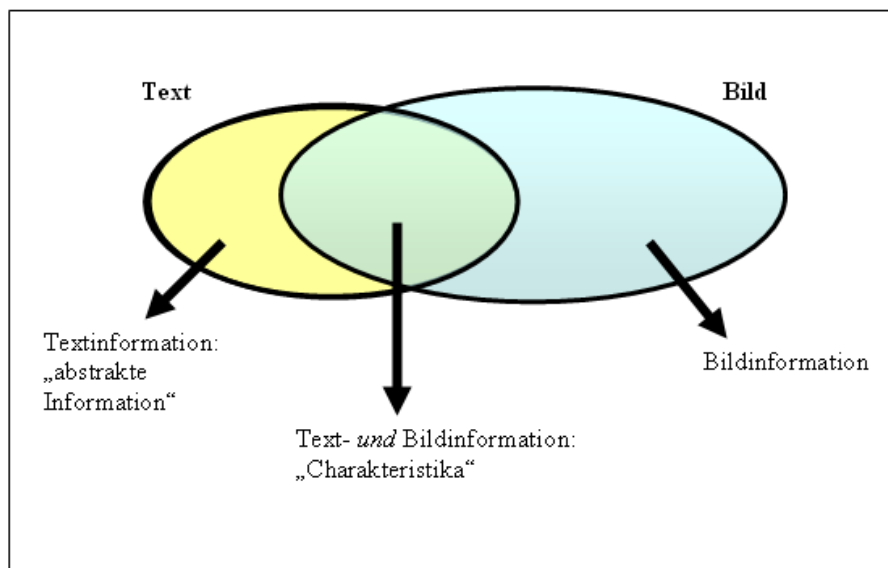


Abbildung 16. Das Verhältnis von Text- und Bildinformation.

Da die Selbststeuerung der Lernphase zu einer Aufhebung des Modalitätseffekts führen kann (Tabbers, 2002), erfolgte die Lernphase in allen drei Experimenten systemgesteuert, d. h. die Steuerung der Präsentationsgeschwindigkeit und der Darbietungsreihenfolge der Lernobjekte erfolgte durch den Computer. Dabei richtete sich die Präsentationsgeschwindigkeit nach der Darbietungsdauer des gesprochenen Texts, was eine in der Literatur übliche Vorgehensweise ist (vgl. z. B. Mayer & Moreno, 1998).

Die *Testphase* im Anschluss an die Lernphase setzte sich aus einem Fragebogen zur Bewertung der Lernphase (vgl. Abschnitt 8.2.3) und der Messung der Erinnerungsleistung zusammen (vgl. Abschnitt 8.2.1).

8 Die Operationalisierung der Variablen

Im folgenden Abschnitt werden die Operationalisierungen der in den drei Experimenten verwendeten Variablen beschrieben. Diese vorgezogene Beschreibung der Operationalisierungen ist deshalb sinnvoll, weil die drei Experimente bezüglich der verwendeten unabhängigen Variablen (UVn) und AVn einen hohen Überlappungsgrad aufwiesen.

8.1 Operationalisierung der unabhängigen Variablen

Im Folgenden wird die Operationalisierung der in allen drei Experimenten verwendeten UVn Textmodalität, Textinhalt und Arbeitsgedächtniskapazität sowie der in Experiment 2 verwendeten UV Zweitaufgabe beschrieben.

8.1.1 *Textmodalität*

Entsprechend den im empirischen Teil der Arbeit zitierten Untersuchungen (vgl. Abschnitt 3) erfolgte die Darbietung der Texte entweder gesprochen oder geschrieben. Bei der gesprochenen Textdarbietung wurde der Text über die Lautsprecher des Computers (Experiment 1) bzw. Kopfhörer (Experimente 2 & 3) dargeboten, bei der geschriebenen Darbietung war der Text unterhalb des Bildes abgebildet. Bei der gesprochenen Darbietung blieb der dem geschriebenen Text entsprechende Platz frei. Abbildung 17 verdeutlicht den Aufbau.

Der auditive Text wurde dabei von einer menschlichen und akzentfreien Stimme gesprochen, was dem so genannten *Voice Principle* entspricht, das folgende Annahme trifft: „... People learn more deeply when narration in a multimedia lesson is spoken in a non-accented human voice rather than in a machine voice or accented human voice” (Mayer, 2005d, S. 201). Zudem handelte es sich um eine weibliche Stimme, da in einem Experiment von Linek, Gerjets und Scheiter (2006) gezeigt werden konnte, dass sowohl männliche als auch weibliche Lerner bessere Leistungen zeigten, wenn die Inhalte von einer weiblichen Stimme präsentiert wurden und sie weibliche Stimmen zudem als attraktiver und motivierender einschätzten als männliche Stimmen. Zwar ist die Vergleichbarkeit mit der vorliegenden Studie eingeschränkt, weil es sich um eine mathematische Domäne handelte, aber es wurde

davon ausgegangen, dass eine weibliche Stimme auch in der vorliegenden konzeptuellen Domäne zumindest keinen negativen Einfluss auf die Ergebnisse nehmen würde.

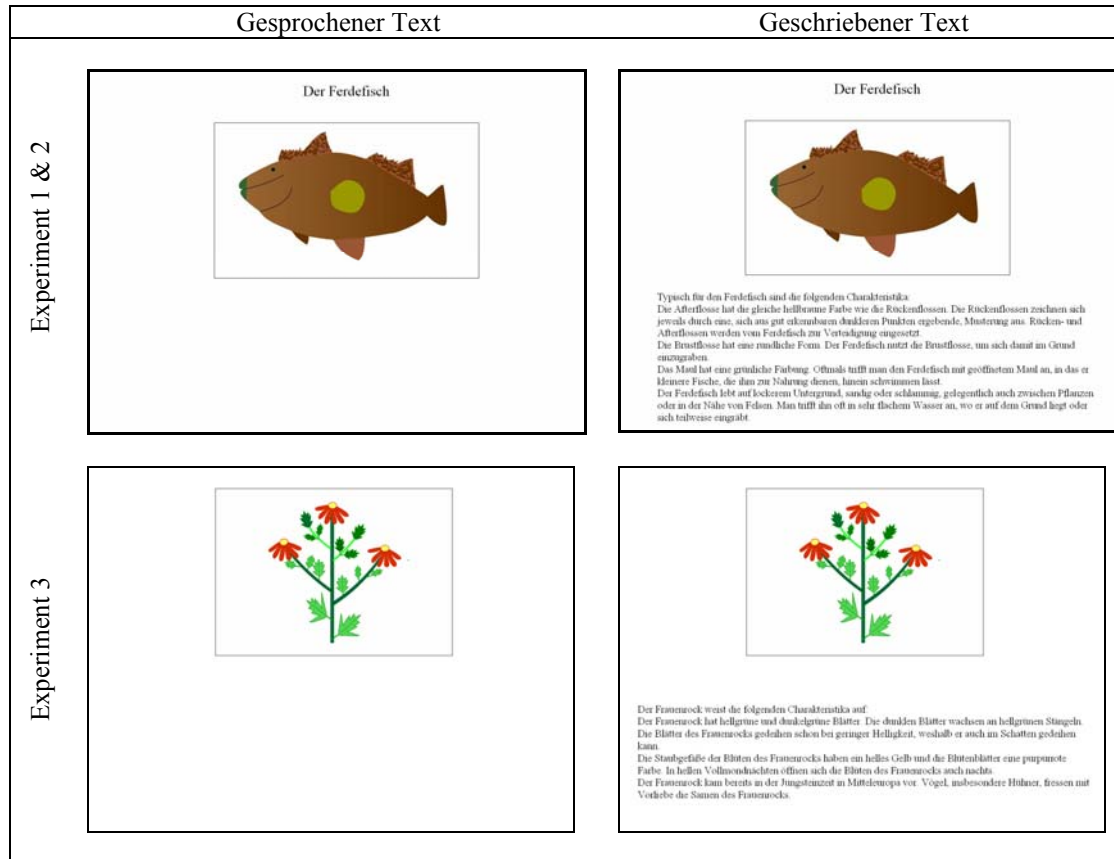


Abbildung 17. Die Bildschirmdarstellung mit gesprochenem und geschriebenem Text.

8.1.2 Textinhalt

Die Inhalte des Texts wurden variiert, indem entweder visuelle Charakteristika – wie die Farbe und Form bestimmter Merkmale – oder räumliche Charakteristika – wie die räumliche Lage oder die Richtung bestimmter Merkmale – im Text beschrieben wurden. In Anlehnung an die Literatur zu visuell-räumlichen Inhalten, wurden visuelle Inhalte über Farbe und Form und räumliche Inhalte über die Lage und Richtung operationalisiert (z. B. Bryant, Tversky & Franklin, 1992; Glass et al., 1985; Knauff & Johnson-Laird, 2002; Phillips et al., 2004). Es sei angemerkt, dass im dritten Experiment visuelle Informationen ausschließlich über Farbinformationen operationalisiert wurden. In Tabelle 7 sind beispielhaft ein Text mit visuellen Inhalten und ein Text mit räumlichen Inhalten gegenüber gestellt. Dabei wurde in

beiden Bedingungen auf die gleichen Merkmale (z. B. Brustflosse, Auge) referiert, um Konfundierungen zwischen den Textinhalten und den unterschiedlichen Merkmalen zu vermeiden. Es sei nochmals darauf verwiesen, dass die beschriebenen visuellen oder räumlichen Charakteristika auch den Bildern zu entnehmen waren. So konnte beispielsweise ein Lerner mit visuellen Textinhalten dem Bild des Ferdefischs die räumliche Lage der Afterflosse entnehmen. Im Unterschied zu Lernern mit räumlichen Textinhalten wurde diese räumliche Lage der Afterflosse aber nicht im Text beschrieben und war somit auch nicht charakteristisch für den abgebildeten Fisch.

Ein mögliches Problem hinsichtlich der Darbietung unterschiedlicher Textinhalte ergab sich durch die Tatsache, dass die beiden Texte nicht nur hinsichtlich der Inhalte, sondern auch hinsichtlich ihrer Schwierigkeit variieren können. Um dies möglichst auszuschließen, wurde bei der Konstruktion der Texte darauf geachtet, dass sie (a) hinsichtlich ihrer Länge und (b) hinsichtlich der Leseschwierigkeiten, gemessen über den Flesch-Index (Flesch, 1948), vergleichbar waren. So ergaben sich im Hinblick auf die Textlänge nur minimale Unterschiede zwischen den beiden Bedingungen. In den ersten beiden Experimenten waren die Texte pro Fisch bei visuellen Textinhalten durchschnittliche 125.83 Wörter lang, bei räumlichen Textinhalten durchschnittlich 127.17 Wörter lang. Somit betrug die durchschnittliche Abweichung 1.34 Wörter, was als praktisch nicht relevant eingestuft werden kann. Die maximale Abweichung bezüglich der Inhaltsdomäne Fische betrug zwischen zwei entsprechenden Texten mit visuellen und räumlichen Inhalten vier Wörter. Bezüglich der Inhaltsdomäne Pflanzen (Experiment 3) waren die Texte pro Pflanze bei visuellen Inhalten durchschnittlich 100.33 Wörter lang, bei räumlichen Inhalten 100.50 Wörter lang. Auch hier wurde davon ausgegangen, dass die durchschnittliche Abweichung von 0.17 Wörtern praktisch nicht bedeutsam ist. Die maximale Abweichung zwischen einem Text mit visuellen und einem Text mit räumlichen Inhalten betrug fünf Wörter. Die Flesch-Indizes (Flesch, 1948), die aussagen, wie leicht ein Text auf Grund seiner Struktur lesbar und verständlich ist, waren ebenfalls vergleichbar (visuelle Textinhalte: 47 [Experiment 1 und 2], 45 [Experiment 3]; räumliche Textinhalte 45 [Experiment 1 und 2], 47 [Experiment 3]). In die Berechnung gehen die durchschnittliche Wortlänge in Silben sowie die durchschnittliche Satzlänge in Wörtern unterschiedlich gewichtet ein. Die

resultierenden Werte reichen von 0-100, wobei der Wert null einen sehr leichten Text indiziert.

Da trotz dieser Maßnahmen Unterschiede in der Textschwierigkeit zwar wenig wahrscheinlich, aber nicht auszuschließen waren, wurde neben der bloßen textbezogenen Erinnerung auch die bildbezogene Erinnerung erfasst (vgl. Abschnitt 8.2.1). Bezüglich dieser ist ein Einfluss der Textschwierigkeit weniger wahrscheinlich, da die Bilder ja in allen Bedingungen gleich waren. Unterschiede in der bildbezogenen Erinnerung sollten somit mit hoher Wahrscheinlichkeit auf Interferenzprozesse rückzuführen sein, während bei Unterschieden in der textbezogenen Erinnerung ein Einfluss der Textschwierigkeit nicht vollkommen ausgeschlossen werden kann.

Tabelle 7

Gegenüberstellung von visuellen und räumlichen Textinhalten.

Visuelle Textinhalte	Räumliche Textinhalte
Typisch für den Ferdefisch sind die folgenden Charakteristika: Die Afterflosse hat die gleiche <i>hellbraune</i> Farbe wie die Rückenflossen.	Typisch für den Ferdefisch sind die folgenden Charakteristika: Die Afterflosse befindet sich <i>zwischen</i> der ersten und zweiten Rückenflosse.
Die Rückenflossen zeichnen sich jeweils durch eine, sich aus gut erkennbaren <i>dunkleren Punkten</i> ergebende, Musterung aus.	Die Rückenflossen zeichnen sich jeweils durch ein <i>nach unten versetztes</i> Vorderteil aus.
Rücken- und Afterflossen werden vom Ferdefisch zur Verteidigung eingesetzt.	Rücken- und Afterflossen werden vom Ferdefisch zur Verteidigung eingesetzt.
Die Brustflosse hat eine <i>rundliche Form</i> .	Die Brustflosse setzt erst <i>hinter</i> der ersten Rückenflosse an.
Der Ferdefisch nutzt die Brustflosse, um sich damit im Grund einzugraben.	Der Ferdefisch nutzt die Brustflosse, um sich damit im Grund einzugraben.
Das Maul hat eine <i>grünliche Färbung</i> .	Das Maul ist bis <i>hinter</i> die Augen gespalten.
Oftmals trifft man den Ferdefisch mit geöffnetem Maul an, in das er kleinere Fische, die ihm zur Nahrung dienen, hinein schwimmen lässt. Der Ferdefisch lebt auf lockerem Untergrund, sandig oder schlammig, gelegentlich auch zwischen Pflanzen oder in der Nähe von Felsen. Man trifft ihn oft in sehr flachem Wasser an, wo er auf dem Grund liegt oder sich teilweise eingräbt.	Oftmals trifft man den Ferdefisch mit geöffnetem Maul an, in das er kleinere Fische, die ihm zur Nahrung dienen, hinein schwimmen lässt. Der Ferdefisch lebt auf lockerem Untergrund, sandig oder schlammig, gelegentlich auch zwischen Pflanzen oder in der Nähe von Felsen. Man trifft ihn oft in sehr flachem Wasser an, wo er auf dem Grund liegt oder sich teilweise eingräbt.

Anmerkung. Kursiv hervorgehoben sind die Operationalisierungen der visuellen und räumlichen Textinhalte

8.1.3 Arbeitsgedächtniskapazität

Kapazität des rVSSP: Die Messung der Spanne des rVSSP erfolgte mit Hilfe des Corsi Block Tests, bei dem die Probanden räumliche Sequenzen erinnern müssen, die der Versuchsleiter ihnen vorgibt (vgl. Abschnitt 2.3.1). Es werden jeweils drei Sequenzen mit gleicher Blockanzahl pro Set dargeboten. Die Länge der Sequenzen des letzten korrekt erinnerten Sets entspricht der räumlichen Gedächtnisspanne. Der in der Untersuchung verwendete Corsi Block Test wurde in Anlehnung an die Beschreibung von Schellig und Hättig (1993) konstruiert. Die Tipp-Sequenzen wurden ebenfalls dieser Publikation entnommen.

Kapazität des vVSSP: Die Spanne des vVSSP wurde mit Hilfe des VPT von Della Sala et al. (1997) erhoben, der das Erinnern abstrakter Muster erfordert (vgl. Abschnitt 2.3.1). In Anlehnung an Della Sala et al. (1997) wurde die durchschnittliche Zellenanzahl der letzten drei korrekt erinnerten Matrizen als Indikator für die visuelle Gedächtnisspanne gewählt.

Kapazität der PL: Die Spanne der PL wurde mit Hilfe des *Digit Span* Tests gemessen, bei dem die Probanden Zahlensequenzen erinnern müssen (vgl. Abschnitt 2.3.1). Es wurden jeweils fünf Zahlensequenzen gleicher Länge in einem Set dargeboten. Die Länge der Zahlensequenzen innerhalb des letzten korrekt erinnerten Sets entspricht der phonologischen Gedächtnisspanne. Zur Standardisierung des Verfahrens wurden die Ziffernsequenzen vor Beginn des Experiments aufgesprochen und jedem Probanden auditiv über den Computer dargeboten. Die Ziffern wurden dabei im Sekundentakt dargeboten, um zu vermeiden, dass die Lerner bedeutungsvolle Einheiten (*Chunks*) aus den Zahlen bildeten, die die Memorierung erleichtern sollten.

8.1.4 Zweitaufgabe

Bezüglich der in Experiment 2 verwendeten Zweitaufgabe zur Belastung des rVSSP wurde variiert, ob die Probanden während der Lernphase eine solche ausführen oder nicht ausführen mussten. Die räumliche Zweitaufgabe bestand im Tippen der neun Tasten eines externen 3×3 Nummernblocks. Dieser Nummernblock wurde unter dem Tisch platziert, so dass die Probanden den Nummernblock während des Tippens nicht sehen konnten (vgl. z. B. Brunyé et al., 2006). Die Probanden wurden instruiert ein Muster zu tippen, dass mit der oberen linken Taste des Nummernblocks begann:

Die Probanden tippten zunächst von links nach rechts, dann in der nächsten unteren Zeile von rechts nach links, dann in der nächsten unteren Zeile von links nach rechts, dann in der nächsten oberen Zeile von rechts nach links usw. Dieses Tippmuster wird auch als furchenwendiges Tippen bezeichnet (z. B. Baddeley & Andrade, 2000). Zum Tippen wurde der Zeigefinger der Schreibhand verwendet. Entsprechend wurde der Nummernblock bei Rechtshändern auf ihrer rechten Seite platziert, während er bei Linkshändern auf der linken Seite platziert wurde. Bezüglich der Tippgeschwindigkeit wurden die Probanden aufgefordert, pro Sekunde eine Taste zu drücken. Vor Beginn der Lernphase wurden das Tippmuster sowie die Tippgeschwindigkeit zusammen mit der Versuchsleiterin eingeübt.

8.2 Operationalisierung der abhängigen Variablen

Als AVn wurden die Erinnerungsleistung sowie die Einschätzung der Lernphase hinsichtlich verschiedener Aspekte gemessen. Im Folgenden wird die Operationalisierung dieser AVn beschrieben.

8.2.1 Erinnerungsleistung

Die Erinnerungsleistung der Lerner wurde einerseits mit Hilfe von *Verifikationsitems* gemessen, bei denen eine Aussage oder ein Bild unter einer bestimmten Aufgabenstellung verifiziert werden mussten, und andererseits mit Hilfe von *Fragen mit offenem Antwortformat*. Dabei war ein Ziel der vorliegenden Arbeit, die Erinnerungsleistung so differenziert wie möglich zu erfassen, da nach der ECTML davon ausgegangen wurde, dass sich die vorhergesagten Effekte nur hinsichtlich bestimmter AVn zeigen sollten, nämlich solcher AVn, die eine Erinnerung räumlicher Inhalte erforderlich machten. Hinsichtlich AVn, die keine Erinnerung räumlicher Inhalte erforderten, sondern die Erinnerung visueller oder abstrakter Inhalte, wurden nach der ECTML keine Effekte erwartet. Zudem war für die vorliegende Arbeit die Trennung von text- und bildbezogener Erinnerung von Bedeutung, d. h. die Erinnerung der im Text dargebotenen Inhalte bzw. die Erinnerung der im Bild dargebotenen Inhalte. Diese Trennung war wichtig, da nicht ausgeschlossen werden konnte, dass Unterschiede hinsichtlich der textbezogenen Erinnerungsleistung auch auf unterschiedliche Schwierigkeiten der Texte zurückgingen, während hinsichtlich der bildbezogenen Erinnerungsleistung ein

Einfluss der Textschwierigkeit wenig wahrscheinlich war. Aufgrund der aufgeführten Differenzierungen hinsichtlich der Erinnerungsleistung resultierten die in Tabelle 8 dargestellten Indizes. Bezüglich der *bildbezogenen Erinnerungsleistung* wurde differenziert, ob visuelle oder räumliche Bildinformationen erinnert werden mussten. Bezüglich der *textbezogenen Erinnerungsleistung* wurde differenziert, ob die experimentell variierte Textinformation, also Charakteristika, oder abstrakte Textinformationen erinnert werden mussten.

Im Folgenden wird nun die Erfassung dieser Indizes mit Hilfe der Verifikationsitems und daran anschließend mit Hilfe der offenen Fragen beschrieben.

Tabelle 8

Indizes und ihre Bedeutung

Index	Erklärung
<i>bildbezogene Erinnerungsleistung</i>	
visuelle Bildinhalte	Erinnern visueller Bildinhalte
räumliche Bildinhalte	Erinnern räumlicher Bildinhalte
<i>textbezogene Erinnerungsleistung</i>	
Charakteristika	Erinnern experimentell variiertes Textinhalte (= Charakteristika)
abstrakte Textinhalte	Erinnern abstrakter Textinhalte

Verifikationsitems

Zur getrennten Messung der textbezogenen und der bildbezogenen Erinnerungsleistung mit Hilfe der Verifikationsitems wurden am Bildschirm verbale Aussagen (verbale Items) oder Bilder (bildhafte Items) dargeboten, die die Lerner unter einer bestimmten Fragestellung als richtig oder falsch bewerten sollten (vgl. Abbildung 18). Bildhaft und verbal dargebotene Items wurden dabei am Ende zusammengefasst, da keine unterschiedlichen Effekte für bildhafte vs. verbale Items erwartet wurden. Die Messung der Leistung in Form von Verifikationsitems wurde gewählt, da es sich dabei um eine in der Literatur zur Verarbeitung visuell-räumlicher Inhalte häufig verwendete Methode handelt (vgl. z. B. Eddy & Glass, 1981; Glass et al., 1985; Glenberg & Langston, 1992; Gyselinck et al., 2007; Jahn, 2004; Pazzaglia et al., 2007; Taylor & Tversky, 1992). Dabei setzten sich die Verifikationsitems aus relevanten Zielitems zur Erfassung der text- bzw. bildbezogenen Erinnerungsleistung sowie aus Distraktoritems zusammen. Als Distraktoren dienten Items, die nicht ausschließlich mit Hilfe des Bildes oder des

Textes sondern mit Hilfe von Text *und* Bild beantwortbar waren oder bei denen eine Trennung in visuelle vs. räumliche vs. abstrakte Inhalte nicht möglich war.

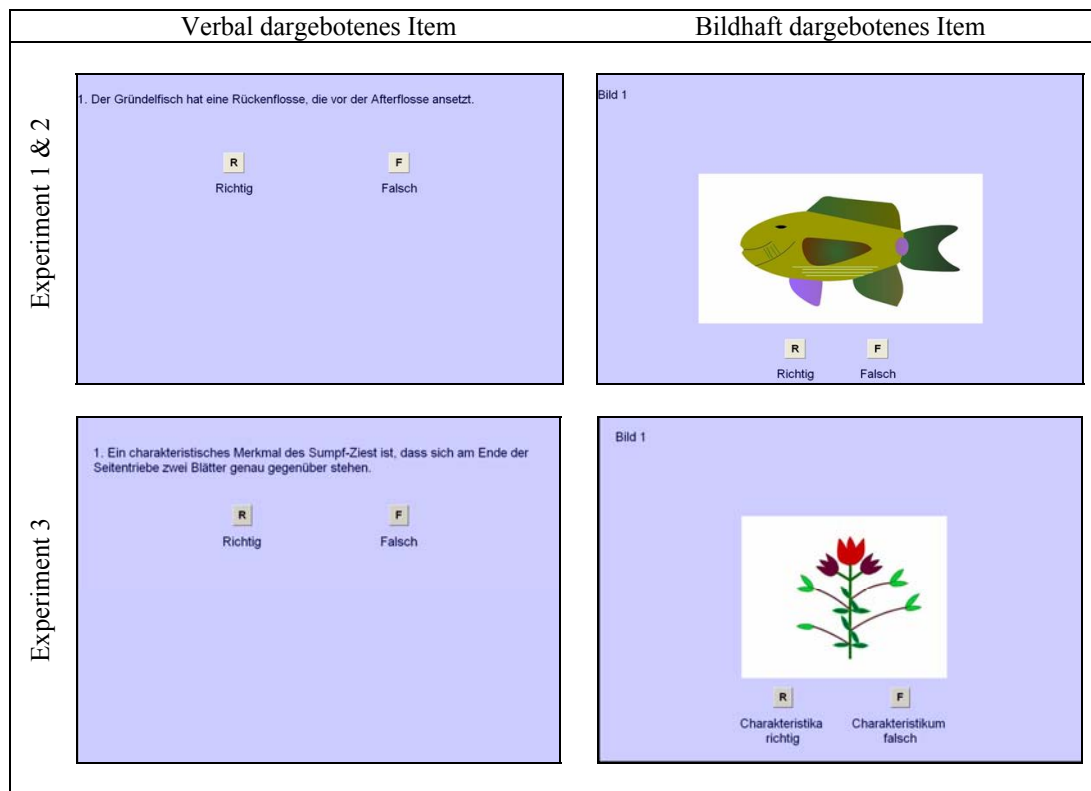


Abbildung 18. Verbale und bildhafte Verifikationsitems.

Erfassung der bildbezogenen Erinnerungsleistung. Zur Erfassung der bildbezogenen Erinnerungsleistung wurden die Lerner aufgefordert zu entscheiden, ob die Inhalte der verbalen oder bildhaften Items generell in der Lernphase dargeboten worden waren (vgl. Tabelle 9). Als Zielitems dienten solche Items, die nur mit Hilfe des Bildes beantwortet werden konnten. Der Inhalt der verbalen und bildhaften Items bezog sich auf visuelle (Farben oder Formen bei Fischen, Farben bei Pflanzen) oder auf räumliche Aspekte (Lage, Position, Richtung), so dass die in Tabelle 8 aufgeführten Indizes der visuellen bzw. räumlichen bildbezogenen Erinnerungsleistung resultierten.

Beispiele für *verbale* Zielitems sind z. B. die Folgenden: „Der Ferdefisch hat einen Schwanz, der die Form eines Dreiecks hat“ (visuelle Iteminhalte) und „Beim Ferdefisch endet die zweite Rückenflosse vor dem Schwanzansatz“ (räumliche Iteminhalte). Um diese Items zu verifizieren mussten die Lerner das Bild erinnern. Des Weiteren wurde unterschieden, ob die dargebotenen verbalen Aussagen den Inhalten der Lernphase entsprachen, also konsistent waren, oder ob sie den Inhalten

der Lernphase nicht entsprachen, also inkonsistent waren. Ein Beispiel für ein verbales inkonsistentes Zielitem mit räumlichen Inhalten ist z. B.: „Der Körper des Ferdefisches ist im hinteren Bereich höher als im vorderen Bereich.“ Dieses Item ist inkonsistent zur Lernphase, da der Körper des Ferdefisches im hinteren Bereich niedriger war als im vorderen Bereich (vgl. Abbildung 17). Verbale Zielitems mit zur Lernphase konsistenten Inhalten mussten somit von den Lernern als „richtig“ beurteilt werden (Hit), verbale Zielitems mit inkonsistenten Inhalten als „falsch“ (korrekte Zurückweisung). Als verbale Distraktoritems dienten verbale Aussagen, die mit Hilfe von Text und Bild beantwortbar waren, weil die Information nicht nur im Bild sondern auch im Text dargeboten worden war.

Tabelle 9

Instruktionen zur Erfassung der bildbezogenen Erinnerungsleistung (Verifikationsitems)

Instruktion verbale Items	Instruktion bildhafte Items
Im Folgenden werden Dir verschiedene Aussagen zu den Fischen der Lernphase präsentiert.	Im Folgenden werden Dir 18 verschiedenen Bilder der Fische präsentiert.
Bitte entscheide dieses Mal, ob die Aussage generell auf den in der Lernphase dargestellten Fisch zutrifft, d. h. ob die Inhalte im Text genannt wurden und/oder im Bild abgebildet wurden. Dabei ist es unerheblich, ob es sich bei dem Inhalt der Aussage um ein Charakteristikum des Fisches handelt, entscheidend ist nur, dass die Aussage auf den in der Lernphase dargestellten Fisch generell zutrifft.	Du sollst nun bei jedem Bild entscheiden, ob es sich bei dem Bild um das Originalbild aus der Lernphase handelt oder ob das Bild verändert wurde, sei es hinsichtlich der Charakteristika oder hinsichtlich anderer Aspekte. Größe und Schärfe des Bildes spielen dabei keine Rolle.
Wenn Du der Meinung bist, dass eine Aussage den Fisch korrekt beschreibt, dann klicke bitte mit der Maus auf den RICHTIG-Button.	Wenn Du der Meinung bist, dass es sich bei dem Bild um das Bild aus der Lernphase handelt, dann klicke bitte mit der Maus auf den RICHTIG-Button.
Wenn Du der Meinung bist, dass eine Aussage den Fisch nicht korrekt beschreibt, dann klicke bitte mit der Maus auf den FALSCH-Button.	Wenn Du der Meinung bist, dass es sich bei dem Bild nicht um das Bild aus der Lernphase handelt, dann klicke bitte mit der Maus auf den FALSCH-Button.

Bildhafte Zielitems zur Erfassung der bildbezogenen Erinnerungsleistung waren immer inkonsistent zur Lernphase. Dies war darin begründet, dass als konsistente bildhafte Items nur die Originalbilder dienen konnten – hier war eine Trennung in visuelle und räumliche Aspekte aber nicht möglich. Daher dienten die Originalbilder als Distraktoren. Bei bildhaften Zielitems wurden visuelle oder räumliche Merkmale des Fisches, wie z. B. die Farbe oder die Position eines Merkmals, verändert. Die bildhaften Zielitems mussten somit von den Lernern als „falsch“ beurteilt werden („korrekte Zurückweisung“).

Die Zuordnung, inwieweit es sich bei einem Item um ein Zielitem, das nur mit Hilfe des Bildes beantwortet werden konnte, oder um einen Distraktor, der mit Hilfe von Text und Bild beantwortet werden konnte, handelte, erfolgte zwischen den beiden Gruppen mit visuellen und räumlichen Textinhalten teilweise verschieden, da die Text-Bild Information je nach Gruppe variierte. So war z. B. das verbale Item „Der Ferdefisch hat ein Maul, das grünlich gefärbt ist“ für Lerner mit visuellen Textinhalten ein Distraktoritem, da sie es mit Text und Bild beantworten konnten. Für Lerner mit räumlichen Textinhalten war es hingegen ein Zielitem, da sie es nur mit Hilfe des Bildes beantworten konnten, da sie diese (visuelle) Information hinsichtlich der Farbe des Fisches in der Lernphase nicht im Text dargeboten bekamen sondern nur im Bild.

Erfassung der textbezogenen Erinnerungsleistung. Zur Erfassung der textbezogenen Erinnerungsleistung wurden die Lerner aufgefordert zu entscheiden, ob die Inhalte der verbalen bzw. bildhaften Items im Text gestanden hatten (vgl. Tabelle 10). Als Zielitems dienten solche Items, die nur mit Hilfe des Textes beantwortbar waren. Der Inhalt der verbalen und bildhaften Zielitems bezog sich auf die Charakteristika, d. h. visuelle Inhalte bzw. räumliche Inhalte in Abhängigkeit des Textinhalts, oder auf die abstrakten Textinformationen, so dass die in Tabelle 8 aufgeführten Indizes der visuellen bzw. räumlichen („Charakteristika“) und abstrakten textbezogenen Erinnerungsleistung resultierten. Es sei darauf verwiesen, dass die Charakteristika zwar in Text und Bild in der Lernphase dargestellt wurden (vgl. Abbildung 16), dass aber die textbezogene Erinnerung an die Charakteristika so operationalisiert wurde, dass zur Beantwortung *allein* der Text erinnert werden musste. Die Erinnerung des Bildes brachte für die Beantwortung textbezogener Items keinen Vorteil, weil man nur durch den Text wissen konnte, ob die im Bild dargestellten Informationen Charakteristika des Fisches waren oder nicht. Da in Abhängigkeit der Bedingung unterschiedliche Charakteristika in den Texten beschrieben wurden, mussten bezüglich der Charakteristika unterschiedliche Kodierungen erfolgen: Lerner mit räumlichen Textinhalten erhielten für die Zustimmung zu einem Zielitem, das ein räumliches spezifisches Charakteristikum korrekt beschrieb, einen Punkt („Treffer“), während Lerner mit visuellen Textinhalten für die Zustimmung zu eben diesem Zielitem keinen Punkt erhielten („falscher Alarm“), da bei visuellen Textinhalten keine räumlichen Charakteristika im Text beschrieben wurden. Entsprechend erhielten Lerner mit räumlichen

Textinhalten für die Ablehnung eines Zielitems, das ein räumliches spezifisches Charakteristikum korrekt beschrieb, keinen Punkt („Auslasser“), während Lerner mit visuellen Textinhalten für die Ablehnung eben dieses Zielitems einen Punkt erhielten („korrekte Zurückweisung“). Diese Zuordnung gilt analog für die Darbietung von Items mit visuellen Inhalten. Da sowohl Treffer als auch korrekte Zurückweisungen Indikatoren für die Erinnerungsleistung sind, wurden die Werte zu einem Index („Charakteristika“, vgl. Tabelle 10) zusammengefasst. Die bezüglich dieses Index erreichte Leistung gibt somit Auskunft darüber, wie gut der jeweilige Textinhalt erinnert werden konnte.

Tabelle 10

Instruktionen zur Erfassung der textbezogenen Erinnerungsleistung (Verifikationsitems)

Instruktion verbale Items	Instruktion bildhafte Items
<p>Im Folgenden werden Dir 30 verschiedene Aussagen präsentiert, die sich auf die im Text genannten Informationen beziehen. Dabei werden sowohl falsche als auch richtige Aussagen bezüglich der Charakteristika der Fische getroffen. Bitte entscheide bei jeder Aussage, ob die Aussage tatsächlich im Text genannt wurde, d. h. ob sie ein charakteristisches Merkmal des Fisches beschreibt. Bedenke dabei, dass nicht alle Aspekte, die auf dem Bild zu sehen waren, Charakteristika sind! Wenn Du der Meinung bist, dass eine Aussage einem Charakteristikum des Fisches entspricht, also im Text genannt wurde, dann klicke bitte mit der Maus den RICHTIG-Button. Wenn Du der Meinung bist, dass eine Aussage nicht den Charakteristika entspricht, also nicht im Text genannt wurde, dann klicke bitte mit der Maus den FALSCH-Button.</p>	<p>Im Folgenden werden Dir nacheinander 18 Bilder der Fische präsentiert. Du sollst nun für jedes Bild entscheiden, ob auf dem Bild die Charakteristika, also die im Text genannten Informationen, weiterhin so abgebildet sind, dass der Fisch seiner Art zugeordnet werden kann. Dabei ist es unerheblich, ob ANDERE Aspekte in dem Bild verändert wurden. Entscheidend ist nur, ob die Charakteristika weiterhin so abgebildet sind, dass der Fisch seiner Art zugeordnet werden kann. Veränderungen im Bild, die NICHT die Charakteristika, d. h. im Text genannte Informationen, betreffen, sind also unerheblich! Wenn Du der Meinung bist, dass die Charakteristika auf dem Bild weiterhin richtig abgebildet sind, d. h. also den im Text beschriebenen Charakteristika entsprechen, dann klicke bitte mit der Maus auf den CHARAKTERISTIKA RICHTIG-Button. Wenn Du der Meinung bist, dass ein Charakteristikum auf dem Bild falsch abgebildet ist, d. h. einem im Text beschriebenen Charakteristikum nicht mehr entspricht, dann klicke bitte mit der Maus auf den CHARAKTERISTIKUM FALSCH-Button.</p>

Richtige *verbale* Zielitems waren somit solche Items, die konsistent zu den im Text beschriebenen Charakteristika oder den abstrakten Informationen waren. Falsche verbale Zielitems waren somit zum Einen solche Items, die sich auf Informationen bezogen, die nicht im Text dargestellt worden waren, die aber durchaus konsistent zu den Bildinhalten sein konnten, und zum Anderen solche Zielitems, die die im Text enthaltenen Charakteristika oder die abstrakten Informationen falsch beschrieben,

also inkonsistent zur Lernphase waren. Items, die nicht ausschließlich mit Hilfe des Texts zu beantworten waren, sondern die zusätzlich mit Hilfe des Bildes beantwortet werden konnten, dienten als Distraktoren. Ein Beispiel für einen solchen verbalen Distraktor ist beispielsweise: „Ein charakteristisches Merkmal des Ferdefisches ist, dass sein Körper eine hellgrüne Färbung hat“. Hier konnten die Lerner zum Einen den Text erinnern um zu erkennen, dass diese Information nicht im Text stand, zum Anderen konnten sie das Bild erinnern um zu erkennen, dass der Ferdefisch eine bräunliche Färbung hatte und somit die Aussage falsch sein musste.

Richtige *bildhafte* Zielitems waren solche Items, die die im Text beschriebenen Charakteristika weiterhin korrekt abbildeten. Dabei konnten diese bildhaften Zielitems durchaus inkonsistent zur Lernphase sein, entscheidend war nur, dass die im Text beschriebenen Charakteristika auf dem Bild nicht verändert wurden (andere Merkmale konnten hingegen durchaus verändert werden). Falsche bildhafte Zielitems waren entsprechend solche Items, bei denen die im Text beschriebenen Charakteristika verändert wurden. Im Falle der bildhaften Items zur Erfassung der textbezogenen Erinnerungsleistung gab es keine Distraktoren, da zur Beantwortung der Frage, ob auf dem Bild ein im Text beschriebenes Charakteristikum verändert wurde, immer der Text erinnert werden musste.

Für Experiment 1 und 3 wurden jeweils 48 Items generiert. Für Experiment 2 mussten aufgrund des Versuchsdesigns weitere 18 Items generiert werden, so dass 60 Items resultierten.

Neben den bereits beschriebenen Aspekten wurden auch die folgenden Dinge bei der Itementwicklung beachtet: (1) Verbal und bildhaft wurden unterschiedliche Informationen abgefragt, d. h. ein verbal abgefragtes Merkmal des Fisches wurde nicht auch bildhaft abgefragt, (2) die Oberfläche der verbalen Items wurde so verändert, dass sie nicht der Oberflächenstruktur des Lerntexts entsprach, um Konfundierungen mit dem Wiedererkennen der Oberflächenstruktur zu verhindern, d. h. die Formulierungen wurden verändert und (3) die verbalen Items waren hinsichtlich ihrer Länge vergleichbar.

Scoring. Die Werte der Indizes wurden wie folgt errechnet: Für jedes richtig beantwortete Zielitem (Treffer oder korrekte Zurückweisung) gab es einen Punkt, für jedes falsch beantwortete Zielitem (falscher Alarm oder Auslasser) null Punkte. Nach der Darbietung jedes Items gaben die Probanden auf einer fünfstufigen Skala an, wie sicher sie sich bei ihrer Entscheidung gewesen waren. Die beiden Endpole der Skala

waren mit „geraten“ und „sehr sicher“ gelabelt. Zur Berechnung der Leistung wurde jede Zielitemantwort mit den entsprechenden subjektiven Einschätzungen gewichtet (vgl. z. B. Jahn, 2004). Dabei entsprach „geraten“ null Punkten, während „sehr sicher“ vier Punkten entsprach. Somit erhielt ein Proband, der ein Zielitem korrekt beantwortete, aber angab, geraten zu haben, für das Zielitem null Punkte, während ein Proband, der bei einem richtig beantworteten Zielitem angab, sehr sicher gewesen zu sein, vier Punkte erhielt. Anschließend wurden die jeweiligen Zielitems der Indizes aufaddiert und durch die Anzahl eingehender Items dividiert. Diese Relativierung an der Gesamtzahl eingehender Items war notwendig, um eine Vergleichbarkeit zwischen den Indizes zu gewährleisten, da, wie erwähnt, je nach Index und Textinhalt in der Lernphase unterschiedliche Anzahlen an Items eingingen (vgl. hierzu die Reliabilitätsanalysen in den Abschnitten 9.2.1, 10.3.1 und 11.3.1). Ein Wert von vier besagte somit, dass der Lerner alle Items richtig beantwortete und sich bei jedem Item sehr sicher gewesen war. Ein Wert von null besagte, dass der Lerner die Items falsch beantwortete und/oder dass er bei jedem Item geraten hatte. Bei allen Werten zwischen null und vier konnten die Werte zum Einen dadurch zustande kommen, dass ein Lerner nur einige wenige Items richtig beantwortete, aber sich jedes Mal sehr sicher war, sie konnten aber auch dadurch zustande kommen, dass ein Lerner sehr viele Items richtig beantwortete, sich aber jedes Mal sehr unsicher war.

Fragen mit offenem Antwortformat

Im zweiten und dritten Experiment mussten die Probanden neben den Verifikationsitems jeweils vier offene Fragen beantworten. Dieses Maß wurde gewählt, da in der traditionellen Multimediaforschung häufig Items mit offenem Antwortformat verwendet werden. Daher stellte sich die Frage, ob sich die Ergebnisse, die sich bezüglich der Verifikationsitems beobachten ließen, bei offenem Recall widerspiegeln. Zwei der Aufgaben testeten die *bildbezogene Erinnerungsleistung*. Sie bestanden jeweils in der Erstellung einer Zeichnung eines der Lernobjekte, z. B.: „Bitte zeichne im Folgenden den Drückerkfisch, so wie er in der Lernphase dargeboten wurde! Nutze dafür die beigelegten Buntstifte!“. Für diese Zeichenaufgaben standen den Probanden jeweils 3.5 Minuten zur Verfügung. Die Auswertung der Zeichnungen erfolgte anhand eines ausgearbeiteten Lösungsschlüssels, der alle bildhaften visuellen bzw. räumlichen Aspekte auflistete.

Für jeden gezeichneten Aspekt gab es einen Punkt. Die bildbezogene Erinnerungsleistung wurde durch die Erinnerung von Informationen definiert, die in der Lernphase ausschließlich im Bild dargeboten wurden. Es wurde unterschieden, ob die gezeichneten Informationen visueller oder räumlicher Natur waren, so dass auch hier die den Verifikationsitems entsprechenden Indizes resultierten (vgl. Tabelle 10). Aufgrund der unterschiedlichen Textinhalte während der Lernphase unterschied sich die Zusammensetzung der Skalen zur Messung der bildbezogenen Erinnerungsleistung zwischen Lernern mit visuellen und räumlichen Textinhalten.

Zwei weitere Aufgaben testeten die *textbezogene Erinnerungsleistung*. Die Aufgabe der Lerner bestand darin, jeweils bezüglich eines der Lernobjekte die im Text genannten Informationen aufzuschreiben, z. B.: „Bitte schreibe im Folgenden alle im Text genannten Informationen bezüglich des Lotserfischs auf, an die Du Dich erinnern kannst!“. Für die Beantwortung dieser Fragen standen den Probanden jeweils drei Minuten zur Verfügung. Zur Auswertung wurden alle Sinneinheiten des Originaltexts der Lernphase mit visuellen bzw. räumlichen Textinhalten aufgelistet. Für jede genannte Sinneinheit gab es einen Punkt. Es wurde differenziert, ob abstrakte Textinhalte oder Charakteristika erinnert wurden, so dass die den Verifikationsitems entsprechenden Indizes resultierten (vgl. Tabelle 10).

Die Antworten der Probanden wurden von zwei unabhängigen Ratern, die die Versuchsbedingungen nicht kannten, ausgewertet und es wurde die Interrater-Reliabilität berechnet. Unstimmigkeiten zwischen den zwei Ratern wurden durch nachträglichen Konsensus gelöst.

8.2.2 *Bewertung der Lernphase*

Im direkten Anschluss an die Lernphase wurden die Probanden aufgefordert, diese hinsichtlich bestimmter Aspekte zu bewerten. Dabei wurden Aspekte wie Anstrengung, Schwierigkeit und Konzentration erfasst. Tabelle 11 gibt eine Übersicht über die verwendeten Items. Die Items wurden auf einer siebenstufigen Antwortskala beantwortet, so dass Werte zwischen 1 und 7 Punkten pro Item resultierten, wobei hohe Werte einer hohen Ausprägung im Sinne der Itemformulierung entsprachen. Die beiden Endpole der Skala waren mit „überhaupt nicht“ und „sehr“ gelabelt. Alle Items wurden am Computer erfasst, wobei jedes Item einzeln dargeboten wurde. Nach der Beantwortung eines Items wurde

automatisch das nächste Item aufrufen. Im zweiten Experiment wurden nur die Items fünf und sieben dargeboten (vgl. Tabelle 11).

Tabelle 11

Fragebogen zur Bewertung der Lernphase

Item	Messabsicht
1 Wie anstrengend war die Lernaufgabe für Dich?	Anstrengung
2 Wie schwierig war es für Dich, die Inhalte zu verstehen?	Schwierigkeiten
3 Wie sehr hast Du Dich beim Lernen auf das Bild konzentriert?	Konzentration
4 Wie sehr hast Du Dich beim Lernen auf den Text konzentriert?	Konzentration
5 Wie sehr hast Du Dich während der Lernphase konzentriert?	Konzentration
6 Wie sehr hast Du Dich angestrengt, um das Lernmaterial zu verstehen?	Anstrengung
7 Wie viel Mühe hast Du in die Lernphase investiert?	Anstrengung

8.3 Personenbezogene Variablen

Zur Beschreibung der Stichprobe wurden personenbezogene Variablen erhoben. Diese Variablen wurden zu Beginn der Untersuchung über eine Eingabemaske erhoben, in die die Versuchspersonen die Daten eintrugen. Es handelt sich somit um selbstberichtete Daten.

(1) *Soziodemographische Daten:* Es wurden die Variablen Alter, Geschlecht, Studienfach und Semesteranzahl erfasst. Biologie- und Medizinstudenten wurden für die spätere Analyse als Studierende eines naturwissenschaftlichen Faches eingestuft.

(2) *Vorwissen:* Da es sich um artifizielles Material handelte, war ein Einfluss des Vorwissens wenig wahrscheinlich. Dennoch wurde zur Beschreibung der Stichprobe erhoben, ob Biologie 4-stündig oder als Leistungskurs in der Schule belegt worden war und es wurde zudem nach der letzten Biologienote gefragt. Des Weiteren wurde nach den Vorkenntnissen bezüglich Fischarten bzw. Pflanzenarten gefragt, was auf einer vierstufigen Skala, die von „keine“ bis „hoch“ reichte, eingeschätzt werden musste.

(3) *Farbsehschwäche:* Vor Beginn der Lernphase wurden die Teilnehmer gefragt, ob sie eine Farbsehschwäche hätten (z. B. Rot-Grün-Blindheit), da die Farbinformationen eine Operationalisierung visueller Bild- und Textinhalte darstellten. Falls die Teilnehmer dieses bejahten, öffnete sich automatisch eine Seite, auf der sie aufgefordert wurden, sich beim Versuchsleiter zu melden.

9 Experiment 1

Mit Hilfe des ersten Experiments wurden die Annahmen der ECTML gegen die Annahmen der CTML geprüft.

Die *Textmodalität-Hypothese* bezog sich auf den Einfluss der Textmodalität beim Lernen mit Multimedia. Nach der CTML wurde ein genereller Haupteffekt der Textmodalität für alle AVn erwartet. Für Lerner mit höherer Kapazität des VSSP sollte dieser Modalitätseffekt geringer ausfallen. Nach der ECTML wurde nur bei der Darbietung räumlicher Textinhalte ein Modalitätseffekt für die Erinnerung räumlicher Informationen, d. h. der räumlichen Bildinhalte und der Charakteristika (vgl. Tabelle 10), erwartet. Für die Erinnerung nicht-räumlicher Informationen, d. h. visueller Bildinhalte und abstrakter Textinhalte, sollte es hingegen keinen Modalitätseffekt geben. Für Lerner mit hoher rVSSP Kapazität wurde ein abgeschwächter Modalitätseffekt erwartet.

Die *Textinhalt-Hypothese* bezog sich auf den Einfluss der Textinhalte beim Lernen mit Multimedia. Nach der CTML wurde kein Effekt des Textinhalts erwartet. Nach der ECTML wurde eine schlechtere Erinnerung räumlicher Inhalte, d. h. räumlicher Bildinhalte und der Charakteristika (vgl. Tabelle 10), bei der Darbietung räumlicher Textinhalte als bei der Darbietung visueller Textinhalte erwartet. Bei der Darbietung räumlicher Textinhalte wurde ein positiver Einfluss der rVSSP Kapazität auf das Lernergebnis erwartet, bei der Darbietung visueller Textinhalte wurde ein positiver Einfluss der vVSSP Kapazität erwartet.

Die *Arbeitsgedächtnis-Hypothese* ging im Einklang mit der CTML als auch der ECTML davon aus, dass hohe Kapazitäten der Komponenten des Arbeitsgedächtnisses zu besserer Leistung führen sollten (für eine ausführliche Beschreibung der Hypothesen siehe auch Abschnitt 6).

9.1 Methode

9.1.1 Versuchsdesign

Die in Abschnitt 6 aufgestellten psychologischen Hypothesen wurden durch experimentelle Variationen in der *Lernphase* der Untersuchung (vgl. Abschnitt 7) mit einem zweifaktoriellen gemischten Versuchsplan getestet (vgl. Tabelle 12). Als

Lernobjekte dienten in Experiment 1 die verschiedenen Fischarten. Die erste UV war die Darbietungsmodalität des verbalen Materials (gesprochen vs. geschrieben), die zweite UV war der Textinhalt (visuell vs. räumlich). Beide Faktoren wurden interindividuell variiert. Pro Zelle wurden $n = 15$ Personen getestet, insgesamt belief sich die Stichprobe also auf $N = 60$ Personen.

Tabelle 12

Darstellung des Versuchsplans mit den Faktoren Textmodalität und Textinhalt (Experiment 1)

		Textmodalität	
		gesprochen	geschrieben
Textinhalt	visuell	$n = 15$	$n = 15$
	räumlich	$n = 15$	$n = 15$

Als kontinuierliche Variablen wurden die Werte im VPT, im Corsi Block und im *Digit Span* Test aufgenommen.

AVn waren die Verifikationsitems sowie die Bewertung der Lernphase (vgl. Abschnitt 8.2). Als personenbezogene Daten wurden die in Abschnitt 8.3 beschriebenen Variablen erfasst.

9.1.2 Versuchspersonen

Insgesamt nahmen 61 Studenten der Eberhard Karls Universität Tübingen an der Untersuchung teil. Als Aufwandsentschädigung für die Teilnahme erhielten sie entweder 10 Euro oder 1.5 Versuchspersonenstunden. Zwei Personen mussten von der Analyse ausgeschlossen werden, da sie angaben, farbenblind zu sein. Die verbleibenden 59 Personen (16 männliche und 43 weibliche Teilnehmer) waren im Durchschnitt 23.76 Jahre alt ($SD = 3.85$ Jahre). Die Zuteilung der Versuchspersonen auf die vier Bedingungen erfolgte randomisiert.

9.1.3 Geräte und Hilfsmittel

Die computerbasierte Lernphase wurde auf einem Notebook mit 15 Zoll Bildschirm dargeboten. Zur Darbietung wurde das Programm Medialab v2008 (Empirisoft, New York) genutzt. Die Zahlenblöcke für den *Digit Span* Test wurden auf einem zweiten Notebook in Form von Audiodateien im .wmv Format dargeboten.

9.1.4 *Versuchsablauf*

Der Versuch wurde über einen Zeitraum von vier Wochen zwischen 9 und 18 Uhr in Einzelsitzungen durchgeführt.

Zunächst bearbeiteten die Probanden die in Abschnitt 7 beschriebene computerbasierte Lernumgebung. Im Anschluss an die Lernphase wurden die in Abschnitt 8.2 beschriebenen Items in der folgenden Reihenfolge dargeboten: Zunächst beantworteten die Versuchspersonen den Fragebogen zur Bewertung der Lernphase. Dann erfolgte die Erfassung der Erinnerungsleistung mit Hilfe der Verifikationsitems: Als erstes wurden die verbalen Items zur Messung der textbezogenen Erinnerungsleistung dargeboten, daran anschließend die verbalen Items zur Messung der bildbezogenen Erinnerungsleistung. Die Anordnung der verbalen Items erfolgte dabei jeweils entsprechend der Reihenfolge der Lernobjekte in der Lernphase, d. h. es wurden zunächst alle verbalen Items bezüglich des ersten Lernobjekts der Lernphase dargeboten, dann die verbalen Items bezüglich des zweiten Lernobjekts der Lernphase usw. Dieses Vorgehen wurde gewählt, um die Beantwortung der Items nicht unnötig zu erschweren. Innerhalb jedes Blocks wurden die Items randomisiert dargeboten. Daran anschließend erfolgte die Darbietung der bildhaften Items zur Abfrage der textbezogenen Erinnerungsleistung und zum Schluss die bildhaften Items zur Abfrage der bildbezogenen Erinnerungsleistung. Bei bildhaften Items erfolgte die Darbietung ebenfalls entsprechend der Reihenfolge der Lernobjekte. Allerdings wurden die Bilder alternierend dargeboten, also hintereinander sechs Bilder der Lernobjekte in der Abfolge der Lernphase, um dann wieder beim ersten Lernobjekt zu beginnen usw. Die Bilder zu einem Lernobjekt wurden nicht wie bei den verbalen Items blockweise dargeboten, um die Probanden nicht für die Unterschiede zwischen den Bildern zu sensibilisieren.

Im Anschluss an die Lernumgebung wurden der Corsi Block Test, der VPT und der *Digit Span* Test durchgeführt. Nach der Durchführung dieser drei Tests war die Untersuchung beendet und die Probanden erhielten die Aufwandsentschädigung. Insgesamt dauerte die Untersuchung circa 90 Minuten.

9.2 Ergebnisse

Im vorliegenden Abschnitt werden die Ergebnisse des ersten Experiments vorgestellt. Alle Signifikanzangaben beruhen auf zweiseitiger Testung, die Irrtumswahrscheinlichkeit wurde mit $\alpha = .05$ festgesetzt.

Die Überprüfung der Hypothesen erfolgte mit Hilfe multivariater Varianzanalysen mit den Zwischensubjektfaktoren Textinhalt und Textmodalität und den kontinuierlichen, z-standardisierten Testwerten im VPT, Corsi Block und *Digit Span* Test. Die *Digit Span* ging als einfache Kovariate in die Analyse ein, während hinsichtlich des VPT und Corsi Block Maß auch die Interaktionen mit den experimentellen Bedingungen berücksichtigt wurden. Nicht berücksichtigt wurden die Interaktionen zwischen den Arbeitsgedächtnismaßen, insofern handelt es sich bei den MANCOVAs um nicht gesättigte Modelle. In den MANCOVAs wurden jeweils solche AVn zusammengefasst, für die nach der ECTML entweder (a) ein Einfluss der experimentellen Variation erwartet wurde oder (b) kein Einfluss der experimentellen Variation erwartet wurde. Die Ergebnisse dieser MANCOVAs werden nach Hypothesen getrennt aufgeführt.

9.2.1 Reliabilitätsanalysen

Bevor die Daten ausgewertet wurden, wurden die Reliabilitäten der einzelnen Indizes berechnet (vgl. Tabelle 13). Dabei ist zu beachten, dass aufgrund der unterschiedlichen Lerninhalte der Lerner mit visuellen und räumlichen Textinhalten die Itemzusammensetzung der Indizes zwischen den beiden Gruppen variieren konnte. Dadurch erklärt sich die in Tabelle 13 bei einigen Indizes vorgenommene Differenzierung der Reliabilitäten in „visuell“ und „räumlich“. Des Weiteren ist in Tabelle 13 dargestellt, welche Reliabilitäten resultierten, falls ein Itemausschluss zur Erhöhung von Cronbach's α notwendig war.

Wie Tabelle 13 aufzeigt, variierten die Reliabilitäten der verschiedenen Indizes stark. Dabei sollten nach Field (2009) akzeptable Reliabilitäten mindestens ein Cronbach's $\alpha = .70$ aufweisen. Dieser Wert wurde auch nach Itemausschluss nur auf wenigen der Skalen erreicht. Dabei ist ein Problem, dass der Reliabilitätsindex auch von der Anzahl der Items abhängt, die in die Skala eingehen (vgl. Bortz & Döring, 2002). Da die Indizes des vorliegenden Experiments sich häufig aus weniger als zehn

Items zusammensetzten, kann dies die geringen Reliabilitäten zumindest teilweise erklären.

Tabelle 13

Cronbach's α der einzelnen Indizes vor und nach Itemausschluss

	Cronbach's α vor Itemausschluss			Cronbach's α nach Itemausschluss		
	Gleicher Index	Ungleicher Index		Gleicher Index	Ungleicher Index	
		visuell	räumlich		visuell	räumlich
<i>bildbezogene Erinnerungsleistung</i>						
visuelle Bildinhalte		$\alpha = .32$ $i = 9$	$\alpha = .55$ $i = 18$		$\alpha = .40$ $i = 8$	$\alpha = .58$ $i = 17$
räumliche Bildinhalte		$\alpha = .68$ $i = 18$	$\alpha = .17$ $i = 9$		$\alpha = .71$ $i = 17$	$\alpha = .38$ $i = 7$
<i>textbezogene Erinnerungsleistung</i>						
Charakteristika	$\alpha = .74$ $i = 30$			$\alpha = .78$ $i = 28$		
abstrakte Textinhalte	$\alpha = .39$ $i = 6$			$\alpha = .43$ $i = 5$		

Anmerkung. i = Anzahl der eingehenden Items.

9.2.2 Datenexploration und Überprüfung inferenzstatistischer Analysevoraussetzungen

Die Daten wurden auf Ausreißer, Normalverteilung mit Hilfe des Kolmogorov-Smirnov Tests (vgl. Anhang C) und Varianzhomogenität mit Hilfe des Levene-Tests (vgl. Anhang D) geprüft. Nur der Kolmogorov-Smirnov Test zeigte für einige Variablen statistisch bedeutsame Abweichungen von der Normalverteilung (vgl. Tabelle 14), die durch Transformationen nicht behoben werden konnten. Für die später berechneten (multivariaten) Varianzanalysen mit kontinuierlichen Faktoren war eine adäquate Absicherung bei Abweichungen von der Normalverteilung mit non-parametrischen Verfahren leider nicht möglich. Da die Gruppengrößen jedoch vergleichbar und die Freiheitsgrade des Fehlerterms größer 20 waren, konnte trotz der Abweichungen davon ausgegangen werden, dass der F -Test robust ist (vgl. hierzu auch Field, 2009). Die Prüfung der Homogenität der Kovarianzmatrizen konnte ebenfalls vernachlässigt werden, da bei vergleichbaren Gruppengrößen der zur Prüfung verwendete Box Test instabil ist und Pillai's Spur (V) als robust gilt (vgl. Field, 2009).

Tabelle 14

Variablen mit Abweichungen von der Normalverteilung sowie ungleichen Fehlervarianzen

Abweichungen von Normalverteilung	<i>Kontinuierliche unabhängige Variablen</i> Corsi Block Test, <i>Digit Span</i> Test <i>Erinnerungsleistung</i> Verifikationsitems: Erinnerung an abstrakte Textinhalte <i>Bewertungsitems</i> Alle Variablen
-----------------------------------	---

In die MANCOVAs ging die *Digit Span* als einfache Kovariate ein. Die hierfür notwendige Vergleichbarkeit der Regressionsgeraden innerhalb der Bedingungen war gegeben (vgl. Anhang E), so dass die *Digit Span* als Kovariate in alle Analysen aufgenommen werden konnte.

9.2.3 Beschreibung der Stichprobe

Tabelle 15 gibt einen Überblick über die Mittelwerte und Standardabweichungen der intervallskalierten Kontrollvariablen.

Tabelle 15

Mittelwerte und Standardabweichungen (in Klammern) der intervallskalierten Variablen in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	geschrieben <i>n</i> = 16	gesprochen <i>n</i> = 15	geschrieben <i>n</i> = 14	gesprochen <i>n</i> = 14
Alter (in Jahren)	24.81 (3.99)	22.00 (2.27)	24.43 (5.53)	23.79 (2.52)
Semesteranzahl	7.25 (5.67)	5.27 (2.89)	6.00 (4.64)	6.07 (3.58)
Letzte Biologienote (0 = ungenügend, 15 = sehr gut plus)	10.06 (2.91)	10.27 (2.81)	9.86 (2.32)	7.92 (3.25)
Vorkenntnisse bzgl. Fischen (1 = keine bis 4 = hoch)	1.81 (0.58)	1.60 (0.63)	1.71 (0.73)	1.79 (0.58)
Corsi Block Test	5.63 (0.62)	6.00 (0.85)	5.64 (0.84)	5.50 (1.09)
<i>Visual Pattern</i> Test	7.52 (1.85)	7.33 (1.96)	7.21 (1.34)	7.45 (1.58)
<i>Digit Span</i> Test	6.00 (1.10)	6.67 (1.11)	5.93 (1.07)	5.64 (1.22)

Bezüglich der erhobenen nominalen Daten (Geschlecht, Biologie 4-stündig/Leistungskurs, Studienfach) gibt Tabelle 16 einen Überblick über die beobachteten Häufigkeiten.

Tabelle 16

Die Verteilung der nominalen Daten in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	geschrieben <i>n</i> = 16	gesprachen <i>n</i> = 15	geschrieben <i>n</i> = 14	gesprachen <i>n</i> = 14
<i>Geschlecht</i>				
männlich	6	5	3	2
weiblich	10	10	11	12
<i>Biologie 4-stündig/Leistungskurs</i>				
Ja	5	3	4	0
Nein	11	12	10	14
<i>Studienfach</i>				
naturwissenschaftlich	1	0	0	2
sonstiges	15	15	14	12

Die Analyse auf Gleichheit der Eingangsbedingungen zeigte für keine der erhobenen Variablen kritische Unterschiede oder Ungleichverteilungen zwischen den Gruppen (vgl. Anhang F).

Die Stichprobe lässt sich daher abschließend wie folgt beschreiben: Nach Ausschluss zweier Versuchspersonen aufgrund von Farbblindheit verblieben 59 teilnehmende Personen, davon 43 Frauen und 16 Männer. Das Durchschnittsalter betrug $M = 23.67$ Jahre ($SD = 3.48$), die durchschnittliche Semesteranzahl belief sich auf 6.15 Semester ($SD = 4.10$). Die letzte Biologienote betrug im Durchschnitt $M = 9.75$ Punkte ($SD = 2.87$), was in etwa einem „befriedigend plus“ entspricht. Zwölf Teilnehmer hatten in der Schule Biologie als Vertiefungsfach oder als Leistungskurs belegt, aber nur drei Teilnehmer studierten aktuell ein naturwissenschaftliches Fach wie Medizin oder Biologie. Das Vorwissen bezüglich Fischen war mit einem durchschnittlichen Wert von $M = 1.72$ und einer geringen Standardabweichung von $SD = 0.67$ (auf einer Skala von 1 = keine Vorkenntnisse bis 4 = hohe Vorkenntnisse) als niedrig einzuschätzen. Die Corsi Block Spanne lag bei durchschnittlich 5.69 Blöcken ($SD = 0.86$), d. h. die Probanden konnten Sequenzen mit durchschnittlich 5.69 Blöcken erinnern. Dies ist vergleichbar mit Ergebnissen

von Kessels, van Zandvoort, Postma, Kappelle und De Haan (2000). Die VPT Spanne lag bei $M = 7.38$ ($SD = 1.67$), d. h. die Probanden konnten Muster erinnern, die sich aus durchschnittlich 7.38 ausgefüllten Kästchen in einer Matrix zusammensetzten. Dieser Wert liegt etwas unter dem von Della Sala et al. (1997) berichteten Normwerten für den VPT Test von $M = 9.08$ ($SD = 2.25$). Beim *Digit Span* Test zeigten die Teilnehmer eine durchschnittliche Erinnerungsspanne von 6.07 Zahlen pro Sequenz ($SD = 1.16$). Dieser Wert entspricht in etwa den Normwerten des Untertests „Zahlenspanne“ des Wechsler Gedächtnistest – Revidierte Fassung (WMS-R; Härtling et al., 2000), bei dem Probanden im Alter von 20-34 Jahre durchschnittlich circa 7 Zahlen erinnern⁷. Dabei ist allerdings anzumerken, dass sich der Zahlenspannen-Test des WMS-R insofern von dem hier durchgeführten *Digit Span* Test unterscheidet, als dass im WMS-R pro Block immer nur zwei Zahlensequenzen gleicher Länge dargeboten werden und nicht fünf Zahlensequenzen wie im vorliegenden Experiment.

9.2.4 Deskriptive Daten

Im Folgenden werden die deskriptiven Daten der einzelnen AVn dargestellt. Dabei handelt es sich um adjustierte Daten, da der Einfluss der individuellen Arbeitsgedächtniskapazität (PL, vVSSP und rVSSP) herausgerechnet wurde. Die adjustierten Mittelwerte und Standardfehler für die Verifikationsitems sind in Tabelle 17 dargestellt und die adjustierten Mittelwerte und Standardfehler für die Bewertungsitems in Tabelle 18.

⁷ Diese Anzahl an Zahlen wurde aus dem im Manual angegebenen Rohwert, der sich aus Leistungspunkten zusammensetzt, rückgerechnet.

Tabelle 17

Adjustierte Mittelwerte und Standardfehler (in Klammern) der Verifikationsitems in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	geschrieben <i>n</i> = 15	gesprochen <i>n</i> = 16	geschrieben <i>n</i> = 14	gesprochen <i>n</i> = 14
<i>bildbezogene Erinnerungsleistung</i>				
visuelle Bildinhalte	1.42 (0.12)	1.81 (0.14)	1.14 (0.13)	1.18 (0.13)
räumliche Bildinhalte	1.29 (0.13)	1.18 (0.15)	0.76 (0.14)	0.79 (0.14)
<i>textbezogene Erinnerungsleistung</i>				
Charakteristika	1.87 (0.12)	1.75 (0.14)	1.34 (0.13)	1.49 (0.13)
abstrakte Textinhalte	2.47 (0.19)	2.05 (0.22)	2.16 (0.20)	2.35 (0.21)

Anmerkung. Minimum = 0, Maximum = 4.

Tabelle 18

Adjustierte Mittelwerte und Standardfehler (in Klammern) der Bewertungsitems in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	geschrieben <i>n</i> = 15	gesprochen <i>n</i> = 16	geschrieben <i>n</i> = 14	gesprochen <i>n</i> = 14
Wie anstrengend war die Lernaufgabe für Dich?	4.59 (0.30)	4.48 (0.35)	4.34 (0.32)	4.67 (0.33)
Wie schwierig war es für Dich, die Inhalte zu verstehen?	2.27 (0.30)	2.27 (0.30)	2.37 (0.32)	2.62 (0.33)
Wie sehr hast Du Dich beim Lernen auf das Bild konzentriert?	4.31 (0.35)	5.74 (0.41)	4.11 (0.37)	5.30 (0.38)
Wie sehr hast Du Dich beim Lernen auf den Text konzentriert?	5.67 (0.34)	5.01 (0.40)	4.85 (0.36)	4.73 (0.37)
Wie sehr hast Du Dich während der Lernphase konzentriert?	5.80 (0.23)	5.16 (0.26)	5.27 (0.24)	5.42 (0.25)
Wie sehr hast Du Dich angestrengt, um das Lernmaterial zu verstehen?	5.93 (0.23)	5.44 (0.27)	5.44 (0.24)	5.61 (0.25)
Wie viel Mühe hast Du in die Lernphase investiert?	5.62 (0.23)	5.56 (0.27)	5.38 (0.24)	5.25 (0.25)

Anmerkung. Minimum = 1, Maximum = 7. Hohe Werte indizieren eine hohe Ausprägung im Sinne der Itemformulierung.

Im nächsten Schritt erfolgte die Überprüfung der Hypothesen mit Hilfe der drei beschriebenen multivariaten Varianzanalysen mit den Zwischensubjektfaktoren

Textinhalt und Textmodalität und den kontinuierlichen, z-standardisierten Testwerten im VPT, Corsi Block und *Digit Span* Test. Die *Digit Span* ging als einfache Kovariate in die Analyse ein, während hinsichtlich des VPT und Corsi Block Maß auch die Interaktionen mit den experimentellen Bedingungen berücksichtigt wurden (für eine ausführlichere Beschreibung siehe oben).

9.2.5 Prüfung der Textmodalität-Hypothese

Nach der CTML wurde ein Haupteffekt der Textmodalität unabhängig vom Textinhalt sowie eine Moderation des Einflusses der Textmodalität durch die Kapazität des vVSSP, gemessen über den VPT, und des rVSSP, gemessen über den Corsi Block Test, für alle AVn erwartet. Die ECTML machte spezifische Vorhersagen in Abhängigkeit des Textinhalts und der AVn: Nur wenn zur Beantwortung der AVn die Erinnerung an räumliche Inhalte notwendig war (räumliche Bildinhalte, Charakteristika), wurde ein Modalitätseffekt bei der Darbietung räumlicher Textinhalte, aber nicht bei der Darbietung visueller Textinhalte vorhergesagt. Es wurde somit kein genereller Haupteffekt der Textmodalität, sondern eine Interaktion zwischen Textinhalt und Textmodalität erwartet. Diese Interaktion sollte zusätzlich durch die Kapazität des rVSSP, gemessen über den Corsi Block, moderiert werden. Hinsichtlich Variablen, bei denen keine räumlichen Informationen erinnert werden mussten (visuelle Bildinhalte, abstrakte Textinhalte), wurde hingegen kein Einfluss der Textmodalität erwartet.

Zur Erinnerung: Die Charakteristika unterschieden sich zwischen Lernern mit visuellen und räumlichen Textinhalten, so dass Lerner mit visuellen Textinhalten hier visuelle Charakteristika erinnern mussten, Lerner mit räumlichen Textinhalten hingegen räumliche Charakteristika. Die abstrakten Textinhalte unterschieden sich nicht zwischen den Gruppen, d. h. alle Lerner erinnerten die gleichen Informationen. Beide Maße geben Auskunft über die textbezogene Erinnerungsleistung. Die Erinnerung der visuellen und räumlichen Bildinhalte gibt hingegen Auskunft über die reine Bilderinnerung, d. h. hier mussten nicht die Textinhalte, sondern ausschließlich die Bildinhalte erinnert werden. Eine Konfundierung mit der Textschwierigkeit ist somit hinsichtlich der Erinnerung visueller und räumlicher Bildinhalte wenig wahrscheinlich.

Verifikationsitems

In Tabelle 19 sind die für die Prüfung des Einflusses der Textmodalität relevanten Haupteffekte und Interaktionen dargestellt. Die erste MANCOVA bezüglich der Erinnerung an räumliche Bildinhalte und Charakteristika konnte weder die Annahmen nach der CTML noch nach der ECTML bestätigen.

Tabelle 19

Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der MANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die Verifikationsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
räumliche Bildinhalte Charakteristika	TM	$V = .003, F < 1$
	TM \times TI	$V = .02, F < 1$
	TM \times Corsi	$V = .04, F < 1$
	TM \times VPT	$V = .01, F < 1$
	TM \times TI \times Corsi	$V = .06, F(2, 45) = 1.38, p = .26$
	TM \times TI \times VPT	$V = .05, F(2, 45) = 1.09, p = .35$
visuelle Bildinhalte abstrakte Textinhalte	TM	$V = .08, F(2, 45) = 1.81, p = .18$
	TM \times TI	$V = .11, F(2, 45) = 2.64, p = .08$
	TM \times Corsi	$V = .05, F(2, 45) = 1.16, p = .33$
	TM \times VPT	$V = .05, F(2, 45) = 1.26, p = .29$
	TM \times TI \times Corsi	$V = .06, F(2, 45) = 1.58, p = .22$
	TM \times TI \times VPT	$V = .15, F(2, 45) = 3.89, p = .03$

Anmerkung. Das *Digit Span* Maß ging zusätzlich als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Die zweite MANCOVA hinsichtlich der Erinnerung an visuelle Bildinhalte und abstrakte Textinhalte konnte ebenfalls weder die Annahmen nach der CTML noch der ECTML bestätigen. Die nach der ECTML nicht vorhandene Interaktion zwischen Textinhalt und Textmodalität war zwar nicht signifikant, verpasste aber das strenge Kriterium einer Irrtumswahrscheinlichkeit von $\alpha = .20$. Zudem zeigte die MANCOVA eine Dreifachinteraktion zwischen der Textmodalität, dem Textinhalt und dem VPT. Zur Analyse dieser multivariaten Interaktion wurden die entsprechenden univariaten Analysen berechnet. Hinsichtlich der Erinnerung an abstrakte Textinhalte wurde die Interaktion zwischen Textmodalität, Textinhalt und VPT nicht statistisch bedeutsam, $F < 1$. Hinsichtlich der Erinnerung an visuelle Bildinhalte bestätigte die univariate Analyse jedoch die Dreifachinteraktion, $F(1, 46) = 7.92$, $MSE = 0.24$, $p = .01$, $\eta_p^2 = .15$. In Abbildung 19 sind die Regressionsgeraden mit den z-standardisierten VPT Messwerten als Prädiktoren und

der Erinnerung an visuelle Bildinhalte als AVn für die einzelnen Bedingungen dargestellt.

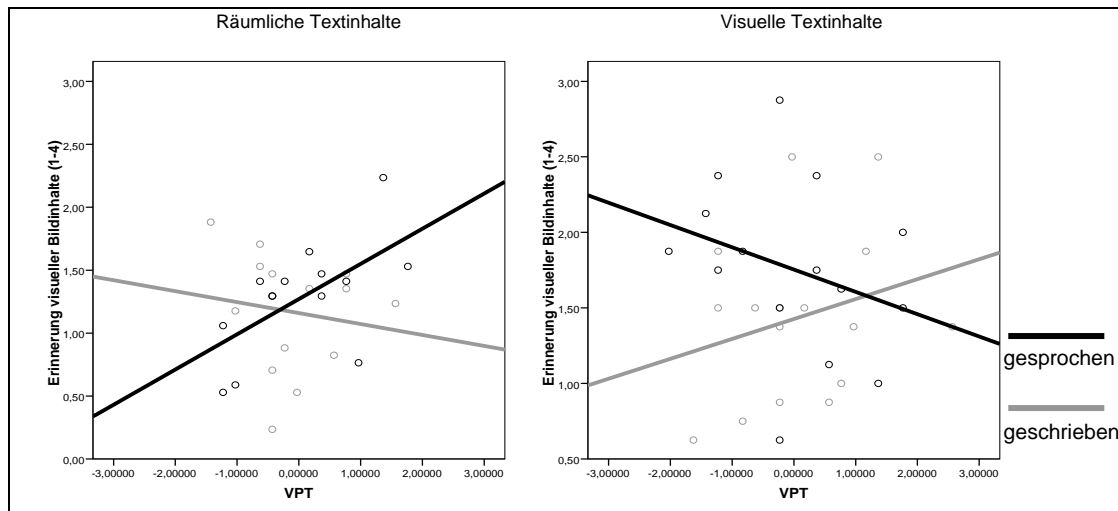


Abbildung 19. Darstellung des Zusammenhangs des VPT Maßes mit der Erinnerung an visuelle Bildinhalte in Abhängigkeit des Textinhalts und der Textmodalität.

Abbildung 19 verdeutlicht, dass der Einfluss des VPT auf die Erinnerung an visuelle Bildinhalte in Abhängigkeit des Textinhalts und der Textmodalität gegenläufig war. Während bei räumlichen Textinhalten Lerner mit gesprochener Textdarbietung mit größerer vVSSP Kapazität mehr visuelle Bildinhalte erinnerten ($pr^{\delta} = .84, p = .001$), gab es bei Lernern mit geschriebener Textdarbietung keinen statistisch bedeutsamen Zusammenhang der Erinnerung an visuelle Bildinhalte mit der vVSSP Kapazität ($pr = -.44, p = .15$). Bei visuellen Textinhalten gab es ebenfalls keine statistisch bedeutsamen Zusammenhänge zwischen Textmodalität und vVSSP Kapazität (gesprochener Text: $pr = -.26, p = .39$; geschriebener Text: $pr = .27, p = .36$).

Bewertungsitems

Die Analysen der Bewertungsitems erfolgten mit der zur Analyse der Erinnerungsleistung äquivalenten MANCOVA (vgl. Tabelle 20).

⁸ Für die Berechnung partieller Korrelationen (pr) gingen jeweils die beiden anderen Arbeitsgedächtnismaße als Kontrollvariablen ein.

Tabelle 20

Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der MANCOVA mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die Bewertungsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	TM	V = 34, $F(7, 40) = 2.51, p = .03$
	TM × TI	V = .16, $F < 1$
	TM × Corsi	V = 21, $F(7, 40) = 1.32, p = .26$
	TM × VPT	V = .22, $F(7, 40) = 1.40, p = .23$
	TM × TI × Corsi	V = .07, $F < 1$
	TM × TI × VPT	V = .19, $F(7, 40) = 1.16, p = .35$

Anmerkung. Das Digit Span Maß ging zusätzlich als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Die MANCOVA zeigte einen statistisch bedeutsamen Haupteffekt der Textmodalität. Die univariaten Analysen bestätigten den Effekt allerdings nur für die Konzentration auf das Bild (vgl. Tabelle 21): Lerner mit gesprochener Textdarbietung ($M = 5.52, SE = 0.27$) konzentrierten sich nach eigenen Angaben mehr auf das Bild als Lerner mit geschriebener Textdarbietung ($M = 4.21, SE = 0.26$).

Tabelle 21

Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der ANCOVAs für den Faktor Textmodalität (TM) für die Bewertungsitems

		Ergebnisse der ANOVA
Wie anstrengend war die Lernaufgabe für Dich? ^a	TM	$F < 1$
Wie schwierig war es für Dich, die Inhalte zu verstehen? ^a	TM	$F < 1$
Wie sehr hast Du Dich beim Lernen auf das Bild konzentriert?	TM	$F(1, 46) = 12.15, MSE = 1.89, p = .001, \eta_p^2 = .21$
Wie sehr hast Du Dich beim Lernen auf den Text konzentriert? ^a	TM	$F(1, 46) = 1.10, MSE = 1.78, p = .30, \eta_p^2 = .02$
Wie sehr hast Du Dich während der Lernphase konzentriert? ^a	TM	$F(1, 46) = 1.02, MSE = 0.80, p = .32, \eta_p^2 = .02$
Wie sehr hast Du Dich angestrengt, um das Lernmaterial zu verstehen? ^a	TM	$F < 1$
Wie viel Mühe hast Du in die Lernphase investiert? ^a	TM	$F < 1$

Anmerkung. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Zusammenfassend kann man somit festhalten, dass die Annahme der CTML einer generellen Überlegenheit gesprochener Textdarbietung nicht bestätigt wurde. Nur im

Hinblick auf die Bewertungssitems zeigte sich der postulierte Modalitätseffekt hinsichtlich der Konzentration der Lerner auf das Bild. Die Annahmen der ECTML wurden durch die Ergebnisse ebenfalls nicht bestätigt, da das vorhergesagte Interaktionsmuster zwischen Textinhalt und Textmodalität für keine der Variablen beobachtet wurde.

9.2.6 Prüfung der Textinhalt-Hypothese

Nach der CTML wurde für keine der AVn ein Einfluss des Textinhalts oder eine Moderation des Einflusses des Textinhalts erwartet. Die ECTML machte hingegen spezifische Vorhersagen hinsichtlich des Einflusses des Textinhalts in Abhängigkeit der AVn: Wenn zur Beantwortung der AVn die Erinnerung an räumliche Inhalte notwendig war (räumliche Bildinhalte, Charakteristika), wurde eine Unterlegenheit der Lerner mit räumlichen Textinhalten erwartet. Es wurde zudem angenommen, dass die Kapazitäten des rVSSP, gemessen über den Corsi Block, und des vVSSP, gemessen über den VPT, die Erinnerung an diese Textinhalte beeinflussen würden. Hinsichtlich Variablen, bei denen keine räumlichen Informationen erinnert werden mussten (visuelle Bildinhalte, abstrakte Textinhalte), wurde kein Einfluss des Textinhalts erwartet. Die Prüfung der Hypothese erfolgte mit Hilfe der oben beschriebenen MANCOVAs.

Verifikationsitems

In Tabelle 22 sind der für die Prüfung des Einflusses des Textinhalts relevante Haupteffekt sowie die Interaktionen mit den Arbeitsgedächtnismaßen des vVSSP und rVSSP dargestellt.

Tabelle 22

Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der MANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) für die Verifikationsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
räumliche Bildinhalte Charakteristika	TI	V = .23, $F(2, 45) = 6.80, p = .003$
	TI × Corsi	V = .02, $F < 1$
	TI × VPT	V = .004, $F < 1$
visuelle Bildinhalte abstrakte Textinhalte	TI	V = .21, $F(2, 45) = 6.00, p = .01$
	TI × Corsi	V = .01, $F < 1$
	TI × VPT	V = .02, $F < 1$

Anmerkung. Die Textmodalität ging zusätzlich als UVn, das *Digit Span* Maß als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Die erste MANCOVA prüfte die Textinhalt-Hypothese hinsichtlich der Erinnerung an räumliche Bildinhalte und Charakteristika. Der von der ECTML vorhergesagte Haupteffekt des Textinhalts wurde statistisch bedeutsam. Die Arbeitsgedächtnismaße nahmen keinen Einfluss auf diesen Haupteffekt. Zur Analyse des multivariaten Haupteffekts wurden die entsprechenden univariaten Analysen berechnet. Bezüglich der Erinnerung an räumliche Bildinhalte schnitten Lerner mit visuellen Textinhalten besser ab ($M = 1.24, SE = 0.10$) als Lerner mit räumlichen Textinhalten ($M = 0.78, SE = 0.10$), $F(1, 42) = 9.98, MSE = 0.27, p = .003, \eta_p^2 = .18$. Bezüglich der Erinnerung an Charakteristika, also der jeweiligen experimentell variierten Textinhalte, zeigte sich das gleiche Muster: Lerner mit visuellen Textinhalten ($M = 1.81, SE = 0.09$) erinnerten die visuellen Textinhalte besser als Lerner mit räumlichen Textinhalten die räumlichen Textinhalte erinnerten ($M = 1.41, SE = 0.09$), $F(1, 46) = 9.35, MSE = 0.22, p = .004, \eta_p^2 = .17$.

Die zweite MANCOVA analysierte Variablen bezüglich derer auch nach der ECTML kein Einfluss des Textinhalts erwartet wurde. Die Analyse zeigte jedoch erneut einen multivariaten Effekt des Textinhalts. Hinsichtlich der Erinnerung an visuelle Bildinhalte zeigte die univariate Analyse einen Einfluss des Textinhalts, $F(1, 46) = 11.31, MSE = 0.25, p = .002, \eta_p^2 = .20$: Lerner mit visuellen Textinhalten erinnerten nicht nur wie oben erwähnt räumliche sondern auch visuelle Bildinhalte besser ($M = 1.61, SE = 0.10$) als Lerner mit räumlichen Textinhalten ($M = 1.16, SE = 0.11$). Hinsichtlich der Erinnerung an abstrakte Textinhalte bestätigte die univariate Analyse die aufgestellten Hypothesen: Der Textinhalt nahm keinen Einfluss auf die Leistung, $F < 1$.

Bewertungsitems

Die Ergebnisse der MANCOVA hinsichtlich der Bewertungsitems sind in Tabelle 23 dargestellt. Es zeigten sich keine statistisch bedeutsamen Effekte hinsichtlich des Textinhalts.

Tabelle 23

Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der MANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) für die Bewertungsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	TI	$V = .12, F < 1$
	TI \times Corsi	$V = .08, F < 1$
	TI \times VPT	$V = .06, F < 1$

Anmerkung. Die Textmodalität ging zusätzlich als UVn, das *Digit Span* Maß als Kovariate ein.

Zusammenfassend kann man somit hinsichtlich der Textinhalt-Hypothese festhalten, dass die Annahmen der ECTML für die Erinnerung an räumliche Bildinhalte und die Charakteristika bestätigt wurden: Lerner mit räumlichen Textinhalten hatten hier jeweils größere Schwierigkeiten die Inhalte zu erinnern als Lerner mit visuellen Textinhalten. Auf Ebene der Bewertungsitems zeigte sich kein Einfluss des Textinhalts. Entgegen den Annahmen der ECTML (und auch der CTML) zeigte sich ein Einfluss des Textinhalts auch auf die Erinnerung an visuelle Bildinhalte: Lerner mit visuellen Textinhalten schnitten hier jeweils besser ab als Lerner mit räumlichen Textinhalten.

9.2.7 Prüfung der Arbeitsgedächtnis-Hypothese

Neben der Prüfung des Einflusses der Textmodalität und des Textinhalts auf das Lernergebnis sowie der Frage, inwiefern die Arbeitsgedächtniskapazität deren Einflüsse moderieren könnte, wurde auch der generelle Einfluss der verschiedenen Arbeitsgedächtniskomponenten auf das Lernergebnis untersucht. Sowohl nach der CTML als auch nach der ECTML sollte das Arbeitsgedächtnis eine entscheidende Rolle beim Lernen mit Multimedia spielen. Daher sollte die individuelle Arbeitsgedächtniskapazität einen Einfluss auf das Lernergebnis nehmen. Die Prüfung dieser Annahme erfolgte im Rahmen der oben beschriebenen MANCOVAs.

Verifikationsitems

In Tabelle 24 sind die relevanten Haupteffekte der *Digit Span*, des VPT und des Corsi Blocks hinsichtlich der Verifikationsitems dargestellt. Keine der drei MANCOVAs zeigte einen generellen Einfluss der drei Arbeitsgedächtnismaße auf das Lernergebnis.

Tabelle 24

Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der MANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und Digit Span für die Verifikationsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
räumliche Bildinhalte Charakteristika	<i>Digit Span</i>	$V = .09, F(2, 45) = 2.15, p = .13$
	Corsi	$V = .03, F < 1$
	VPT	$V = .01, F < 1$
visuelle Bildinhalte abstrakte Textinhalte	<i>Digit Span</i>	$V = .10, F(2, 45) = 2.54, p = .09$
	Corsi	$V = .01, F < 1$
	VPT	$V = .02, F < 1$

Anmerkung. Die Textmodalität und der Textinhalt gingen zusätzlich als UVn in die Analyse ein.

Bewertungsitems

Die Analyse des Einflusses der Arbeitsgedächtniskapazität auf die Bewertung der Lernphase erfolgte äquivalent zu den Analysen hinsichtlich der Erinnerungsleistung (vgl. Tabelle 25). Hinsichtlich der Bewertungsitems zeigte die MANCOVA keinen Einfluss der Arbeitsgedächtnismaße.

Tabelle 25

Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der MANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und Digit Span für die Bewertungsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	<i>Digit Span</i>	$V = .15, F < 1$
	Corsi	$V = .19, F(7, 40) = 1.16, p = .35$
	VPT	$V = .09, F < 1$

Anmerkung. Die Textmodalität und der Textinhalt gingen zusätzlich als UVn in die Analyse ein.

Zusammenfassend kann man somit festhalten, dass entgegen der Annahmen kein genereller Einfluss der Arbeitsgedächtniskapazität zu beobachten war.

9.3 Zusammenfassung und Diskussion

Das Ziel des ersten Experiments war die Prüfung der Vorhersagen der CTML sowie der Vorhersagen der ECTML. Es wurden drei Hypothesen geprüft, nämlich die Textmodalität-Hypothese, die Textinhalt-Hypothese und die Arbeitsgedächtnis-Hypothese. Zur Prüfung der Hypothesen wurde variiert, ob visuelle oder räumliche Textinhalte verwendet wurden und ob der Text gesprochen oder geschrieben dargeboten wurde. Zudem wurde die Kapazität der PL, des vVSSP und des rVSSP erfasst. AVn waren die Erinnerungsleistung sowie die Bewertung der Lernphase hinsichtlich verschiedener Aspekte.

Welchen Einfluss nimmt die Textmodalität beim Lernen mit Multimedia?

Die **Textmodalität-Hypothese** nach der ECTML ging davon aus, dass bei der Darbietung geschriebener Textinhalte zusätzliche Blickbewegungen evoziert werden, was in einer erhöhten Belastung des rVSSP resultiert. Diese erhöhte Belastung sollte jedoch nur dann zu einem Modalitätseffekt führen, wenn der rVSSP zusätzlich durch die Verarbeitung räumlicher Textinhalte belastet wird. Unter dieser Bedingung wurde für die Erinnerung an räumliche bildbezogene Informationen sowie an räumliche textbezogene („Charakteristika“) Informationen ein Modalitätseffekt erwartet. Für die Erinnerung an nicht-räumliche Informationen wurde hingegen kein Einfluss der Textmodalität erwartet. Diese Hypothesen konnten durch die Ergebnisse jedoch nicht bestätigt werden, auch nicht unter Einbeziehung der Kapazität des rVSSP.

Nach der CTML wurde dagegen ein Modalitätseffekt für *alle* AVn erwartet. Diese Annahme konnte ebenfalls nicht bestätigt werden, denn ein Modalitätseffekt wurde für die Erinnerungsleistung nicht beobachtet. Einzig hinsichtlich der subjektiv wahrgenommenen Konzentration auf das Bild zeigte sich ein Modalitätseffekt, da Lerner mit geschriebener Textdarbietung angaben, sich weniger auf das Bild konzentriert zu haben als Lerner mit gesprochener Textdarbietung. Dieser Befund bestätigt im weitesten Sinne die in Abschnitt 3.3.1 vorgenommene Spezifizierung der räumlichen Kontiguitätsklärung für den Modalitätseffekt: Denn statt visuelle Suchprozesse durchzuführen, konzentrierten sich die Lerner auf den geschriebenen Text und vernachlässigten das Bild. Allerdings führte die Vernachlässigung des Bildes auf Performanzebene nicht zu einem Modalitätseffekt für die bildbezogene Erinnerung. Die Annahme der CTML bezüglich einer Moderation des

Modalitätseffekts durch die Kapazität des rVSSP oder vVSSP konnte nicht bestätigt werden.

Zusammenfassend kann man somit festhalten, dass es allein auf der subjektiv wahrgenommenen Ebene Unterschiede zwischen den Gruppen gab, nicht aber hinsichtlich der Erinnerungsleistung. Die Ergebnisse widersprechen somit sowohl den Annahmen der CTML als auch der ECTML.

Neben diesen hypothesenbezogenen Befunden gab es einen weiteren Befund hinsichtlich des Einflusses der Textmodalität: Bei der Erinnerung an visuelle Bildinhalte gab es eine Dreifachinteraktion zwischen Textinhalt, Textmodalität und vVSSP Kapazität. Die vVSSP Kapazität nahm einen positiven Einfluss auf die Erinnerung an visuelle Bildinhalte bei Lernern mit gesprochener Darbietung räumlicher Textinhalte. Da dieser Befund nicht vorhergesagt wurde, sollte zunächst eine Replikation erfolgen, bevor eine Interpretation vorgenommen wird.

Welchen Einfluss nimmt der Textinhalt beim Lernen mit Multimedia?

Die **Textinhalt-Hypothese** nach der CTML ging davon aus, dass der Textinhalt keinen Einfluss auf die Leistung nimmt, unabhängig davon, welche Informationen erinnert werden müssen. Nach der ECTML sollte es hingegen bei der Darbietung räumlicher Textinhalte zu einer erhöhten Belastung des rVSSP kommt. Diese erhöhte Belastung sollte sich im Vergleich zur Darbietung visueller Textinhalte sowohl in einer schlechteren Erinnerung an räumliche bildbezogene Informationen als auch an die Charakteristika niederschlagen.

Die Annahmen der ECTML wurden für die Erinnerung an räumliche Bildinhalte und Charakteristika bestätigt. Des Weiteren zeigte sich entgegen den Annahmen auch ein besseres Abschneiden der Lerner mit visuellen Textinhalten bei der Erinnerung an visuelle Bildinhalte: Lerner mit räumlichen Textinhalten hatten also sowohl größere Schwierigkeiten räumliche als auch visuelle Aspekte des Bildes zu erinnern. Dieser Befund deutet darauf hin, dass bei Bildern eine Trennung in visuelle und räumliche Aspekte nicht sinnvoll ist. Räumliche Textinhalte würden demnach zwar insbesondere mit räumlichen Bildinhalten im rVSSP interferieren, dadurch aber die Verarbeitung des gesamten Bildes verschlechtern. Dieser Befund ist im Einklang mit den in Abschnitt 2.3.2 zitierten Studien, die zeigen, dass die Belastung des rVSSP durch eine räumliche Zweitaufgabe die Bilderinnerung

insgesamt verschlechtert (z. B. Brunyé et al. 2006; Gyselinck et al., 2002; Nam & Pujari, 2005).

Zusammenfassend sprechen diese Befunde für die Annahme der ECTML, dass der Textinhalt einen Einfluss auf die Verarbeitung multimedialer Materialien nehmen kann. Insbesondere hinsichtlich der bildbezogenen Erinnerung sind diese Befunde interessant, da hier ein Einfluss der Textschwierigkeit weitestgehend ausgeschlossen werden kann. Diese Ergebnisse unterstützen somit die Annahme einer Interferenz zwischen räumlichen Textinhalten und der Bildverarbeitung.

Welchen Einfluss nimmt die Arbeitsgedächtniskapazität auf das Lernen mit Multimedia?

Nach der **Arbeitsgedächtnis-Hypothese** sollte die Kapazität des Arbeitsgedächtnisses einen entscheidenden Einfluss auf das Lernen mit Multimedia nehmen. Diese Annahme ist sowohl im Einklang mit der ECTML als auch der CTML. Unerwarteterweise nahmen die drei Arbeitsgedächtnismaße auf die Erinnerungsleistung keinen Einfluss. Mögliche Ursachen für dieses Ergebnis werden in Abschnitt 12.3 diskutiert.

Einschränkungen des ersten Experiments und Implikationen für Folgeexperimente

Wie oben aufgeführt, gibt es erste Evidenz für die Textinhalt-Hypothese nicht jedoch für die Textmodalität-Hypothese nach der ECTML. Dies spricht dafür, dass räumliche Textinhalte den rVSSP belasten. Da aber kein spezifischer Modalitätseffekt bei der Darbietung räumlicher Textinhalte auftrat, scheinen die durch geschriebenen Text evozierten Blickbewegungen zur Belastung des rVSSP nicht soviel beizutragen, dass es zu einem Leistungsunterschied hinsichtlich der räumlichen Erinnerungsleistung im Vergleich zu gesprochenen Texten kommt. Dies könnte darin begründet sein, dass bei der Darbietung multimedialer Materialien Blickbewegungen auch durch die Bildbetrachtung hervorgerufen werden. Denkbar wäre, dass allein die Kontrolle dieser Blickbewegungen zu einer Belastung des rVSSP führt und die Kontrolle zusätzlicher Blickbewegungen bei geschriebenen Texten weniger ins Gewicht fällt. Eine Möglichkeit, auch einen solchen geringen Effekt sichtbar zu machen, ist die Minimierung der Fehlervarianz im Vergleich zur experimentell erzeugten Varianz. Daher wurde im zweiten Experiment die Textmodalität intraindividuell variiert.

Eine weitere Einschränkung betrifft die Interpretierbarkeit der Ergebnisse bezüglich des Einflusses des Textinhalts. Wie bereits in Abschnitt 8.1.2 aufgeführt, kann insbesondere hinsichtlich der Erinnerung der Charakteristika eingewendet werden, dass die Unterschiede zwischen Lernern mit visuellen und räumlichen Textinhalten auf die jeweilige Schwierigkeit des Texts zurückgehen könnten. Hinsichtlich der bildbezogenen Erinnerung ist ein Einfluss der Textschwierigkeit weniger plausibel, es wäre allerdings denkbar, dass Lerner mit räumlichen Textinhalten den Bildern aufgrund einer höheren Textschwierigkeit weniger Aufmerksamkeit widmeten, was wiederum in schlechterer bildbezogener Erinnerungsleistung resultierte. Gegen die Annahme unterschiedlicher Textschwierigkeiten können jedoch die folgenden Argumente aufgeführt werden: (1) Die Texte waren hinsichtlich des Flesch-Index vergleichbar, was für eine vergleichbare Leseschwierigkeit der Texte spricht. (2) Lerner mit räumlichen Textinhalten stuften die Inhalte nicht als schwieriger ein, sie empfanden die Lernaufgabe nicht als anstrengender, sie konzentrierten sich nicht mehr auf den Text bzw. während der gesamten Lernphase, sie investierten nicht mehr Mühe in die Lernphase und sie strengten sich nicht mehr an, um das Lernmaterial zu verstehen. (3) Lerner mit visuellen und räumlichen Textinhalten gaben vergleichbare Werte für die Konzentration auf das Bild an. Somit ist es unwahrscheinlich, dass die Überlegenheit von Lernern mit visuellen Textinhalten bei der Erinnerung bildhafter Information darauf zurückging, dass sie aufgrund einer geringeren Textschwierigkeit dem Bild mehr Aufmerksamkeit widmeten. Aufgrund dieser Argumente wird die Textschwierigkeit als Erklärung für die Unterlegenheit der Lerner mit räumlichen Textinhalten hinsichtlich der Erinnerungsleistung als eher unwahrscheinlich betrachtet.

Für eine objektivere Absicherung, dass räumliche Textinhalte im rVSSP verarbeitet werden und Unterschiede zwischen den Lernern mit verschiedenen Textinhalten nicht auf Schwierigkeitsunterschiede zurückgehen, wurde in Experiment 2 eine weitere UV eingeführt, nämlich die Durchführung einer räumlichen Zweitaufgabe. Ein weiterer Vorteil der Zweitaufgabe war, dass dadurch die spezifische Belastung des rVSSP nochmals erhöht werden konnte, um auch kleine Effekte, wie z. B. einen Modalitätseffekt evoziert durch die Kontrolle von Blickbewegungen, sichtbar zu machen. Zudem konnte die Leistung in der Zweitaufgabe ebenfalls als AV ausgewertet werden. Da dieses Maß als prozessnäher

eingestuft werden kann als z. B. die Erinnerungsleistung, ist anzunehmen, dass sich hier die Arbeitsgedächtnisprozesse eher abbilden. In Experiment 3 wurden des Weiteren die Blickbewegungen aufgezeichnet, um abzusichern, dass sich die Lerner mit visuellen und räumlichen Textinhalten nicht bezüglich der dem Text und dem Bild gewidmeten visuellen Aufmerksamkeit unterscheiden.

Im zweiten und dritten Experiment wurden neben den Verifikationsitems Fragen mit offenem Antwortformat eingeführt. Letzteres Maß erschien vor allem deshalb sinnvoll, weil in der traditionellen Multimediaforschung sehr häufig offene Antwortformate verwendet werden (z. B. Moreno & Mayer, 1999). Daher ließ sich mit Hilfe der offenen Fragen untersuchen, inwiefern sich die Ergebnisse im *cued recall*, d. h. also der Verifikationsitems, auch in der offenen Erinnerung widerspiegelten und somit auch für die angewandte Multimediaforschung von Relevanz sind.

10 Experiment 2

10.1 Erweiterung der Hypothesen

Mit Hilfe des zweiten Experiments wurden erneut die Annahmen der ECTML bezüglich des Einfluss der Textmodalität und des Textinhalts auf das Lernen mit Multimedia gegen die entsprechenden Annahmen der CTML geprüft. Des Weiteren erfolgte die Prüfung des Einfluss der Arbeitsgedächtniskapazität auf das Lernergebnis. Das Design des ersten Experiments wurde in Experiment 2 durch eine räumliche Zweitaufgabe erweitert. Diese diente dazu, die Belastung des rVSSP nochmals spezifisch zu erhöhen und die Prozesse auf Arbeitsgedächtnisebene besser sichtbar zu machen. Des Weiteren konnte mit Hilfe der räumlichen Zweitaufgabe geprüft werden, ob räumliche Textinhalte tatsächlich im rVSSP verarbeitet werden.

Hinsichtlich der in Abschnitt 6 aufgestellten Annahmen bezüglich der Belastung des VSSP in Abhängigkeit des Textinhalts und der Textmodalität (vgl. auch Abbildung 11) kam in Experiment 2 für den rVSSP zusätzlich die Belastung durch die räumliche Zweitaufgabe hinzu. Dies ist in Abbildung 20 aufgeführt.

Die in Abschnitt 6 aufgestellten Hypothesen wurden in Experiment 2 um die Annahmen der Wirkweise der räumlichen Zweitaufgabe erweitert. Dabei entsprechen die Hypothesen für Lerner ohne Zweitaufgabe weiterhin den in Abschnitt 6 beschriebenen Annahmen. Eine Ausnahme bildete die Erinnerung an visuelle Bildinhalte, für die aufgrund der Ergebnisse des ersten Experiments dieselben Vorhersagen getroffen wurden wie für die Erinnerung an räumliche Bildinhalte. Hinsichtlich der AV „Leistung in der Zweitaufgabe“ entsprechen die Hypothesen den im Folgenden getätigten Annahmen bezüglich der Erinnerungsleistung bei Ausführung einer Zweitaufgabe.

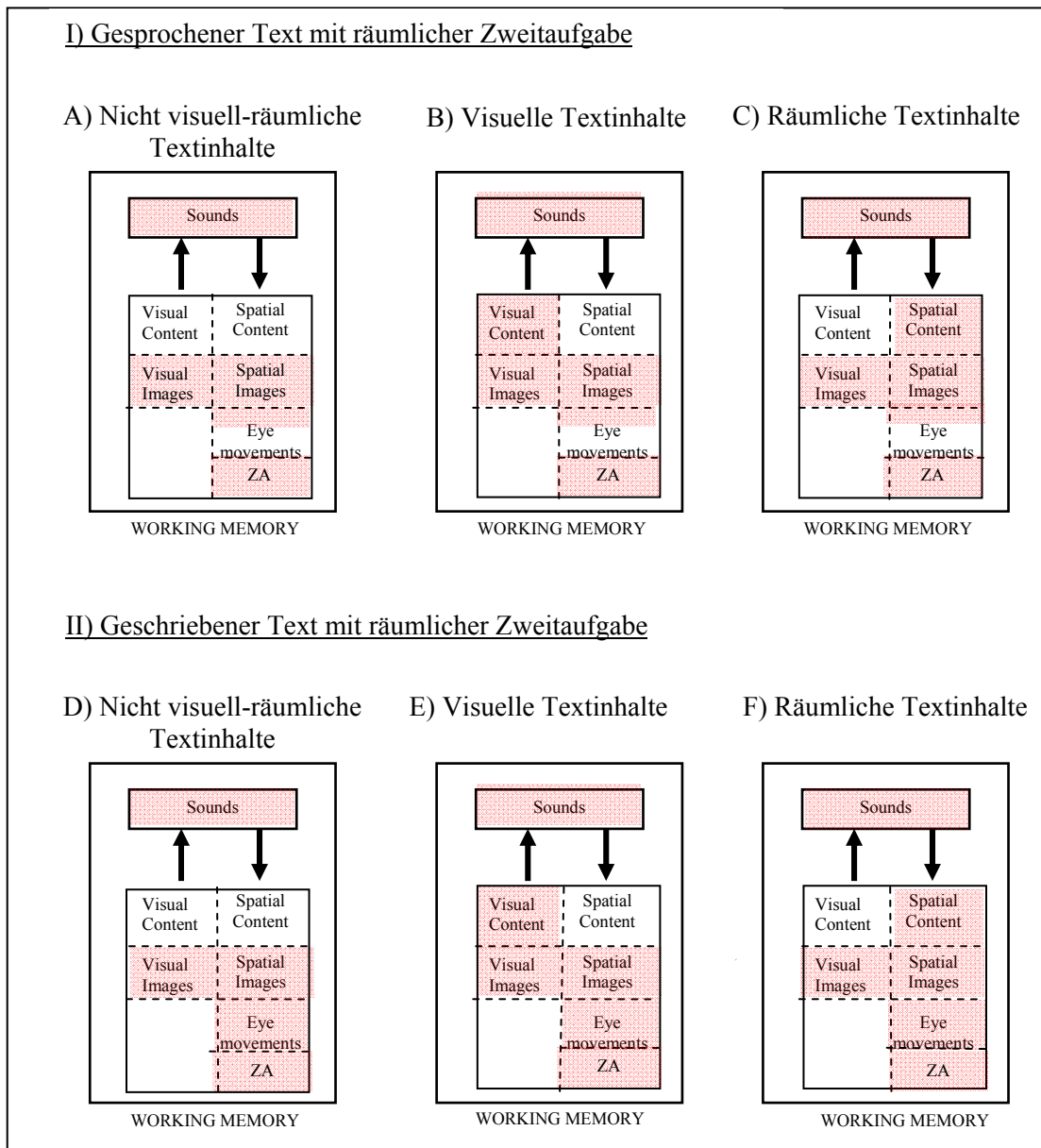


Abbildung 20. Die angenommene Belastung des Arbeitsgedächtnisses bei der Darbietung verschiedener Textmodalitäten und Textinhalte zusammen mit einer räumlichen Zweitaufgabe (ZA). Färbungen indizieren Beanspruchung der Komponenten.

Die **Textmodalität-Hypothese** machte die folgenden Annahmen:

Unter Gültigkeit der CTML wird kongruent zum ersten Experiment ein genereller Haupteffekt der Textmodalität für alle AVn erwartet, sowie eine Moderation des Modalitätseffekts durch die Kapazität des vVSSP und rVSSP. Durch die Ausführung der räumlichen Zweitaufgabe wird der rVSSP stärker

belastet, was zu Interferenzen mit der Verarbeitung geschriebenen Texts und der Bildverarbeitung führen sollte (vgl. Abbildung 21). Statistisch sollte sich diese Interferenz für die Texterinnerung in einer Interaktion zwischen Textmodalität und Zweitaufgabe ausdrücken, da die Leistung bei Lernern mit geschriebenen Texten absinkt, während sie bei Lernern mit gesprochenen Texten nicht absinkt. Hinsichtlich der Bilderinnerung wird hingegen keine spezifische Interaktion der Zweitaufgabe mit der Textmodalität erwartet, da die Zweitaufgabe die Bilderinnerung sowohl bei gesprochener als auch geschriebener Textdarbietung vergleichbar senkt (vgl. Zweitaufgaben-Hypothese).

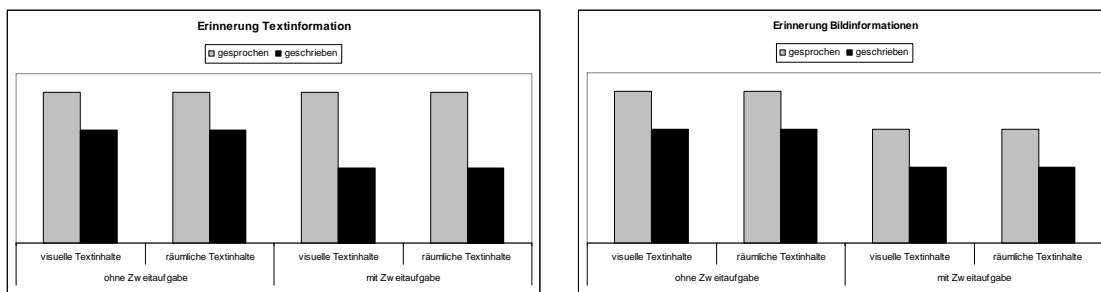


Abbildung 21. Darstellung der erwarteten Ergebnismuster für die Textmodalität-Hypothese nach der CTML in Abhängigkeit verschiedener AVn.

Unter Gültigkeit der ECTML wird der Modalitätseffekt nur erwartet, wenn der rVSSP stark belastet ist und räumliche bzw. bildhafte Informationen erinnert werden müssen. Bei Ausführung einer räumlichen Zweitaufgabe wird der rVSSP zusätzlich belastet, was zu den folgenden Erweiterungen der Hypothese führt: Bezüglich der Erinnerung an abstrakte Textinhalte wird weiterhin kein Modalitätseffekt erwartet, da die zu erinnernden Inhalte nicht im rVSSP verarbeitet werden. Bezüglich der Erinnerung an bildbezogene visuelle und räumliche Inhalt sowie die Charakteristika wird ein Modalitätseffekt für Lerner mit räumlichen Textinhalten erwartet, da die Belastung des rVSSP durch die Kontrolle der Blickbewegungen bei geschriebenen Texten höher ist als bei gesprochenen Texten, was zu einer *Overload*-Situation führt. Dieser Modalitätseffekt sollte größer ausfallen als in der Gruppe, die keine Zweitaufgabe durchführt, da die Zweitaufgabe die Belastung im rVSSP

spezifisch erhöht und somit die *Overload*-Situation verschärft. Für Lerner mit visuellen Textinhalten wird weiterhin kein Modalitätseffekt erwartet.

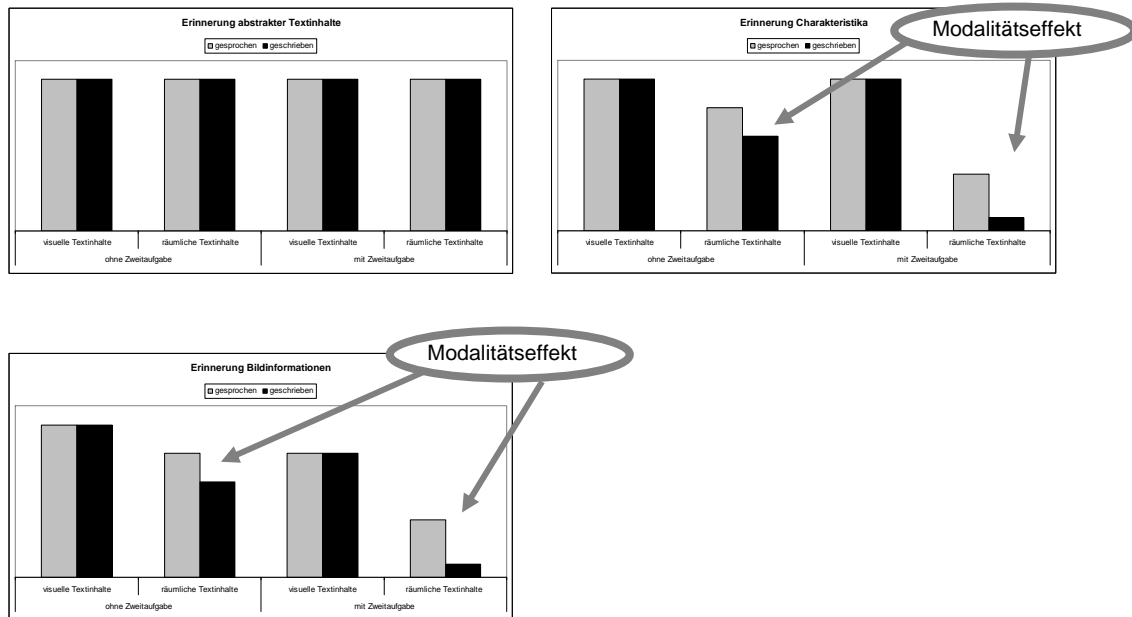


Abbildung 22. Darstellung der erwarteten Ergebnismuster für die Textmodalität-Hypothese nach der ECTML in Abhängigkeit verschiedener AVn.

Statistisch sollte sich diese Hypothese bezüglich der Erinnerung an abstrakte Textinhalte darin ausdrücken, dass weder Haupteffekte noch Interaktionen beobachtet werden (vgl. Abbildung 22, oben links). Bezüglich der Erinnerung an die Bildinformationen und Charakteristika wird eine Dreifachinteraktion erwartet, die darauf zurückgeht, dass es sowohl in der Gruppe ohne Zweitaufgabe als auch mit Zweitaufgabe nur einen Modalitätseffekt für Lerner mit räumlichen Textinhalten gibt, dass dieser Modalitätseffekt aber in der Gruppe ohne Zweitaufgabe im Vergleich zu Lernern mit Zweitaufgabe geringer ausfällt (vgl. Abbildung 22, oben rechts und unten links).

Die *Textinhalt-Hypothese* machte die folgenden Annahmen:

Unter Gültigkeit der CTML wird weiterhin kein Effekt des Textinhalts erwartet.

Unter Gültigkeit der ECTML wird angenommen, dass es bei zusätzlicher Ausführung einer räumlichen Zweitaufgabe weiterhin keine Unterschiede

zwischen Lernern mit visuellen und räumlichen Textinhalten bezüglich der Erinnerung an abstrakte Textinhalte gibt, während es bei der Erinnerung an Charakteristika sowie an räumliche und visuelle Bildinhalte zu Unterschieden kommt: Bezüglich der Erinnerung an Charakteristika wird eine spezifische Interferenz zwischen räumlichen Textinhalten und räumlicher Zweitaufgabe erwartet, da beide den rVSSP belasten. Die Leistung der Lerner mit räumlichen Textinhalten sollte daher im Vergleich zur Bedingung ohne Zweitaufgabe absinken, während dies bei Lernern mit visuellen Textinhalten nicht der Fall ist, da hier die Zweitaufgabe und die Textinhalte in zwei verschiedenen Komponenten des VSSP verarbeitet werden. Bezüglich der Erinnerung an die Bildinformationen wird ein Absinken der Leistung bei Lernern sowohl mit visuellen als auch räumlichen Textinhalten erwartet, da in beiden Gruppen die räumliche Zweitaufgabe mit der Bildverarbeitung interferiert (vgl. Zweitaufgabenhypothese). Das Absinken der Leistung bei Ausführung einer Zweitaufgabe sollte innerhalb der Gruppe mit räumlichen Textinhalten aufgrund der erhöhten Belastung des rVSSP jedoch stärker sein.

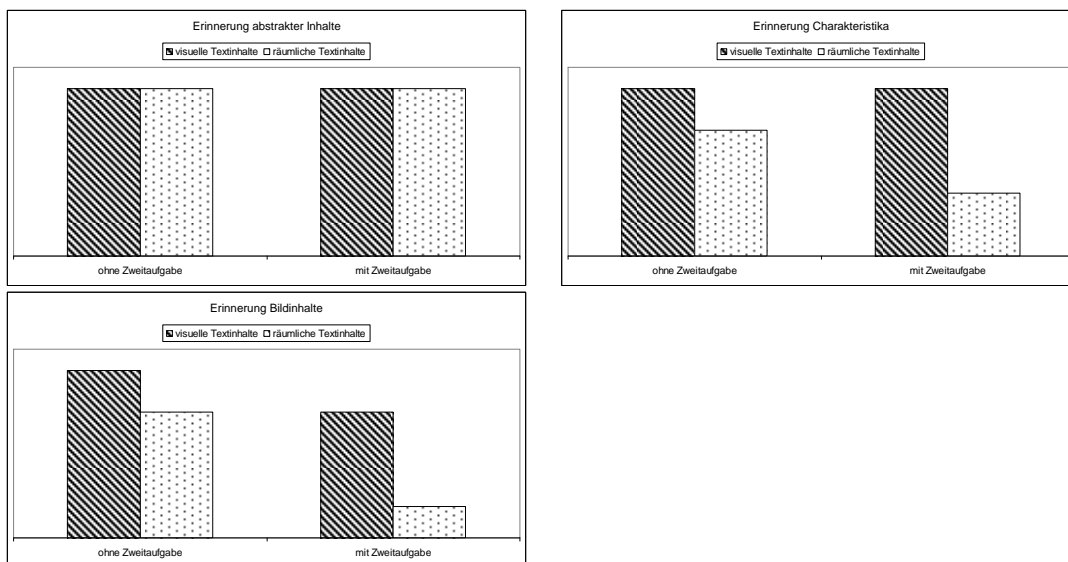


Abbildung 23. Darstellung des erwarteten Ergebnismusters für die Textinhalt-Hypothese.

Statistisch sollte sich diese Hypothese bezüglich der Erinnerung an abstrakte Textinhalte darin ausdrücken, dass weder Haupteffekte noch Interaktionen beobachtet werden (vgl. Abbildung 23, links oben). Bezüglich der Erinnerung an

visuelle vs. räumliche Text- und Bildinhalte wird ein Haupteffekt des Textinhalts erwartet, da Lerner mit räumlichen Textinhalten generell schlechter abschneiden sollten als Lerner mit visuellen Textinhalten. Zudem wird eine Interaktion erwartet, die darauf zurückgeht, dass die Leistungsunterschiede innerhalb der Gruppe mit Zweitaufgabe ausgeprägter sind als in der Gruppe ohne Zweitaufgabe (vgl. Abbildung 23 rechts oben).

Die *Arbeitsgedächtnis-Hypothese* ging weiterhin davon aus, dass hohe Kapazitäten der Komponenten des Arbeitsgedächtnisses zu besserer Leistung führen sollten. Statistisch sollte sich ein Einfluss der Arbeitsgedächtniskapazität in einem Haupteffekt für das vVSSP-Maß, das rVSSP-Maß und das PL-Maß ausdrücken.

Im Hinblick auf die Effekte der Zweitaufgabe wurde die *Zweitaufgaben-Hypothese* formuliert. Sowohl nach der CTML als auch der ECTML ist davon auszugehen, dass die Zweitaufgabe mit der Bildverarbeitung interferieren sollte. Daher wurde nach beiden Theorien ein Haupteffekt der Zweitaufgabe bei der Bilderinnerung erwartet, nicht aber bei der Texterinnerung, was auch im Einklang mit den in Abschnitt 2.3.2 zitierten Studien ist (z. B. Gyselinck et al., 2002; Nam & Pujari; 2005; Brunyé et al. 2006). Zudem wurde angenommen, dass die Höhe der rVSSP Kapazität das Ausmaß des Effekts der räumlichen Zweitaufgabe moderieren kann, da bei hoher rVSSP Kapazität die zusätzliche Belastung durch eine räumliche Zweitaufgabe eher ausgeglichen werden kann.

10.2 Methode

10.2.1 Versuchsdesign

Im vorliegenden Abschnitt wird das der empirischen Untersuchung zugrunde liegende Versuchsdesign vorgestellt sowie ein kurzer Überblick über die Operationalisierung der UVn und AVn gegeben.

Versuchsdesign und Variablen

Die in Abschnitt 10.1 aufgestellten psychologischen Hypothesen wurden durch experimentelle Variationen in der *Lernphase* der Untersuchung (vgl. Abschnitt 7) mit einem dreifaktoriellen, gemischten Versuchsplan getestet. Erste UV war die

Textmodalität (gesprochen vs. geschrieben), zweite UV waren die Textinhalte (visuell vs. räumlich) und dritte UV war die Zweitaufgabe (vorhanden vs. nicht vorhanden). Im Unterschied zum ersten Experiment wurde die Textmodalität intraindividuell variiert. Bei den Textinhalten und der Zweitaufgabe handelte es sich um interindividuell variierte Faktoren. Durch Kreuzung der Faktoren ergab sich das in Tabelle 26 dargestellte $2 \times 2 \times 2$ Design. Die Reihenfolge der Darbietung der Textmodalität wurde kontrolliert, indem entweder zunächst drei Lernobjekte mit gesprochenem Text und dann drei Lernobjekte mit geschriebenem Text dargeboten wurden oder vice versa. Die Reihenfolge der Lernobjekte wurde dabei partiell ausbalanciert. Als Lernobjekte dienten dabei wie in Experiment 1 die verschiedenen Fischarten.

Tabelle 26

Darstellung des dreifaktoriellen Versuchsplans mit den interindividuell variierten Faktoren Textinhalt und Zweitaufgabe und dem intraindividuell variierten Faktor Textmodalität

		Zweitaufgabe	
		nicht vorhanden	vorhanden
Textinhalt	visuell	gesprochen / geschrieben geschrieben / gesprochen	gesprochen / geschrieben geschrieben / gesprochen
	räumlich	gesprochen / geschrieben geschrieben / gesprochen	gesprochen / geschrieben geschrieben / gesprochen

Pro Zelle wurden $n = 18$ Personen getestet, insgesamt belief sich die Stichprobe also auf $N = 72$.

Die Operationalisierung der unabhängigen Variablen

UVn waren die Textinhalte, die Textmodalität, die Zweitaufgabe sowie die Maße der Arbeitsgedächtniskapazität. Die Operationalisierung der UVn wurde bereits in Abschnitt 8.1 beschrieben und wird daher an dieser Stelle nicht weiter ausgeführt.

Die Operationalisierung der abhängigen Variablen

Als AVn wurden die Erinnerungsleistung mit Hilfe der Verifikationsitems und der offenen Fragen, die Bewertung der Lernphase sowie bei Lernern mit Zweitaufgabe die Leistung in der Zweitaufgabe gemessen. Aufgrund des intraindividuellen Designs

waren einige Veränderungen im Gegensatz zum ersten Experiment notwendig, die im Folgenden beschrieben werden. Außerdem wird die Operationalisierung der Leistungsmessung bezüglich der Zweitaufgabe beschrieben.

Bewertung der Lernphase. Die Bewertung der Lernphase erfolgte im zweiten Experiment mit nur drei Items: es wurde die in die Lernphase investierte Mühe („Wie viel Mühe hast Du in die Lernphase investiert?“) und die allgemeine Konzentration während der Lernphase („Wie sehr hast Du Dich während der Lernphase konzentriert“) erfasst. Die Beantwortung und Darbietung der Items am Computer erfolgte äquivalent zum ersten Experiment. Aufgrund des intraindividuell variierten Faktors Textmodalität wurden die Items aber zwei Mal dargeboten, nämlich nach Darbietung der Lerninhalte mit gesprochenem Text und nach Darbietung der Lerninhalte mit geschriebenem Text.

Leistung in der Zweitaufgabe. Zur Messung der Leistung in der Zweitaufgabe wurde zum Einen die Fehlerrate in der Tippsequenz als Leistungsmaß erhoben, zum Anderen die durchschnittliche Tippgeschwindigkeit, jeweils in Abhängigkeit der Textmodalität. Die Tippsequenzen sowie die Tippgeschwindigkeit wurden mit Hilfe eines externen Computers, an den der Nummernblock angeschlossen war, aufgezeichnet. Zur Aufzeichnung wurde das Programm Medialab v2008 (Empirisoft, New York) genutzt. Die Berechnung der Fehlerrate und der Tippgeschwindigkeit erfolgte sowohl für die Darbietung mit gesprochenem als auch für die Darbietung mit geschriebenem Text. Dabei wurde die Fehlerrate errechnet, indem die Fehlerzahl an der Gesamtmenge der Tastendrucke relativiert wurde. Als Fehler wurden zum Einen Auslasser in der Sequenz, zum Anderen doppelte Tastendrucke gewertet. Wurden mehrere Tasten übersprungen, aber die Sequenz dann korrekt fortgeführt, so wurde dieses nur als ein Fehler bewertet. Die durchschnittliche Tippgeschwindigkeit wurde berechnet, indem die Zeiten aller Tastendrucke aufsummiert wurden und durch die Anzahl aller Tastendrucke dividiert wurden. Vor Beginn der Lernphase wurden zudem die Fehlerrate und die Tippgeschwindigkeit ohne zusätzliche Belastung erhoben (*Baseline*). Dazu tippten die Probanden zwei Minuten lang das vorgegebene Muster auf dem Nummernblock, ohne dass sie eine Hauptaufgabe ausführen mussten.

Die Operationalisierung der personenbezogenen Daten

Die personenbezogenen Daten Alter, Geschlecht, Studienfach, Semesteranzahl und Vorwissen wurden, wie in Abschnitt 8.3 beschrieben, operationalisiert.

10.2.2 Versuchspersonen

Insgesamt nahmen 74 Studenten der Universität Tübingen (12 männliche und 62 weibliche) mit einem Durchschnittsalter von 21.89 Jahren ($SD = 3.08$) an der Untersuchung teil. Als Aufwandsentschädigung für die Teilnahme erhielten sie entweder 10 Euro oder 1.5 Versuchspersonenstunden. Ein Ausschlusskriterium für die Teilnahme an Experiment 2 war eine Teilnahme an Experiment 1, da es sich um dieselben Inhalte handelte. Die Zuteilung der Versuchspersonen auf die vier Versuchsgruppen erfolgte randomisiert.

10.2.3 Geräte und Hilfsmittel

Die Lernumgebung wurde auf einem externen 19 Zoll Monitor dargeboten. Zur Steuerung der Lernumgebung und zur Beantwortung der computerbasierten Items stand den Probanden im Gegensatz zu Experiment 1 nicht die Mouse, sondern eine externe Tastatur zur Verfügung. Auf der Tastatur waren die Tasten zur Beantwortung der Items mit „R“ (für richtig) und „F“ (für falsch) mit einem Aufkleber gekennzeichnet.

Auf einem zweiten Notebook wurden die Daten der Zweitaufgabe aufgezeichnet sowie im späteren Verlauf der Untersuchung der *Digit Span* Test dargeboten.

10.2.4 Versuchsablauf

Der Versuch wurde in einem Zeitraum von acht Wochen zwischen 9 und 19 Uhr als Einzelsitzung durchgeführt. Der Versuchsablauf ähnelte stark dem in Abschnitt 9.1.4 beschriebenen Ablauf.

Zunächst bearbeiteten die Versuchspersonen die computerbasierte Lernumgebung. Sie entsprach, bis auf einige kleine Veränderungen, der in Abschnitt 7 beschriebenen Lernumgebung. Die erwähnten Veränderungen betrafen dabei vor

allem die Bedingungen mit Zweitaufgabe. Hier übten die Lerner nach dem Erlernen der Begrifflichkeiten zunächst die Durchführung der Zweitaufgabe zusammen mit der Versuchsleiterin. Anschließend erfolgte die Erhebung der *Baseline* über eine Zeitspanne von zwei Minuten.

Eine weitere Veränderung betraf die Instruktion vor der Lernphase, die um einige wenige Aspekte erweitert wurde (vgl. Anhang G). In der Lernphase wurden zunächst drei Lernobjekte zusammen mit gesprochenem oder geschriebenem Text präsentiert, anschließend folgten auf jeweils einer Seite die zwei Items zur Bewertung der Lernumgebung. Dann wurden die anderen drei Lernobjekte zusammen mit der jeweils anderen Textmodalität und im Anschluss erneut die zwei Items zur Bewertung der Lernumgebung dargeboten. In der Testphase wurden zunächst die Verifikationsitems präsentiert. Im Vergleich zum ersten Experiment wurden zur Messung der bildbezogenen Erinnerung 36 statt 18 Items verwendet, was aufgrund des intraindividuellen Designs notwendig war. Im Anschluss beantworteten die Versuchspersonen die vier offenen, *paper-pencil* basierten Items. Für die schriftlichen Fragen standen den Probanden 3 Minuten, für die Zeichenaufgaben 3.5 Minuten zur Verfügung. Nach Beantwortung der offenen Fragen wurden der Corsi Block Test, der VPT und der *Digit Span* Test durchgeführt. Danach war die Untersuchung beendet und die Probanden erhielten die Aufwandsentschädigung. Insgesamt dauerte die Untersuchung circa 90 Minuten.

10.3 Ergebnisse

Die Überprüfung der Hypothesen des zweiten Experiments erfolgte mit Hilfe (multivariater) Varianzanalysen mit den Zwischensubjektfaktoren Textinhalt und Zweitaufgabe, dem Innersubjektfaktor Textmodalität und den kontinuierlichen, z-standardisierten Testwerten im VPT, Corsi Block und *Digit Span* Test. Die *Digit Span* ging als einfache Kovariate in die Analyse ein, während hinsichtlich des VPT und Corsi Block Maß auch theoretisch ableitbare Interaktionen mit den experimentellen Bedingungen berücksichtigt wurden (vgl. Abschnitt 6 und Abschnitt 10.1) In den (M)ANCOVAs wurden jeweils solche AVn zusammengefasst, für die nach der ECTML oder CTML die gleichen Einflüsse der experimentellen Variation erwartet wurden. Aufgrund der intraindividuellen Variation der Textmodalität war es

nicht möglich hinsichtlich der Textmodalität ausschließlich theoriegeleitete Interaktionen in die Analyse aufzunehmen. Diese für die vorliegende Untersuchung unerheblichen Interaktionen werden im Folgenden nicht berichtet.

10.3.1 Reliabilitätsanalysen

Die Reliabilitäten für die einzelnen Indizes vor und nach Itemausschluss sind in Tabelle 27 dargestellt. Dabei war es aufgrund des intraindividuellen Designs des Experiments notwendig, eine Trennung der Items vorzunehmen: In der Lernphase wurden jeweils zwei Blöcke aus jeweils drei Fischen (Block 1: Gründelfisch, Lotserfisch, Ferdefisch, Block 2: Deuterfisch, Purtzerfisch, Zierterpampel) dargeboten. Die Textmodalität der Blöcke wechselte dabei während der Lernphase, d. h. Block 1 wurde mit gesprochenem Text und Block 2 mit geschriebenem Text dargeboten oder vice versa. Daher mussten die Indizes zunächst nach diesen Blöcken gebildet werden. Anschließend erfolgte dann eine Zuweisung der jeweiligen Textmodalität zu den Indizes, um varianzanalytische Analysen mit Messwiederholung durchführen zu können.

Wie aus Tabelle 27 ersichtlich wird, variierten die Reliabilitäten auch nach Itemausschluss noch stark. Ein nach Field (2009) akzeptables Cronbach's $\alpha = .70$ wurde dabei auf fast keiner der Skalen erreicht.

Die Interrater-Reliabilität für die offenen Fragen betrug bei schriftlich zu beantwortenden Fragen .79 (Cohen's Kappa), bei zeichnerisch zu beantwortenden Fragen .73 (Cohen's Kappa). Somit waren die Reliabilitäten in einem akzeptablen Bereich (vgl. z. B. van Gog, Paas, van Merriënboer & Witte, 2005). Nichtübereinstimmungen zwischen den beiden Ratern wurden durch nachträglichen Konsensus gelöst.

Zur Berechnung der Reliabilitäten der einzelnen Indizes mussten erneut die Reliabilitäten pro offene Frage, d. h. pro Fisch, berechnet werden, da die Textmodalität in der Lernphase wechselte. Nach der Berechnung der Reliabilitäten sowie der Bildung der Indizes erfolgte dann eine Zuweisung der jeweiligen Textmodalität zu den Indizes, um varianzanalytische Analysen mit Messwiederholung durchführen zu können. Die Reliabilitäten der Indizes pro Aufgabe sowie die Zuwächse in Cronbach's α durch Itemausschluss sind in Tabelle 28 dargestellt.

Tabelle 27

Cronbach's α der einzelnen Indizes der Verifikationsitems vor und nach Itemausschluss

	Cronbach's α vor Itemausschluss					Cronbach's α nach Itemausschluss						
	Items GLF		Items DPZ			Items GLF		Items DPZ				
	gleicher Index	ungleicher Index		gleicher Index	ungleicher Index		gleicher Index	ungleicher Index		gleicher Index	ungleicher Index	
	visuell	räumlich		visuell	räumlich		visuell	räumlich		visuell	räumlich	
<i>bildbezogene Erinnerungsleistung</i>												
visuelle Bildinhalte	$\alpha = .09$ $i = 6$	$\alpha = .67$ $i = 12$		$\alpha = .30$ $i = 6$	$\alpha = .61$ $i = 12$		$\alpha = .56$ $i = 4$	$\alpha = .72$ $i = 11$		$\alpha = .62$ $i = 4$	$\alpha = .65$ $i = 11$	
räumliche Bildinhalte	$\alpha = .41$ $i = 12$	$\alpha = .45$ $i = 6$		$\alpha = .20$ $i = 12$	$\alpha = -.30$ $i = 6$		$\alpha = .46$ $i = 11$	$\alpha = .53$ $i = 5$		$\alpha = .41$ $i = 9$	$\alpha = .19$ $i = 3$	
<i>textbezogene Erinnerungsleistung</i>												
Charakteristika	$\alpha = .63$ $i = 15$		$\alpha = .70$ $i = 15$			$\alpha = .65$ $i = 14$			$\alpha = .74$ $i = 14$			
abstrakte Textinhalte	$\alpha = .13$ $i = 3$		$\alpha = .28$ $i = 3$			$\alpha = .23$ $i = 2$			$\alpha = .44$ $i = 2$			

Anmerkung. i = Anzahl der Items. Items GLF = Items bezüglich des Gründel-, Lotser- und Ferdefisches. Items DPZ = Items bezüglich des Deuter-, Putzer- und Zieterfisches

Tabelle 28

Cronbach's α der einzelnen Indizes der offenen Fragen vor und nach Itemausschluss

	Cronbach's α vor Itemausschluss				Cronbach's α nach Itemausschluss				
	gleicher Index	ungleicher Index		gleicher Index	ungleicher Index		gleicher Index	ungleicher Index	
		visuell	räumlich		visuell	räumlich		visuell	räumlich
<i>bildbezogene Erinnerungsleistung^a</i>									
visuelle Bildinhalte	$\alpha = .58^c$ $i = 15$	$\alpha = .62^c$ $i = 18$	$\alpha = .56^d$ $i = 13$	$\alpha = .53^d$ $i = 20$	$\alpha = .59^c$ $i = 14$	$\alpha = .64^c$ $i = 17$	$\alpha = .58^d$ $i = 12$	$\alpha = .56^d$ $i = 19$	
räumliche Bildinhalte	$\alpha = .60^c$ $i = 11$	$\alpha = .52^c$ $i = 6$	$\alpha = .59^d$ $i = 15$	$\alpha = .50^d$ $i = 9$	$\alpha = .61^c$ $i = 10$	$\alpha = .58^c$ $i = 5$	$\alpha = .67^d$ $i = 13$	$\alpha = .52^d$ $i = 8$	
<i>textbezogene Erinnerungsleistung^b</i>									
Charakteristika	$\alpha = .53^e$ $i = 5$	$\alpha = .49^e$ $i = 5$	$\alpha = .34^f$ $i = 7$	$\alpha = .64^f$ $i = 7$	$\alpha = .60^e$ $i = 4$	$\alpha = .52^e$ $i = 4$	$\alpha = .46^f$ $i = 5$	$\alpha = .66^f$ $i = 6$	
abstrakte Textinhalte	$\alpha = .64^c$ $i = 10$		$\alpha = .60^f$ $i = 13$		$\alpha = .65^e$ $i = 9$		$\alpha = .65^f$ $i = 12$		

Anmerkung. i = Anzahl der Items. ^aDie Erfassung erfolgte durch Zeichnen. ^bDie Erfassung erfolgte durch eine schriftliche Aufgabe. ^cFerdefisch, ^dDeuterfisch, ^eLotserfisch, ^fPurtzerfisch

10.3.2 Datenexploration und Überprüfung inferenzstatistischer Analysevoraussetzungen

Aufgrund technischer Schwierigkeiten bei der Aufzeichnung gab es für drei AVn Fälle von *Missing Data* (vgl. Tabelle 29). Die Daten wurden auf Ausreißer, Normalverteilung mit Hilfe des Kolmogorov-Smirnov Tests (vgl. Anhang H) und Varianzhomogenität der einzelnen Variablen mit Hilfe des Levene-Tests (vgl. Anhang I) geprüft. Für einige AVn wurden Abweichungen von der Normalverteilung beobachtet (vgl. Tabelle 29). Bezüglich der Fehlerrate bei der Durchführung der Zweitaufgabe konnte eine teilweise Annäherung an Normalverteilung durch Logarithmieren und Erweiterung um die Konstante 1 erreicht werden. Die Prüfung der Homogenität der Kovarianzmatrizen konnte vernachlässigt werden, da vergleichbare Gruppengrößen vorlagen.

Tabelle 29

Variablen mit Missing Data, Abweichungen von der Normalverteilung sowie ungleichen Fehlervarianzen

<i>Missing Data</i>	<i>Personenbezogene Daten: Biologienote (3 Missing Data) Zweitaufgabe: Fehlerrate, Tippsequenz (jeweils 3 Missing Data)</i>
Abweichungen von Normalverteilung	<i>Kontrollvariablen: Corsi Block Test, Digit Span Test Erinnerungsleistung: Verifikationsitems: abstrakte Textinhalte, räumlicher Bildinhalte Offene Fragen: Charakteristika, abstrakter Textinhalte, räumlicher Bildinhalte Leistung in der Zweitaufgabe: Fehlerrate, Tippsequenz Bewertungsitems: Alle Variablen</i>

Die *Digit Span* ging als Kovariate in die Analysen ein. Die Prüfung der Homogenität der Regressionsgeraden der *Digit Span* innerhalb der vier Bedingungen kann Anhang J entnommen werden. Es lagen keine Verletzungen vor.

10.3.3 Beschreibung der Stichprobe

Nach der Datenexploration wurden die demographischen Daten ausgewertet. Tabelle 30 gibt einen Überblick über die erhobenen intervallskalierten Variablen, Tabelle 31 über die erhobenen nominalen Daten jeweils in Abhängigkeit des Textinhalts und der Zweitaufgabe.

Tabelle 30

Mittelwerte und Standardabweichungen (in Klammern) der intervallskalierten personenbezogenen Daten in Abhängigkeit des Textinhalts und der Zweitaufgabe (ZA)

Textinhalt Zweitaufgabe	visuell		räumlich	
	ohne ZA n = 18	mit ZA n = 19	ohne ZA n = 18	mit ZA n = 19
Alter (in Jahren)	21.94 (3.23)	21.21 (2.39)	22.67 (3.66)	21.79 (3.03)
Semesteranzahl	2.39 (2.85)	1.74 (1.20)	3.44 (3.29)	2.68 (2.75)
Letzte Bionote (in Punkten, Maximum = 15)	11.29 (1.93)	11.65 (1.80)	11.50 ^a (1.62)	10.84 ^a (1.64)
Vorkenntnisse bzgl. Fischen (Maximum = 4)	1.50 (0.51)	1.42 (0.51)	1.56 (0.51)	1.56 (0.50)
Corsi Block Test	5.22 (0.73)	5.63 (0.90)	5.56 (0.62)	5.89 (0.57)
<i>Visual Pattern Test</i>	8.31 (1.72)	8.67 (1.67)	8.39 (1.30)	8.43 (1.22)
<i>Digit Span Test</i>	6.61 (1.04)	6.42 (0.84)	6.11 (0.96)	6.37 (0.83)

Anmerkung. ^an = 17

Tabelle 31

Die Verteilung der nominalen personenbezogenen Daten in Abhängigkeit des Textinhalts und der Zweitaufgabe (ZA)

Textinhalt Zweitaufgabe	visuell		räumlich	
	ohne ZA n = 18	mit ZA n = 19	ohne ZA n = 18	mit ZA n = 19
<i>Geschlecht</i>				
männlich	5	2	2	3
weiblich	13	17	16	16
<i>Biologie 4-stündig/Leistungskurs</i>				
Ja	3	7	6	5
Nein	15	12	12	14
<i>Studienfach</i>				
Naturwissenschaftlich	0	0	0	1
Anderes	18	19	18	18

Die Analyse auf Gleichheit der Eingangsbedingungen zeigte für keine der erhobenen personenbezogenen Variablen kritische Unterschiede oder Ungleichverteilungen zwischen den Gruppen (vgl. Anhang K).

In Tabelle 32 sind die Mittelwerte und Standardabweichungen der *Baseline* der Zweitaufgabe für die prozentuale Fehlerrate und die durchschnittliche Tippgeschwindigkeit in Abhängigkeit des Textinhalts dargestellt.

Tabelle 32

Mittelwerte und Standardabweichungen (in Klammern) der intervallskalierten Kennwerte der Baseline in Abhängigkeit des Textinhalts

Textinhalt	visuell <i>n</i> = 18	räumlich <i>n</i> = 16
Baseline Fehlerrate (%)	3.82 (11.29)	6.30 (14.48)
Baseline Tippgeschwindigkeit (ms)	986.25 (272.15)	928.11 (316.07)

Zusammenfassend lässt sich die Stichprobe somit wie folgt beschreiben: An der Untersuchung nahmen 74 Personen teil, davon 62 Frauen und 12 Männer. Im Durchschnitt waren die Teilnehmer $M = 21.89$ Jahre ($SD = 3.08$) alt und hatten 2.55 Semester ($SD = 2.65$) studiert. Die letzte Biologienote betrug durchschnittlich 11.31 Punkte ($SD = 1.74$), was in etwa einem „gut“ entspricht. 21 Teilnehmer hatten in der Schule Biologie als Vertiefungsfach oder als Leistungskurs belegt, aber nur ein Teilnehmer studierte aktuell ein naturwissenschaftliches Fach. Das Vorwissen bezüglich Fischen betrug auf einer Skala von eins (keine Vorkenntnisse) bis vier (hohe Vorkenntnisse) durchschnittlich $M = 1.51$ ($SD = 0.50$) und war somit als niedrig einzuschätzen. Die Corsi Block Spanne lag im Mittel bei 5.58 Blöcken ($SD = 0.74$), die VPT Spanne bei $M = 8.46$ ($SD = 1.47$) und die *Digit Span* bei 6.38 Zahlen pro Sequenz ($SD = 0.92$). Während die Werte hinsichtlich des Corsi Block Tests und der *Digit Span* vergleichbar mit den Werten in Experiment 1 waren, lag der durchschnittliche Wert im VPT über dem in Experiment 1 erreichten Wert von $M = 7.38$. Damit näherte sich in Experiment 2 der Durchschnittswert im VPT den von Della Sala et al. (1997) berichteten Werten an. Die Fehlerrate bei der Zweitaufgabe betrug im Mittel 4.99% ($SD = 12.75$), die durchschnittliche Tippgeschwindigkeit lag bei 958.89 Millisekunden pro Tastendruck ($SD = 272.15$).

10.3.4 Deskriptive Daten

Im Folgenden werden die deskriptiven Daten der einzelnen AVn dargestellt. Dabei handelt es sich erneut um adjustierte Daten, da der Einfluss der individuellen Arbeitsgedächtniskapazität (PL, vVSSP und rVSSP) herausgerechnet wurde. Die adjustierten Mittelwerte und Standardfehler für die Verifikationsitems sind in Tabelle 33 dargestellt, die adjustierten Mittelwerte und Standardfehler für die offenen Fragen in Tabelle 34, die adjustierten Mittelwerte und Standardfehler für Leistung in der Zweitaufgabe in Tabelle 35 und die adjustierten Mittelwerte und Standardfehler für die Bewertungsitems in Tabelle 36.

Tabelle 33

Adjustierte Mittelwerte und Standardfehler (in Klammern) der Verifikationsitems in Abhängigkeit des Textinhalts, der Zweitaufgabe (ZA) und der Textmodalität

Textinhalt Zweitaufgabe Textmodalität	visuell				räumlich			
	ohne ZA		mit ZA		ohne ZA		mit ZA	
	gesp <i>n</i> = 18	gesch <i>n</i> = 18	gesp <i>n</i> = 19	gesch <i>n</i> = 19	gesp <i>n</i> = 18	gesch <i>n</i> = 18	gesp <i>n</i> = 19	gesch <i>n</i> = 19
<i>bildbezogene Erinnerungsleistung</i>								
visuelle Bildinhalte	1.99 (0.24)	1.95 (0.23)	1.62 (0.22)	1.32 (0.20)	1.60 (0.31)	1.62 (0.29)	2.00 (0.35)	1.43 (0.33)
räumliche Bildinhalte	1.49 (0.19)	1.41 (0.17)	1.08 (0.17)	1.21 (0.15)	1.15 (0.24)	1.35 (0.21)	1.22 (0.27)	1.41 (0.24)
<i>textbezogene Erinnerungsleistung</i>								
Charakteristika	2.03 (0.13)	1.97 (0.13)	1.83 (0.11)	1.35 (0.12)	1.46 (0.16)	1.30 (0.17)	1.42 (0.18)	1.21 (0.19)
abstrakte Textinhalte	1.80 (0.31)	2.36 (0.32)	2.44 (0.27)	1.66 (0.28)	1.90 (0.28)	2.32 (0.29)	2.09 (0.31)	2.02 (0.32)

Anmerkung. Minimum = 0, Maximum = 4. Gesp = gesprochen, gesch = geschrieben.

Tabelle 34

Adjustierte Mittelwerte und Standardfehler (in Klammern) der offenen Fragen in Prozent in Abhängigkeit des Textinhalts, der Zweitaufgabe (ZA) und der Textmodalität

Textinhalt	visuell				räumlich			
	ohne ZA		mit ZA		ohne ZA		mit ZA	
Zweitaufgabe	gesp	gesch	gesp	gesch	gesp	gesch	gesp	gesch
Textmodalität	<i>n</i> = 18	<i>n</i> = 18	<i>n</i> = 19	<i>n</i> = 19	<i>n</i> = 18	<i>n</i> = 18	<i>n</i> = 19	<i>n</i> = 19
<i>bildbezogene Erinnerungsleistung</i>								
visuelle Bildinhalte	56.84 (5.30)	52.48 (4.75)	52.08 (4.62)	47.07 (4.14)	33.93 (4.78)	33.07 (4.29)	35.55 (5.34)	25.27 (4.79)
räumliche Bildinhalte	52.22 (5.82)	43.41 (6.54)	38.98 (5.07)	42.54 (5.70)	43.78 (5.25)	57.62 (5.90)	40.17 (5.87)	31.84 (6.60)
<i>textbezogene Erinnerungsleistung</i>								
Charakteristika	34.10 (7.70)	36.08 (7.04)	29.64 (6.71)	29.28 (6.14)	21.48 (6.95)	21.15 (6.36)	11.99 (7.77)	17.70 (7.10)
abstrakte Textinhalte	10.13 (3.75)	21.43 (4.86)	13.94 (3.27)	10.11 (4.24)	14.48 (3.39)	15.56 (4.39)	8.93 (3.78)	16.74 (4.90)

Anmerkung. Minimum = 0, Maximum = 4. Gesp = gesprochen, gesch= geschrieben

Tabelle 35

Adjustierte Mittelwerte und Standardfehler (in Klammern) der Leistung in der Zweitaufgabe in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	gesprochen <i>n</i> = 18	geschrieben <i>n</i> = 18	gesprochen <i>n</i> = 16	geschrieben <i>n</i> = 16
Textmodalität				
Fehlerrate (%)	8.39 (4.48)	8.77 (4.46)	7.62 (4.95)	10.23 (4.93)
Tippgeschwindigkeit (ms)	1092.04 (58.27)	1064.76 (48.91)	837.80 (53.64)	926.08 (63.91)

Tabelle 36

Adjustierte Mittelwerte und Standardfehler (in Klammern) der Bewertungsitems in Abhängigkeit des Textinhalts, der Zweitaufgabe (ZA) und der Textmodalität

Textinhalt Zweitaufgabe	visuell				räumlich			
	ohne ZA		mit ZA		ohne ZA		mit ZA	
Textmodalität	gesp <i>n</i> = 18	gesch <i>n</i> = 18	gesp <i>n</i> = 19	gesch <i>n</i> = 19	gesp <i>n</i> = 18	gesch <i>n</i> = 18	gesp <i>n</i> = 19	gesch <i>n</i> = 19
Wie sehr hast Du Dich während der Lernphase konzentriert?	2.59 (0.38)	3.18 (0.41)	3.17 (0.33)	3.95 (0.35)	2.79 (0.34)	3.58 (0.37)	3.53 (0.38)	5.15 (0.41)
Wie viel Mühe hast Du in die Lernphase investiert?	4.41 (0.37)	4.86 (0.35)	4.99 (0.32)	4.99 (0.30)	5.42 (0.33)	5.26 (0.31)	5.11 (0.37)	5.03 (0.35)

Anmerkung. Minimum = 1, Maximum = 7. Hohe Werte indizieren eine hohe Ausprägung im Sinne der Itemformulierung.

Gesp = gesprochen, gesch = geschrieben.

10.3.5 Baseline der Zweitaufgabe: Vergleichbarkeit zwischen den Gruppen und Einfluss in der Lernphase

Hinsichtlich der Leistung in der Zweitaufgabe wurde überprüft, inwiefern die *Baseline* zwischen den Gruppen vergleichbar war. Zur Überprüfung wurden zwei univariate ANOVAs mit dem Textinhalt als interindividuell variiertem Faktor und der Fehlerrate bzw. der Tippgeschwindigkeit als AVn berechnet. Sowohl für die Fehlerrate als auch für die Tippgeschwindigkeit zeigten sich keine Unterschiede zwischen den Gruppen, beide $F_s < 1^9$. Somit waren die Gruppen hinsichtlich der *Baseline* bei der Ausführung der Zweitaufgabe vergleichbar.

In der späteren Analyse der Leistung in der Zweitaufgabe wurde die *Baseline* als Kovariate aufgenommen. Die Analysen zeigten erwartungsgemäß einen statistisch bedeutsamen Einfluss der Kovariate auf die Ergebnisse (Fehlerrate: $F(1, 26) = 14.43$, $MSE = 1.83$, $p = .001$, $\eta_p^2 = .36$; Tippsequenz: $F(1, 26) = 7.49$, $MSE = 90925.30$, $p = .01$, $\eta_p^2 = .22$). Lerner mit hohen Fehlerraten in der *Baseline* hatten auch hohe Fehlerraten während der Lernphase (gesprochener Text: $pr = .55$, $p = .001$; geschriebener Text: $pr = .44$, $p = .01$), und Lerner mit hohen Tippgeschwindigkeiten in der *Baseline* hatten auch eine hohe Tippgeschwindigkeit während der Lernphase (gesprochener Text: $pr = .52$, $p = .003$; geschriebener Text: $pr = .49$, $p = .01$).

⁹ Eine Absicherung der Analyse mit dem Mann-Whitney Test bestätigte das Ergebnis.

10.3.6 Prüfung der Textmodalität-Hypothese

Nach der CTML wurde ein Einfluss der Textmodalität sowie eine Moderation des Einflusses der Textmodalität durch die Kapazität des vVSSP, gemessen über den VPT, und des rVSSP, gemessen über den Corsi Block Test, für alle AVn erwartet. Des Weiteren wurde für die textbezogene Erinnerungsleistung eine Interaktion zwischen Textmodalität und Zweitaufgabe erwartet. Diese Effekte sollte durch die Kapazität des rVSSP moderiert werden, da die Zweitaufgabe den rVSSP belastet. Bezüglich der bildbezogenen Erinnerungsleistung wurde keine Interaktion zwischen Textmodalität und Zweitaufgabe erwartet. Nach der ECTML wurden bezüglich der Erinnerung an abstrakte Textinhalte keine Effekte erwartet. Hinsichtlich der anderen AVn (räumliche Bildinhalte, visuelle Bildinhalte, Charakteristika) wurde eine Dreifachinteraktion zwischen Textinhalt, Zweitaufgabe und Textmodalität erwartet.

Verifikationsitems

In Tabelle 37 sind die für die Prüfung des Einflusses der Textmodalität relevanten Haupteffekte und Interaktionen dargestellt. Die MANCOVA hinsichtlich der visuellen und räumlichen Bilderinnerung zeigte keine statistisch bedeutsamen Effekte.

Die ANCOVA bezüglich der Erinnerung an Charakteristika zeigte im Einklang mit der CTML, dass Lerner mit gesprochener Textdarbietung ($M = 1.73$, $SE = .06$) Charakteristika besser erinnerten als Lerner mit geschriebener Textdarbietung ($M = 1.50$, $SE = .07$).

Hinsichtlich der abstrakten Inhalte zeigte sich entsprechend den Erwartungen der CTML eine Interaktion zwischen Textmodalität und Zweitaufgabe. Eine Bonferroni-adjustierte Analyse der Interaktion zeigte jedoch keine statistisch bedeutsamen Unterschiede innerhalb der Gruppen, allerdings deuten die absoluten Werte darauf hin, dass geschriebener Text mit der räumlichen Zweitaufgabe tendenziell interferierte ($p = .10$; ohne Zweitaufgabe: $M = 2.24$, $SE = 0.21$; mit Zweitaufgabe: $M = 1.84$, $SE = 0.21$), während es bei Lernern mit gesprochener Textdarbietung keine Interferenz mit der Zweitaufgabe gab ($p = .15$; ohne Zweitaufgabe: $M = 1.84$, $SE = 0.20$; mit Zweitaufgabe: $M = 2.26$, $SE = 0.21$). Das Absinken der Leistung bei geschriebener Textdarbietung mit Zweitaufgabe führte dazu, dass innerhalb der Gruppe mit Zweitaufgabe ein tendenzieller Modalitätseffekt zu beobachten war ($p = .10$), während in der Gruppe ohne Zweitaufgabe tendenziell

ein invertierter Modalitätseffekt auftrat ($p = .10$). Während die Interferenz zwischen geschriebener Textdarbietung und Zweitaufgabe den Annahmen der CTML entsprach, widerspricht der Befund eines tendenziell invertierten Modalitätseffekts, wenn keine Zweitaufgabe ausgeführt wurde, ihren Annahmen.

Tabelle 37

Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Verifikationsitems

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
visuelle Bildinhalte räumliche Bildinhalte	TM	$V = .02, F < 1$
	TM \times VPT	$V = .02, F < 1$
	TM \times Corsi	$V = .02, F < 1$
	TM \times ZA	$V = .01, F < 1$
	TM \times TI \times ZA	$V = .01, F < 1$
	TM \times ZA \times Corsi	$V = .01, F < 1$
	TM \times TI \times ZA \times Corsi	$V = .02, F < 1$
Charakteristika	TM	$F(1, 63) = 7.47, MSE = 0.22, p = .01, \eta_p^2 = .11$
	TM \times VPT	$F < 1$
	TM \times Corsi	$F < 1$
	TM \times ZA	$F(1, 63) = 3.28, MSE = 0.22, p = .08, \eta_p^2 = .05$
	TM \times TI \times ZA	$F < 1$
	TM \times ZA \times Corsi	$F < 1$
	TM \times TI \times ZA \times Corsi	$F < 1$
abstrakte Textinhalte	TM	$F < 1$
	TM \times VPT	$F(1, 63) = 2.04, MSE = 1.37, p = .16, \eta_p^2 = .03$
	TM \times Corsi	$F < 1$
	TM \times ZA	$F(1, 63) = 4.94, MSE = 1.37, p = .03, \eta_p^2 = .07$
	TM \times TI \times ZA	$F(1, 63) = 1.07, MSE = 1.37, p = .31, \eta_p^2 = .02$
	TM \times ZA \times Corsi	$F < 1$
	TM \times TI \times ZA \times Corsi	$F < 1$

Anmerkung. Das Digit Span Maß ging zusätzlich als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Offene Fragen

Die Analysen bezüglich der offenen Fragen (vgl. Tabelle 38) erfolgten äquivalent zu den Analysen der Verifikationsitems. Bei keiner der Analysen zeigten sich statistisch bedeutsame Einflüsse der Textmodalität. Hinsichtlich des von der ECTML vorhergesagten Nulleffekts für die Interaktion zwischen Textmodalität, Textinhalt und Zweitaufgabe bei der Erinnerung an abstrakte Textinhalte wurde das strenge Kriterium von $\alpha = .20$ nicht erreicht.

Tabelle 38

Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die offenen Fragen

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
visuelle Bildinhalte räumliche Bildinhalte	TM	$V = .06, F(2, 62) = 1.82, p = .17$
	TM \times VPT	$V = .03, F < 1$
	TM \times Corsi	$V = .06, F(2, 62) = 1.80, p = .17$
	TM \times ZA	$V = .01, F < 1$
	TM \times TI \times ZA	$V = .07, F(2, 62) = 2.36, p = .10$
	TM \times ZA \times Corsi	$V = .02, F < 1$
	TM \times TI \times ZA \times Corsi	$V = .04, F(2, 62) = 1.17, p = .32$
Charakteristika	TM	$F < 1$
	TM \times VPT	$F < 1$
	TM \times Corsi	$F < 1$
	TM \times ZA	$F < 1$
	TM \times TI \times ZA	$F < 1$
	TM \times ZA \times Corsi	$F < 1$
	TM \times TI \times ZA \times Corsi	$F < 1$
abstrakte Textinhalte	TM	$F(1, 63) = 1.68, MSE = 321.02, p = .20, \eta_p^2 = .03$
	TM \times VPT	$F(1, 63) = 3.77, MSE = 321.02, p = .06, \eta_p^2 = .06$
	TM \times Corsi	$F < 1$
	TM \times ZA	$F < 1$
	TM \times TI \times ZA	$F(1, 63) = 2.92, MSE = 321.02, p = .09, \eta_p^2 = .04$
	TM \times ZA \times Corsi	$F < 1$
	TM \times TI \times ZA \times Corsi	$F < 1$

Anmerkung. Das Digit Span Maß ging zusätzlich als Kovariate ein.

Leistung in der Zweitaufgabe

Für Lerner mit Zweitaufgabe wurde neben den Performanzdaten und der Bewertung der Lernphase auch die Leistung in der Zweitaufgabe analysiert. Die Annahme dabei war, dass sich die Arbeitsgedächtnisprozesse durch ein prozessnahes Maß wie die Zweitaufgabenleistung besser abbilden lassen. Es wurden jedoch keine statistisch bedeutsamen Effekte beobachtet (vgl. Tabelle 39).

Tabelle 39

Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Leistung in der Zweitaufgabe

AVn der ANCOVA	Effekte der ANCOVA	Ergebnisse der ANCOVA
Fehlerrate	TM	$F(1, 26) = 3.30, MSE = 0.94, p = .08, \eta_p^2 = .11$
	TM \times TI	$F < 1$
	TM \times VPT	$F < 1$
	TM \times Corsi	$F < 1$
	TM \times TI \times VPT	$F(1, 26) = 1.16, MSE = 0.94, p = .29, \eta_p^2 = .04$
	TM \times TI \times Corsi	$F < 1$
Tipp- geschwindigkeit	TM	$F(1, 26) = 1.24, MSE = 7067.53, p = .28, \eta_p^2 = .05$
	TM \times TI	$F(1, 26) = 2.46, MSE = 7067.53, p = .13, \eta_p^2 = .09$
	TM \times VPT	$F(1, 26) = 3.52, MSE = 7067.53, p = .07, \eta_p^2 = .12$
	TM \times Corsi	$F(1, 26) = 2.62, MSE = 7067.53, p = .12, \eta_p^2 = .09$
	TM \times TI \times VPT	$F(1, 26) = 2.70, MSE = 7067.53, p = .11, \eta_p^2 = .09$
	TM \times TI \times Corsi	$F(1, 26) = 2.47, MSE = 7067.53, p = .13, \eta_p^2 = .09$

Anmerkung. Das Digit Span Maß ging zusätzlich als Kovariate ein.

Bewertungsitems

Die Analysen der Bewertungsitems (vgl. Tabelle 40) erfolgten mit einer zur Analyse der Performanz äquivalenten MANCOVA.

Tabelle 40

Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Bewertungsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	TM	$V = .25, F(2, 62) = 6.90, p = .049$
	TM \times VPT	$V = .10, F(2, 62) = 2.82, p = .09$
	TM \times Corsi	$V = .02, F < 1$
	TM \times ZA	$V = .03, F < 1$
	TM \times TI \times ZA	$V = .02, F < 1$
	TM \times ZA \times Corsi	$V = .02, F < 1$
	TM \times TI \times ZA \times Corsi	$V = .08, F(2, 62) = 1.64, p = .19$

Anmerkung. Das Digit Span Maß ging zusätzlich als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Die MANCOVA zeigte einen Haupteffekt für die Textmodalität. Die univariaten Analysen bestätigten, dass Lerner mit geschriebener Textdarbietung ($M = 5.44, SE = 0.15$) mehr Mühe in die Lernphase investierten als Lerner mit gesprochener Textdarbietung ($M = 5.13, SE = 0.16$), $F(1, 63) = 5.30, MSE = 86.07, p = .03, \eta_p^2 = .08$. Für die Konzentration während der Lernphase wurde der Einfluss der Textmodalität nicht bestätigt, $F < 1$.

Zusammenfassend kann man somit festhalten, dass die Annahme der CTML einer generellen Überlegenheit gesprochener Textdarbietung für die Erinnerung an Charakteristika gemessen über die Verifikationsitems sowie für die Konzentration während der Lernphase bestätigt wurde. Für alle anderen Variablen wurde keine Überlegenheit der gesprochenen Textdarbietung beobachtet. Für die Erinnerung an abstrakte Textinhalte gemessen über die Verifikationsitems zeigte sich bei Lernen ohne Zweitaufgabe sogar eine tendenzielle Überlegenheit geschriebener Textdarbietung, welche aber durch eine Interferenz zwischen geschriebener Textdarbietung und räumlicher Zweitaufgabe aufgehoben wurde. Die von der ECTML vorhergesagte Dreifachinteraktion zwischen Textmodalität, Textinhalt und Zweitaufgabe wurde durch keine der Analysen bestätigt.

10.3.7 Prüfung der Textinhalt-Hypothese

Nach der CTML wurde für keine der AVn ein Einfluss des Textinhalts oder eine Moderation des Einflusses des Textinhalts erwartet. Die ECTML machte hingegen spezifische Vorhersagen hinsichtlich des Einflusses des Textinhalts in Abhängigkeit der AVn: Hinsichtlich der Erinnerung an bildhafte Informationen (räumliche Bildinhalte, visuelle Bildinhalte) wurde eine Interaktion zwischen Textinhalt und Zweitaufgabe erwartet, zusätzlich aber auch ein Haupteffekt der Zweitaufgabe (vgl. Abschnitt 10.1). Hinsichtlich der Erinnerung an Charakteristika wurde eine Interaktion zwischen Textinhalt und Zweitaufgabe erwartet, die darauf zurückging, dass die Zweitaufgabe mit den räumlichen Textinhalten nicht aber mit den visuellen Textinhalten interferieren sollte. Bezüglich der Erinnerung an abstrakte Textinhalte wurden keine Effekte erwartet, da abstrakte Information nicht im VSSP verarbeitet wird. Es wurde zudem angenommen, dass die Kapazität des rVSSP, gemessen über den Corsi Block, und des vVSSP, gemessen über den VPT, die Erinnerung an Textinhalte beeinflussen kann. Die Prüfung der Hypothesen erfolgte mit Hilfe der beschriebenen (M)ANCOVAs.

Verifikationsitems

In Tabelle 41 sind der für die Prüfung des Einflusses des Textinhalts relevante Haupteffekt sowie die Interaktionen dargestellt.

Tabelle 41

Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Verifikationsitems

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
visuelle Bildinhalte räumliche Bildinhalte	TI	$V = .01, F < 1$
	TI \times ZA	$V = .01, F < 1$
	TI \times Corsi	$V = .02, F < 1$
	TI \times VPT	$V = .004, F < 1$
	TI \times ZA \times Corsi	$V = .01, F < 1$
Charakteristika	TI	$F(1, 63) = 13.20, MSE = 0.33, p = .001, \eta_p^2 = .17$
	TI \times ZA	$F(1, 63) = 1.03, MSE = 0.33, p = .31, \eta_p^2 = .02$
	TI \times Corsi	$F(1, 63) = 1.04, MSE = 0.33, p = .31, \eta_p^2 = .02$
	TI \times VPT	$F(1, 63) = 5.98, MSE = 0.33, p = .02, \eta_p^2 = .08$
	TI \times ZA \times Corsi	$F < 1$
abstrakte Textinhalte	TI	$F < 1$
	TI \times ZA	$F < 1$
	TI \times Corsi	$F < 1$
	TI \times VPT	$F(1, 63) = 2.43, MSE = 1.42, p = .12, \eta_p^2 = .04$
	TI \times ZA \times Corsi	$F < 1$

Anmerkung. Die Textmodalität ging zusätzlich als UVn, das Digit Span Maß als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Hinsichtlich der Erinnerung an bildbezogene Informationen zeigte die MANCOVA keinen statistisch bedeutsamen Einfluss der Textinhalte.

Hinsichtlich der Erinnerung an Charakteristika wurden der Haupteffekt des Textinhalts sowie die Interaktion zwischen Textinhalt und vVSSP Kapazität bedeutsam. Entsprechend den Annahmen der ECTML erinnerten Lerner mit visuellen Textinhalten ($M = 1.80; SE = 0.07$) Charakteristika besser als Lerner mit räumlichen Textinhalten ($M = 1.43, SE = 0.07$). Die partiellen Korrelationen zwischen vVSSP Kapazität und Erinnerungsleistung wurden jedoch weder für Lerner mit visuellen Textinhalten ($pr = -.25, p = .15$) noch für Lerner mit räumlichen Textinhalten ($pr = .28, p = .10$) statistisch bedeutsam, daher ist die Interaktion nicht interpretierbar.

Hinsichtlich der Erinnerung an abstrakte Textinhalte zeigten sich erwartungsgemäß keine statistisch bedeutsamen Unterschiede, allerdings wurde das strenge Kriterium von $\alpha = .20$ für die Interaktion zwischen Textinhalt und Corsi Block Maß nicht erreicht.

Offene Fragen

Die Analysen bezüglich der offenen Fragen (vgl. Tabelle 42) erfolgten äquivalent zu den Analysen der Verifikationsitems.

Im Hinblick auf die bildbezogene Erinnerungsleistung konnte der von der ECTML angenommene Haupteffekt des Textinhalts hinsichtlich der Erinnerung an visuelle Bildinhalte bestätigt werden, $F(1, 60) = 42.94$, $MSE = 312.31$, $p < .001$, $\eta_p^2 = .42$: Lerner mit visuellen Textinhalten ($M = 52.10$, $SE = 2.10$) erinnerten visuelle Bildinhalte besser als Lerner mit räumlichen Textinhalten ($M = 31.31$, $SE = 2.29$). Für die Erinnerung an räumliche Bildinhalte wurden keine Unterschiede zwischen den Lernern beobachtet, $F < 1$.

Hinsichtlich der Erinnerung an Charakteristika schnitten Lerner mit visuellen Textinhalten ($M = 32.49$, $SE = 3.41$) besser ab als Lerner mit räumlichen Textinhalten ($M = 19.36$, $SE = 3.41$). Des Weiteren wurde eine Interaktion zwischen Textinhalt und vVSSP Kapazität beobachtet: Lerner mit visuellen Textinhalten zeigten mit zunehmender vVSSP Kapazität bessere Erinnerungsleistungen ($pr = .43$, $p = .01$), während bei Lernern mit räumlichen Textinhalten kein Zusammenhang zwischen vVSSP Kapazität und Erinnerung vorherrschte ($pr = -.19$, $p = .27$).

Entsprechend den Annahmen der ECTML zeigte die Analyse keinen Einfluss des Textinhalts bei der Erinnerung an abstrakte Textinhalte (auch wenn das strenge Kriterium von $\alpha = .20$ für die Interaktion von Textinhalt, Zweitaufgabe und Corsi Block Maß verfehlt wurde), aber bei der Erinnerung an Charakteristika und Bildinhalte.

Tabelle 42

Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die offenen Fragen

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
visuelle Bildinhalte räumliche Bildinhalte	TI	$V = .47, F(3, 62) = 27.15, p < .001$
	TI \times ZA	$V = .03, F < 1$
	TI \times Corsi	$V = .04, F(3, 62) = 1.15, p = .32$
	TI \times VPT	$V = .004, F < 1$
	TI \times ZA \times Corsi	$V = .03, F < 1$
Charakteristika	TI	$F(1, 63) = 8.45, MSE = 764.25, p = .01, \eta_p^2 = .12$
	TI \times ZA	$F < 1$
	TI \times Corsi	$F < 1$
	TI \times VPT	$F(1, 63) = 4.56, MSE = 764.25, p = .04, \eta_p^2 = .07$
	TI \times ZA \times Corsi	$F < 1$
abstrakte Textinhalte	TI	$F < 1$
	TI \times ZA	$F < 1$
	TI \times Corsi	$F < 1$
	TI \times VPT	$F(1, 63) = 1.56, MSE = 218.29, p = .22, \eta_p^2 = .02$
	TI \times ZA \times Corsi	$F(1, 63) = 2.21, MSE = 218.29, p = .14, \eta_p^2 = .03$

Anmerkung. Die Textmodalität ging zusätzlich als UVn, das Digit Span Maß als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Leistung in der Zweitaufgabe

In Tabelle 43 sind die Ergebnisse hinsichtlich der Leistung in der Zweitaufgabe dargestellt.

Tabelle 43

Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Leistung in der Zweitaufgabe

AVn der ANCOVA	Effekte der ANCOVA	Ergebnisse der ANCOVA
Fehlerrate	TI	$F < 1$
	TI \times Corsi	$F < 1$
	TI \times VPT	$F(2, 26) = 1.01, MSE = 1.83, p = .33, \eta_p^2 = .04$
Tippgeschwindigkeit	TI	$F(2, 26) = 4.02, MSE = 90925.30, p = .06, \eta_p^2 = .13$
	TI \times Corsi	$F < 1$
	TI \times VPT	$F(2, 26) = 1.33, MSE = 90925.30, p = .26, \eta_p^2 = .05$

Anmerkung. Die Textmodalität ging zusätzlich als UVn, das Digit Span Maß als Kovariate ein.

Hinsichtlich der Leistung in der Zweitaufgabe zeigten sich entgegen den Annahmen der ECTML keine Unterschiede in Abhängigkeit des Textinhalts.

Bewertungsitems

Die Analysen der Bewertungsitems (vgl. Tabelle 44) erfolgten mit der zur Analyse der Erinnerungsleistung äquivalenten MANCOVA.

Tabelle 44

Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), Zweitaufgabe (ZA), VPT und Corsi Block (Corsi) für die Bewertungsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	TI	$V = .09, F(2, 62) = 2.08, p = .11$
	TI \times ZA	$V = .04, F < 1$
	TI \times Corsi	$V = .04, F < 1$
	TI \times VPT	$V = .06, F(2, 62) = 1.30, p = .28$
	TI \times ZA \times Corsi	$V = .06, F(2, 62) = 3.97, p = .04$

Anmerkung. Die Textmodalität ging zusätzlich als UVn, das *Digit Span* Maß als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Hinsichtlich der Bewertungsitems zeigte die MANCOVA eine Dreifachinteraktion zwischen Textinhalt, Zweitaufgabe und rVSSP Kapazität. Die univariaten Analysen bestätigten den Effekt jedoch nicht für das Ausmaß an Konzentration während der Lernphase, $F < 1$. Für die investierte Mühe in die Lernphase wurde der Effekt hingegen statistisch bedeutsam, $F(1, 63) = 4.55, MSE = 2.42, p = .04, \eta_p^2 = .07$. Die Interaktion ist in Abbildung 24 dargestellt. Bei Lernern mit visuellen Textinhalten nahm die rVSSP Kapazität weder mit Zweitaufgabe ($pr = .16, p = .55$) noch ohne Zweitaufgabe ($pr = .12, p = .67$) einen Einfluss. Bei Lernern mit räumlichen Textinhalten und Zweitaufgabe stieg die investierte Mühe mit zunehmender rVSSP Kapazität tendenziell ($pr = .41, p = .10$), während sie bei Lernern mit räumlichen Textinhalten und ohne Zweitaufgabe mit zunehmender rVSSP Kapazität tendenziell sank ($pr = -.44, p = .09$).

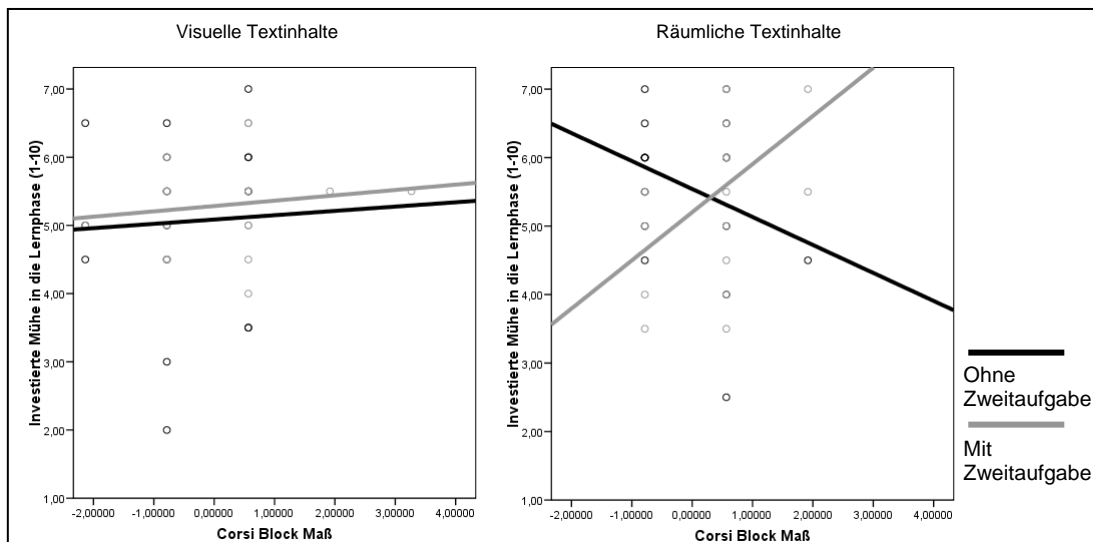


Abbildung 24. Darstellung des Zusammenhangs des Corsi Block Maßes mit der in die Lernphase investierten Mühe in Abhängigkeit der Zweitaufgabe und des Textinhalts.

Zusammenfassend kann man somit festhalten, dass die Annahmen der ECTML eines negativen Einfluss räumlicher Textinhalte auf das Lernergebnis für die Erinnerung an visuelle Bildinhalte gemessen über die offenen Fragen und die Erinnerung an Charakteristika gemessen über die Verifikationsitems und über die offenen Fragen bestätigt wurden. Zudem wurde bestätigt, dass die Erinnerung an abstrakte Textinhalte nicht durch die anderen Textinhalte moderiert wurde. Hinsichtlich der Erinnerung an Charakteristika gemessen über die offenen Fragen nahm auch die vVSSP Kapazität Einfluss: Lerner mit hoher vVSSP Kapazität konnten visuelle Textinhalte besser erinnern als Lerner mit niedriger vVSSP Kapazität, während in der Gruppe mit räumlichen Textinhalten die vVSSP Kapazität erwartungsgemäß keinen Einfluss nahm. Hinsichtlich der Leistung in der Zweitaufgabe zeigten die Analysen überraschenderweise keinen Einfluss des Textinhalts.

Die nach der ECTML erwartete Interaktion zwischen Textinhalt und Zweitaufgabe wurde für keine der Variablen bestätigt. Im Hinblick auf die in die Lernphase investierte Mühe zeigte die Analyse eine durch die rVSSP Kapazität moderierte Interaktion zwischen Textinhalt und Zweitaufgabe. Interessant ist, dass Lerner bei der Durchführung einer Zweitaufgabe und der Darbietung räumlicher Textinhalte mit zunehmender rVSSP Kapazität mehr Mühe investierten. Dies legt die Vermutung nahe, dass die Lerner durch mehr Anstrengung der höheren Belastung entgegen steuerten.

10.3.8 Prüfung der Arbeitsgedächtnis-Hypothese

Sowohl nach der CTML als auch nach der ECTML wurde ein Einfluss der Arbeitsgedächtniskapazität auf das Lernen mit Multimedia erwartet. Die Prüfung dieser Annahme erfolgte im Rahmen der beschriebenen MANCOVAs. In Tabelle 45 sind die relevanten Haupteffekte der *Digit Span*, des Corsi Blocks und des VPT hinsichtlich der Verifikationsitems dargestellt. Der Einfluss der Arbeitsgedächtniskapazität konnte nur für die Erinnerung an Charakteristika sowie an abstrakte Textinhalte bestätigt werden: Hier nahm die Kapazität der PL, gemessen über die *Digit Span*, jeweils einen statistisch bedeutsamen Einfluss auf die Erinnerungsleistung. Mit zunehmender Kapazität der PL erinnerten die Lerner Charakteristika ($pr = .27$, $p = .02$) und abstrakte Textinhalte ($pr = .37$, $p = .001$) besser.

Tabelle 45

Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Verifikationsitems

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
visuelle Bildinhalte räumliche Bildinhalte	<i>Digit Span</i> Corsi VPT	$V = .004$, $F < 1$ $V = .01$, $F < 1$ $V = .002$, $F < 1$
Charakteristika	<i>Digit Span</i> Corsi VPT	$F(1, 63) = 5.53$, $MSE = 0.33$, $p = .02$, $\eta_p^2 = .08$ $F(2, 63) = 1.03$, $MSE = 0.33$, $p = .32$, $\eta_p^2 = .02$ $F < 1$
abstrakte Textinhalte	<i>Digit Span</i> Corsi VPT	$F(2, 63) = 12.93$, $MSE = 1.42$, $p = .004$, $\eta_p^2 = .13$ $F < 1$ $F < 1$

Anmerkung. Die Textmodalität, die Textinhalte und die Zweitaufgabe gingen zusätzlich als UVn in die Analyse ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Offene Fragen

Hinsichtlich der offenen Fragen erfolgten die zu den Verifikationsitems äquivalenten Analysen. Die MANCOVAs zeigten keine Effekte der Arbeitsgedächtnismaße auf die Erinnerung der offenen Fragen (vgl. Tabelle 46).

Tabelle 46

Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Erinnerung der offenen Fragen

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
visuelle Bildinhalte räumliche Bildinhalte	<i>Digit Span</i> Corsi VPT	$V = .02, F < 1$ $V = .04, F(2, 62) = 1.18, p = .31$ $V < .001, F < 1$
Charakteristika	<i>Digit Span</i> Corsi VPT	$F(1, 63) = 3.42, MSE = 764.25, p = .07, \eta_p^2 = .05$ $F < 1$ $F < 1$
abstrakte Textinhalte	<i>Digit Span</i> Corsi VPT	$F < 1$ $F(1, 63) = 1.27, MSE = 218.29, p = .26, \eta_p^2 = .02$ $F(1, 63) = 2.76, MSE = 218.29, p = .10, \eta_p^2 = .04$

Anmerkung. Die Textmodalität, die Textinhalte und die Zweitaufgabe gingen zusätzlich als UVn in die Analyse ein.

Leistung in der Zweitaufgabe

Auch hinsichtlich der Leistung in der Zweitaufgabe wurde der Einfluss der Arbeitsgedächtniskapazität geprüft (vgl. Tabelle 47). Die Analyse zeigte einen Einfluss der vVSSP Kapazität, gemessen über den VPT, auf die Tippgeschwindigkeit. Die entsprechende Korrelation wurde jedoch nicht statistisch bedeutsam ($pr = -.23, p = .21$).

Tabelle 47

Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Leistung in der Zweitaufgabe

AVn der ANCOVA	Effekte der ANCOVA	Ergebnisse der ANCOVA
Fehlerrate	<i>Digit Span</i> Corsi VPT	$F < 1$ $F < 1$ $F < 1$
Tippgeschwindigkeit	<i>Digit Span</i> Corsi VPT	$F < 1$ $F < 1$ $F(2, 26) = 4.88, MSE = 90925.30, p = .04, \eta_p^2 = .16$

Anmerkung. Die Textmodalität, die Textinhalte und die Zweitaufgabe gingen zusätzlich als UVn in die Analyse ein.

Statistisch bedeutsame Effekte sind grau hervorgehoben.

Bewertungsitems

Die Analyse des Einflusses der Arbeitsgedächtniskapazität auf die Bewertung der Lernphase erfolgte äquivalent zu den Analysen hinsichtlich der Erinnerungsleistung. Es wurde kein Einfluss der Arbeitsgedächtnismaße beobachtet (vgl. Tabelle 48).

Tabelle 48

Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Bewertungsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	<i>Digit Span</i>	$V = .11, F(2, 62) = 2.47, p = .07$
	Corsi	$V = .02, F < 1$
	VPT	$V = .06, F(2, 62) = 1.19, p = .32$

Anmerkung. Die Textmodalität, die Textinhalte und die Zweitaufgabe gingen zusätzlich als UVn in die Analyse ein.

Zusammenfassend kann man somit festhalten, dass der Einfluss der Arbeitsgedächtniskapazität auch im zweiten Experiment sehr gering ausfiel. Nur bei der Erinnerung an verbale Informationen gemessen über die Verifikationsitems (Charakteristika, abstrakte Textinhalte) konnte ein Einfluss der PL nachgewiesen werden: Je höher die PL Kapazität war, desto besser erinnerten die Lerner die Inhalte. Der Einfluss der vVSSP Kapazität auf die Tippgeschwindigkeit konnte zwar varianzanalytisch nachgewiesen werden, war aber aufgrund der nicht vorhandenen Korrelation zwischen VPT Maß und Tippgeschwindigkeit nicht interpretierbar.

10.3.9 Prüfung der Zweitaufgaben-Hypothese

Sowohl nach der CTML als auch nach der ECTML wurde ein genereller Einfluss der Zweitaufgabe auf die Bildverarbeitung, d. h. bei der Erinnerung an Bildinformationen (visuelle Bildinhalte, räumliche Bildinhalte), erwartet. Bei der Erinnerung an Textinformationen (abstrakte Textinhalte, Charakteristika) wurde hingegen kein Haupteffekt der Zweitaufgabe erwartet. Für die Erinnerung an Bildinformationen wurde zudem eine Moderation der Wirkung der Zweitaufgabe durch die rVSSP Kapazität, gemessen über den Corsi Block, erwartet.

Verifikationsitems

In Tabelle 49 sind der für die Prüfung des Einflusses der Zweitaufgabe relevante Haupteffekt und die Interaktion dargestellt. Hinsichtlich der Bilderinnerung bestätigte die MANCOVA den von beiden Theorien vorhergesagten Haupteffekt der Zweitaufgabe nicht, da das Signifikanzniveau verfehlt wurde. Die Analysen zeigten einen Haupteffekt der Zweitaufgabe hinsichtlich der Erinnerung an Charakteristika: Lerner mit Zweitaufgabe erinnerten Charakteristika ($M = 1.46$, $SE = 0.07$) schlechter als Lerner ohne Zweitaufgabe ($M = 1.77$, $SE = 0.07$). Dieser Haupteffekt entsprach weder den Annahmen der ECTML noch der CTML.

Tabelle 49

Zur Prüfung der Zweitaufgaben-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Zweitaufgabe (ZA) und Corsi Block (Corsi) für die Verifikationsitems

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
visuelle Bildinhalte räumliche Bildinhalte	ZA	$V = .08$, $F(2, 62) = 2.64$, $p = .08$
	ZA \times Corsi	$V = .01$, $F < 1$
Charakteristika	ZA	$F(1, 63) = 9.04$, $MSE = 0.33$, $p = .004$, $\eta_p^2 = .13$
	ZA \times Corsi	$F < 1$
abstrakte Textinhalte	ZA	$F < 1$
	ZA \times Corsi	$F < 1$

Anmerkung. Die Textmodalität, die Textinhalte und das VPT Maß gingen zusätzlich als UVn, das Digit Span Maß als Kovariate in die Analyse ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Offene Fragen

Die Analysen bezüglich des Einflusses der Zweitaufgabe bei den offenen Fragen (vgl. Tabelle 50) erfolgten äquivalent zu den Analysen der Verifikationsitems. Erwartungsgemäß zeigten die Analysen einen Einfluss der Zweitaufgabe auf die Bilderinnerung. Die univariaten Analysen bestätigten den Effekt nicht für die Erinnerung an visuelle Bildinhalte, $F(1, 60) = 1.62$, $MSE = 312.31$, $p = .21$, $\eta_p^2 = .03$, aber für die Erinnerung an räumliche Bildinhalte, $F(1, 60) = 7.94$, $MSE = 508.64$, $p = .01$, $\eta_p^2 = .12$: Hier interferierte die Zweitaufgabe mit der Bilderinnerung (ohne Zweitaufgabe: $M = 49.59$, $SE = 2.89$; mit Zweitaufgabe: $M = 38.19$, $SE = 2.84$).

Tabelle 50

Zur Prüfung der Zweitaufgaben-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Zweitaufgabe (ZA) und Corsi Block (Corsi) für die offenen Fragen

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
visuelle Bildinhalte räumliche Bildinhalte	ZA	V = .11, $F(2, 62) = 3.76, p = .03$
	ZA × Corsi	V = .02, $F < 1$
Charakteristika	ZA	$F(1, 63) = 1.53, MSE = 764.25, p = .22, \eta_p^2 = .02$
	ZA × Corsi	$F < 1$
abstrakte Textinhalte	ZA	$F(1, 63) = 1.29, MSE = 218.29, p = .26, \eta_p^2 = .02$
	ZA × Corsi	$F < 1$

Anmerkung. Die Textmodalität, die Textinhalte und das VPT Maß gingen zusätzlich als UVn, das *Digit Span* Maß als Kovariate in die Analyse ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Bewertungsitems

Die Analysen der Bewertungsitems erfolgten mit der zur Analyse der Erinnerungsleistung äquivalenten MANCOVA (vgl. Tabelle 51). Es zeigten sich keine statistisch bedeutsamen Effekte.

Tabelle 51

Zur Prüfung der Zweitaufgaben-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Zweitaufgabe (ZA) und Corsi Block (Corsi) für die Bewertungsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	ZA	V = 01, $F < 1$
	ZA × Corsi	V = .07, $F(2, 62) = 2.67, p = .11$

Anmerkung. Die Textmodalität, die Textinhalte und das VPT Maß gingen zusätzlich als UVn, das *Digit Span* Maß als Kovariate in die Analyse ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Zusammenfassend kann man somit festhalten, dass die Annahmen der CTML und der ECTML einer Interferenz zwischen Bildverarbeitung und räumlicher Zweitaufgabe für die Erinnerung an räumliche Bildinhalte (gemessen über die offenen Fragen) bestätigt wurden. Hinsichtlich der Erinnerung an Charakteristika war kein Einfluss der Zweitaufgabe vorhergesagt worden. Die Analyse zeigte jedoch einen generellen Einfluss der Zweitaufgabe auf die Erinnerung an die experimentell variierte Textinformation, während sich hinsichtlich der Erinnerung an abstrakte

Textinhalte keine Effekte zeigten. Es scheint daher, als ob die räumliche Zweitaufgabe sowohl mit räumlichen als auch visuellen Textinhalten interferierte.

10.4 Zusammenfassung und Diskussion

Das Ziel des zweiten Experiments war die Prüfung der Vorhersagen der CTML sowie der Vorhersagen der ECTML unter Zuhilfenahme einer räumlichen Zweitaufgabe. Es wurden vier Hypothesen geprüft, nämlich die Textmodalität-Hypothese, die Textinhalt-Hypothese, die Arbeitsgedächtnis-Hypothese und die Zweitaufgaben-Hypothese. Zur Prüfung der Hypothesen wurde variiert, ob visuelle oder räumliche Textinhalte dargeboten wurden, ob die Lerner eine räumliche Zweitaufgabe durchführten oder nicht und ob der Text gesprochen oder geschrieben präsentiert wurde. Zudem wurde die Kapazität der PL, des vVSSP und des rVSSP erfasst. AVn waren die Erinnerungsleistung gemessen über Verifikationsitems und offene Fragen, die Leistung in der räumlichen Zweitaufgabe, sowie die Bewertung der Lernphase hinsichtlich verschiedener Aspekte.

Welchen Einfluss nimmt die Textmodalität beim Lernen mit Multimedia?

Die **Textmodalität-Hypothese** nach der ECTML ging davon aus, dass nur bei starker Belastung des rVSSP ein Modalitätseffekt auftritt. In den Gruppen ohne und mit Zweitaufgabe wurde daher ein Modalitätseffekt nur für Lerner mit räumlichen Textinhalten bei der Erinnerung an bildhafte oder verbale räumliche Informationen erwartet. In der Gruppe mit räumlicher Zweitaufgabe wurde die Belastung des rVSSP jedoch erhöht, so dass hier ein stärkerer Modalitätseffekt für Lerner mit räumlichen Textinhalten erwartet wurde. Diese Annahmen der ECTML konnten allerdings für keines der Maße bestätigt werden, auch nicht unter Einbeziehung der rVSSP Kapazität. Für die Erinnerung an abstrakte Informationen wurden keine Einflüsse der experimentellen Manipulation erwartet, was sich bestätigte, wenn auch das strenge Kriterium von $\alpha = .20$ in einem Fall verfehlt wurde.

Die Textmodalität-Hypothese nach der CTML sagte einen generellen Haupteffekt der Textmodalität für *alle* AVn vorher. Die Annahme eines generellen Modalitätseffekts konnte jedoch nur für die textbezogene Erinnerung an Charakteristika gemessen über die Verifikationsitems bestätigt werden. Somit ist die Evidenz stark eingeschränkt, da der Modalitätseffekt eben nicht für eine Vielzahl von

Variablen vorhergesagt werden konnte. Es sei darauf verwiesen, dass dieser Effekt auch nicht mit der spezifizierten räumlichen Kontiguitätserklärung erklärt werden kann, da nach dieser ein Modalitätseffekt für die bildbezogene Erinnerung erwartet worden wäre, nicht aber für die textbezogene Erinnerung. Bei der Bewertung der Lernumgebung zeigten sich Modalitätseffekte hinsichtlich der in die Lernphase investierten Mühe. Die angenommene Moderation des Modalitätseffekts durch die Kapazität des vVSSP und rVSSP wurde für keine der Variablen bestätigt. Des Weiteren wurde eine Interferenz zwischen der Ausführung der räumlichen Zweitaufgabe und der Verarbeitung geschriebenen Texts für die textbezogene Erinnerungsleistung erwartet, welche durch die rVSSP Kapazität moderiert werden konnte. Eine solche Interferenz konnte für die Erinnerung an abstrakte Textinhalte nachgewiesen werden, nicht aber für die Erinnerung an die Charakteristika. Letzteres Ergebnis entspricht den Befunden von Gyselinck et al. (2008), die ebenfalls keine Interferenz zwischen räumlicher Zweitaufgabe und geschriebener Textdarbietung fanden. Hinsichtlich der Erinnerung abstrakter Inhalte nahm die Zweitaufgabe entgegen den Annahmen der ECTML bei geschriebener Darbietung tendenziell einen negativen Einfluss, was dafür spricht, dass geschriebener Text in diesem Fall den rVSSP belastete. Dieses Ergebnis ist konsistent mit den Annahmen der CTML, nach der ein Abbild des geschriebenen Texts unabhängig von der Art des Textinhalts mit der räumlichen Zweitaufgabe interferieren sollte. Problematisch ist jedoch, dass diese spezifischen Interferenzen nur tendenziell statistisch bedeutsam waren. Daher ist die Interpretierbarkeit des Ergebnisses eingeschränkt.

Die Annahmen des zweiten Experiments, dass der Modalitätseffekt mit Hilfe der spezifischen Belastungserhöhung des rVSSP durch die Ausführung einer Zweitaufgabe sowie durch die intraindividuelle Variation der Textmodalität sichtbar gemacht werden würde, konnten somit nicht bestätigt werden. Auch hinsichtlich des prozessnahen Maßes, nämlich der Leistung in der Zweitaufgabe, ließ sich die Textmodalität-Hypothese nach der ECTML nicht bestätigen.

Welchen Einfluss nimmt der Textinhalt beim Lernen mit Multimedia?

Die **Textinhalt-Hypothese** nach der CTML ging davon aus, dass der Textinhalt keinen Einfluss auf das Lernergebnis nimmt. Diese Annahme konnte nicht bestätigt werden, denn die Analysen bestätigten die Annahmen der ECTML, die in Abhängigkeit der AVn spezifische Vorhersagen machte: Für die Erinnerung an

abstrakte Textinhalte wurden keine Unterschiede zwischen den Lernern mit visuellen und räumlichen Textinhalten erwartet, was die Analysen bestätigten, wenn auch das strenge Kriterium von $\alpha = .20$ in zwei Fällen verfehlt wurde. Für die Erinnerung an bildhafte Informationen sowie die Charakteristika wurde ein Haupteffekt des Textinhalts erwartet. Diese Annahme konnte teilweise bestätigt werden: Lerner mit räumlichen Textinhalten erinnerten visuelle Bildinhalte (gemessen über die offenen Fragen) sowie Charakteristika (gemessen über die Verifikationsitems und die offenen Fragen) schlechter als Lerner mit visuellen Textinhalten. Wie in Experiment 1 zeigten sich die Unterschiede zwischen den Lernern also nicht nur bezüglich der textbezogenen Erinnerungsleistung, sondern auch bezüglich der bildbezogenen Erinnerungsleistung. Dies unterstützt die Annahme, dass sich eine durch räumliche Textinhalte induzierte Belastung auch auf die Bildverarbeitung auswirken kann. Hinsichtlich der Erinnerung an Charakteristika gemessen über die offenen Fragen wurde des Weiteren eine Moderation durch die vVSSP Kapazität bei visuellen Textinhalten beobachtet: In der Gruppe mit visuellen Textinhalten führte eine zunehmende vVSSP Kapazität zu besserer Leistung, während sie bei räumlichen Textinhalten keinen Einfluss nahm. Dies unterstützt die Annahme, dass visuelle Textinhalte im vVSSP verarbeitet werden, während räumliche Textinhalte hier nicht verarbeitet werden.

Des Weiteren wurde eine Interaktion zwischen Textinhalt und Zweitaufgabe erwartet, da innerhalb der Gruppe mit Zweitaufgabe die Unterschiede zwischen visuellen und räumlichen Textinhalten ausgeprägter sein sollten als in der Gruppe ohne Zweitaufgabe, da die Zweitaufgabe insbesondere die Leistung bei räumlichen Textinhalten senken sollte. Diese Annahme konnte nicht bestätigt werden. Hinsichtlich der in die Lernphase investierten Mühe wurde jedoch eine Interaktion zwischen Textinhalt, Zweitaufgabe und rVSSP Kapazität beobachtet: Lerner mit räumlichen Textinhalten investierten mit zunehmender rVSSP Kapazität mehr Mühe, wenn sie eine räumliche Zweitaufgabe durchführen mussten, während sie ohne Durchführung einer Zweitaufgabe bei zunehmender rVSSP Kapazität weniger Mühe investierten. Dies deutet darauf hin, dass der erhöhten Belastung des rVSSP entgegengesteuert werden konnte, indem – falls vorhanden – mehr Kapazitäten für das Lernen investiert wurden.

Zusammengefasst sprechen diese Ergebnisse leider nicht eindeutig dafür, dass der räumliche Textinhalt im Vergleich zum visuellen Textinhalt einen

Belastungszuwachs im rVSSP induzierte, da sich bezüglich der Erinnerung an Charakteristika nicht die postulierte spezifische Interferenz zwischen räumlichen Textinhalten und Zweitaufgabe zeigte. Es kann jedoch nicht ausgeschlossen werden, dass Lerner mit räumlichen Textinhalten der erhöhten Belastung entgegensteuern konnten, so dass sich daher auf Performanzebene die postulierte Interaktion nicht zeigte.

Welchen Einfluss nimmt die Arbeitsgedächtniskapazität auf das Lernen mit Multimedia?

Nach der **Arbeitsgedächtnis-Hypothese** sollte die Kapazität des Arbeitsgedächtnisses einen entscheidenden Einfluss auf das Lernen mit Multimedia nehmen. Diese Annahme ist sowohl im Einklang mit der ECTML als auch der CTML. Erneut zeigte sich allerdings ein eher zu vernachlässigender Einfluss der Arbeitsgedächtniskapazität, denn nur bei der Erinnerung an verbale Informationen gemessen über die Verifikationsitems nahm die PL Kapazität Einfluss auf die Leistung: Mit zunehmender Kapazität der PL erinnerten die Lerner die verbalen Inhalte besser. Dieses Ergebnis ist im Einklang mit den Annahmen der ECTML und CTML, nach denen in der PL verbale Informationen verarbeitet werden. Die vVSSP und rVSSP Kapazität nahmen hingegen keinen (interpretierbaren) Einfluss auf die Ergebnisse. Hier hätte man insbesondere im Hinblick auf die Bilderinnerung Effekte erwartet.

Somit bestätigten die Ergebnisse von Experiment 2 die Befunde aus Experiment 1, dass die Arbeitsgedächtniskapazität keinen bzw. nur einen sehr geringen Einfluss auf die Ergebnisse nimmt.

Welchen Einfluss nimmt eine räumliche Zweitaufgabe auf das Lernen mit Multimedia?

Nach der **Zweitaufgaben-Hypothese** gingen sowohl die CTML als auch die ECTML von einem generellen Einfluss der räumlichen Zweitaufgabe auf die Bildverarbeitung aus. Diese Annahme konnte bestätigt werden, da bei der Erinnerung räumliche Bildinhalte (gemessen über die offenen Fragen) Interferenzen mit der Zweitaufgabe nachgewiesen wurden. Dieses Ergebnis bestätigt somit die Annahme, dass eine erhöhte Belastung des rVSSP, induziert über eine räumliche Zweitaufgabe oder über räumliche Textinhalte, der Bildverarbeitung abträglich sein kann.

Hinsichtlich der Textverarbeitung wurde kein genereller Einfluss der Zweitaufgabe erwartet, da die räumliche Zweitaufgabe nur mit den räumlichen Textinhalten interferieren sollte, nicht aber mit visuellen oder abstrakten Textinhalten (s. o.). Hinsichtlich der Erinnerung an abstrakte Textinhalte wurde diese Annahme bestätigt, was zeigt, dass abstrakte Textinhalte nicht im rVSSP verarbeitet wurden. Hinsichtlich der Erinnerung an visuelle und räumliche Textinhalte wurde jedoch ein genereller Einfluss der Zweitaufgabe beobachtet, d. h. die Zweitaufgabe senkte sowohl die Erinnerung an visuelle als auch räumliche Textinhalte (gemessen über die Verifikationsitems). Dies deutet darauf hin, dass nicht nur räumliche Textinhalte, sondern auch visuelle Textinhalte im rVSSP verarbeitet wurden. Da es im Hinblick auf die Erinnerung an abstrakte Textinhalte keinen Einfluss der Zweitaufgabe gab, kann ausgeschlossen werden, dass die Zweitaufgabe generell mit der Verarbeitung verbaler Informationen interferiert. Dies spricht dafür, dass die in der vorliegenden Studie als „visuell“ deklarierten Textinhalte ebenfalls im rVSSP verarbeitet wurden. Eine Erklärung hierfür könnte sein, dass die Operationalisierung der visuellen Inhalte über Formen zu dieser Interferenz führte. Zwar ist die Zuordnung der Verarbeitung von Formen im vVSSP aus theoretischer Sicht unbestritten (vgl. z. B. Zimmer, 2010), empirisch deutet jedoch ein Befund von Eddy und Glass (1981) darauf hin, dass Formen eventuell auch im rVSSP verarbeitet werden könnten (vgl. Abschnitt 4.3.2). Daher wurde in Experiment 3 die Operationalisierung der visuellen Textinformation verändert: Statt der Operationalisierung über Farben und Formen wurden hier nur noch Farbinformationen dargeboten, da die Zuordnung von Farbinformationen zur visuellen Komponente eindeutiger ist.

Einschränkungen des zweiten Experiments und Implikationen für Folgeexperimente

Im Hinblick auf die Frage, ob unterschiedliche Schwierigkeiten der Textinhalte die Ergebnisse ebenfalls erklären können, muss auch in diesem Experiment festgehalten werden, dass eine eindeutige Beantwortung dieser Frage nicht möglich ist. Für unterschiedliche Textschwierigkeiten sprechen die Ergebnisse bezüglich einiger der Variablen, bei denen der Textinhalt einen Einfluss nahm, aber nicht mit der Zweitaufgabe interagierte. Beispielsweise erinnerten Lerner mit visuellen Textinhalten bei Abfrage über die offenen Fragen Charakteristika besser als Lerner

mit räumlichen Textinhalten – die Zweitaufgabe nahm aber keinen Einfluss auf das Ergebnis.

Gegen die Annahme, dass allein die Textschwierigkeit die Ergebnisse erklären kann, spricht jedoch wie bereits in Experiment 1 der Befund, dass die Textinhalte nicht nur Einfluss auf die Texterinnerung nahmen, sondern auch auf die Verarbeitung der Bilder. Da deren Verarbeitung im VSSP stattfindet, wie es durch die Interferenzen mit der Zweitaufgabe belegt wurde, und nicht in der PL, scheint es fraglich, ob die Textschwierigkeit einen so starken Einfluss auf die Bildverarbeitung nehmen kann. Die mögliche Alternativerklärung, nämlich dass schwierigen Texten mehr Zeit und dem Bild entsprechend weniger Zeit gewidmet wird, was wiederum zu schlechterer Verarbeitung des Bildes führt, kann im vorliegenden Experiment nicht geklärt werden. In Experiment 3 wurden daher die Blickbewegungen der Lerner aufgezeichnet, um die den Texten und den Bildern gewidmete Aufmerksamkeit zu erfassen. Des Weiteren konnte so geprüft werden, ob das Ausmaß der Blickbewegungen die Erinnerungsleistung beeinflussen kann wie es die ECTML vorhersagen würde.

In Experiment 3 wurden die visuellen Inhalte nur noch über Farbinformationen operationalisiert, um eine mögliche Konfundierung von Formen und der Verarbeitung im rVSSP auszuschließen.

Da in den ersten beiden Experimenten die gleiche Domäne und die gleichen Inhalte verwendet wurden, wurde in Experiment 3 eine neue Domäne verwendet. Dadurch war es möglich, die Generalisierbarkeit der Befunde der ersten beiden Experimente zu untersuchen.

Als UVn dienten die Textinhalte sowie die Textmodalität. Da das intraindividuelle Design der Textmodalität nicht wie erwartet die Effekte der Textmodalität sichtbar gemacht hatte, wurde in Experiment 3 die Textmodalität erneut interindividuell variiert.

11 Experiment 3

11.1 Erweiterung der Hypothesen

Mit Hilfe des dritten Experiments wurden erneut die Annahmen der CTML sowie der ECTML geprüft. Dabei bestand das Ziel des dritten Experiments zum Einen darin, die Hypothesen in Bezug auf eine neue konzeptuelle Inhaltsdomäne zu untersuchen um zu prüfen, inwiefern die Hypothesen hier bestätigt wurden bzw. inwiefern die bisherigen Ergebnisse hier repliziert werden konnten. Ein weiteres Ziel des dritten Experiments war es, die nach der ECTML erwartete Interferenz zwischen Blickbewegungen und Verarbeitung räumlicher Textinhalte direkt zu prüfen. Des Weiteren diente Experiment 3 dazu die Text und Bild gewidmete visuelle Aufmerksamkeit der Lernergruppen zu vergleichen, um auszuschließen, dass die in den ersten beiden Experimenten beobachteten Ergebnisse auf unterschiedliche Verteilungen der visuellen Aufmerksamkeit in der Lernphase zurückgingen.

Die *Textmodalität-Hypothese*, die *Textinhalt-Hypothese* und die *Arbeitsgedächtnis-Hypothese* entsprachen den in Abschnitt 6 aufgestellten Hypothesen. Eine Ausnahme bildeten die Annahmen bezüglich der in Experiment 3 erhobenen Blickverhaltensindizes (Fixationen auf Bild, Fixationen auf Text und Anzahl der Sakkaden). Hier machten die CTML als auch die ECTML die gleichen Vorhersagen: Hinsichtlich der Textmodalität wurde erwartet, dass Lerner mit gesprochenen Texten dem Bild mehr Aufmerksamkeit schenken sollten und weniger Sakkaden ausführen sollten als Lerner mit geschriebenen Texten, was sich statistisch jeweils in einem Haupteffekt für die Textmodalität ausdrücken sollte. Im Hinblick auf die Textinhalte und die Arbeitsgedächtniskapazität wurde kein Einfluss auf das Blickverhalten erwartet.

Zusätzlich wurde eine vierte Hypothese formuliert, welche Erwartungen hinsichtlich des Zusammenhangs von Textinhalt und Ausmaß an Blickbewegungen aufstellte und deshalb im Folgenden als *Blickbewegungs-Hypothese* bezeichnet wird. Sie machte die folgenden Annahmen:

Unter Gültigkeit der CTML wird kein Effekt des Ausmaßes der Blickbewegungen auf die Leistung erwartet, da Blickbewegungen keine spezifische Rolle beim Lernen mit Multimedia zugeschrieben wird. Der

Modalitätseffekt wird entsprechend auf Interferenzen zwischen den Abbildern des geschriebenen Texts und des Bilds im VSSP zurückgeführt und nicht auf die mit dem Lesen assoziierten Blickbewegungen.

Unter Gültigkeit der ECTML wird hingegen ein Einfluss des Ausmaß an Blickbewegungen auf die Leistung erwartet: Bei der Darbietung räumlicher Textinhalte kommt es zu einer Belastung des rVSSP. Da hier auch die Kontrolle der Blickbewegungen lokalisiert ist, sollte mit zunehmendem Ausmaß an Blickbewegungen die Belastung des rVSSP weiter steigen. Es wird daher erwartet, dass bei Darbietung räumlicher Textinhalte mit zunehmenden Blickbewegungen die Erinnerung an bildbezogene Informationen sowie Charakteristika aufgrund der erhöhten Belastung des rVSSP sinkt, während bei Darbietung visueller Textinhalte die Erinnerung an bildbezogene Informationen sowie Charakteristika durch zunehmende Blickbewegungen nicht beeinflusst wird. Auch bezüglich der Erinnerung an abstrakte Textinhalte sollten die Blickbewegungen keinen Einfluss nehmen. Hinsichtlich der Interferenz zwischen räumlichen Textinhalten und Ausmaß an Blickbewegungen wird zudem eine Moderation durch die rVSSP Kapazität angenommen, da mit zunehmender rVSSP Kapazität die Belastung durch räumliche Textinhalte und Blickbewegungen besser handhabbar wird.

Statistisch sollte sich die Blickbewegungs-Hypothese nach der ECTML in einer Zweifachinteraktion zwischen Ausmaß der Blickbewegungen und Textinhalt bzw. einer Dreifachinteraktion zwischen Ausmaß der Blickbewegungen, Textinhalt und rVSSP Kapazität ausdrücken, wenn Bildinformationen oder Charakteristika erinnert werden müssen. Bei der Erinnerung an abstrakte Textinhalte wurden hingegen keine Unterschiede erwartet (vgl. Abbildung 25).

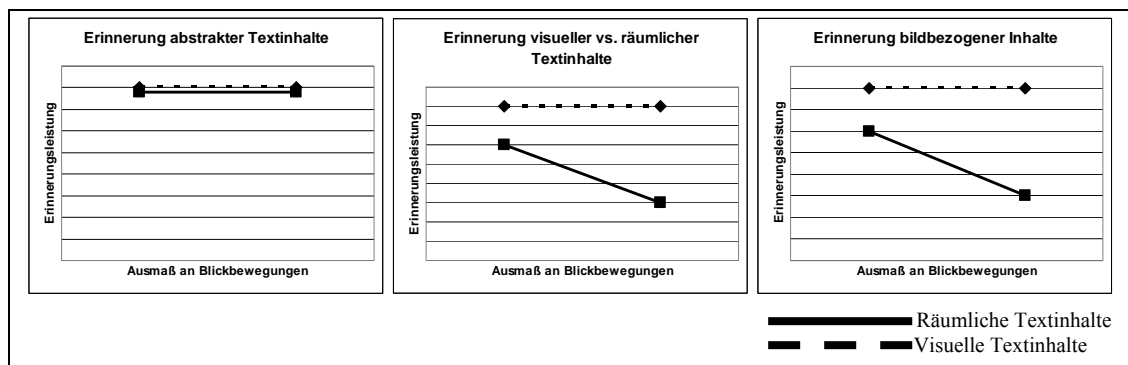


Abbildung 25. Darstellung des erwarteten Ergebnismusters für die Textinhalt-Hypothese.

Die Prüfung der Blickbewegungs-Hypothese erfolgt getrennt für Lerner mit gesprochenen und geschriebenen Textdarbietungen, da die Textmodalität und das Ausmaß der Blickbewegungen konfundiert sind: So sollte geschriebener Text in einem hohen Ausmaß an Blickbewegungen resultieren, während gesprochener Text ein vergleichsweise niedrigeres Ausmaß an Blickbewegungen zur Folge haben sollte. Es wurde angenommen, dass eine Zunahme an Blickbewegungen sowohl bei gesprochener als auch geschriebener Textdarbietung mit der Verarbeitung räumlicher Textinhalte interferieren sollte, so dass eine Berechnung der Analysen getrennt nach Textmodalität sinnvoll ist.

11.2 Methode

Die formulierten psychologischen Hypothesen (vgl. Abschnitt 6 und Abschnitt 11.1) wurden erneut mit Hilfe einer computerbasierten, multimedialen Lernumgebung getestet. Die verwendete Domäne war strukturell mit den ersten beiden Experimenten identisch, unterschied sich aber inhaltlich, da als Lernobjekte artifizielle Pflanzenarten dienten. Für eine genaue Beschreibung der Gestaltung der Lernumgebung vgl. Abschnitt 7.

11.2.1 *Versuchsdesign*

Die Testung der aufgestellten psychologischen Hypothesen (Abschnitt 6 und Abschnitt 11.1) erfolgte wie in den ersten beiden Experimenten durch experimentelle Variationen während der *Lernphase* (vgl. Abschnitt 7) mit Hilfe eines zweifaktoriellen Versuchsplans. Erste UV war die Darbietungsmodalität des verbalen Materials (gesprochen vs. geschrieben), zweite UV waren die Textinhalte (visuell vs. räumlich). Beide Faktoren wurden interindividuell variiert. Zur Testung der Blickbewegungshypothese wurde das Ausmaß der Blickbewegungen als kontinuierlicher Faktor in das Design aufgenommen.

Pro Versuchsbedingung wurden $n = 20$ Personen getestet, insgesamt also $N = 80$ Personen.

Die Operationalisierung der unabhängigen Variablen

Die Operationalisierung der zwei kategorialen UVn „Textmodalität“ und „Textinhalt“ wurde bereits in Abschnitt 8.1 beschrieben und wird daher nicht weiter

ausgeführt. Im Hinblick auf die Variable Textinhalt gab es jedoch einen Unterschied zu den ersten beiden Experimenten: Visuelle Textinformationen wurden ausschließlich über die Farbe und nicht mehr über die Form operationalisiert. Diese Variation wurde vorgenommen, da das zweite Experiment für einige Variablen Interferenzen zwischen räumlicher Zweitaufgabe und visuellen Textinhalten gezeigt hatte. Eine mögliche Erklärung hierfür ist, dass Formen – über die visuelle Textinhalte operationalisiert wurden – auch im rVSSP verarbeitet wurden. Dies ist zwar aus theoretischer Sicht unwahrscheinlich, kann aber aufgrund des empirischen Befunds von Eddy und Glass (1981) nicht vollkommen ausgeschlossen werden. Des Weiteren wurde die Arbeitsgedächtniskapazität der Lerner erhoben (vgl. Abschnitt 8.1.3).

Die kontinuierliche Variable „Ausmaß an Blickbewegungen“ wurde über die Anzahl an Sakkaden operationalisiert. Zur Aufzeichnung der Blickbewegungen wurde der iView XTM Hi-Speed 500 der SensoMotoric Instruments (SMI) GmbH benutzt. Der High-Speed 500 nutzt zur Berechnung der Blickbewegungen einen auf Sakkaden basierenden Algorithmus: Ab einer – vom Nutzer zu definierenden – Geschwindigkeit wird die Blickbewegung als Sakkade eingestuft. Alle darunter liegenden Blickgeschwindigkeiten werden folglich als Fixationen bzw. als Rauschen innerhalb der Fixationen definiert. Welche Geschwindigkeit als Schwelle für die Sakkaden definiert werden sollte, wird kontrovers diskutiert (vgl. Holmqvist, in Vorbereitung). Im vorliegenden Experiment wurde eine Geschwindigkeit von $50^\circ/\text{sec}$ als Schwelle für eine Sakkade festgelegt (Holmqvist, persönliche Mitteilung). Alle Bewegungen unterhalb dieser Schwelle wurden somit als Fixation betrachtet. Mit Hilfe der Auswertungssoftware BeGaze™ 2.1 Software wurde die Anzahl an Sakkaden pro Versuchsperson berechnet. Diese ging dann als kontinuierliche Variable in die Auswertung ein.

Die Operationalisierung der abhängigen Variablen

AVn waren im dritten Experiment die Erinnerungsleistung gemessen über Verifikationsitems und offene Fragen, die Bewertung der Lernphase sowie – im Unterschied zu den ersten beiden Experimenten – das Blickverhalten der Versuchsteilnehmer.

Bezüglich des Blickverhaltens wurden die folgenden Parameter betrachtet: (a) Die Gesamtanzahl an Sakkaden während der Lernphase – diese Variable diente somit

nicht nur als kontinuierlicher Faktor, sondern es wurde auch verglichen, inwiefern sich die Gruppen diesbezüglich unterschieden, (b) die Fixationsdauer auf dem Bild und (c) für Lerner mit geschriebenen Texten die Fixationsdauer auf dem Text. Zur Berechnung der Variablen (b) und (c) wurden *Areas of Interests* (AOIs) definiert, also Bildschirmbereiche, die für die Analyse von besonderem Interesse waren. Bei diesen AOIs handelte es sich in allen Bedingungen um das Bild und bei Lernern mit geschriebenen Texten zudem um den Text (vgl. Abbildung 26). Die Fixationszeiten waren dabei als die Gesamtzeit aller in die jeweilige AOI fallenden Fixationen definiert. Sakkaden wurden nicht in die Fixationszeit eingerechnet, da während einer Sakkade die Informationsaufnahme unterdrückt ist (Rayner, 1998).

Die Operationalisierung der personenbezogenen Daten

Es wurden soziodemographische Daten wie Alter, Geschlecht, Studienfach und Semesteranzahl erfasst sowie das Vorwissen erhoben. Eine genaue Operationalisierung ist Abschnitt 8.3 zu entnehmen.

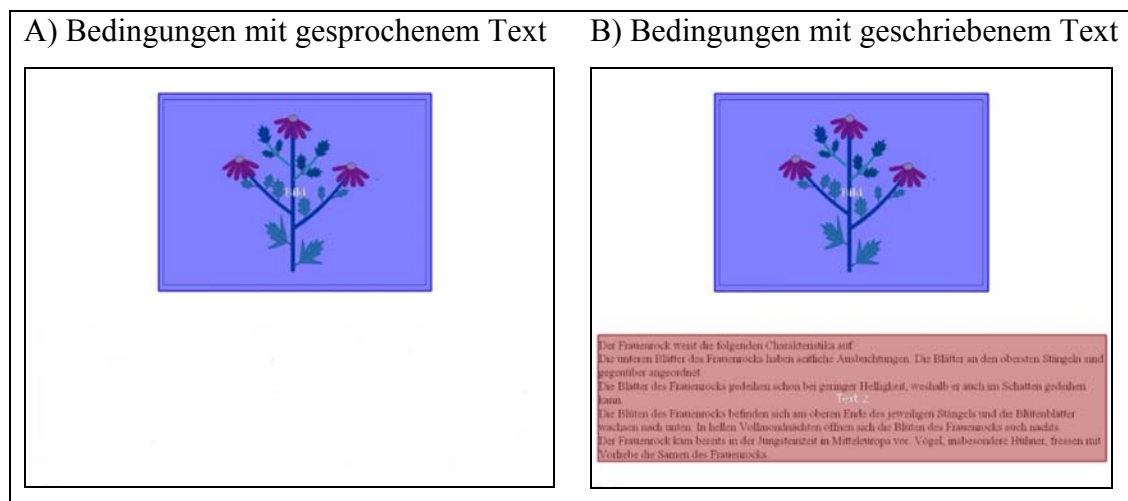


Abbildung 26. AOIs in der Lernphase. Blau = bildbezogenen AOIs, rot = textbezogene AOI.

11.2.2 Versuchspersonen

Insgesamt nahmen 80 Studenten der Universität Tübingen an der Untersuchung teil. Als Aufwandsentschädigung für die Teilnahme erhielten die Probanden entweder 10 Euro oder 1.5 Versuchspersonenstunden. Die Probanden waren zwischen 19 und 30 Jahren alt ($M = 23.97$ Jahre, $SD = 2.68$). Es nahmen 64 Frauen (80%) und 16 Männer

(20%) an der Untersuchung teil. Die Zuteilung der Versuchspersonen auf die vier Bedingungen erfolgte randomisiert.

11.2.3 *Geräte und Hilfsmittel*

Die computerbasierte Einführungsphase und Testphase wurden auf einem Notebook mit 15 Zoll Bildschirm dargeboten. Zur Darbietung wurde das Programm Medialab v2008 (Empirisoft, New York) genutzt. Aufgrund der Aufzeichnung der Blickbewegungen wechselten die Probanden für die Lernphase an einen zweiten Computer. Hier wurden die Inhalte der Lernphase auf einem TFT 19 Zoll Monitor mit einer Bildschirmauflösung von 1280×1024 Pixel und einer Bildschirmaktualisierungsrate von 60Hz dargeboten. Der Bildschirm war in einem Abstand von circa 65 cm vor den Probanden platziert. Die Inhalte der Lernphase wurden videobasiert im .avi Format dargeboten. Die Videos hatten Auflösung von 1280×1024 Pixel und eine Bildfrequenz von 15 Bildern/Sekunde. Diese relativ niedrige Bildfrequenz war unproblematisch, da es sich um die Darstellung statischer Inhalte handelte. Die Erstellung der Videos erfolgte durch das Abfilmen einer PowerPoint Präsentation der Lerninhalte mit Hilfe des Programms Camtasia Studio 5.1 (TechSmith Corporation; East Lansing, Michigan). Die Präsentation der Lerninhalte in einem videobasierten Format war notwendig, da die verwendete Präsentationssoftware (SMI Experiment Center 2.1) audiovisuelle Inhalte nur in Form von .avi Dateien darbieten kann.

Zur Aufzeichnung der Blickbewegungen wurde der iView XTM Hi-Speed 500 der SensoMotoric Instruments (SMI) GmbH sowie die Software SMI iView 2.1 benutzt. Bei dem Hi-Speed 500 handelt es sich um ein säulenbasiertes System, bei dem der Kopf des Probanden durch eine Kinn- und Stirnstütze fixiert wird. Dadurch werden Kopfbewegungen des Probanden, die die Datenqualität beeinträchtigen können, minimiert. Die Aufnahme der Blickbewegungen erfolgte monokular durch Aufnahme des rechten Auges mit einer zeitlichen Auflösung von 500 Hz. Die Berechnung der Blickdaten sowie die Definition der AOIs (vgl. Abschnitt 11.2.1) erfolgte mit Hilfe der SMI BeGazeTM 2.1 Software.

11.2.4 *Versuchsablauf*

Der Versuch wurde über einen Zeitraum von fünf Wochen zwischen 8 und 19 Uhr als Einzelsitzung durchgeführt.

Zunächst unterschrieben die Teilnehmer eine Einverständniserklärung zur freiwilligen Teilnahme an der Untersuchung.

Dann begannen sie mit der Einführungsphase der computerbasierten Lernumgebung (vgl. Abschnitt 7). Im Anschluss an die Einführungsphase wechselten die Probanden für die Lernphase an den Eyetracker. Es wurde eine automatische 13 Punkt Kalibrierung durchgeführt. Hierbei wurden den Probanden in zwei aufeinanderfolgenden Durchgängen jeweils 13 Punkte auf dem Bildschirm an unterschiedlichen Positionen präsentiert. Ihre Aufgabe bestand darin, den jeweiligen Punkt zu fixieren. Anschließend wurden als Maß für die Güte der Kalibrierung die Abweichungen bei der ersten und zweiten Betrachtung eines jeden Punktes berechnet. Ggf. wurden mehrere Kalibrierungen durchgeführt, um möglichst geringe Abweichungen zu erhalten. Probanden, deren Abweichungen auch nach mehreren Versuchen größer als ein Grad waren, wurden von der Analyse der Blickbewegungen ausgeschlossen. Unabhängig von der Kalibrierungsqualität wurde die Lernphase jedoch am Eyetracker präsentiert. Nach der Kalibrierungsphase wurde die Lernumgebung gestartet. Nach Abschluss der Lernphase begann die dritte Phase, die Testphase. Hierfür wechselten die Teilnehmer zurück an das Notebook. Der Aufbau der Testphase war äquivalent zum ersten Experiment, nur dass im Anschluss an die Verifikationsitems weitere offene Fragen dargeboten wurden. Nach der computerbasierten Testphase beantworteten die Probanden die vier offenen Fragen hinsichtlich der Text- und Bilderinnerung, dann fand die Durchführung der Tests zur Messung der Arbeitsgedächtniskapazität mit Hilfe des Corsi Block Tests, des VPTs sowie des *Digit Span* Tests statt (vgl. Abschnitt 2.3.1). Nach der Durchführung dieser drei Tests war die Untersuchung beendet. Die Probanden erhielten eine Aufklärung über den Versuch und die Aufwandsentschädigung. Insgesamt dauerte die Untersuchung circa 90 Minuten.

11.3 Ergebnisse

Die Überprüfung der Hypothesen des dritten Experiments erfolgte mit Hilfe (multivariater) Varianzanalysen mit den Zwischensubjektfaktoren Textinhalt und Textmodalität und den kontinuierlichen, z-standardisierten Testwerten im VPT, Corsi Block und *Digit Span* Test. Die *Digit Span* ging als einfache Kovariate in die Analyse ein, während hinsichtlich des VPT und Corsi Block Maß auch die Interaktionen mit den experimentellen Bedingungen berücksichtigt wurden. Nicht

berücksichtigt wurden die Interaktionen zwischen den Arbeitsgedächtnismaßen. In den Analysen wurden jeweils solche AVn zusammengefasst, für die nach der ECTML oder CTML entweder (a) ein Einfluss der experimentellen Variation erwartet wurde (Charakteristika, visuelle Bildinhalte, räumliche Bildinhalte), oder (b) kein Einfluss der experimentellen Variation erwartet wurde (abstrakte Textinhalte). Die Ergebnisse dieser Analysen werden nach Hypothesen getrennt aufgeführt.

11.3.1 Reliabilitätsanalysen

Die Analyse der Reliabilitäten und die Indexbildung erfolgten analog zum ersten Experiment. Die Reliabilitäten der einzelnen Indizes sowie die nach Itemausschluss resultierenden Indizes sind in Tabelle 52 dargestellt.

Tabelle 52

Cronbach's α der einzelnen Indizes vor und nach Itemausschluss

	Cronbach's α vor Itemausschluss		Cronbach's α nach Itemausschluss	
	gleicher Index	ungleicher Index visuell räumlich	gleicher Index	ungleicher Index visuell räumlich
<i>bildbezogene Erinnerungsleistung</i>				
visuelle Bildinhalte	$\alpha = .39$ $i = 9$	$\alpha = .65$ $i = 18$	$\alpha = .51$ $i = 7$	$\alpha = .69$ $i = 16$
räumliche Bildinhalte	$\alpha = .57$ $i = 18$	$\alpha = .58$ $i = 9$	$\alpha = .59$ $i = 17$	$\alpha = .61$ $i = 8$
<i>textbezogene Erinnerungsleistung</i>				
Charakteristika	$\alpha = .81$ $i = 30$		k.A.	
abstrakte Textinhalte	$\alpha = .33$ $i = 6$		$\alpha = .38$ $i = 5$	

Anmerkung. i = Anzahl der Items. K.A. = kein Ausschluss von Items notwendig.

Die Reliabilitätsanalyse zeigte, dass auch im dritten Experiment die Mehrheit der Indizes die geforderte Reliabilität von mindestens Cronbach's $\alpha = .70$ nicht aufwiesen. Dies kann teilweise, wie bereits erwähnt, an der geringen Itemanzahl liegen.

Die Auswertung der offenen Fragen erfolgte erneut durch zwei Rater, die die jeweilige Versuchsbedingung nicht kannten. Die Interrater-Reliabilität lag in einem akzeptablen Bereich, mit Cohen's Kappa = .77 für die schriftlichen Fragen und Cohen's Kappa = .77 für die Zeichenaufgaben. Nichtübereinstimmungen zwischen den beiden Ratern wurden durch nachträglichen Konsensus gelöst. Die Reliabilitäten

der einzelnen Indizes der offenen Fragen sowie die Zuwächse in Cronbach's α durch Itemausschluss sind in Tabelle 53 dargestellt.

Tabelle 53

Cronbach's α der einzelnen Indizes der offenen Fragen vor und nach Itemausschluss

	Cronbach's α		Cronbach's α nach Itemausschluss	
	gleicher Index	ungleicher Index visuell räumlich	gleicher Index	ungleicher Index visuell räumlich
<i>bildbezogene Erinnerungsleistung</i>				
visuelle Bildinhalte	$\alpha = .45$ $i = 6$	$\alpha = .72$ $i = 16$	$\alpha = .48$ $i = 5$	$\alpha = .76$ $i = 15$
räumliche Bildinhalte	$\alpha = .51$ $i = 17$	$\alpha = .47$ $i = 8$	$\alpha = .62$ $i = 15$	$\alpha = .51$ $i = 7$
<i>textbezogene Erinnerungsleistung</i>				
Charakteristika	$\alpha = .36$ $i = 8$	$\alpha = .37$ $i = 8$	$\alpha = .57$ $i = 6$	$\alpha = .42$ $i = 6$
abstrakte Textinhalte	$\alpha = .47$ $i = 18$		$\alpha = .54$ $i = 16$	

Anmerkung. Die Erfassung der textbezogenen Erinnerungsleistung erfolgte schriftlich, die Erfassung der bildbezogenen Erinnerungsleistung erfolgte zeichnerisch. i = Anzahl der Items

11.3.2 Datenexploration und Überprüfung inferenzstatistischer Analysevoraussetzungen

Aufgrund von technischen Schwierigkeiten oder schlechter Kalibrierungsqualität verblieben bei der Analyse der Blickbewegungsdaten 71 Probanden, davon 57 weiblichen und 14 männlichen Geschlechts ($M = 24.00$ Jahre, $SD = 2.69$ Jahre).

Die Daten wurden auf Ausreißer, Normalverteilung mit Hilfe des Kolmogorov-Smirnov Tests (vgl. Anhang L) und Varianzhomogenität der einzelnen Variablen mit Hilfe des Levene-Tests (vgl. Anhang M) geprüft. Tabelle 54 zeigt Verletzungen der Normalverteilung und Abweichungen von der Fehlervarianz auf. Bezüglich der Erinnerung an abstrakte Textinhalte konnte durch Ziehen der Quadratwurzel die ungleichen Fehlervarianzen behoben werden.

Die Prüfung der Homogenität der Regressionsgeraden der Kovariaten *Digit Span* innerhalb der vier Bedingungen kann Anhang N entnommen werden. Eine Verletzung der Annahme vergleichbarer Regressionsgraden innerhalb der Bedingungen lag nicht vor.

Tabelle 54

Variablen mit Missing Data, Abweichungen von der Normalverteilung sowie ungleichen Fehlervarianzen

Abweichungen von Normalverteilung	<i>Kontrollvariablen:</i> Corsi Block Test, <i>Digit Span</i> Test <i>Erinnerungsleistung:</i> Offene Fragen: Charakteristika, abstrakte Textinhalte, visuelle Bildinhalte, räumliche Bildinhalte <i>Blickbewegungsdaten:</i> Ausmaß an Sakkaden <i>Bewertungsitems:</i> Alle Variablen
Ungleiche Fehlervarianzen	<i>Erinnerungsleistung:</i> Offene Fragen: abstrakte Textinhalte <i>Blickbewegungsdaten:</i> Fixationszeit auf Bild

11.3.3 Beschreibung der Stichprobe

Nach der Datenexploration wurden die demographischen Daten ausgewertet. Die Mittelwerte und Standardabweichungen der als intervallskaliert behandelten Daten sind in Tabelle 55 in Abhängigkeit der Zwischensubjektfaktoren Textmodalität und Textinhalt dargestellt.

Tabelle 55

Mittelwerte und Standardabweichungen (in Klammern) der personenbezogenen Daten in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	gesprochen <i>n</i> = 20	geschrieben <i>n</i> = 20	gesprochen <i>n</i> = 20	geschrieben <i>n</i> = 20
Alter (in Jahren)	24.40 (2.64)	23.35 (2.64)	24.35 (2.66)	23.19 (2.83)
Semesteranzahl	7.80 (3.83)	6.30 (4.38)	6.60 (4.83)	6.10 (2.99)
Letzte Bionote (0 = ungenügend, 15 = sehr gut plus)	9.45 (3.85)	11.00 ^a (2.49)	10.35 (3.25)	10.00 (2.36)
Vorkenntnisse bzgl. Pflanzen (1 = keine bis 4 = hoch)	2.05 (0.76)	1.80 (0.62)	2.25 (0.64)	1.95 (0.69)
Corsi Block Test	5.75 (0.85)	5.45 (0.69)	5.75 (0.72)	5.40 (0.88)
<i>Visual Pattern</i> Test	7.82 (1.79)	8.58 (1.36)	8.85 (1.58)	8.53 (1.56)
<i>Digit Span</i> Test	6.25 (1.12)	6.20 (1.11)	6.30 (1.26)	6.55 (0.94)

Anmerkung. ^a*n* = 17 da ein *Missing Data*.

Tabelle 56 gibt einen Überblick bezüglich der Häufigkeitsverteilung der nominalen Variablen in den vier Bedingungen.

Tabelle 56

Die Verteilung der nominalen personenbezogenen Daten in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	gesprochen <i>n</i> = 20	geschrieben <i>n</i> = 20	gesprochen <i>n</i> = 20	geschrieben <i>n</i> = 20
<i>Geschlecht</i>				
männlich	5	3	2	6
weiblich	15	17	18	14
<i>Biologie 4-stündig/Leistungskurs</i>				
Ja	8	4	3	3
Nein	12	16	17	17
<i>Studienfach</i>				
Naturwissenschaftlich	0	1	2	0
Anderes	20	19	20	19

Die Analyse auf Gleichheit der Eingangsbedingungen zeigte für keine der erhobenen Variablen kritische Unterschiede oder Ungleichverteilungen zwischen den Gruppen (vgl. Anhang O).

Die Gesamtstichprobe lässt sich daher wie folgt beschreiben: Insgesamt nahmen 80 Personen an der Untersuchung teil, davon 64 Frauen und 16 Männer im Alter von 19 bis 30 Jahren. Das Durchschnittsalter betrug $M = 23.97$ Jahre ($SD = 2.68$), die durchschnittliche Semesteranzahl belief sich auf 6.70 Semester ($SD = 4.04$). Die letzte Biologienote betrug im Durchschnitt $M = 10.19$ Punkte ($SD = 3.16$), was in etwa einem „gut“ entspricht. Achtzehn Teilnehmer hatten in der Schule Biologie als Vertiefungsfach oder als Leistungskurs belegt, aber nur zwei Teilnehmer studierten aktuell ein naturwissenschaftliches Fach wie Medizin oder Biologie. Das Vorwissen bezüglich Pflanzen lag mit einem durchschnittlichen Wert von $M = 2.01$ und einer Standardabweichung von $SD = 0.68$ im mittleren Bereich der Skala, die von „1 = keine Vorkenntnisse“ bis „4 = hohe Vorkenntnisse“ reichte. Die Corsi Block Spanne lag bei durchschnittlich 5.59 Blöcken ($SD = 0.79$), die Leistung im VPT lag bei $M = 8.45$ ($SD = 1.60$) und beim *Digit Span* Test zeigten die Teilnehmer eine durchschnittliche Erinnerungsspanne von 6.33 Zahlen pro Sequenz ($SD = 1.10$). Die Werte hinsichtlich des Corsi Block Tests sowie der *Digit Span* sind vergleichbar mit den Ergebnissen der ersten beiden Experimente. Der

Durchschnittswert im VPT lag wie auch in Experiment 2 über dem in Experiment 1 erreichten Wert.

11.3.4 Deskriptive Daten

Im Folgenden werden die deskriptiven Daten der einzelnen AVn dargestellt. Dabei handelt es sich um adjustierte Daten, da der Einfluss der individuellen Arbeitsgedächtniskapazität (PL, vVSSP und rVSSP) herausgerechnet wurde. Die adjustierten Mittelwerte und Standardfehler für die Verifikationsitems sind in Tabelle 57 dargestellt, die adjustierten Mittelwerte und Standardfehler für die offenen Fragen in Tabelle 58, die adjustierten Mittelwerte und Standardfehler für die Blickverhaltensindizes in Tabelle 59 und die adjustierten Mittelwerte und Standardfehler für die Bewertungsitems in Tabelle 60.

Tabelle 57

Adjustierte Mittelwerte und Standardfehler (in Klammern) der Verifikationsitems in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	geschrieben <i>n</i> = 20	gesprochen <i>n</i> = 20	geschrieben <i>n</i> = 20	gesprochen <i>n</i> = 20
<i>bildbezogene Erinnerungsleistung</i>				
visuelle Bildinhalte	1.33 (0.15)	1.28 (0.17)	1.39 (0.15)	1.37 (0.15)
räumliche Bildinhalte	1.24 (0.13)	1.21 (0.14)	1.94 (0.12)	1.63 (0.12)
<i>textbezogene Erinnerungsleistung</i>				
Charakteristika	1.90 (0.13)	2.00 (0.15)	1.65 (0.13)	1.52 (0.13)
abstrakte Textinhalte	2.49 (0.17)	2.32 (0.19)	2.23 (0.17)	2.07 (0.17)

Anmerkung. Minimum = 0, Maximum = 4.

Tabelle 58

Adjustierte Mittelwerte und Standardfehler (in Klammern) der offenen Fragen in Prozent in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	geschrieben <i>n</i> = 20	gesprochen <i>n</i> = 20	geschrieben <i>n</i> = 20	gesprochen <i>n</i> = 20
<i>bildbezogene Erinnerungsleistung</i>				
visuelle Bildinhalte	40.45 (5.13)	48.89 (5.86)	47.15 (5.14)	45.96 (5.25)
räumliche Bildinhalte	43.85 (4.44)	37.27 (5.07)	29.50 (4.44)	30.99 (4.54)
<i>textbezogene Erinnerungsleistung</i>				
Charakteristika	32.38 (5.25)	45.84 (6.00)	28.56 (5.26)	33.97 (5.37)
abstrakte Textinhalte	6.04 (2.09)	9.34 (2.38)	12.66 (2.09)	13.92 (2.13)

Tabelle 59

Adjustierte Mittelwerte und Standardfehler (in Klammern) der Blickverhaltensindizes in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	gesprochen <i>n</i> = 17	geschrieben <i>n</i> = 18	gesprochen <i>n</i> = 17	geschrieben <i>n</i> = 19
Anzahl der Sakkaden	432.20 (42.02)	964.32 (33.90)	470.92 (36.90)	1029.83 (33.23)
Fixationszeit auf Bild (s)	253.75 (6.70)	53.61 (5.41)	247.31 (5.69)	56.11 (5.20)
Fixationszeit auf Text (s)	—	180.17 (5.38)	—	173.61 (5.23)

Tabelle 60

Adjustierte Mittelwerte und Standardfehler (in Klammern) der Bewertungsitems in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	geschrieben <i>n</i> = 15	gesprochen <i>n</i> = 16	geschrieben <i>n</i> = 14	gesprochen <i>n</i> = 14
Wie anstrengend war die Lernaufgabe für Dich?	4.26 (0.31)	3.65 (0.35)	4.68 (0.31)	4.15 (0.32)
Wie schwierig war es für Dich die Inhalte zu verstehen?	2.51 (0.28)	2.41 (0.32)	3.02 (0.28)	2.60 (0.29)
Wie sehr hast Du Dich beim Lernen auf das Bild konzentriert?	4.28 (0.30)	5.65 (0.34)	4.47 (0.25)	5.64 (0.31)
Wie sehr hast Du Dich beim Lernen auf den Text konzentriert?	5.25 (0.29)	5.06 (0.33)	4.82 (0.29)	5.26 (0.29)
Wie sehr hast Du Dich während der Lernphase konzentriert?	5.69 (0.25)	5.84 (0.29)	5.90 (0.24)	5.58 (0.26)
Wie sehr hast Du Dich angestrengt, um das Lernmaterial zu verstehen?	5.50 (0.27)	5.41 (0.31)	5.76 (0.27)	5.63 (0.28)
Wie viel Mühe hast Du in die Lernphase investiert?	5.17 (0.27)	5.56 (0.30)	5.65 (0.27)	5.41 (0.27)

Anmerkung. Minimum = 1, Maximum = 7. Hohe Werte indizieren eine hohe Ausprägung im Sinne der Itemformulierung.

11.3.5 Prüfung der Textmodalität-Hypothese

Nach der CTML wurde ein Haupteffekt der Textmodalität unabhängig vom Textinhalt sowie eine Moderation des Einflusses der Textmodalität durch die Kapazität des vVSSP, gemessen über den VPT, und des rVSSP, gemessen über den Corsi Block Test, für alle AVn erwartet. Die ECTML machte spezifische Vorhersagen in Abhängigkeit der AVn: Wenn zur Beantwortung der AVn die Erinnerung an bildhafte oder räumliche verbale Inhalte notwendig war (visuelle Bildinhalte, räumliche Bildinhalte, Charakteristika), wurde ein Modalitätseffekt bei der Darbietung räumlicher Textinhalte, aber nicht bei der Darbietung visueller Textinhalte erwartet. Es wurde somit kein genereller Haupteffekt der Textmodalität erwartet. Diese Interaktion sollte durch die Kapazität des rVSSP, gemessen über den Corsi Block, moderiert werden. Hinsichtlich Variablen, bei denen keine bildhaften oder räumlich verbalen Informationen erinnert werden mussten (abstrakte Textinhalte), wurde hingegen kein Einfluss der Textmodalität erwartet.

Hinsichtlich des Blickverhaltensverhaltens wurde sowohl nach der CTML als auch der ECTML ein Haupteffekt der Textmodalität erwartet, da Lerner mit

gesprochenen Texten das Bild länger anschauen sollten und ein geringeres Ausmaß an Augenbewegungen als Lerner mit geschriebenen Texten zeigen sollten. Es wurden keine weiteren Effekte erwartet.

Verifikationsitems

In Tabelle 61 sind die für die Prüfung des Einflusses der Textmodalität relevanten Haupteffekte und Interaktionen dargestellt.

Tabelle 61

Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die Verifikationsitems

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
räumliche Bildinhalte visuelle Bildinhalte Charakteristika	TM	$V = .04, F < 1$
	TM \times TI	$V = .07, F(3, 56) = 1.44, p = .24$
	TM \times Corsi	$V = .06, F(3, 56) = 1.13, p = .35$
	TM \times VPT	$V = .15, F(3, 56) = 3.15, p = .03$
	TM \times TI \times Corsi	$V = .02, F < 1$
	TM \times TI \times VPT	$V = .05, F(3, 56) = 1.08, p = .37$
abstrakte Textinhalte	TM	$F < 1$
	TM \times TI	$F < 1$
	TM \times Corsi	$F(1, 79) = 1.13, MSE = 0.53, p = .30, \eta_p^2 = .02$
	TM \times VPT	$F < 1$
	TM \times TI \times Corsi	$F(1, 79) = 3.69, MSE = 0.53, p = .06, \eta_p^2 = .05$
	TM \times TI \times VPT	$F < 1$

Anmerkung. Das *Digit Span* Maß ging als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Die erste MANCOVA konnte weder die Annahmen der ECTML noch der CTML bestätigen. Statistisch bedeutsam wurde die Interaktion zwischen Textmodalität und vVSSP Kapazität. Die univariaten Analysen bestätigten diese Zweifachinteraktion jedoch weder für die Erinnerung an visuelle Bildinhalte, $F < 1$, noch für die Erinnerung an Charakteristika, $F(1, 58) = 1.45, MSE = 0.31, p = .23, \eta_p^2 = .03$. Für die Erinnerung an räumliche Bildinhalte konnte die Interaktion hingegen bestätigt werden, $F(1, 58) = 4.36, MSE = 0.30, p = .04, \eta_p^2 = .07$. Entgegen den Annahmen der CTML sank bei Lernern mit geschriebener Textdarbietung mit zunehmender vVSSP Kapazität die Erinnerung an räumliche Bildinhalte ($pr = -.43, p = .01$), während es bei Lernern mit gesprochener Textdarbietung keinen Zusammenhang gab ($pr = .13, p = .47$).

Die univariate Analyse hinsichtlich der Erinnerung an abstrakte Textinhalte bestätigte die Annahmen der ECTML, da die Textmodalität keinen Einfluss nahm. Allerdings wurde für die Interaktion Textmodalität, Textinhalt und vVSSP Kapazität das strenge Kriterium von $\alpha = .20$ verfehlt.

Offene Fragen

Die Analysen bezüglich der offenen Fragen erfolgten äquivalent zu den Analysen der Verifikationsitems (vgl. Tabelle 62). Die Analysen konnten weder den nach der CTML vorhergesagten Haupteffekt der Textmodalität noch die nach der ECTML vorhergesagte Interaktion zwischen Textmodalität und Textinhalt bestätigen.

Tabelle 62

Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die offenen Fragen

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
räumliche Bildinhalte visuelle Bildinhalte Charakteristika	TM	$V = .09, F(4, 55) = 1.30, p = .28$
	TM \times TI	$V = .07, F < 1$
	TM \times Corsi	$V = .03, F < 1$
	TM \times VPT	$V = .04, F < 1$
	TM \times TI \times Corsi	$V = .07, F(4, 55) = 1.08, p = .38$
	TM \times TI \times VPT	$V = .04, F < 1$
abstrakte Textinhalte ^a	TM	$F < 1$
	TM \times TI	$F < 1$
	TM \times Corsi	$F < 1$
	TM \times VPT	$F < 1$
	TM \times TI \times Corsi	$F(1, 79) = 2.33, MSE = 2.99, p = .13, \eta_p^2 = .03$
	TM \times TI \times VPT	$F(1, 79) = 2.30, MSE = 2.99, p = .13, \eta_p^2 = .03$

Anmerkung. Das *Digit Span* Maß ging als Kovariate ein. ^amit transformierten Werten berechnet

Blickverhaltensindizes

Die Analyse der Blickverhaltensindizes erfolgte äquivalent zu den Performanzdaten (vgl. Tabelle 63). Die Analysen bestätigten die Annahmen der CTML und der ECTML: Lerner mit gesprochenen Texten blickten länger auf die Bilder ($M = 249.32$ s, $SE = 4.42$) als Lerner mit geschriebenen Texten ($M = 54.66$ s, $SE = 3.75$) und sie führten weniger Blickbewegungen durch ($M = 451.56$, $SE = 27.96$) als Lerner mit geschriebenen Texten ($M = 997.07$, $SE = 27.96$). Für die

Anzahl an Sakkaden wurde zudem die Interaktion zwischen rVSSP Kapazität und Textmodalität statistisch bedeutsam: Bei Lernern mit gesprochener Textdarbietung stieg mit zunehmender rVSSP Kapazität die Anzahl an Sakkaden ($pr = .38, p = .03$), während es bei Lernern mit geschriebener Textdarbietung keinen Zusammenhang gab ($pr = -.14, p = .42$).

Tabelle 63

Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die Blickverhaltensindizes

AVn der ANCOVA	Effekte der ANCOVA	Ergebnisse der ANCOVA
Fixationen auf Bild	TM	$F(1, 71) = 1149.66, MSE = 489847535.91, p < .001, \eta_p^2 = .95$
	TM \times TI	$F < 1$
	TM \times Corsi	$F < 1$
	TM \times VPT	$F < 1$
	TM \times TI \times Corsi	$F < 1$
	TM \times TI \times VPT	$F(1, 71) = 1.15, MSE = 489847535.91, p = .35, \eta_p^2 = .02$
Anzahl der Sakkaden	TM	$F(1, 71) = 221.10, MSE = 19241.82, p < .001, \eta_p^2 = .80$
	TM \times TI	$F < 1$
	TM \times Corsi	$F(1, 71) = 4.88, MSE = 19241.82, p = .03, \eta_p^2 = .08$
	TM \times VPT	$F < 1$
	TM \times TI \times Corsi	$F < 1$
	TM \times TI \times VPT	$F < 1$

Anmerkung. Das *Digit Span* Maß ging als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Bewertungsitems

Die Analysen der Bewertungsitems erfolgten mit der zur Analyse der Erinnerungsleistung äquivalenten MANCOVA (vgl. Tabelle 64). Die Analyse zeigte keinen statistisch bedeutsamen Einfluss der Textmodalität.

Tabelle 64

Zur Prüfung der Textmodalität-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textmodalität (TM), Textinhalt (TI), VPT und Corsi Block (Corsi) für die Bewertungsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	TM	V = .30, $F(7, 61) = 2.92, p = .002$
	TM × TI	V = .15, $F(7, 61) = 1.36, p = .23$
	TM × Corsi	V = .07, $F < 1$
	TM × VPT	V = .04, $F < 1$
	TM × TI × Corsi	V = .09, $F < 1$
	TM × TI × VPT	V = .10, $F < 1$

Anmerkung. Das Digit Span Maß ging als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Die MANCOVA zeigte einen statistisch bedeutsamen Haupteffekt der Textmodalität. Die univariaten Analysen bestätigten den Effekt allerdings nur für die Konzentration auf das Bild: Lerner mit gesprochener Textdarbietung ($M = 5.64, SE = 0.23$) konzentrierten sich nach eigenen Angaben mehr auf das Bild als Lerner mit geschriebener Textdarbietung ($M = 4.38, SE = 0.21$), $F(1, 67) = 27.32, MSE = 1.65, p < .001, \eta_p^2 = .20$ (alle anderen F s < 1 , außer für „Wie anstrengend war die Lernaufgabe für Dich?“, $F(1, 67) = 5.47, MSE = 1.65, p = .08, \eta_p^2 = .04$).

Zusammenfassend kann man somit festhalten, dass die Annahme der CTML einer generellen Überlegenheit gesprochener Textdarbietung sowie die Annahmen der ECTML einer Interaktion zwischen Textinhalt und Textmodalität für keine der Variablen bestätigt wurden. Es wurde jedoch eine Interaktion zwischen der Kapazität des vVSSP und der Erinnerung an räumliche Bildinhalte beobachtet: Entgegen den Annahmen der CTML profitierten Lerner mit geschriebenen Texten jedoch nicht von einer hohen Kapazität, sondern die Erinnerungsleistung sank mit zunehmender vVSSP Kapazität. Hinsichtlich der Bewertungsitems wurde, wie bereits in Experiment 1, eine höhere Konzentration auf das Bild bei Lernern mit gesprochener Textdarbietung beobachtet.

Hinsichtlich des Blickverhaltens bestätigten die Analysen sowohl im Sinne der ECTML als auch der CTML einen Einfluss der Textmodalität: Lerner mit geschriebenen Texten widmeten den Bildern weniger visuelle Aufmerksamkeit und sie führten mehr Blickbewegungen durch. Letzteres kann damit erklärt werden, dass das Lesen eine Vielzahl an Blickbewegungen erfordert. Zudem nahm mit

zunehmender rVSSP Kapazität die Anzahl an Sakkaden bei gesprochenem Text, nicht aber bei geschriebenem Text zu.

11.3.6 Prüfung der Textinhalt-Hypothese

Nach der CTML wurde für keine der AVn ein Einfluss des Textinhalts oder eine Moderation des Einflusses des Textinhalts erwartet. Die ECTML machte hingegen spezifische Vorhersagen hinsichtlich des Einflusses des Textinhalts in Abhängigkeit der AVn: Wenn zur Beantwortung der AVn die Erinnerung an bildhafte oder räumliche verbale Inhalte notwendig war (visuelle Bildinhalte, räumliche Bildinhalte, Charakteristika), wurde eine Unterlegenheit der Lerner mit räumlichen Textinhalten erwartet. Es wurde zudem angenommen, dass die Kapazitäten des rVSSP, gemessen über den Corsi Block, und des vVSSP, gemessen über den VPT, die Erinnerung an Textinhalte beeinflussen konnten. Hinsichtlich Variablen, bei denen keine räumlichen Informationen erinnert werden mussten (abstrakte Textinhalte), wurde kein Einfluss des Textinhalts erwartet. Hinsichtlich des Blickverhaltens wurde sowohl nach der ECTML als auch der CTML kein Einfluss des Textinhalts erwartet. Die Prüfung der Hypothese erfolgte mit Hilfe der beschriebenen MANCOVAs.

Verifikationsitems

In Tabelle 65 sind der für die Prüfung des Einflusses des Textinhalts relevante Haupteffekt sowie die Interaktionen mit den Arbeitsgedächtnismaßen des vVSSP und rVSSP dargestellt. Die erste MANCOVA prüfte die Textinhalt-Hypothese hinsichtlich der Erinnerung an visuelle Bildinhalte, räumliche Bildinhalte und Charakteristika. Entgegen den Annahmen der CTML zeigten die Analysen einen multivariaten Effekt des Textinhalts. Zur Analyse des multivariaten Haupteffekts wurden univariate Analysen berechnet. Hinsichtlich der Erinnerung an visuelle Bildinhalte zeigten sich keine Unterschiede zwischen Lernern mit unterschiedlichen Textinhalten, $F < 1$. Hinsichtlich der Erinnerung an räumliche Bildinhalte, zeigte die univariate Analyse ein den Annahmen der ECTML entgegengesetztes Muster: Lerner mit räumlichen Textinhalten ($M = 1.78$, $SE = 0.09$) erinnerten räumliche Bildinhalte besser als Lerner mit visuellen Textinhalten ($M = 1.22$, $SE = 0.90$), $F(1, 64) = 20.31$, $MSE = 0.26$, $p < .001$, $\eta_p^2 = .24$. Hinsichtlich der Charakteristika zeigte sich, dass

Lerner mit visuellen Textinhalten ($M = 1.95$, $SE = 0.10$) diese besser erinnert als Lerner mit räumlichen Textinhalten ($M = 1.58$, $SE = 0.09$), $F(1, 64) = 7.24$, $MSE = 0.31$, $p = .01$, $\eta_p^2 = .10$.

Die univariate Analyse hinsichtlich der Erinnerung an abstrakte Textinhalte bestätigte die Annahmen der ECTML, da kein Einfluss des Textinhalts beobachtet wurde. Allerdings wurde für den Haupteffekt das strenge Kriterium von $\alpha = .20$ verfehlt.

Tabelle 65

Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) für die Verifikationsitems

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
räumliche Bildinhalte	TI	V = .36, $F(3, 53) = 9.92$, $p < .001$
visuelle Bildinhalte	TI × Corsi	V = .04, $F < 1$
Charakteristika	TI × VPT	V = .12, $F(3, 53) = 2.34$, $p = .08$
abstrakte Textinhalte	TI	$F(1, 79) = 2.07$, $MSE = 0.53$, $p = .16$, $\eta_p^2 = .03$
	TI × Corsi	$F(1, 79) = 1.02$, $MSE = 0.53$, $p = .32$, $\eta_p^2 = .02$
	TI × VPT	$F < 1$

Anmerkung. Die Textmodalität ging zusätzlich als UVn, das *Digit Span* Maß als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben.

Offene Fragen

Die Analysen bezüglich der offenen Fragen erfolgten äquivalent zu den Analysen der Verifikationsitems (vgl. Tabelle 66).

Tabelle 66

Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) bei der Erinnerung der offenen Fragen

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
räumliche Bildinhalte	TI	$V = .30, F(4, 55) = 5.94, p < .001$
visuelle Bildinhalte	TI \times Corsi	$V = .04, F < 1$
Charakteristika	TI \times VPT	$V = .02, F < 1$
abstrakte Textinhalte ^a	TI	$F(1, 79) = 5.91, MSE = 2.99, p = .02, \eta_p^2 = .08$
	TI \times Corsi	$F < 1$
	TI \times VPT	$F < 1$

Anmerkung: Die Textmodalität ging zusätzlich als UVn, das *Digit Span* Maß als Kovariate ein. Statistisch bedeutsame Effekte sind grau hervorgehoben. ^amit transformierten Werten berechnet

Die MANCOVA hinsichtlich der Erinnerung an visuelle Bildinhalte, räumliche Bildinhalte und Charakteristika zeigte einen multivariaten Effekt des Textinhalts. Die univariaten Analysen bestätigten den Effekt nicht für die Erinnerung an visuelle Bildinhalte, $F < 1$, oder an Charakteristika, $F(1, 67) = 2.07, MSE = 509.24, p = .09, \eta_p^2 = .04$, jedoch für die Erinnerung an räumliche Bildinhalte, $F(1, 67) = 4.99, MSE = 364.01, p = .03, \eta_p^2 = .07$. Entsprechend den Annahmen der ECTML erinnerten Lerner mit visuellen Textinhalten ($M = 40.56, SE = 3.37$) räumliche Bildinhalte besser als Lerner mit räumlichen Textinhalten ($M = 30.25, SE = 3.16$).

Hinsichtlich der Erinnerung an abstrakte Textinhalte zeigte die Analyse entgegen den Annahmen der ECTML einen Effekt des Textinhalts: Lerner mit räumlichen Textinhalten ($M = 13.20, SE = 1.48$) schnitten besser ab als Lerner mit visuellen Textinhalten ($M = 7.69, SE = 1.58$).

Blickverhaltensindizes

In Tabelle 67 sind die Ergebnisse hinsichtlich der Blickverhaltensindizes dargestellt. Entsprechend den Annahmen sowohl der ECTML als auch der CTML beeinflusste der Textinhalt das Blickverhalten nicht.

Tabelle 67

Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) für die Blickverhaltensindizes

AVn der ANCOVA	Effekte der ANCOVA	Ergebnisse der ANCOVA
Fixationen auf Bild	TI	$F < 1$
	TI \times Corsi	$F < 1$
	TI \times VPT	$F < 1$
Fixationen auf Text	TI	$F < 1$
	TI \times Corsi	$F < 1$
	TI \times VPT	$F < 1$
Ausmaß an Blickbewegungen	TI	$F(1, 71) = 2.22, MSE = 19241.82, p = .14, \eta_p^2 = .04$
	TI \times Corsi	$F < 1$
	TI \times VPT	$F < 1$

Anmerkung. Die Textmodalität ging zusätzlich als UVn, das Digit Span Maß als Kovariate ein.

Bewertungsitems

Die Ergebnisse der MANCOVA hinsichtlich der Bewertungsitems sind in Tabelle 68 dargestellt. Es zeigten sich keine statistisch bedeutsamen Effekte hinsichtlich des Textinhalts.

Tabelle 68

Zur Prüfung der Textinhalt-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Textinhalt (TI), VPT und Corsi Block (Corsi) für die Bewertungsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANOVA
Bewertungsitems	TI	$V = .08, F < 1$
	TI \times Corsi	$V = .18, F(7, 61) = 1.62, p = .14$
	TI \times VPT	$V = .14, F(7, 61) = 1.23, p = .30$

Anmerkung. Die Textmodalität ging zusätzlich als UVn, das Digit Span Maß als Kovariate ein.

Zusammenfassend kann man somit festhalten, dass entgegen den Annahmen der CTML der Textinhalt erneut Einfluss auf die Leistung der Lerner nahm. Lerner mit visuellen Textinhalten erinnerten räumliche Bildinhalte gemessen über die offenen Fragen sowie Charakteristika gemessen über die Verifikationsitems besser als Lerner mit räumlichen Textinhalten. Bei der Erinnerung an abstrakte Textinhalte gemessen über die Verifikationsitems nahm der Textinhalt hingegen keinen Einfluss. Entgegen den Annahmen der ECTML zeigte sich aber auch ein positiver Einfluss räumlicher Textinhalte, nämlich bei der Erinnerung räumlicher Bildinhalte gemessen über die

Verifikationsitems sowie der Erinnerung an abstrakte Textinhalte gemessen über die offenen Fragen. Die Analyse des Blickverhaltens bestätigte die Annahme, dass das Blickverhalten nicht durch die Textinhalte beeinflusst wurde, d. h. Lerner mit visuellen und räumlichen Textinhalten schenkten den Texten und Bildern die gleiche Aufmerksamkeit und ihre Anzahl an Sakkaden war vergleichbar. Auf die Bewertung der Lernphase nahm der Textinhalt ebenfalls keinen Einfluss.

11.3.7 Prüfung der Arbeitsgedächtnis-Hypothese

Sowohl nach der CTML als auch nach der ECTML sollte die individuelle Arbeitsgedächtniskapazität einen Einfluss auf das Lernergebnis nehmen.

Verifikationsitems

In Tabelle 69 sind die relevanten Haupteffekte der *Digit Span*, des VPT und des Corsi Blocks hinsichtlich der Verifikationsitems dargestellt. Entgegen den Annahmen zeigten sich keine statistisch bedeutsamen Effekte der Arbeitsgedächtniskapazität.

Tabelle 69

Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Verifikationsitems

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
räumliche Bildinhalte	<i>Digit Span</i>	$V = .062, F < 1$
visuelle Bildinhalte	Corsi	$V = .02, F < 1$
Charakteristika	VPT	$V = .05, F < 1$
abstrakte Textinhalte	<i>Digit Span</i>	$F(1, 79) = 1.79, MSE = 0.53, p = .19, \eta_p^2 = .04$
	Corsi	$F(1, 79) = 2.93, MSE = 0.53, p = .09, \eta_p^2 = .04$
	VPT	$F < 1$

Anmerkung. Die Textmodalität und die Textinhalte gingen zusätzlich als UVn in die Analyse ein.

Offene Fragen

Hinsichtlich der offenen Fragen erfolgten die zu den Verifikationsitems äquivalenten Analysen (vgl. Tabelle 70). Erneut zeigten sich keine statistisch bedeutsamen Unterschiede.

Tabelle 70

Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Erinnerung der offenen Fragen

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
räumliche Bildinhalte	<i>Digit Span</i>	$V = .06, F < 1$
visuelle Bildinhalte	Corsi	$V = .14, F(4, 55) = 2.23, p = .08$
Charakteristika	VPT	$V = .08, F(4, 55) = 1.23, p = .31$
abstrakte Textinhalte ^a	<i>Digit Span</i>	$F(1, 79) = 1.24, MSE = 2.99, p = .27, \eta_p^2 = .02$
	Corsi	$F < 1$
	VPT	$F < 1$

Anmerkung. Die Textmodalität und die Textinhalte gingen zusätzlich als UVn in die Analyse ein. ^amit transformierten Werten berechnet

Blickverhaltensindizes

Auch hinsichtlich des Blickverhaltens zeigte sich kein Einfluss der Arbeitsgedächtniskapazität (vgl. Tabelle 71).

Tabelle 71

Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Blickverhaltensindizes

AVn der ANCOVA	Effekte der ANCOVA	Ergebnisse der ANCOVA
Fixationen auf Bild ¹	<i>Digit Span</i>	$F < 1$
	Corsi	$F < 1$
	VPT	$F < 1$
Fixationen auf Text ²	<i>Digit Span</i>	$F(1, 29) = 1.13, MSE = 529649036.71, p = .30, \eta_p^2 = .04$
	Corsi	$F < 1$
	VPT	$F < 1$
Ausmaß an Blickbewegungen ¹	<i>Digit Span</i>	$F(1, 71) = 1.62, MSE = 19241.82, p = .21, \eta_p^2 = .03$
	Corsi	$F < 1$
	VPT	$F(1, 71) = 1.17, MSE = 19241.82, p = .29, \eta_p^2 = .02$

Anmerkung. ¹Die Textmodalität und die Textinhalte gingen zusätzlich als UVn in die Analyse ein. ²Die Textinhalte gingen zusätzlich als UVn in die Analyse ein

Bewertungsitems

Die Analyse des Einflusses der Arbeitsgedächtniskapazität auf die Bewertung der Lernphase erfolgte äquivalent zu den Analysen hinsichtlich der Erinnerungsleistung.

Erneut zeigte sich kein Einfluss der Arbeitsgedächtnismaße auf die Leistung (vgl. Tabelle 72).

Tabelle 72

Zur Prüfung der Arbeitsgedächtnis-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren VPT, Corsi Block (Corsi) und der Digit Span für die Bewertungsitems

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	<i>Digit Span</i>	$V = .13, F(7, 61) = 1.13, p = .36$
	Corsi	$V = .08, F < 1$
	VPT	$V = .21, F(7, 61) = 1.98, p = .06$

Anmerkung. Die Textmodalität und die Textinhalte gingen zusätzlich als UVn in die Analyse ein.

Zusammenfassend kann man somit festhalten, dass der von der ECTML und CTML vorhergesagte Einfluss der Arbeitsgedächtniskapazität für keine der AVn bestätigt wurde.

11.3.8 Prüfung der Blickbewegungs-Hypothese

Die Blickbewegungs-Hypothese testete den Zusammenhang der Anzahl an Sakkaden mit der Verarbeitung der Textinhalte. Nach der CTML wurde kein Einfluss der Anzahl an Sakkaden mit der Verarbeitung der Textinhalte erwartet. Nach der ECTML wurde hingegen erwartet, dass Lerner mit räumlichen Textinhalten mit zunehmender Anzahl an Sakkaden schlechtere Erinnerung an bildhafte Informationen und räumliche Textinformationen zeigen sollten, da die Belastung des rVSSP steigen sollte. Es wurde zudem eine Moderation dieser Interaktion durch die rVSSP Kapazität erwartet. Bei visuellen Textinhalten wurde ein solcher Zusammenhang nicht erwartet. Zur Prüfung der Blickbewegungs-Hypothese wurden ungesättigte Analysen berechnet, in die nur die theoretisch ableitbaren Interaktionen eingingen sowie zusätzlich die Haupteffekte und die *Digit Span* und das VPT Maß als Kovariaten. Im Folgenden werden ausschließlich die im vorliegenden Abschnitt interessierenden Ergebnisse berichtet, d. h. die Interaktion zwischen Textinhalt und Anzahl an Sakkaden sowie die mögliche Moderation dieser Interaktion durch die rVSSP Kapazität. Zusätzlich wird der Haupteffekt der Anzahl an Sakkaden berichtet, um zu prüfen, inwiefern Blickbewegungen einen generellen Einfluss auf das Lernergebnis hatten. Die Analysen erfolgten dabei getrennt für Lerner mit

gesprochenen und geschriebenen Textdarbietungen, da es als problematisch erachtet wurde, die mit gesprochenem und geschriebenem Text assoziierten unterschiedlichen Ausmaße an Blickbewegungen in eine Analyse aufzunehmen. Aufgrund von *Missing Data* bei der Aufzeichnung der Blickbewegungen gingen in die Analyse bei geschriebener Textdarbietung nur $N = 37$ Probanden, in die Analyse bei gesprochener Textdarbietung $N = 34$ Probanden ein.

Verifikationsitems

Hinsichtlich der Verifikationsitems konnte die Blickbewegungshypothese nach der ECTML weder für geschriebenen noch gesprochenen Text bestätigt werden (vgl. Tabelle 73). Hinsichtlich der Erinnerung abstrakter Inhalte zeigte sich zwar erwartungsgemäß keine signifikanten Effekte, das Kriterium der strengen Testung von Allerdings wurde für den Haupteffekt das strenge Kriterium von $\alpha = .20$ wurde aber teilweise verfehlt.

Tabelle 73

Zur Prüfung der Blickbewegungs-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Anzahl an Sakkaden (AS), Textinhalt (TI) und Corsi Block (Corsi) für die Verifikationsitems

	AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
Gesprochene Textdarbietung	räumliche Bildinhalte	AS	$V = .21, F(3, 23) = 2.04, p = .14$
	visuelle Bildinhalte	AS \times TI	$V = .07, F < 1$
	Charakteristika	AS \times TI \times Corsi	$V = .07, F < 1$
	abstrakte Textinhalte	AS	$F(1, 25) = 3.71, MSE = 0.50, p = .07, \eta_p^2 = .13$
		AS \times TI	$F(1, 25) = 2.45, MSE = 0.50, p = .13, \eta_p^2 = .09$
		AS \times TI \times Corsi	$F < 1$
Geschriebene Textdarbietung	räumliche Bildinhalte	AS	$V = .13, F(3, 26) = 1.28, p = .30$
	visuelle Bildinhalte	AS \times TI	$V = .04, F < 1$
	Charakteristika	AS \times TI \times Corsi	$V = .17, F < 1$
	abstrakte Textinhalte	AS	$F(1, 28) = 2.54, MSE = 0.47, p = .12, \eta_p^2 = .08$
		AS \times TI	$F < 1$
		AS \times TI \times Corsi	$F(1, 28) = 1.56, MSE = 0.47, p = .23, \eta_p^2 = .10$

Anmerkung. Das VPT und *Digit Span* Maß gingen zusätzlich in die Analyse ein.

Offene Fragen

Die Analysen bezüglich der offenen Fragen erfolgten äquivalent zu den Analysen der Verifikationsitems. Die erste MANCOVA innerhalb der Gruppe der Lerner mit gesprochener Textdarbietung zeigte eine multivariate Zweifachinteraktion zwischen

Anzahl an Sakkaden und Textinhalt. Die univariaten Analysen bestätigten sie für die Erinnerung an Charakteristika, $F(1, 25) = 3.36$, $MSE = 475.03$, $p = .08$, $\eta_p^2 = .09$: Es zeigten sich jedoch weder bei räumlichen Textinhalten ($pr = -.23$, $p = .42$) noch bei visuellen Textinhalten ($pr = .42$, $p = .14$) Zusammenhänge zwischen der Anzahl an Sakkaden und der Erinnerung an Charakteristika. Daher war die Zweifachinteraktion nicht interpretierbar. Für die Erinnerung an visuelle Bildinhalte, $F(1, 25) = 1.76$, $MSE = 424.67$, $p = .20$, $\eta_p^2 = .07$, und räumliche Bildinhalte, $F(1, 25) = 2.58$, $MSE = 277.54$, $p = .12$, $\eta_p^2 = .09$, wurde die Zweifachinteraktion nicht statistisch bedeutsam.

Tabelle 74

Zur Prüfung der Blickbewegungs-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Anzahl an Sakkaden (AS), Textinhalt (TI) und Corsi Block (Corsi) für die offenen Fragen

	AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
Gesprochene Textdarbietung	räumliche Bildinhalte	AS	$V = .31$, $F(4, 22) = 2.52$, $p = .07$
	visuelle Bildinhalte	AS \times TI	$V = .37$, $F(4, 22) = 3.33$, $p = .03$
	Charakteristika	AS \times TI \times Corsi	$V = .39$, $F(8, 46) = 1.39$, $p = .23$
	abstrakte Textinhalte ^a	AS AS \times TI AS \times TI \times Corsi	$F < 1$ $F(1, 25) = 1.65$, $MSE = 3.70$, $p = .21$, $\eta_p^2 = .06$ $F < 1$
Geschriebene Textdarbietung	räumliche Bildinhalte	AS	$V = .20$, $F(4, 25) = 1.56$, $p = .22$
	visuelle Bildinhalte	AS \times TI	$V = .24$, $F(4, 25) = 1.96$, $p = .13$
	Charakteristika	AS \times TI \times Corsi	$V = .42$, $F(8, 52) = 1.75$, $p = .11$
	abstrakte Textinhalte ^a	AS AS \times TI AS \times TI \times Corsi	$F(1, 28) = 10.19$, $MSE = 1.47$, $p = .003$, $\eta_p^2 = .27$ $F(1, 28) = 4.61$, $MSE = 1.47$, $p = .04$, $\eta_p^2 = .14$ $F(2, 28) = 3.85$, $MSE = 1.47$, $p = .06$, $\eta_p^2 = .30$

Anmerkung. Das VPT und Digit Span Maß gingen zusätzlich in die Analyse ein. Statistisch bedeutsame Effekte sind grau hervorgehoben. ^amit transformierten Werten berechnet

Die univariate Analyse der Erinnerung an abstrakte Textinhalte innerhalb der Gruppen mit geschriebenen Textdarbietungen zeigte entgegen der Annahmen der ECTML einen Haupteffekt der Anzahl der Sakkaden und eine Zweifachinteraktion zwischen Anzahl an Sakkaden und Textinhalt (vgl. Tabelle 74). Ein höheres Ausmaß an Sakkaden führte zu besserer Erinnerung der abstrakten Textinhalte ($pr = .44$, $p = .01$). Dieser Zusammenhang wurde durch den Textinhalt moderiert: Während bei der Darbietung räumlicher Textinhalte kein Zusammenhang zu beobachten war

($pr = .31$, $p = .24$), führte bei visuellen Textinhalten eine höhere Anzahl an Sakkaden zu besserer Erinnerung an abstrakte Textinhalte ($pr = .68$, $p = .01$).

Bewertungsitems

Die Analysen der Bewertungsitems erfolgten mit der zur Analyse der Erinnerungsleistung äquivalenten MANCOVAs getrennt für Lerner mit geschriebener und gesprochener Textdarbietung. Die Annahmen der ECTML wurden nicht bestätigt (vgl. Tabelle 75).

Tabelle 75

Zur Prüfung der Blickbewegungs-Hypothese relevante Ergebnisse der (M)ANCOVAs mit den Faktoren Anzahl an Sakkaden (AS), Textinhalt (TI) und Corsi Block (Corsi) für die Bewertungsitems

	AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
Gesprochene Textdarbietung	Bewertungsitems	AS	$V = .28, F < 1$
		AS \times TI	$V = .34, F(7, 19) = 1.16, p = .37$
		AS \times TI \times Corsi	$V = .27, F < 1$
Geschriebene Textdarbietung	Bewertungsitems	AS	$V = .18, F < 1$
		AS \times TI	$V = .15, F < 1$
		AS \times TI \times Corsi	$V = .73, F(14, 40) = 1.56, p = .12$

Anmerkung. Das VPT und *Digit Span* Maß gingen zusätzlich in die Analyse ein.

Zusammenfassend kann man somit festhalten, dass die Blickbewegungs-Hypothese nach der ECTML für keine der AVn bestätigt wurde. Die Analysen zeigten stattdessen einen Einfluss der Anzahl an Sakkaden bei der Erinnerung an abstrakte Textinhalte gemessen über die offenen Fragen: Hier führte bei visuellen Textinhalten eine hohe Anzahl an Sakkaden zu besserer Erinnerung an abstrakte Textinhalte. Dieses Ergebnis wurde weder von der CTML noch der ECTML vorhergesagt.

11.4 Zusammenfassung und Diskussion

Das Ziel des dritten Experiments war die Prüfung der Vorhersagen der CTML sowie der Vorhersagen der ECTML innerhalb einer neuen Inhaltsdomäne. Es wurden vier Hypothesen geprüft, nämlich die Textmodalität-Hypothese, die Textinhalt-Hypothese, die Arbeitsgedächtnis-Hypothese sowie die Blickbewegungs-Hypothese. Wie in den ersten beiden Experimenten wurde variiert, ob visuelle oder räumliche Textinhalte verwendet wurden und ob der Text gesprochen oder geschrieben dargeboten wurde. Zudem wurden erneut die Kapazitäten der PL, des vVSSP und des rVSSP erfasst. Im Gegensatz zu Experiment 1 und 2 wurden zudem die Blickbewegungen in der Lernphase aufgezeichnet. Dies erlaubte zum Einen das Blickverhalten der Lerner zwischen den Bedingungen zu vergleichen, zum Anderen die von der ECTML vorhergesagte Interferenz zwischen hohem Ausmaß an Blickbewegungen und der Verarbeitung räumlicher Textinhalte zu prüfen.

Welchen Einfluss nimmt die Textmodalität beim Lernen mit Multimedia?

Die **Textmodalität-Hypothese** nach der CTML ging davon aus, dass gesprochener Text generell zu besserer Leistung führt als geschriebener Text. Diese Annahmen konnte in Experiment 3 für keine der AVn bestätigt werden. Dies zeigt erneut, dass die Forderung der CTML Bilder mit gesprochenen Texten darzubieten nur unter bestimmten Bedingungen gültig ist.

Die Textmodalität-Hypothese nach der ECTML ging davon aus, dass durch zusätzliche Blickbewegungen bei der Darbietung geschriebener Textinhalte eine erhöhte Belastung des rVSSP resultiert. Diese erhöhte Belastung sollte dann zu einem Modalitätseffekt führen, wenn der rVSSP zusätzlich durch die Verarbeitung räumlicher Textinhalte belastet wird. Unter dieser Bedingung wurde für die Erinnerung an bildbezogene sowie räumliche textbezogene Informationen ein Modalitätseffekt erwartet. Für die Erinnerung an abstrakte Textinformationen wurde hingegen kein Einfluss der Textmodalität erwartet. Diese Hypothesen konnten durch die Ergebnisse jedoch nicht bestätigt werden, denn es wurde kein Modalitätseffekt in Abhängigkeit des Textinhalts beobachtet.

Im Gegensatz zu Experiment 1 und 2 wurden im vorliegenden Experiment visuelle Textinhalte nur über Farbinformationen operationalisiert, da trotz der theoretisch eindeutigen Zuordnung von Formen zum vVSSP (z. B. Zimmer, 2010), ein empirischer Befund aus der Studie von Eddy und Glass (1981) darauf hindeutet,

dass Formen eventuell auch den rVSSP belasten. Dennoch konnten die Annahmen der ECTML nicht bestätigt werden. Dies deutet darauf hin, dass der Textinhalt in multimedialen Lernumgebungen nicht die Wirkweise geschriebenen und gesprochenen Texts moderiert.

Des Weiteren zeigte die Analyse eine Interaktion zwischen Textmodalität und vVSSP Kapazität für die Erinnerung räumlicher Bildinhalte, die weder von der ECTML noch der CTML vorhergesagt worden war: Bei geschriebener Textdarbietung sank die Leistung mit zunehmender vVSSP Kapazität, während es bei gesprochener Textdarbietung keinen Zusammenhang gab. Dieses Befundmuster ist schwierig zu erklären, da es aus keiner der beiden Theorien ableitbar ist. Daher sollte der Befund zunächst repliziert werden, bevor Interpretationsversuche unternommen werden.

Im Hinblick auf das Blickverhalten sagten sowohl die ECTML als auch die CTML einen Haupteffekt der Textmodalität vorher: Die Analysen zeigten entsprechend, dass Lerner mit gesprochenen Texten dem Bild mehr visuelle Aufmerksamkeit widmeten als Lerner mit geschriebenen Texten und dass Lerner mit geschriebenen Texten mehr Blickbewegungen durchführten. Letzteres kann durch die Tatsache erklärt werden, dass das Lesen geschriebener Texte mehr Blickbewegungen erfordert als das Hören gesprochener Texte. Interessanterweise wurde im Hinblick auf die Blickbewegungen auch ein Zusammenhang zwischen rVSSP Kapazität und Textmodalität beobachtet: Lerner mit gesprochenem Text führten mit zunehmender rVSSP Kapazität mehr Blickbewegungen durch, was bei Lernern mit geschriebener Textdarbietung nicht der Fall war. Dies kann damit erklärt werden, dass die Freiheitsgrade für Lerner mit geschriebenen Texten hinsichtlich der Anzahl an Blickbewegungen begrenzt sind: Das Lesen des Texts macht Blickbewegungen erforderlich, so dass die interindividuelle Varianz bei geschriebenen Texten vermutlich relativ gering ist. Bei gesprochenen Texten ist die Ausführung von Blickbewegungen hingegen nicht unbedingt erforderlich, so dass hier je nach individueller Kapazität ein Spielraum vorhanden ist. Dass Lerner mit niedriger rVSSP Kapazität eher weniger Blickbewegungen durchführten, stützt die Annahme, dass diese den rVSSP belasten und eventuell aus eben diesem Grunde bei geringer rVSSP Kapazität weniger ausgeführt bzw. vermieden werden.

Welchen Einfluss nimmt der Textinhalt beim Lernen mit Multimedia?

Die **Textinhalt-Hypothese** nach der ECTML ging davon aus, dass es bei der Darbietung räumlicher Textinhalte zu einer erhöhten Belastung des rVSSP kommt. Diese erhöhte Belastung sollte sich im Vergleich zur Darbietung visueller Textinhalte sowohl in einer schlechteren Erinnerung an bildbezogene als auch räumliche textbezogene Informationen niederschlagen. Nach der CTML sollte der Textinhalt hingegen keinen Einfluss auf die Leistung nehmen, unabhängig davon, welche Informationen erinnert werden mussten. Erneut wurden die Annahmen der ECTML für die Erinnerung an Charakteristika und räumliche Bildinhalte (gemessen über die offenen Fragen) bestätigt.

Für zwei der AVn wurde jedoch ein den Annahmen der ECTML gegenläufiges Muster beobachtet: Lerner mit räumlichen Textinhalten konnten Verifikationsitems bezüglich räumlicher Bildinformationen sowie offene Fragen bezüglich abstrakter Inhalte besser beantworten. Diese Ergebnisse sind insofern schwierig zu erklären, als dass sie den Befunden, die im Einklang mit der ECTML waren, widersprechen. So erinnerten Lerner mit räumlichen Textinhalten räumliche Bildinhalte schlechter, wenn sie über offene Fragen abgefragt wurden, jedoch besser, wenn sie über Verifikationsitems abgefragt wurden. Da in den ersten beiden durchgeführten Experimenten die Effekte auch für die Verifikationsitems nachgewiesen wurden, bleibt unklar, warum in Experiment 3 konträre Befunde bezüglich der Verifikationsitems beobachtet wurden.

Welchen Einfluss nimmt die Arbeitsgedächtniskapazität auf das Lernen mit Multimedia?

Nach der **Arbeitsgedächtnis-Hypothese** sollte die Kapazität des Arbeitsgedächtnisses einen entscheidenden Einfluss auf das Lernen mit Multimedia nehmen. Diese Annahme war sowohl im Einklang mit der ECTML als auch der CTML. Erneut zeigte sich allerdings, dass die Arbeitsgedächtniskapazität gemessen über den *Digit Span* Test, den Corsi Block Test sowie den VPT keinen generellen Einfluss auf die Leistung nahm. Somit untermauern die Ergebnisse von Experiment 3 erneut die Ergebnisse der ersten beiden Experimente, die zeigten, dass der Einfluss der Arbeitsgedächtniskapazität auf das Lernergebnis zumindest in der vorliegenden Experimentalreihe vernachlässigbar ist.

Welchen Einfluss nehmen Blickbewegungen auf das Lernen mit Multimedia?

Im dritten Experiment wurde des Weiteren der Einfluss der Blickbewegungen auf die Verarbeitung des multimedialen Materials direkt geprüft. Nach der **Blickbewegungs-Hypothese** der CTML wurde kein Einfluss der Blickbewegungen auf die Verarbeitung visueller vs. räumlicher Textinhalte erwartet. Nach der ECTML wurde eine Interaktion zwischen Textinhalt und Anzahl an Sakkaden erwartet, da räumliche Textinhalte sowie Blickbewegungen den rVSSP belasten sollten, so dass mit einer Zunahme an Blickbewegungen für Lerner mit räumlichen Textinhalten schlechtere Leistungen erwartet wurden. Die Blickbewegungs-Hypothese war somit eine direktere Prüfung der Textmodalität-Hypothese nach der ECTML, da hier die Blickbewegungen direkt in die Analyse eingingen und nicht nur indirekt über die Textmodalität operationalisiert wurden. Um Konfundierungen mit der UV Textmodalität zu vermeiden, erfolgte die Prüfung der Blickbewegungshypothese für Lerner mit gesprochenen und geschriebenen Texten getrennt.

Die Ergebnisse konnten die Annahmen der ECTML jedoch nicht bestätigen: Es zeigte sich keine (interpretierbare) Interaktion im Sinne der ECTML zwischen Textinhalt und Anzahl an Sakkaden, auch nicht unter Berücksichtigung der rVSSP Kapazität. Des Weiteren wurde auch kein genereller Einfluss des Ausmaßes an Blickbewegungen auf die Leistung beobachtet, d. h. ein Einfluss der Blickbewegungen sowohl auf visuelle als auch räumliche Textinhalte kann ebenfalls ausgeschlossen werden. Einzig bei der Erinnerung abstrakter Inhalte gemessen über die offenen Fragen führte eine hohe Anzahl an Sakkaden zu *besserer* Erinnerungsleistung, allerdings nur, wenn visuelle Textinhalte in geschriebener Form dargeboten wurden. Dieser Befund deutet darauf hin, dass Blickbewegungen nicht nur eine zusätzliche Belastung des rVSSP darstellen und somit dem Lernen abträglich sein sollten, sondern dass Blickbewegungen auch dem Lernerfolg förderlich sind, insbesondere bei geschriebener Textdarbietung. So ist ein hohes Ausmaß an Blickbewegungen auch mit einer intensiveren Verarbeitung des Materials assoziiert. Es ist somit denkbar, dass eine absolute Aussage wie „Blickbewegungen schaden“ bzw. „Blickbewegungen fördern“ nicht möglich ist. Dies spiegelt sich auch darin wider, dass das Ausmaß der Blickbewegungen auch generell weder einen positiven noch einen negativen Einfluss auf die Lernleistung hatte.

Einschränkungen des dritten Experiments

Im Hinblick auf die Diskussion, ob unterschiedliche Schwierigkeiten der Textinhalte die Ergebnisse ebenfalls erklären können, deuten die Analysen des Blickverhaltens darauf hin, dass dies nicht der Fall ist. So widmeten die Lerner unabhängig vom Textinhalt Text und Bild die gleiche visuelle Aufmerksamkeit. Die vergleichbare Aufmerksamkeit auf Text und Bild zeigt, dass Lerner mit räumlichen Textinhalten keine größeren Schwierigkeiten hatten die Inhalte zu verstehen als Lerner mit visuellen Textinhalten, denn wäre dies der Fall, hätte man eine stärkere visuelle Aufmerksamkeit auf den Text und eine entsprechend geringere visuelle Aufmerksamkeit auf das Bild bei räumlichen Textinhalten erwartet. Auch die Anzahl an Sakkaden war zwischen Lernern mit visuellen und räumlichen Textinhalten vergleichbar, was darauf hindeutet, dass räumliche Textinhalte nicht schwieriger waren, da die Textschwierigkeit Einfluss auf die Anzahl der Sakkaden nimmt (Rayner, 1998). Diese Befunde sprechen dafür, dass nicht die Textschwierigkeit, sondern die unterschiedlichen Textinhalte die Ergebnisse beeinflussten.

12 Abschließende Zusammenfassung und Diskussion

Hinsichtlich des Lernens mit Texten und Bildern wurde in den letzten Jahren eine Reihe an Prinzipien postuliert, wie Texte und Bilder dargeboten werden sollten, um den Lernerfolg zu fördern. Die führende Theorie zum Lernen mit Multimedia, die CTML, leitet diese Empfehlungen unter anderem aus älteren Konzeptionen der Struktur des Arbeitsgedächtnisses nach Baddeley (z. B. 1992, 1999) ab.

Dass das Arbeitsgedächtnis eine entscheidende Rolle beim Lernen mit Multimedia spielt, ist nicht nur aus theoretischer Sicht plausibel, sondern konnte auch empirisch nachgewiesen werden (z. B. Brunyé et al., 2006; Gyselinck et al., 2002; Kruley et al., 1994; vgl. auch Abschnitt 2.3.2). Problematisch ist allerdings die Interpretation des Arbeitsgedächtnismodells im Rahmen der CTML hinsichtlich der Verarbeitung geschriebenen Texts: So geht Mayer (2009) in der CTML davon aus, dass geschriebener Text zunächst wie bildhafte Information behandelt wird und ein Abbild des geschriebenen Texts im VSSP verarbeitet wird. Daraus resultiert nach der CTML ein Engpass im VSSP bei der Darbietung geschriebenen Texts mit Bildern, welcher bei der Darbietung gesprochenen Texts nicht vorherrscht. Mit Hilfe dieser theoretischen Annahme erklärt Mayer (2009) den sogenannten Modalitätseffekt, d. h. den empirischen Befund, dass gesprochener Text mit Bild zu besseren Lernergebnissen führt als geschriebener Text mit Bild.

Einen alternativen Erklärungsmechanismus für den Modalitätseffekt stellt die Annahme fehlender räumlicher bzw. zeitlicher Kontiguität bei geschriebenem Text dar. Aufgrund der empirischen Evidenz ist die Annahme der räumlichen Kontiguitätserklärung, dass ein hohes Ausmaß an visuellen Suchprozessen bei geschriebenem Text und Bild den Modalitätseffekt verursacht, allerdings wenig wahrscheinlich. Es scheint eher so zu sein, dass bei der Darbietung geschriebenen Texts das Bild vernachlässigt wird, was wiederum in schlechterer bildbezogener Erinnerung resultiert (Schüler et al., 2008). Daher ist der Geltungsbereich der räumlichen Kontiguitätserklärung eingeschränkt. Auch der Geltungsbereich der zeitlichen Kontiguitätserklärung ist eingeschränkt, da ihr Wirkmechanismus auf die Bildung eines integrierten mentalen Modells beschränkt ist. Daher kommt sie als Erklärung für den Modalitätseffekt – im Gegensatz zur der von Mayer (2009) präferierten Erklärung – nur dann in Frage, wenn die Lerner das Gelernte

transferieren müssen. Da es eine Reihe an Untersuchungen gibt, die den Modalitätseffekte generell bei der Erinnerung von Faktenwissen nachweisen konnten (z. B. Brünken et al., 2005; Mayer & Moreno, 1998), scheint es plausibel, dass es einen weiteren Mechanismus für den Modalitätseffekt gibt, der einen Modalitätseffekt nicht nur für die bildbezogene Erinnerung wie die fehlende räumliche Kontiguität bei geschriebenem Text, sondern generell für die Erinnerungsleistung erklären kann.

Ein solcher alternativer Erklärungsansatz für den Modalitätseffekt wurde erstmals in der vorliegenden Forschungsarbeit vor dem Hintergrund der VSSP-Struktur generiert. Demnach unterteilt sich der VSSP in eine räumliche und eine visuelle Komponente. In diesen Komponenten werden räumliche und visuelle Bildinformationen sowie visuelle und räumliche Textinhalte verarbeitet. Zudem erfolgt die Kontrolle von Blickbewegungen im rVSSP. Diese strukturellen Annahmen zum VSSP wurden in der vorliegenden Arbeit in die CTML integriert, so dass ein erweitertes Modell, die ECTML, resultierte. Mit Hilfe der ECTML wird ein Modalitätseffekt in Abhängigkeit des Textinhalts vorhersagt: Demnach sollte der Modalitätseffekt nur auftreten, wenn räumliche Textinhalte dargeboten werden, da die Verarbeitung räumlicher Textinhalte, räumlicher Bildinhalte und die Kontrolle der Blickbewegungen miteinander im rVSSP interferieren und es zu einer *Overload*-Situation kommt. Wie im Theorieteil aufgezeigt wurde (vgl. Abschnitt 4.3), konnte diese Annahme einer Moderation des Modalitätseffekts durch die Textinhalte in einigen monomedialen und multimedialen Untersuchungen bestätigt werden (z. B., Beech, 1977; Brooks, 1967; Eddy & Glass; 1981; Glass et al., 1985; Hörnig et al., 2000; Hörnig et al., 1999; Kürschner & Schnotz, 2007; Kürschner, Schnotz & Eid, 2006; Kürschner et al., 2007; Langston et al., 1998; Unnava et al., 1996).

Aus der ECTML leitet sich neben der Vorhersage eines Modalitätseffekts in Abhängigkeit des Textinhalts auch ein genereller Einfluss des Textinhalts beim Lernen mit Multimedia ab. Demnach sollten räumliche Textinhalte mit der Bildverarbeitung, insbesondere räumlichen Bildinhalten, interferieren, während nicht-räumliche Textinhalte nicht zu Interferenzen führen sollten. Ein solcher Einfluss räumlicher Textinhalte auf das Lernen mit Multimedia lässt sich aus der CTML nicht vorhersagen.

Ziel der vorliegenden Arbeit war die Prüfung der Vorhersagen der ECTML im Vergleich zu den Vorhersagen der CTML. Durch die differenzierte Erfassung der

Erinnerungsleistung wurden die Einflüsse der zeitlichen und räumlichen Kontiguität kontrolliert. Es wurden vier Grundhypothesen gebildet, die Textmodalität-Hypothese, die Textinhalt-Hypothese, die Arbeitsgedächtnis-Hypothese und die Zweitaufgaben-Hypothese. Hinsichtlich der Textmodalität-Hypothese wurde nach der CTML ein Modalitätseffekt für alle AVn erwartet, während nach der ECTML ein Modalitätseffekt nur bei Darbietung räumlicher Textinhalte und dem Abruf von bildhaften bzw. räumlichen verbalen Informationen erwartet wurde. Hinsichtlich der Textinhalt-Hypothese wurden nach der CTML keine Effekte erwartet, während nach der ECTML schlechtere Leistungen bei Darbietung räumlicher Textinhalte und dem Abruf von bildhaften bzw. räumlichen verbalen Informationen erwartet wurden. Hinsichtlich der Arbeitsgedächtnis-Hypothese wurden sowohl nach der CTML als auch der ECTML bessere Leistungen bei hoher Kapazität der PL, des vVSSP sowie des rVSSP erwartet. Hinsichtlich der Zweitaufgaben-Hypothese wurde nach der CTML eine Interferenz der Zweitaufgabe mit der Erinnerung bildhafter Informationen erwartet. Nach der ECTML wurde eine Interferenz der Zweitaufgabe mit bildhaften Informationen sowie räumlichen verbalen Informationen erwartet, da letztere den rVSSP ebenfalls belasten sollten (vgl. auch Tabelle 5).

Im empirischen Teil der Arbeit erfolgte die Prüfung der Annahmen mit Hilfe dreier Experimente. Die Ergebnisse bezüglich der Erinnerungsleistung hinsichtlich der unterschiedlichen Vorhersagen der beiden Theorien sind in Tabelle 76, die Ergebnisse hinsichtlich gleicher Vorhersagen in Tabelle 77 zusammengefasst.

In den Tabellen sind dabei nur solche Variablen aufgeführt, für die die univariaten Analysen nach einem signifikanten multivariaten Effekt die statistische Bedeutsamkeit bestätigten. Zur Einordnung der Ergebnisse sind im Folgenden nochmals alle performanzbasierten AVn aufgeführt, die in den drei durchgeführten Experimenten gemessen wurden:

- Verifikationsitems (Experiment 1-3): Charakteristika, abstrakte Textinhalte, visuelle Bildinhalte, räumliche Bildinhalte.
- Offene Fragen (Experiment 2-3): Charakteristika, abstrakte Textinhalte, visuelle Bildinhalte, räumliche Bildinhalte.

Im Folgenden werden die empirischen Befunde hinsichtlich der Hypothesen zusammengefasst und diskutiert.

Tabelle 76

Zusammenfassung der empirischen Befunde hinsichtlich unterschiedlicher Vorhersagen der CTML und ECTML

		CTML	ECTML	
Hypothesen	I			
	Textmodalität	Exp 1	nicht bestätigt	nicht bestätigt
		Exp 2	<i>HE Textmodalität:</i> VI: Charakteristika <i>Textmodalität x Zweitaufgabe:</i> VI: abstrakte Textinhalte	nicht bestätigt
		Exp3	nicht bestätigt	nicht bestätigt
			<u>Konträre Befunde:</u> <i>Textmodalität × vVSSP Kapazität</i> VI: räumliche Bildinhalte	
	II			
	Textinhalt	Exp 1	keine Effekte erwartet	<i>HE Textinhalt</i> VI: räumliche Bildinhalte, visuelle Bildinhalte, Charakteristika
		Exp 2	keine Effekte erwartet	<i>HE Textinhalt</i> VI: Charakteristika OF: visuelle Bildinhalte, Charakteristika <i>Textinhalt × vVSSP</i> OF: Charakteristika
		Exp 3	keine Effekte erwartet	<i>HE Textinhalt</i> VI: Charakteristika, OF: räumliche Bildinhalte
			<u>Konträre Befunde:</u> <i>HE Textinhalt</i> VI: räumliche Bildinhalte OF: abstrakte Textinhalte	
	III			
	Zweitaufgabe	Exp 2	s. Tabelle 77	<u>Konträre Befunde:</u> <i>HE Zweitaufgabe</i> VI: Charakteristika
IV				
Blickbewegungen	Exp 3	keine Effekte erwartet	nicht bestätigt	
			<u>Konträre Befunde:</u> <i>Textinhalt × Anzahl an Sakkaden</i> OF: abstrakte Textinhalte	

Anmerkung. VI = Verifikationsitems, OF = Offene Fragen.

Tabelle 77

Zusammenfassung der empirischen Befunde hinsichtlich gleicher Vorhersagen der CTML und ECTML

I		
Hypothesen	Arbeitsgedächtnis	
	Exp 1	nicht bestätigt
	Exp 2	<i>HE PL Kapazität</i> VI: Charakteristika, abstrakte Textinhalte <i>HE vVSSP Kapazität</i> Tippgeschwindigkeit
	Exp 3	nicht bestätigt
II		
Zweitaufrage		
	Exp 2	<i>HE Zweitaufrage</i> OF: Räumliche Bildinhalte

Anmerkung. VI = Verifikationsitems, OF = Offene Fragen.

12.1 Welchen Einfluss nimmt die Textmodalität beim Lernen mit Multimedia?

Nach der CTML wurde eine generelle Überlegenheit gesprochener Textdarbietung über geschriebene Textdarbietung erwartet, da die Abbilder geschriebenen Texts mit der Bildverarbeitung interferieren sollten (z. B. Mayer, 2009). Des Weiteren wurde für Lerner mit höherer Kapazität des VSSP ein abgeschwächter Modalitätseffekt erwartet, da bei erhöhter Kapazität des VSSP der kognitive Engpass bei der Darbietung geschriebener Texte mit Bild geringer sein sollte. Diese Vorhersagen wurden in Experiment 2 für eine einzelne Variable bestätigt, in Experiment 1 und 3 wurde jedoch kein Einfluss der Textmodalität beobachtet (vgl. Tabelle 76, Abschnitt I). Diese Befunde widersprechen insofern den Annahmen der CTML, als dass aufgrund des postulierten Ressourcenkonflikts zwischen geschriebenem Text und Bild ein starker Effekt der Textmodalität für die Erinnerungsleistung gemessen über verschiedene Variablen erwartet worden wäre. Die Ergebnisse widersprechen somit auch dem von Ginns (2005) in seiner Metaanalyse berichteten Befund eines mittleren bis hohen Effekts der Textmodalität.

Wie in Abschnitt 3.3.1 aufgeführt, ist die Annahme eines generellen Ressourcenkonflikts zwischen geschriebenem Text und Bild im VSSP auch aus theoretischer Sicht zweifelhaft, da geschriebener Text nach Baddeley (1992) keineswegs wie ein Abbild behandelt wird, sondern direkt in die PL eingelesen wird. Diese Annahme wurde in Experiment 2 indirekt bestätigt, denn die räumliche Zweitaufgabe interferierte nicht mit der Leistung bei geschriebener Textdarbietung (einzige Ausnahme bildete die Erinnerung abstrakter Textinhalte gemessen über die Verifikationsitems, allerdings war der Effekt auch hier nur tendenziell statistisch bedeutsam). Dieser Befund ist im Einklang mit den in Abschnitt 3.3.1 berichteten Ergebnissen von Gyselinck et al. (2008), die ebenfalls keine Hinweise darauf fanden, dass geschriebener Text das VSSP so stark belastet, dass es zu Interferenzen mit einer räumlichen Zweitaufgabe kommt.

In den vorliegenden Untersuchungen zeigten sich vereinzelt Einflüsse der Textmodalität hinsichtlich der Bewertung der Lernphase: So konzentrierten sich Lerner mit geschriebenen Texten nach eigenen Angaben weniger auf das Bild (Experiment 1 und 3) und sie investierten mehr Mühe in die Lernphase (Experiment 2). Dies spricht dafür, dass die Textmodalität zwar in der Wahrnehmung der Lerner einen Einfluss auf das Lernen nimmt, dass sich dieses aber auf Performanzebene nicht widerspiegelt. Hinsichtlich der Konzentration auf das Bild wurde die subjektive Beurteilung der Lerner durch die Analyse des Blickverhaltens (Experiment 3) bestätigt, d. h. Lerner mit gesprochenen Textdarbietungen widmeten dem Bild mehr visuelle Aufmerksamkeit als Lerner mit geschriebener Textdarbietung. Lerner mit geschriebener Textdarbietung widmeten ihre visuelle Aufmerksamkeit hingegen hauptsächlich dem Text und vernachlässigten das Bild. Dies entspricht den in Abschnitt 3.3.1 getätigten Spezifizierungen der räumlichen Kontiguitätsannahme. Auf Performanzebene zeigten sich jedoch auch hier keine Unterschiede. Dass sich bei der Erinnerung des Faktenwissens kein genereller Modalitätseffekt zeigte, bestätigt die in der vorliegenden Arbeit getätigten Annahmen, dass (a) die theoretischen Überlegungen der CTML fehlerhaft sind, dass (b) die im Rahmen der räumlichen Kontiguitätsklärung postulierten extensiven visuellen Suchprozesse zwischen geschriebenem Text und Bild nicht stattfinden, und dass (c) die zeitliche Kontiguität nicht auf die Erinnerung von Faktenwissen wirkt. Würde eine dieser Annahmen *nicht* zutreffen, hätte man einen generellen Modalitätseffekt für eine Vielzahl an Variablen erwartet. Die getätigte Annahme,

dass Lerner sich explizit auf den geschriebenen Text konzentrieren und das Bild vernachlässigen, konnte mit Hilfe der subjektiven Selbsteinschätzungen der Lerner bestätigt werden, allerdings wirkte sich dieses nicht auf die bildbezogene Erinnerungsleistung aus. Daher scheint die praktische Relevanz der spezifizierten räumlichen Kontiguitätsklärung des Modalitätseffekts eingeschränkt zu sein.

Eine alternative Erklärung für den Modalitätseffekt stellt nach der ECTML der verwendete Textinhalt dar. Demnach wird ein Modalitätseffekt bei der Erinnerung von Faktenwissen nur dann erwartet, wenn der rVSSP sehr stark belastet ist. Dies sollte der Fall sein, wenn zusätzlich zur Bildverarbeitung viele Blickbewegungen durchgeführt werden müssen, wie z. B. beim Lesen, und außerdem räumliche Textinhalte dargeboten werden. Empirisch wurde diese Annahme sowohl in Experimenten mit monomedialen Materialien (z. B. Beech, 1977; Brooks, 1967; Eddy & Glass, 1981; Glass et al., 1985; Kürschner, Seufert et al., 2006) als auch in Experimenten mit multimedialen Materialien bestätigt (z. B. Langston et al., 1998; Kürschner et al., 2007). In der vorliegenden Studie wurde zusätzlich stärker differenziert, für welche AVn die Effekte erwartet wurden. Dies war in der Annahme begründet, dass bei einer sehr geringen Belastung des rVSSP, z. B. bei der Verarbeitung abstrakter Textinhalte, kein Modalitätseffekt resultieren sollte, hingegen schon bei hoher Belastung des rVSSP, wie z. B. bei der Verarbeitung räumlicher bildhafter Informationen. Aufgrund der Ergebnisse des ersten durchgeführten Experiments (vgl. Tabelle 76, Abschnitt I) wurde zudem erwartet, dass die schlechtere Verarbeitung räumlicher Bildinhalte die gesamte Bildverarbeitung stören sollte. Diese Annahme ist im Einklang mit den in Abschnitt 2.3.2 zitierten Befunden (z. B. Brunyé et al., 2006; Gyselinck et al., 2002; Nam & Pujari, 2005), die zeigten, dass die zusätzliche Belastung des rVSSP durch Ausführung einer räumlichen Zweitaufgabe die gesamte Bildverarbeitung störte. (Allerdings muss kritisch angemerkt werden, dass in diesen Studien keine explizite Trennung zwischen visueller und räumlicher Bilderinnerung erfolgte. Somit wäre es theoretisch denkbar, dass ausschließlich räumliche Bildinhalte abgefragt wurden). Die a priori Annahme der ECTML, dass Lerner mit geschriebener Textdarbietung mehr Blickbewegungen durchführen als Lerner mit gesprochener Textdarbietung, konnte bestätigt werden (Experiment 3). Einen Hinweis, dass die Blickbewegungen den rVSSP tatsächlich belasten, stellte der Befund dar, dass bei der Darbietung gesprochener Texte Lerner mit niedriger rVSSP weniger Sakkaden durchführten als

Lerner mit hoher rVSSP Kapazität (Experiment 3). Dennoch konnten die weiteren empirischen Ergebnisse der vorliegenden Experimentalreihe die Annahmen der ECTML nicht bestätigen, auch nicht unter Berücksichtigung der individuellen VSSP Kapazitäten. In Experiment 3 wurde daher zusätzlich das Ausmaß an Sakkaden als Prädiktor verwendet, um zu prüfen, ob mit zunehmender Sakkadenanzahl die Leistung bei der Darbietung räumlicher Textinhalte sinkt. Auch diese direkte Prüfung des Zusammenspiels von Blickbewegungen und räumlichen Textinhalten konnte die Annahmen der ECTML nicht bestätigen (vgl. Tabelle 76, Abschnitt IV). Zusammengenommen sprechen die Ergebnisse somit nicht für eine erhöhte Belastung des rVSSP bei der Darbietung räumlicher Textinhalte in geschriebener Form, zumindest nicht in der vorliegenden Studie.

Eine mögliche Erklärung hierfür können solche empirischen Studien bieten, die die postulierte Interferenz zwischen räumlichen Textinhalten und Textmodalität ebenfalls nicht bestätigen konnten (z. B. De Beni & Moè, 1997, Experiment 1; Jahn, 2004; Kürschner, Schnotz et al., 2006; Kürschner et al., 2007, Experiment 2). Interessanterweise führen alle Autoren den Befund einer fehlenden Interferenz zwischen geschriebenem Text und Textinhalt auf die zur Verfügung stehende Lernzeit bzw. die Darbietung des geschriebenen Texts als Fließtext, also die Darbietung längerer Textabschnitte, zurück. Demnach ermöglichen längere Lernzeiten und Fließtextdarstellung die Anwendung von Textverarbeitungsstrategien wie z. B. verlangsamtes oder erneutes Lesen (z. B. Byrne & Curtis, 2000; Danks & End, 1987; Furnham, De Siena & Gunter, 2002; Kozma, 1991). Die Annahme ist, dass durch diese Möglichkeiten zur adaptiven Verarbeitung geschriebenen Texts die erhöhte Belastung des Arbeitsgedächtnisses kompensiert werden kann, so dass als Resultat vergleichbare Leistungen bei geschriebenem und gesprochenem Text resultieren. Dieser Annahme entsprechend fanden De Beni und Moè (1997) die Interferenz zwischen räumlichem Textinhalt und geschriebenem Text bei einer segmentierten Darbietung des geschriebenen Texts, nicht aber bei Fließtextdarstellungen. Es könnte daher sein, dass in der vorliegenden Studie, in der zwar die Lernzeit begrenzt war, die Darbietungsweise des geschriebenen Texts als Fließtext dennoch die Anwendung von Textverarbeitungsstrategien erlaubte. Problematisch ist allerdings, dass eine segmentierte Darbietung geschriebenen Texts das mit geschriebenem Text assoziierte Ausmaß an Blickbewegungen gering halten sollte. Da das Ausmaß an Blickbewegungen aber nach der ECTML eine

entscheidende Rolle spielt, wurde in der vorliegenden Studie Fließtext als Darbietungsform gewählt.

Eine weitere Erklärung der fehlenden Interaktion zwischen Textinhalt und Textmodalität könnte in der Darbietung der Bilder zu sehen sein. Demnach wäre die Interferenz zwischen geschriebenem Text und räumlichen Textinhalten bei einer monomedialen Darbietung darauf zurückzuführen, dass die Lerner mental ein Bild zu den beschriebenen verbalen Inhalten generieren. Dies entspricht den Annahmen der doppelten Kodierung, nach der eine zusätzliche Gedächtnisspur durch *Imagens*, d. h. die Aktivierung interner nonverbaler Repräsentationen, generiert werden kann (Paivio, 1990; vgl. Abschnitt 2.2.1). Es könnte sein, dass dieser Prozess das VSSP stärker belastet als die Verarbeitung eines extern dargebotenen Bildes, wie es in den vorliegenden Experimenten erforderlich war. Entsprechend dieser Vermutung fanden z. B. Langstone et al. (1998), dass die Bilddarbietung einen Einfluss auf die Stärke des Modalitätseffekts bei der Darbietung räumlicher Textinhalte hatte: Der Modalitätseffekt war ausgeprägter, wenn kein Bild dargeboten wurde.

In der vorliegenden Untersuchung führten daher eventuell die Freiheitsgrade bei der Verarbeitung geschriebenen Texts sowie die Bilddarbietung dazu, dass die durch Blickbewegungen induzierte Belastung des rVSSP gering blieb bzw. kompensiert werden konnte, so dass die vorhergesagte Interferenz zwischen hohem Ausmaß an Blickbewegungen – also geschriebenem Text – und räumlichen Textinhalten nicht bestätigt wurde. In Folgeuntersuchungen sollten daher die Darbietungsweise des geschriebenen Texts sowie die Bilddarbietung stärker in den Fokus rücken um zu untersuchen, inwiefern diese beiden Faktoren die Entstehung des Modalitätseffekts beeinflussen können.

Zusammenfassend kann man festhalten, dass, obwohl die Präsentation der Inhalte systemgesteuert (Atkinson, 2002; Tabbers, 2002) erfolgte und das Material artifizieller Natur war, so dass das Vorwissen der Lerner eine geringe Rolle spielen sollte (Kalyuga et al., 2000), die Textmodalität in der vorliegenden Untersuchung kaum einen Einfluss auf die Ergebnisse nahm, auch nicht bei einer intraindividuellen Darbietung geschriebener und gesprochener Texte (Experiment 2) oder bei der Verwendung prozessnaher Maße wie der Leistung in einer Zweitaufgabe (Experiment 2). Dies spricht dafür, dass der Modalitätseffekt nur unter bestimmten Bedingungen zum Tragen kommt. Denkbar wäre, dass der Modalitätseffekt bei Transferleistung, die eine mentale Modellbildung erforderlich macht, auf die

fehlende zeitliche Kontiguität bei geschriebenem Text zurückgeht, während bei der Erinnerung von Faktenwissen, die keine mentale Modellbildung erforderlich macht, andere, noch zu klärende, Mechanismen zugrunde liegen. Die Ergebnisse der vorliegenden Arbeit widersprechen somit den Annahmen der CTML und der ECTML, die sich beide auf der mentalen Modellbildung vorgeschaltete Prozesse konzentrieren und somit einen Modalitätseffekt (in Abhängigkeit bestimmter Randbedingungen) nicht für die Transferleistung, sondern auch für die Erinnerung von Faktenwissen postulieren. Die Ergebnisse der vorliegenden Arbeit widersprechen jedoch nicht der Annahme, dass die fehlende zeitliche Kontiguität bei geschriebenem Text und Bild zu einem Modalitätseffekt hinsichtlich der Transferleistung führt, da diese in der vorliegenden Arbeit nicht erhoben wurde.

12.2 Welchen Einfluss nimmt der Textinhalt beim Lernen mit Multimedia?

Nach der CTML wurde kein Effekt des Textinhalts erwartet, da hier keine Annahmen zu einer spezifischen Verarbeitung unterschiedlicher Textinhalte gemacht werden. Nach der ECTML spielen Textinhalte hingegen aufgrund der Struktur des VSSP durchaus eine Rolle beim Lernen mit Multimedia. So wird nach der ECTML erwartet, dass räumliche Textinhalte den rVSSP belasten, während visuelle Textinhalte den vVSSP belasten. Da im rVSSP zusätzlich die Kontrolle von Blickbewegungen lokalisiert ist, wurde nach der ECTML ein Absinken der Leistung bei der Darbietung räumlicher Textinhalte in multimedialen Lernumgebungen im Vergleich zur Darbietung visueller Textinhalte oder abstrakter Textinhalte erwartet.

Diese Annahmen der ECTML konnten in den hier berichteten drei Experimenten größtenteils bestätigt werden: Während sich die Gruppen hinsichtlich der Erinnerung abstrakter Textinhalte zumindest auf einem $\alpha = .05$ Level nicht unterschieden – eine Ausnahme bildete Experiment 3 – zeigten sich hinsichtlich der Bilderinnerung und der Erinnerung visueller vs. räumlicher Textinhalte konsistent Unterschiede zwischen den Gruppen. Dabei ist insbesondere festzuhalten, dass sich die Unterschiede zwischen Lernern mit visuellen und räumlichen Textinhalten nicht nur hinsichtlich der Texterinnerung selber manifestierten, was unter Umständen auf die Textschwierigkeit zurückgeführt werden könnte, sondern dass auch die Bilderinnerung schlechter wurde, wenn räumliche Textinhalte dargeboten wurden

(vgl. Tabelle 76, Abschnitt II). Räumliche Textinhalte scheinen dem Lernen mit Multimedia also abträglich sein, was den Befunden von Scheiter und Schmidt-Weigand (2008) entspricht, die eine Aufhebung des Multimediaeffekts für Erinnerungsmaße fanden, wenn der Text ein hohes Ausmaß an räumlicher Information enthielt.

Nach der ECTML ist die Interferenz zwischen räumlichen Textinhalten und Bild in einer erhöhten Belastung des rVSSP begründet. Ob dieses tatsächlich der zugrundeliegende Prozess ist, konnte in den vorliegenden Experimenten leider nicht eindeutig geklärt werden. Problematisch ist, dass die Verarbeitung räumlicher Textinhalte im rVSSP nicht belegt werden konnte (vgl. Experiment 2). Stattdessen wurde eine Interferenz der räumlichen Zweitaufgabe bei der Erinnerung der Charakteristika sowohl mit visuellen als auch räumlichen Textinhalten beobachtet. Dies kann darauf zurückgehen, dass die Operationalisierung der visuellen Textinhalte in Experiment 1 und 2 auch über Formaspekte erfolgte. Trotz dieser aus theoretische Sich eindeutigen Zuordnung (vgl. Zimmer, 2010) von Formen zum vVSSP und nicht zum rVSSP, kann aufgrund eines empirischen Befunds von Glass et al. (1985) nicht ausgeschlossen werden, dass Formen auch im rVSSP verarbeitet werden, was wiederum die Interferenzen zwischen räumlicher Zweitaufgabe und visuellen Textinhalten erklären kann. In Experiment 3 wurden daher visuelle Textinhalte nicht mehr über Formaspekte, sondern ausschließlich über Farbinformationen operationalisiert, um so eine eindeutigere Trennung zwischen vVSSP und rVSSP zu erreichen. Jedoch konnte auch in Experiment 3 nicht nachgewiesen werden, dass ausschließlich räumliche Textinhalte im rVSSP verarbeitet werden, denn das Ausmaß an Blickbewegungen interagierte nicht mit der Verarbeitung räumlicher Textinhalte, was man erwartet hätte, wenn Blickbewegungen *und* räumliche Textinhalte im rVSSP verarbeitet werden. Es wurde jedoch auch kein genereller Einfluss des Ausmaß an Blickbewegungen auf die Verarbeitung visueller oder räumlicher Textinhalte beobachtet, was man erwarten würde, wenn sowohl visuelle als auch räumliche Textinhalte im rVSSP verarbeitet werden. Es ist insofern eher plausibel anzunehmen, dass das Ausmaß an Blickbewegungen nicht per se als negativer Faktor einzuschätzen ist, sondern dass Blickbewegungen als Indikator für Lese- bzw. Bildverarbeitungsstrategien auch dem Lernen förderlich sein können. Dies erklärt auch, warum der vorhergesagte Modalitätseffekt bei räumlichen Textinhalten nicht bestätigt werden konnte: Das

Ausmaß an Blickbewegungen hat keinen *zusätzlichen* Einfluss auf die negative Wirkung räumlicher Textinhalte, entweder weil Blickbewegungen per se nicht als negativ einzuschätzen sind oder aber weil das unterschiedliche Ausmaß an Blickbewegungen induziert über gesprochenen oder geschriebenen Text praktisch nicht bedeutsam ist. Demnach würden räumliche Textinhalte und Blickbewegungen – unabhängig von deren Ausmaß – generell zu einer erhöhten Belastung des rVSSP führen, was wiederum die negative Wirkweise räumlicher Textinhalte erklären könnte.

Eine alternative Erklärung für den negativen Einfluss der räumlichen Textinhalte auf die Text- und Bildverarbeitung stellt die Annahme dar, dass räumlichen Textinhalten eine höhere Schwierigkeit inhärent ist: Demnach würden räumliche Textinhalte die Text- und ggf. auch die Bildverarbeitung negativ beeinflussen, weil sie schwieriger zu verstehen sind. Gegen die Annahme unterschiedlicher Textschwierigkeiten von visuellen und räumlichen Textinhalten in den vorliegenden Experimenten sprechen jedoch sowohl objektive als auch subjektive Kriterien: Erstens entsprachen sich die Texte in Wortanzahl, Silbenanzahl und Satzlänge, was sich in einem vergleichbaren Flesch-Index für die Lesbarkeit der Texte (Flesch, 1948) äußerte. Zweitens zeigten sich im Blickverhalten der Lerner mit visuellen und räumlichen Textinhalten keine Unterschiede, d. h. sie widmeten Bild und Text ein vergleichbares Ausmaß an visueller Aufmerksamkeit. Drittens unterschieden sich die subjektiven Einschätzungen der Lernphase durch Lerner mit visuellen und räumlichen Textinhalten nicht voneinander, d. h. Lerner mit räumlichen Textinhalten stuften die Inhalte nicht als schwieriger ein, sie empfanden die Lernaufgabe nicht als anstrengender, sie konzentrierten sich nicht mehr auf den Text oder das Bild bzw. während der gesamten Lernphase, sie investierten nicht mehr Mühe in die Lernphase und sie strengten sich nicht mehr an, um das Lernmaterial zu verstehen. Alle diese Aspekte sprechen dafür, dass die Textschwierigkeit vergleichbar war und somit keine Konfundierung zwischen Textinhalt und Textschwierigkeit vorlag. Die Frage, ob die Textinhalte tatsächlich mit der Bildverarbeitung interferieren oder ob die Schwierigkeit der Texte einen negativen Einfluss auf die Bildverarbeitung hatte, sollte in zukünftigen Studien näher untersucht werden. Denkbar wäre z. B. die Wirkung der Textinhalte in einer Bedingung ohne Bilddarbietung zu vergleichen. Hier sollte sich die Leistung

hinsichtlich der Texterinnerung nicht unterscheiden, wenn die Schwierigkeit der Texte vergleichbar ist.

Unabhängig davon, welcher Prozess der negativen Wirkweise räumlicher Textinhalte zugrunde liegt, bleibt festzuhalten, dass der Einfluss des Textinhalts auf die Bildverarbeitung ein Aspekt ist, der von der CTML nicht vorhergesagt wird. Es ist durchaus denkbar, dass die Textinhalte ähnlich wie Zweitaufgaben mit der Bildverarbeitung interferieren können. Zukünftige Forschungsarbeiten sollten den Fokus dabei aber auch auf die generelle Unterscheidung zwischen nicht visuell-räumlichen und visuell-räumlichen Textinhalten legen. Es wird angenommen, dass sich hier die Wirkweise noch stärker nachweisen lässt, da der Nachweis einer strikten Verarbeitungstrennung von visuellen und räumlichen Textinhalten im VSSP problematisch ist. Eine Herausforderung bei dem Vergleich nicht visuell-räumlicher mit visuell-räumlichen Textinhalten besteht allerdings darin, dass sie den gleichen Überlappungsgrad mit dem Bild aufweisen müssen, um eine Konfundierung mit eben diesem ausschließen zu können. Des Weiteren ist es denkbar, dass nicht nur die Textinhalte, sondern ganz generell das Thema der Lernaufgabe einen Einfluss auf die Effektivität multimedialen Materials nimmt, da in Abhängigkeit des Themas die Notwendigkeit der Bilddarbietung unterschiedlich sein kann. Dies ist ein wichtiger Aspekt, der in zukünftigen Forschungsarbeiten ebenfalls stärker Berücksichtigung finden sollte.

12.3 Welchen Einfluss nimmt das Arbeitsgedächtnis beim Lernen mit Multimedia?

Sowohl die CTML als auch die ECTML gingen im Sinne der *Arbeitsgedächtnis-Hypothese* davon aus, dass hohe Kapazitäten der Komponenten des Arbeitsgedächtnisses zu besserer Leistung führen sollten. Diese Annahme konnte empirisch nur für wenige der Variablen bestätigt werden. In Experiment 2 nahm die PL Kapazität Einfluss auf die Erinnerung der verbalen Inhalte, was im Einklang mit der Annahme ist, dass in der PL verbales Material verarbeitet wird. Für den rVSSP gemessen über den Corsi Block konnte in keinem der drei Experimente ein genereller Einfluss auf das Lernergebnis nachgewiesen werden. Dies ist insofern überraschend, als dass in der vorliegenden Studie dem rVSSP eine entscheidende

Rolle zugeschrieben wurde, so dass eine Überlegenheit von Lernern mit hoher rVSSP Kapazität zu erwarten gewesen wäre.

Eine mögliche Erklärung für die sehr geringe Evidenz für eine Involviertheit des Arbeitsgedächtnisses bei Messung der individuellen Arbeitsgedächtniskapazität ist die Tatsache, dass die verwendeten Maße *Digit Span*, VPT und Corsi Block nur die Speicherkapazität der PL und des VSSP messen, nicht aber die Kapazität der zentralen Exekutive, die beim Lernen komplexerer Inhalte eine entscheidende Rolle spielt (vgl. Brunyé et al., 2006), vermutlich da sie die Verarbeitungsprozesse koordiniert. Eine Möglichkeit die Beteiligung zentral-exekutive Prozesse zu erfassen, bieten sogenannte komplexe Arbeitsgedächtnismaße, wie z. B. die *Reading Span Task* (Daneman & Carpenter, 1980) oder die *Operational Span Task* (Turner & Engle, 1989). Diesen Maßen ist gemein, dass sie gleichzeitig die Speicherkapazität eines der Subsysteme und die Ausführung zentraler Prozesse messen. Im Falle der *Reading Span Task* müssen die Probanden beispielsweise Wörter speichern, was die PL belasten sollte, und parallel dazu Sätze auf ihren Wahrheitsgehalt hin beurteilen, was zentrale Prozesse involvieren sollte. Solche komplexe Maße zeigten in verschiedenen Studien einen Bezug zum Lernen mit Multimedia (z. B. Austin, 2009; Doolittle & Mariano, 2008; Lusk et al., 2009), problematisch ist allerdings die eindeutige Erklärung der Effekte, da sie sowohl auf die Kapazität der involvierten Arbeitsgedächtniskomponente als auch auf die zentralen Prozesse als auch auf beides zurückgeführt werden können.

Des Weiteren ist nicht auszuschließen, dass bei der Verarbeitung angewandter Materialien neben exekutiven Funktionen weitere zusätzliche Einflussfaktoren, wie z. B. Vorwissen oder Lernstrategien, ins Spiel kommen. Es ist denkbar, dass solche Faktoren den Beitrag der einzelnen Arbeitsgedächtniskomponenten, der sich insbesondere im Speichern von Informationen über einige Sekunden zeigen sollte, überdecken. Grundsätzlich stellt sich somit die Frage, ob Speicherprozesse auf Ebene der PL und des VSSP auf die Lernleistung mit komplexen Materialien einen praktisch relevanten Einfluss nehmen können.

Trotz dieser Problematik ist es wichtig, dass in zukünftigen Forschungsarbeiten viel stärker individuelle Unterschiede der Arbeitsgedächtniskapazität berücksichtigt werden, da sich in Abhängigkeit der individuellen Kapazität unterschiedliche Darbietungsweisen von Text und Bild als effektiv herausstellen können. Während es in anderen Forschungsbereichen bereits seit langer Zeit üblich ist,

Arbeitsgedächtnismaße gezielt einzusetzen (vgl. z. B. Alloway, Gathercole, Adams & Willis, 2005; Pickering, 2006), zeigt sich hier in Multimediaforschung ein Mangel an entsprechender Forschungsaktivität.

Während durch die Messungen der individuellen Arbeitsgedächtniskapazität kein genereller Einfluss der Arbeitsgedächtniskomponenten auf das Lernergebnis nachgewiesen werden konnte, konnte mit Hilfe der räumlichen Zweitaufgabe ein Einfluss des rVSSP auf die Bildverarbeitung gezeigt werden (vgl. Tabelle 77, Abschnitt II). Dies spricht dafür, dass die Maße der *Digit Span*, des Corsi Blocks und des VPTs zu störanfällig sind, um einen Einfluss der Arbeitsgedächtniskapazität nachzuweisen, während prozessnahe Maße wie Zweitaufgaben eher dazu geeignet sind, Arbeitsgedächtnisprozesse widerzuspiegeln. Des Weiteren interferierte die räumliche Zweitaufgabe nicht mit der Verarbeitung abstrakter Textinhalte, wohl aber mit der Verarbeitung visueller und räumlicher Textinhalte. Dies zeigt, dass Textinhalte, wenn sie visuelle und räumliche Informationen enthalten, im VSSP verarbeitet werden.

12.4 Stärken und Schwächen der empirischen Studien

Nach der Zusammenfassung und Diskussion der Hypothesen sollen im Folgenden kurz die Schwächen aber auch die die Stärken der durchgeführten Experimente aufgeführt werden.

Einschränkungen der empirischen Studien

Die Darbietung unterschiedlicher Textinhalte brachte methodische Probleme mit sich, da die Lerner unterschiedliche Inhalte lernen mussten.

Es ergab sich zum Einen die Problematik der Messung der gelernten Inhalte. In der vorliegenden Studie wurde davon abgesehen den Lernern unterschiedliche Items darzubieten, da dieses eine zusätzliche Variation zwischen den Gruppen bedeutet hätte. Daher beantworteten alle Lerner die gleichen Items. Aufgrund der unterschiedlichen Inhalte der Lernphase mussten diese Items jedoch in Abhängigkeit des Textinhalts zu unterschiedlichen Indizes zusammengefasst werden. So waren z. B. einige Items für Lerner mit räumlichen Textinhalten mit Hilfe von Text *und* Bild beantwortbar, da die Information im Text genannt wurde und im Bild zu sehen war, während Lerner mit visuellen Textinhalten diese Items nur mit Hilfe des Bildes

beantworten konnten, da die Information nicht im Text genannt wurde. Somit resultierten für die Bilderinnerung Indizes, in die in Abhängigkeit des Textinhalts unterschiedliche Items eingingen. Die Lösung für dieses Problem stellt ein generelles Leistungsmaß dar, in das alle Items eingehen, ohne eine Differenzierung in die zugrundeliegende Quelle zur Itembeantwortung. Mit einem solchen generellen Maß wäre es allerdings nicht möglich gewesen, die von der ECTML vorhergesagten spezifischen Interferenzen zwischen Textinhalt und Text- bzw. Bildverarbeitung aufzuzeigen.

Ein weiteres Problem der Darbietung unterschiedlicher Textinhalte lag in der möglichen Konfundierung mit unterschiedlichen Schwierigkeitsgraden der Texte. Zwar sprechen Kriterien wie die Textstruktur, der Flesch-Index, das Blickverhalten sowie die subjektiven Einschätzungen der Lerner gegen diese alternative Erklärung, es kann jedoch nicht nachgewiesen werden, dass visuellen oder räumlichen Inhalten nicht von Natur aus eine höhere Schwierigkeit inhärent ist.

Eine weitere Einschränkung der Studie ist die Vergleichbarkeit des verwendeten Materials mit den üblicherweise in der Multimedialiteratur verwendeten Materialien. Während in Studien zum Multimedialernen typischerweise kausale Inhalte vermittelt werden, wurden in der vorliegenden Studie konzeptuelle Inhalte vermittelt. Dieses Vorgehen wurde gewählt, da hier (a) die Operationalisierung vergleichbarer visueller und räumlicher Inhalte möglich war, (b) die Erinnerung von Faktenwissen differenziert erfasst werden konnte und (c) die Bildung eines integrierten mentalen Modells hier weniger notwendig erscheint, so dass zeitliche Kontiguität eine geringere Rolle spielen sollte. Kritisch ist zum letzten Punkt allerdings anzumerken, dass es vereinzelt empirische Evidenz dafür gibt, dass auch bei eher als konzeptuell einzuschätzenden Inhalten ein mentales Modell gebildet wird (z. B. bei Szenebeschreibungen, Gentner & Loftus, 1979), so dass der Einfluss zeitlicher Kontiguität nicht vollkommen ausgeschlossen werden kann. Allerdings sollte sich dieser explizit bei der Transferleistung, nicht aber bei der in den vorliegenden Studien erfassten Erinnerungsleistung zeigen. Dass auch die räumliche Kontiguität keine Rolle in der vorliegenden Studie spielte, zeigt sich darin, dass keine generellen Modalitätseffekte für die bildbezogenen Erinnerungsmaße beobachtet wurden.

Eine weitere Einschränkung der Vergleichbarkeit mit anderen Studien aus der Multimedialiteratur stellen die verwendeten AVn dar. Während die

Erinnerungsleistung ein in der Multimediaforschung übliches Maß darstellt, werden prozessnahe Maße in der Multimedialiteratur üblicherweise stark vernachlässigt. Stattdessen wird oftmals die Transferleistung gemessen, in der Annahme, dass für erfolgreichen Transfer die Bildung eines integrierten mentalen Modells die Voraussetzung ist. In der vorliegenden Studie wurde die Transferleistung nicht gemessen, da davon ausgegangen wurde, dass sich die Prozesse auch in der Erinnerungsleistung widerspiegeln sollten, und da hinsichtlich der Transferleistung eine Konfundierung mit der zeitlichen Kontiguität bei gesprochenem Text und Bild nicht ausgeschlossen werden konnte. Stattdessen wurde in der vorliegenden Studie der Fokus auf eine differenzierte Erfassung der Erinnerungsleistung sowie die Erfassung prozessnaher Maße gelegt, da diese Maße häufig auch in Studien verwendet wurden, die den Einfluss des Textinhalts auf die Leistung untersuchten (vgl. Abschnitt 4.2.3). Dies schränkt die Vergleichbarkeit mit typischen Studien der Multimedialiteratur ein. Daher sollte in Folgeuntersuchungen die Gültigkeit der Befunde auch für kausale Inhalte sowie Transferleistungen untersucht werden.

Stärken der empirischen Studien

Ein erklärtes Ziel der vorliegenden Forschungsarbeit war eine systematische Vorgehensweise bei der Prüfung der Hypothesen. Diese Systematik zeigte sich insbesondere bei der Materialentwicklung und unterscheidet die vorliegende Arbeit von üblicherweise in der Multimedialiteratur durchgeführten Studien. Die Systematik bei der Materialentwicklung äußerte sich insbesondere in den folgenden Punkten:

- Die Entwicklung artifiziellen Materials um Vorwissenseffekte auszuschließen, die den Modalitätseffekt beeinflussen können (z. B. Kalyuga, 2005).
- Die Entwicklung zweier vergleichbarer Lernsets mit unterschiedlichen Lernobjekten, um die Generalisierbarkeit der Befunde zu prüfen.
- Die systematische Entwicklung der AVn unter Berücksichtigung verschiedener Dimensionen, die Einfluss auf das Testergebnis nehmen konnten, nämlich das Symbolsystem der Itemdarbietung, die Iteminhalte und die Kongruenz zur Lernphase.
- Die Darbietung der Items unter verschiedenen Instruktionen, um die textbezogene Erinnerung und die bildbezogene Erinnerung differenziert erfassen zu können.

Insbesondere die Differenzierung in textbezogene und bildbezogene Erinnerungsleistung findet im Rahmen von Multimediauntersuchungen kaum Beachtung. Sie ist aber unerlässlich, um die dem Lernen mit Multimedia zugrundeliegende Prozesse zu untersuchen. So zeigten z. B. Levie und Lentz (1982) in ihrer Übersicht über Multimediastudien, dass der Multimediaeffekt nur dann nachweisbar war, wenn illustrierte Informationen abgefragt wurden.

Eine weitere Stärke der vorliegenden Studie liegt darin, dass visuelle und räumliche Textinhalte den gleichen Überlappungsgrad zum Bild aufwiesen, so dass die Redundanz zwischen Text und Bild vergleichbar war. Hinsichtlich des Einflusses der Redundanz von Text und Bild gehen sowohl die Theorie der dualen Kodierung (Paivio, 1990) sowie die CTML (Mayer, 1999) davon aus, dass die Darbietung redundanter verbaler und bildhafter Informationen dem Lernen zuträglich ist. Durch die systematische Entwicklung des Materials war sichergestellt, dass das schlechtere Abschneiden von Lernern mit räumlichen Textinhalten nicht auf niedrigere Redundanz zum Bild zurückging, was in anderen empirischen Studien teilweise nicht kontrolliert wurde (z. B. Scheiter & Schmidt-Weigand, 2008).

Eine weitere Stärke der vorliegenden Forschungsarbeit ist der Ansatz, Befunde aus der Grundlagenforschung auf das Lernen mit Multimedia zu übertragen und empirisch zu überprüfen. Insbesondere neuere Befunde der Grundlagenforschung, wie z. B. hinsichtlich der Struktur des VSSP, werden häufig in der angewandten Forschung nicht wahrgenommen, obwohl sie wichtige Implikationen für die Anwendung haben können. Stattdessen werden in der Anwendung Effekte verschiedener Text-Bilddarstellungen getestet, ohne dass ein theoretischer Bezug erkennbar wäre. Dadurch wird es schwierig Effekte eindeutig zu interpretieren. Andererseits profitiert auch die Grundlagenforschung von der Übertragung ihrer Annahmen in angewandte Kontexte, da so die externe Validität erhöht werden kann.

Eine weitere Stärke der vorliegenden Untersuchungen liegt in der Anwendung prozessnaher Methoden, die in der Grundlagenforschung häufig verwendet werden, wie z. B. dem Zweitaufgabenparadigma, der Messung der individuellen Arbeitsgedächtniskapazität und der Analyse prozessnaher Daten. Erst durch die Anwendung dieser prozessnahen Methoden kann auch eine Absicherung prozessnaher Erklärungen erfolgen, die auf der Struktur des Arbeitsgedächtnisses beruhen.

Der Austausch zwischen Grundlagenforschung und angewandter Forschung ist somit vielversprechend und sollte auch in Zukunft weiterhin betrieben werden.

13 Fazit und Ausblick

Ein Charakteristikum der CTML ist, dass sie allgemeine Befunde der Kognitionspsychologie, nämlich das Drei-Speicher-Modell von Atkinson und Shiffrin (1968), die Theorie der dualen Kodierung von Paivio (1990) und das Arbeitsgedächtnismodell nach Baddeley (1992), nutzt, um daraus ihre Annahmen bezüglich der Verarbeitung multimedialen Materials abzuleiten. Dieser Rückgriff auf kognitive Grundlagen zur Bildung einer Theorie zur Verarbeitung multimedialen Materials ist durchaus sinnvoll. Problematisch ist allerdings zum Einen die teilweise fehlerhafte Interpretation der kognitiven Grundlagentheorien, sowie zum Anderen die oberflächliche Anwendung der kognitiven Grundlagentheorien auf das Lernen mit Multimedia bis hin zur völligen Vernachlässigung bestimmter Aspekte.

In der vorliegenden Forschungsarbeit wurde daher die ECTML entwickelt, in der die fehlerhaften Annahmen der CTML bezüglich der Verarbeitung geschriebenen Texts korrigiert wurden und die oberflächliche Adaption des VSSP spezifiziert wurde. So wird nach der ECTML die Verarbeitung geschriebenen Texts von Beginn an in der PL verortet, und es wird nicht angenommen, dass Abbilder des geschriebenen Texts im VSSP verarbeitet werden. Unabhängig von ihrer Textmodalität sollten Texte somit die PL in gleichem Ausmaß von Anfang an belasten. Die in der ECTML getroffenen Spezifizierungen des VSSP sehen vor, dass eine Unterteilung des VSSP in eine visuelle und eine räumliche Komponente vorzunehmen ist und dass im VSSP nicht nur bildhafte Informationen verarbeitet werden, sondern dass verbale Beschreibungen visueller oder räumlicher Konfigurationen das VSSP ebenfalls belasten können. Des Weiteren findet die Steuerung von Blickbewegungen im rVSSP statt. Es sei angemerkt, dass demnach geschriebene Texte den rVSSP stärker belasten sollten als gesprochene Texte, da erstere mit einem höheren Ausmaß an Blickbewegungen assoziiert sind. Es ist aber wichtig festzuhalten, dass (a) auch bei der Darbietung gesprochener Texte der rVSSP belastet wird, da die Lerner auch beim Hören Blickbewegungen durchführen, und dass sich (b) hinsichtlich der Belastung der PL bei geschriebenem und gesprochenem Text nach der ECTML keine Unterschiede ergeben. Die empirischen Befunde

zeigen, dass eine Erweiterung der CTML durchaus sinnvoll ist. So scheinen die Textinhalte eine wichtige Rolle hinsichtlich der Effektivität der Bildverarbeitung zu spielen, was auf Interferenzen im VSSP zurückgeführt werden kann. Es stellt sich aber rückblickend die Frage, inwiefern eine strikte Trennung in ein vVSSP und ein rVSSP praktikabel ist. Die Interferenz der Zweitaufgabe sowohl mit visuellen als auch räumlichen Textinhalten nicht aber mit abstrakten Textinhalten legt nahe, dass die Unterschiede zwischen visuellen und räumlichen verbalen Beschreibungen fließend sein können. Dieser Befund entspricht den Ergebnissen von Pazzaglia und Cornoldi (1999), die ebenfalls einen Einfluss der räumlichen Zweitaufgabe auf die Verarbeitung visueller Textinhalte fanden (vgl. Abschnitt 4.2.3). Insofern erscheint es zunächst sinnvoller, eine Trennung des VSSP in Komponenten zur Verarbeitung visuell-räumlicher Textinhalte, visuell-räumlicher Bildinhalte sowie zur Kontrolle von Augenbewegungen vorzunehmen.

In der vorliegenden Forschungsarbeit lag der Fokus insbesondere auf der Struktur des VSSP. Die oberflächliche Anwendung der kognitiven Grundlagentheorien durch die CTML lässt sich allerdings auch anhand anderer Strukturen aufzeigen, z. B. der zentralen Exekutive oder dem episodischen Buffer. Während der episodische Buffer keinerlei Verwendung in der Theorie findet, wird die zentrale Exekutive in der CTML nur am Rande erwähnt. Demnach fällt in ihre Verantwortlichkeit die Selektion, Organisation und Integration der multimedialen Information (Mayer, 2005a). Eine weitere Spezifizierung dieses zentralen Konzepts erfolgt aber nicht. Dies ist insofern erstaunlich, als dass die Prozesse, die nach Smith und Jonides (1999) über die zentrale Exekutive gesteuert werden (*Attention and Inhibition, Task Management, Planning, Monitoring, Coding*; vgl. Abschnitt 2.2.1), bei der Auswahl von Text- und Bildinformationen eine wichtige Rolle spielen sollten. Auch empirisch konnte die Involviertheit der zentralen Exekutive beim Lernen mit Multimedia bereits nachgewiesen werden (vgl. Brunyé et al., 2006). Ebenso stellt der episodische Buffer eine für das Lernen mit Multimedia vielversprechende Struktur dar. Baddeley integrierte den episodischen Buffer in sein Modell, um das *binding problem*, also die Frage, wo verbale und visuell-räumliche Informationen im Arbeitsgedächtnis miteinander integriert werden, zu lösen. Dieser Integrationsprozess stellt neben der dualen Kodierung eine zentrale Annahme multimedialen Lernens in der CTML dar. Die bisherige Konzeption der CTML bietet hierfür aber keine entsprechende Komponente. Es erscheint somit sinnvoll die

CTML um die Strukturen der zentralen Exekutive sowie des episodischen Buffers zu erweitern und aus dieser Erweiterung ableitbare Vorhersagen empirisch zu überprüfen.

Generell erscheint es notwendig die empirische Überprüfung der kognitiven Grundlagen multimedialen Lernens viel stärker als bisher in den Fokus zu rücken. Die vorliegende Arbeit stellt einen ersten Schritt in diese Richtung dar, aber zukünftige Forschungsarbeiten sind notwendig, um die empirische Vorgehensweise zu optimieren. Zur Messung der Involviertheit des Arbeitsgedächtnisses sollten das Zweitaufgabenparadigma sowie individuelle Arbeitsgedächtnismaße verwendet werden. Insbesondere hinsichtlich der Messung individueller Unterschiede sollte dabei nicht nur die Speicherkapazität der einzelnen Komponenten berücksichtigt werden, sondern auch komplexe Arbeitsgedächtnismaße, wie z. B. die *Reading Span Task* (Daneman & Carpenter, 1980) oder die *Operational Span Task* (Turner & Engle, 1989). Ein weiterer bisher vernachlässigter Aspekt ist die Differenzierung verschiedener Lernmaße. Gerade im Hinblick auf multimediales Lernen erscheint es sinnvoll zwischen Texterinnerung, Bilderinnerung sowie ggf. einem integrierten Maß zu unterscheiden. Prozessnahe Maße wie Blickbewegungsdaten (vgl. Van Gog & Scheiter, 2010), Laut-Denken-Protokolle (z. B. Koning, Tabbers, Rikers & Paas, 2010) oder Antwortlatenzen bieten zudem die Möglichkeit die zugrundeliegenden Prozesse zu untersuchen.

So wichtig die Anwendung der Grundlagenforschung auf die Multimediaforschung auch ist, so bringt sie auch Herausforderungen mit sich. Eine Herausforderung besteht darin, dass es für die empirische Überprüfung theoretischer Konzepte teilweise noch keine adäquaten Methoden zur Verfügung stehen. So gibt es z. B. bis heute kein etabliertes Verfahren zur Untersuchung des episodischen Buffers. Eine weitere Problematik besteht darin, dass auch die Grundlagenforschung selber ständig im Wandel ist. So gibt es zum Einen alternative Konzeptionen des Arbeitsgedächtnisses (für eine Übersicht vgl. Myiake & Shah, 1999), zum Anderen gibt es auch teilweise widersprüchliche Befunde hinsichtlich des Arbeitsgedächtnismodells nach Baddeley (1992). So schlagen einige Autoren vor den VSSP in eine aktive und eine passive Komponente zu unterteilen (z. B. Cornoldi, Rigoni, Venneri & Vecchi, 2000; Vecchi & Cornoldi, 1999), während andere Autoren sogar von einer Dreiteilung des VSSP in (a) eine visuelle Komponente, (b) eine simultan-räumliche Komponente zur Verarbeitung von simultan präsentierten

räumlichen Informationen und (c) eine sequentiell-räumlichen Komponente, zur Erinnerung von nacheinander präsentierten räumlichen Informationen ausgehen (Mammarella, Pazzaglia & Cornoldi, 2008). Diese unterschiedlichen Annahmen machen deutlich, dass eine Anforderung an die angewandte Forschung darin besteht, den *state-of-the-art* und die damit verbundenen Zweifel und Schwierigkeiten zu akzeptieren, da sonst keine Übertragung der Grundlagenforschung auf die angewandte Forschung möglich ist.

Die Erforschung der kognitiven Grundlagen des multimedialen Lernens bildet die Voraussetzung für eine adäquate theoretische Beschreibung des Lernens mit Multimedia. Nur durch eine Überprüfung der impliziten Annahmen zu Prozessen beim Lernen mit Multimedia können fehlerhafte Annahmen korrigiert werden und korrekte Empfehlungen hinsichtlich der Text-Bildgestaltung abgeleitet werden. Erst wenn nicht mehr die Effekte selber, sondern ihre Ursachen in den Fokus rücken, ist der von Mayer (2009) postulierte Anspruch erfüllt, dass die Multimediaforschung nicht nur evidenzbasiert sondern auch theoriebasiert erfolgen sollte.

14 Literaturverzeichnis

- Allen, R. H., Hitch, G. J. & Baddeley, A. D. (2009). Cross-modal binding and working memory. *Visual Cognition*, 17, 83-102. doi:10.1080/13506280802281386
- Alloway, T. P., Gathercole, S. E., Adams, A.-M. & Willis, C. (2005). Working memory abilities in children with special educational needs. *Educational & Child Psychology*, 22, 56-67.
- Andrade, J. (2001). *Working memory in perspective*. Hove, England: Psychology Press.
- Andrade, J., Kemps, E., Werniers, Y., May, J. & Szmalec, A. (2002). The insensitivity of visual short-term memory to irrelevant visual information. *The Quarterly Journal of Experimental Psychology*, 55A, 753-774.
- Anglin, G. J., Towers, R. L. & Levie, W. (1996). Visual message design and learning: The role of static and dynamic illustrations. In D. Jonasson (Ed.), *Handbook of research for educational communications and technology* (pp. 755-794). New York: Macmillan.
- 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* (Vol. 2) (pp. 89-195). New York: Academic Press.
- Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology*, 94, 416-427. doi:10.1037/0022-0663.94.2.416
- Austin, K. A. (2009). Multimedia learning: Cognitive individual differences and *Digit Span* lay design techniques predict transfer learning with multimedia learning modules. *Computers & Education*, 53, 1339-1354. doi:10.1016/j.compedu.2009.06.017
- Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *The Quarterly Journal of Experimental Psychology*, 18A, 302-309.
- Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Oxford University Press.
- Baddeley, A. D. (1992). Working memory. *Science*, 225, 556-559. doi:10.1126/science.1736359

- Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, *49*, 5-28. doi:10.1080/027249896392784
- Baddeley, A. D. (1999). *Essentials of human memory*. Hove: Psychology Press.
- Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*, 417-423. doi:10.1016/S1364-6613(00)01538-2
- Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*, 198-208. doi:10.1016/S0021-9924(03)00019-4
- Baddeley, A. D. (2006). Working memory: An overview. In S. Pickering (Ed.), *Working memory and education* (pp. 1-31). New York: Academic Press. doi:10.1016/B978-012554465-8/50003-X
- Baddeley, A. D. (2007). Notes and comments: The phonological loop and the irrelevant speech effect: Some comments on Neath (2000). *Psychonomic Bulletin & Review*, *7*, 544-549.
- Baddeley, A. D. & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129*, 126-145. doi:10.1037/0096-3445.129.1.126
- Baddeley, A. D., Elridge, M. & Lewis, V. (1981). The role of subvocalisation in reading. *Quarterly Journal of Experimental Psychology*, *33*, 439-454.
- Baddeley, A. D., Emslie, H., Kolodny, J. & Duncan, J. (1998). Random generation and the executive control of working memory. *The Quarterly Journal of Experimental Psychology*, *51A*, 819-852. doi:10.1080/027249898391413
- Baddeley, A. D., Gathercole S. E. & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158-173. doi:10.1037/0033-295X.105.1.158
- Baddeley, A. D., Grant, S., Wight, E. & Thomson, N. (1973). Imagery and visual working memory. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V* (pp. 205-217). London: Academic Press.
- Baddeley, A. D. & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.
- Baddeley, A. D., Thomson, N. & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behaviour*, *14*, 575-589. doi:10.1016/S0022-5371(75)80045-4

- Baggett, P. & Ehrenfeucht, A. (1983). Encoding and retaining information in the visuals and verbals of an educational movie. *Educational Communications and Technology Journal*, 31, 23-32.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617-645. doi:10.1146/annurev.psych.59.103006.093639
- Beech, J. R. (1977). Effect of selective visual interference on visualization. *Perception and Motor Skills*, 45, 951-954.
- Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler* (3. überarbeitete Aufl.). Berlin, Germany: Springer.
- Brooks, L. R. (1967). The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology*, 19A, 289-299.
- Brooks, L. R. (1970). An extension of the conflict between visualization and reading. *The Quarterly Journal of Experimental Psychology*, 22A, 91-96.
- Brünken, R., Plass, J. L. & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual task methodology: Auditory load and modality effects. *Instructional Science*, 32, 115-132. doi:10.1023/B:TRUC.0000021812.96911.c5
- Brünken, R., Seufert, T. & Zander, S. (2005). Förderung der Kohärenzbildung beim Lernen mit multiplen Repräsentationen. *German Journal of Educational Psychology*, 19, 61-75.
- 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. doi:10.1027//1618-3169.49.2.109
- Brunyé, T. T., Taylor, H. A., Rapp, D. N. & Spiro, A. B. (2006). Learning procedures: The role of working memory in multimedia learning experiences. *Applied Cognitive Psychology*, 20, 917-940. doi:10.1002/acp.1236
- Bryant, D. J., Tversky, B. & Franklin, N. (1992). Internal and external spatial frameworks for representation described scenes. *Journal of Memory and Language*, 31, 74-98. doi:10.1016/0749-596X(92)90006-J
- Byrne, M. & Curtis, R. (2000). Designing health communication: Testing the explanations for the impact of communication medium on effectiveness. *British Journal of Health Psychology*, 5, 189- 199. doi:10.1348/135910700168856

- Carney, R. N. & Levin, R. N. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14, 5-26. doi:10.1023/A:1013176309260
- Chung, K. K. H. (2007). What effect do mixed sensory modes instructional formats have on both novice and experienced learners of Chinese characters? *Learning and Instruction*, 18, 96-108. doi:10.1016/j.learninstruc.2007.01.001
- Clark, R. E. & Feldon, D. F. (2005). Five common but questionable principles of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (97-116). Cambridge: University Press.
- Cocchini, G., Logie, R. H., Della Sala, S., McPherson, S. & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory and Cognition*, 30, 1086-1095.
- Colle, H. A. & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior*, 15, 17-31. doi:10.1016/S0022-5371(76)90003-7
- Cornoldi, C., Rigoni, F., Venneri, A. & Vecchi, T. (2000). Passive and active processes in visuo-spatial memory: Double dissociation in developmental learning disabilities. *Brain and Cognition*, 43, 117-120.
- Courtney, S. M., Ungerleider, L. G., Keil, K. & Haxby, J. V. (1996). Objects in spatial working memory activate separate neural systems in human cortex. *Cerebral Cortex*, 6, 39-49. doi:10.1093/cercor/6.1.39
- Craig, S. D., Gholson, B. & Driscoll, D. M. (2002). Animated pedagogical agents in educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, 94, 428-434. doi:10.1037/0022-0663.94.2.428
- Daneman, M. & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Learning and Verbal Behavior*, 19, 450-466. doi:10.1016/S0022-5371(80)90312-6
- Danks, J. H. & End, L. J. (1987). Processing strategies for reading and listening. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 271-294). San Diego: Academic press.
- Darling, S., Della Sala, S. & Logie, R. H. (2007). Behavioural evidence for separating components of visuo-spatial working memory. *Cognitive Processing*, 8, 175-181. doi:10.1007/s10339-007-0171-1

- Darling, S., Della Sala, S., Logie, R. H. & Canagello, A. (2006). Neuropsychological evidence for separating components of visuo-spatial working memory. *Journal of Neurology*, 253, 176-180. doi:10.1007/s00415-005-0944-3
- De Beni, R. & Moè, A. (1997). Learning from texts or lectures: Loci mnemonics can interfere with reading but not with listening. *European Journal of Cognitive Psychology*, 9, 401-415. doi:10.1080/713752565
- De Beni, R., Pazzaglia, F., Gyselinck, V. & Meneghetti, C. (2005). Visuospatial working memory and mental representation of spatial description. *European Journal of Cognitive Psychology*, 17, 77-95. doi:10.1080/09541440340000529
- Della Sala, S., Gray, C., Baddeley, A. D., Allamano, N. & Wilson, L. (1999). Pattern spans: A tool for unwelding visuo-spatial memory. *Neuropsychologia*, 37, 1189-1199. doi:10.1016/S0028-3932(98)00159-6
- Della Sala, S., Gray, C., Baddeley, A. D. & Wilson, L. (1997). *Visual Pattern Test: a test of short-term visual recall*. London: Harcourt Assessment.
- Deyzack, E., Logie, R. H. & Denis, M. (2006). Visuospatial working memory and the processing of spatial description. *British Journal of Psychology*, 97, 217-243. doi:10.1348/000712605X67484
- Doolittle, P. E. & Mariano, G. J. (2008). Working memory capacity and mobile multimedia learning environments: Individual differences in learning while mobile. *Journal of Educational Multimedia and Hypermedia*, 17, 511-530.
- Dunsworth, Q. & Atkinson, R. K. (2007). Fostering multimedia learning of science: Exploring the role of an animated agent's image. *Computers & Education*, 49, 677-690. doi:10.1016/j.compedu.2005.11.010
- Eddy, J. & Glass, A. (1981). Reading and listening to high and low imagery sentences. *Journal of Verbal Learning and Verbal Behavior*, 20, 333-345. doi:10.1016/S0022-5371(81)90483-7
- Farah, M. J., Hammond, K. M., Levine, D. N. & Calvanio, R. (1988). Visual and spatial mental imagery: dissociable systems of representation. *Cognitive Psychology*, 20, 439-462. doi:10.1016/0010-0285(88)90012-6
- Farmer, E. F., Berman, J. V. F. & Fletcher, Y. L. (1986). Evidence for a visuo-spatial scratch-pad in working memory. *Quarterly Journal of Experimental Psychology*, 38, 675-688.
- Ferguson, E. L. & Hegarty, M. (1994). Properties of cognitive maps constructed from texts. *Memory & Cognition*, 22, 455-473.

- Field, A. (2009). *Discovering Statistics Using SPSS (and sex and drugs and rock'n'roll)*. Third edition. Los Angeles: Sage.
- Fischer, M. H. & Zwaan, R. A. (2008). Embodied language – A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology*, *61*, 825-850. doi:10.1080/17470210701623605
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*, 221-233. doi:10.1037/h0057532
- Furnham, A., De Siena, S. & Gunter, B. (2002). Children's and adults recall of children's news stories in both print and audio-visual presentation modalities. *Applied Cognitive Psychology*, *16*, 191-210. doi:10.1002/acp.777
- Gathercole, S. E. & Baddeley, A. D. (1993). *Working memory and language*. Hove: Lawrence Erlbaum Associates Ltd.
- Gentner, D. & Loftus, E. F. (1979). Integration of verbal and visual information as evidenced by distortions in picture memory. *American Journal of Psychology*, *92*, 363-375. doi:10.2307/1421930
- Ginns, P. (2005). Meta-analysis of the modality effect. *Learning & Instruction*, *15*, 313-331. doi:10.1016/j.learninstruc.2005.07.001
- Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, *16*, 511-525. doi:10.1016/j.learninstruc.2006.10.001
- Glass, A., Millen, D., Beck, L. & Eddy, J. (1985). Representations of images in sentence verification. *Journal of Memory and Language*, *24*, 442-465. doi:10.1016/0749-596X(85)90039-7
- 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. doi:10.1016/0749-596X(92)90008-L
- Griffin, M. M. & Robinson, D. H. (2000). Role of mimeticism and spatiality in textual recall. *Contemporary Educational Psychology*, *25*, 124-149. doi:10.1006/ceps.1999.0997
- Griffin, M. M. & Robinson, D. H. (2005). Does spatial and visual information in text facilitate text recall? Reconsidering the conjoint retention hypothesis. *Educational Technology and Research and Development*, *53*, 23-36. doi:10.1007/BF02504855

- Gyselinck, V., Cornoldi, C., Dubois, V., De Beni, R. & Ehrlich, M. F. (2002). Visuospatial memory and phonological loop in learning from multimedia. *Applied Cognitive Psychology, 16*, 665-685. doi:10.1002/acp.823
- Gyselinck, V., De Beni, R., Pazzaglia, F., Meneghetti, C. & Mondoloni, A. (2007). Working memory components and imagery instruction in the elaboration of a spatial mental model. *Psychological Research, 71*, 373-382. doi:10.1007/s00426-006-0091-1
- Gyselinck, V., Ehrlich, M. F., Cornoldi, C., De Beni, R. & Dubois, V. (2000). Visuospatial working memory in learning from multimedia systems. *Journal of Computer Assisted Learning, 16*, 166-176. doi:10.1046/j.1365-2729.2000.00128.x
- Gyselinck, V., Jamet, E. & Dubois, V. (2008). The role of working memory components in multimedia comprehension. *Applied Cognitive Psychology, 22*, 353-374. doi:10.1002/acp.1411
- Hamilton, C. J., Coates, R. O. & Heffernan, T. (2003). What develops in visuospatial working memory development? *European Journal of Cognitive Psychology, 15*, 43-69. doi:10.1080/09541440303597
- Harskamp, E. G., Mayer, R. E. & Suhre, C. (2007). Does the modality principle for multimedia learning apply to science classrooms? *Learning and Instruction, 17*, 465-477. doi:10.1016/j.learninstruc.2007.09.010
- Härting, C., Markowitsch, H. N., Neufeld, H., Calabrese, P., Deisinger, K. & Kessler, J. (2000). *WMS-R. Wechsler Gedächtnistest – Revidierte Fassung. Deutsche Adaptation der revidierten Fassung der Wechsler Memory Scale. Manual*. Bern: Huber.
- Hauk, O., Johnsrude, I. & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*, 301-307. doi:10.1016/S0896-6273(03)00838-9
- Hecker, R. & Mapperson, B. N. (1997). Dissociation of visual and spatial processing in working memory. *Neuropsychologia, 35*, 599-603. doi:10.1016/S0028-3932(96)00106-6
- Hegarty, M. (1992). The mechanics of comprehension and comprehension of mechanics. In K. Rayner (Ed.), *Eye movements and visual cognition: scene perception and reading*. (pp. 428-443). New York: Springer.

- Holmqvist, K. *Eye-tracking data and dependent variables*. Manuskript in Vorbereitung.
- Hörnig, R., Claus, B. & Eyfert, K. (2000). In search for an overall organizing principle in spatial mental models. In S. O. Nualláin (Ed.), *Spatial Cognition - Foundations and application, Advances in consciousness research*, 26 (pp. 69-81). Amsterdam: John Benjamins.
- Hörnig, R., Eyfert, K. & Claus, B. (1999). Egozentrische Inferenz von Objektpositionen beim Lesen und Hören. *Zeitschrift für Experimentelle Psychologie*, 46, 140-151.
- Jahn, G. (2004). Three turtles in danger: Spontaneous construction of causally relevant spatial situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 969-987. doi:10.1037/0278-7393.30.5.969
- Jeung, H., Chandler, P. & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology*, 17, 329-343. doi:10.1080/0144341970170307
- Kalyuga, S. (2005). Prior knowledge in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 325-338). Cambridge: University Press.
- Kalyuga, S., Chandler, P. & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, 92, 126-136. doi:10.1037/0022-0663.92.1.126
- Kessels, R. P. C., Postma, A. & De Haan, E. H. F. (1999). P and M channel-specific interference in the what and where pathway. *Neuroreport*, 10, 3765-3767. doi:10.1097/00001756-199912160-00008
- Kessels, R. P. C., Van Zandvoort, M. J. E., Postma, A., Kappelle, L. J. & De Haan, E. H. F. (2000). The Corsi Block-Tapping Task: Standardization and normative data. *Applied Neuropsychology*, 7, 252-258. doi:10.1207/S15324826AN0704_8
- Kintsch, W., Healy, A., Hegarty, M., Pennington, B. & Salthouse, T. (1999). Models of working memory: Eight questions and some general answers. In A. Miyake & P. Shah (Eds.), *Models of Working Memory* (pp. 412-441). Cambridge: University Press.
- Knauff, M. & Johnson-Laird, P. N. (2002). Visual imagery can impede reasoning. *Memory & Cognition*, 30, 363-371.

- Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P. & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction, 20*, 111-122. doi:10.1016/j.learninstruc.2009.02.010
- Kozma, R. B. (1991). Learning with media. *Review of Educational Research, 61*, 179-211.
- Kruley, P., Sciama, S. C. & Glenberg, A. M. (1994). On-line processing of textual illustrations in the visuospatial sketchpad: Evidence from dual-task studies. *Memory & Cognition, 22*, 261-272.
- Kürschner, C. & Schnotz, W. (2007). Konstruktion mentaler Repräsentationen bei der Verarbeitung von Text und Bildern. *Unterrichtswissenschaft, 35*, 48-67.
- Kürschner, C., Schnotz, W. & Eid, M. (2006). Konstruktion mentaler Repräsentationen beim Hör- und Leseverstehen. *Zeitschrift für Medienpsychologie, 18*, 48-59. doi:10.1026/1617-6383.18.2.48
- Kürschner, C., Schnotz, W. & Eid, M. (2007). Welchen Einfluss haben die Präsentationsmodalität und Repräsentationsmodalität auf die kognitive Verarbeitung von Text mit Bildern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 39*, 70-83. doi:10.1026/0049-8637.39.2.70
- Kürschner, C., Seufert, T., Hauck, G., Schnotz, W. & Eid, M. (2006). Konstruktion visuell-räumlicher Repräsentationen beim Hör- und Leseverstehen. *Zeitschrift für Psychologie, 412*, 117-132. doi:10.1026/0044-3409.214.3.117
- Kulhavy, R. W., Lee, J. B. & Caterino, L. C. (1985). Conjoint retention of maps and related discourse. *Contemporary Educational Psychology, 10*, 28-37. doi:10.1016/0361-476X(85)90003-7
- Langston, W. E., Kramer, D. C. & Glenberg, A. M. (1998). The representation of space in mental models derived in text. *Memory & Cognition, 26*, 247-262.
- Lawrence, B. M., Myerson, J., Oonk, H. M. & Abrams, R. A. (2001). The effect of eye and limb movement on working memory. *Memory, 9*, 433-444. doi:10.1080/09658210143000047
- Leahy, W., Chandler, P. & Sweller, J. (2003). When auditory presentations should and should not be a component of multimedia instruction. *Applied Cognitive Psychology, 17*, 401-418. doi:10.1002/acp.877
- Levie, W. & Lentz, R. (1982). Effects of text illustrations: A review of research. *Educational Communication and Technology Journal, 30*, 195-232.

- Levin, J. R., Anglin, G. J. & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows & H. A. Houghton (Eds.), *The psychology of illustration* (pp. 51-85). New York: Springer.
- Levin, J. R. & Lesgold, A. M. (1978). On pictures in prose. *Educational Communication and Technology Journal*, 26, 233-243.
- Linek, S., Gerjets, P. & Scheiter, K. (2006). Speaker/Gender Effect: Impact of the speaker's gender on learning with narrated animations. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 483-488). Mahwah, NY: Erlbaum.
- Logie, R. H. (1986). Visuo-spatial processing in short-term memory. *The Quarterly Journal of Experimental Psychology*, 38A, 229-247.
- Logie, R. H. (1995). *Visuo-spatial working memory*. Aberdeen: Lawrence Erlbaum.
- Logie, R. H. & Marchetti, H. (1991). Visuo-spatial working memory: Visuo, spatial or central executive? In R. H. Logie & M. Denis (Eds.), *Images in human cognition* (pp.105-117). Amsterdam: North-Holland. doi:10.1016/S0166-4115(08)60507-5
- Logie, R. H. & Pearson, D. G. (1997). The inner eye and the inner scribe of visuo-spatial working memory: Evidence from developmental fractionation. *European Journal of Cognitive Psychology*, 9, 241-257. doi:10.1080/713752559
- Lusk, D. L., Evans, A. D., Jeffrey, T. R., Palmer, K. R., Wikstrom, C. S. & Doolittle, P. F. (2009). Multimedia learning and individual differences: Mediating the effects of working memory capacity with segmentation. *British Journal of Developmental Psychology*, 40, 636-651.
- Mammarella, I. C., Pazzaglia, F. & Cornoldi, C. (2006). The assessment of imaginary and visuo-spatial working memory functions in children and adults. In V. Tomaso & G. Bottini (Eds.), *Imagery and Spatial Cognition* (pp. 15-38). Amsterdam: John Benjamin.
- Mammarella, I. C., Pazzaglia, F. & Cornoldi, C. (2008). Evidence for different components in children's visuospatial working memory. *British Journal of Developmental Psychology*, 26, 337-355. doi:10.1348/026151007X236061
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge: University Press.
- Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge: University Press.

- Mayer, R. E. (2005a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge: University Press.
- Mayer, R. E. (2005b). Introduction to multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 1-16). Cambridge: University Press.
- Mayer, R. E. (2005c). Principles for managing essential processing in multimedia learning: segmentation, pretraining, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 169-182). Cambridge: University Press.
- Mayer, R. E. (2005d). Principles of multimedia learning based on social cues: personalization, voice, and image principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 201-212). Cambridge: University Press.
- Mayer, R. E. (2009). *Multimedia Learning. Second edition*. Cambridge: University Press.
- Mayer, R. E., Dow, G. T. & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology*, 95, 806-813. doi:10.1037/0022-0663.95.4.806
- 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. doi:10.1037/0022-0663.90.2.312
- McConnel, J. & Quinn, J. G. (2000). Interference in visual working memory. *The Quarterly Journal of Experimental Psychology*, 53A, 53-67. doi:10.1080/027249800390664
- Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin*, 27, 272-277.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H. & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cognitive Psychology* 41, 49-100. doi:10.1006/cogp.1999.0734
- Miyake, A. & Shah, P. (1999). *Models of Working Memory. Mechanisms of active maintenance and executive control*. Cambridge: University Press.

- Moreno, R. & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*, 358-368. doi:10.1037/0022-0663.91.2.358
- Moreno, R. & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology*, *94*, 598-610. doi:10.1037/0022-0663.94.3.598
- Moreno, R., Mayer, R. E., Spires, H. A. & Lester, J. C. (2001). The case for social agency in computer based multimedia learning: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, *19*, 177-214. doi:10.1207/S1532690XCI1902_02
- Mousavi, S. Y., Low, R. & Sweller, J. (1995). Reducing cognitive load by micing auditory and visual presentation modes. *Journal of Educational Psychology*, *87*, 319-334. doi:10.1037/0022-0663.87.2.319
- Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology*, *21*, 263-276. doi:10.1037/h0082978
- Nam, C. S. & Pujari, A. (2005). The role of working memory in multimedia learning. *Proceedings of the 11th International Conference on Human-Computer Interaction (HCI'05)*. Las Vegas: Mira Digital Publishing (CD-Rom).
- Noordzij, M. L., van der Lubbe, R. H. J., Neggers, S. F. W. & Postma, A. (2004). Spatial tapping interferes with the processing of linguistic spatial relations. *Canadian Journal of Experimental Psychology*, *58*, 259-271. doi:10.1037/h0087449
- Owen, A. M., Milner, B., Petrides, M. & Evans, A. (1996). Memory for object features versus memory for object location: A positron emission tomography study of encoding and retrieval processes. *Proceedings of the National Academy of Sciences*, *93* (pp. 9212–9217). doi:10.1073/pnas.93.17.9212
- Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired-associate learning. *Journal of Verbal Learning & Verbal Behavior*, *4*, 32–38. doi:10.1016/S0022-5371(65)80064-0
- Paivio, A. (1986). *Mental representations: A Dual Coding approach*. New York: Oxford University Press.
- Paivio, A. (1990). *Mental representations*. New York: Oxford University Press. doi:10.1093/acprof:oso/9780195066661.001.0001

- Pazzaglia, F. (2008). Text and picture integration in comprehending and memorizing spatial descriptions. In R. L. R.-F. Rouet and W. Schnotz (Eds.), *Understanding multimedia documents*. Heidelberg: Springer. doi:10.1007/978-0-387-73337-1_3
- Pazzaglia, F. & Cornoldi, C. (1999). The role of distinct components of visuo-spatial working memory in the processing of texts. *Memory & Cognition*, 7, 19-41.
- Pazzaglia, F., De Beni, R. & Meneghetti, C. (2007). The effects of verbal and spatial interference in the encoding and retrieval of spatial and nonspatial texts. *Psychological Research*, 71, 484-494. doi:10.1007/s00426-006-0045-7
- Pearson, P. D. & Sahraie, A. (2003). Oculomotor control and the maintenance of spatially and temporally distributed events in visuo-spatial working memory. *The Quarterly Journal of Experimental Psychology*, 56A, 1089-1111.
- Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, 17, 398-422.
- Phillips, C. E., Jarrold, C., Baddeley, A. D., Grant, J. & Karmiloff-Smith, A. (2004). Comprehension of spatial language terms in Williams Syndrome: Evidence of an interaction between domains of strength and weakness. *Cortex*, 40, 85-101. doi:10.1016/S0010-9452(08)70922-5
- Pickering, S. J. (2001). Cognitive approaches to the fractionation of visuo-spatial working memory. *Cortex*, 37, 457-473. doi:10.1016/S0010-9452(08)70587-2
- Pickering, S. J. (2006). Assessment of working memory in children. In S. J. Pickering (Ed.). *Working memory and education* (pp. 242-273). Amsterdam: Academic Press. doi:10.1016/B978-012554465-8/50011-9
- Postle, B. R., Idzikowski, Z., Della Sala, S., Logie R. H. & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. *The Quarterly Journal of Experimental Psychology*, 59, 100-120. doi:10.1080/17470210500151410
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6, 576-582. doi:10.1038/nrn1706
- Quinn, J. G. & McConnell, J. (1996). Irrelevant pictures in visual working memory. *The Quarterly Journal of Experimental Psychology*, 49A, 200-215. doi:10.1080/027249896392865
- Quinn, J. G. & Ralstone, G. E. (1986). Movement and attention in visual working memory. *The Quarterly Journal of Experimental Psychology*, 38A, 689-702.

- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422. doi:10.1037/0033-2909.124.3.372
- Rinck, M. & Glowalla, U. (1996). Die multimediale Darstellung quantitativer Daten. *Zeitschrift für Psychologie*, *204*, 383-399.
- Rummer, R., Fürstenberg A. & Schweppe, J. (2008). Lernen mit Texten und Bildern. Der Anteil akustisch sensorischer Informationen am Zustandekommen des Modalitätseffekts. *Zeitschrift für Pädagogische Psychologie*, *22*, 37-45. doi:10.1024/1010-0652.22.1.37
- Rummer, R., Schweppe, J., Scheiter, K. & Gerjets, P. (2008). Lernen mit Multimedia: Die kognitiven Grundlagen des Modalitätseffekts. *Psychologische Rundschau*, *59*, 98-107. doi:10.1026/0012-1924.59.2.98
- Scheiter, K. & Schmidt-Weigand, F. (2008). The influence of spatial text information on learning with visualizations. In A. Maes & S. Ainsworth (Eds.), *Proceedings EARLI Special Interest Group Text and Graphics: Exploiting the opportunities - Learning with textual, graphical, and multimodal representations* (pp. 123-126). Tilburg, NL: Tilburg University.
- Schellig, D. & Hättig, H. A. (1993). Die Bestimmung der visuellen Merkspanne mit dem Block-Board. *Zeitschrift für Neuropsychologie*, *4*, 104-112.
- Schmidt-Weigand, F., Kohnert, A. & Glowalla, U. (2009). Explaining the modality and contiguity effects: New insights from investigating students' viewing behavior. *Applied Cognitive Psychology, Early View*. Retrieved November, 10, 2006 from <http://www3.interscience.wiley.com/cgi-bin/fulltext/122263493/PDFSTART>.
- Schmidt-Weigand, F., Kohnert, A. & Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning & Instruction*, *20*, 100-110. doi:10.1016/j.learninstruc.2009.02.011
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49-69). Cambridge: University Press.
- Schüler, A., Scheiter, K., Gerjets, P. & Rummer, R. (2008). Does a lack of contiguity with visual text cause the modality effect in multimedia learning? In B. C. Love, K. McRae & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual*

- Conference of the Cognitive Science Society* (pp. 2353-2358). Austin, TX: Cognitive Science Society.
- Segers, E., Verhoeven, L. & Hulstijn-Hendrikse, N. (2008). Cognitive processes in children's multimedia text learning. *Applied Cognitive Psychology*, 22, 375-387. doi:10.1002/acp.1413
- Smith, E. E., Jonides, J., Koeppel, R. A., Awh, E., Schumacher, E. H. & Minoshima, S. (1995). Spatial versus object working memory: PET investigations. *Journal of Cognitive Neuroscience*, 7, 337-356. doi:10.1162/jocn.1995.7.3.337
- Smith, E. E. & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, 283, 1657-1661. doi:10.1126/science.283.5408.1657
- Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, England: Australian Council for Educational Research (ACER).
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19-30). Cambridge: University Press.
- Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251-296. doi:10.1023/A:1022193728205
- Tabbers, H. K. (2002). *The modality of text in multimedia instructions - refining the design guidelines*. Unveröffentlichte Dissertation, Open University of the Netherlands, Heerlen. Retrieved December 16, 2004 from the World Wide Web: <http://www.ou.nl/Docs/Expertise/OTEC/Publicaties/huib%20tabbers/doctoral%20dissertation%20Huib%20Tabbers%20-%20web%20version.pdf>
- Taylor, A. & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, 31, 261-292. doi:10.1016/0749-596X(92)90014-O
- Tiene, D. (2000). Sensory mode and information load: Examining the effects of timing on multisensory processing. *International Journal of Instructional Media*, 27, 183- 197.
- Tindall-Ford, S., Chandler, P. & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, 3, 257-287. doi:10.1037/1076-898X.3.4.257
- Tresch, M. C., Sinnamon, H. M. & Seamon, J. G. (1993). Double dissociation of spatial and object visual memory: Evidence from selective interference in intact

- human subjects. *Neuropsychologia*, *31*, 211-219. doi:10.1016/0028-3932(93)90085-E
- Turner, M. L. & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127-154. doi:10.1016/0749-596X(89)90040-5
- Unnava, H. R., Agarwal, S. & Haugtvedt, C. P. (1996). Interactive effects of presentation modality and message-generated imagery on recall of advertising information. *Journal of Consumer Research*, *23*, 81-88. doi:10.1086/209468
- Van Gerven, P. W. M. (2002). *Efficient complex skills training into old age: Exploring the benefits of cognitive load theory*. Unveröffentlichte Dissertation, Open University of the Netherlands, Heerlen. Retrieved January 03, 2006 from the World Wide Web: http://www.personeel.unimaas.nl/p.vangerven/Index_files/Pascal%20van%20Gerven%20-%20Dissertation.pdf.
- Van Gog, T., Paas, F., Van Merriënboer, J. J. G. & Witte, P. (2005). Uncovering expertise-related difference in troubleshooting performance: Combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology*, *19*, 205-221. doi:10.1002/acp.1112
- Van Gog, T. & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, *20*, 95-99. doi:10.1016/j.learninstruc.2009.02.009
- Vecchi, T. & Cornoldi, C. (1999). Passive storage and active manipulation in visuo-spatial working memory: Further evidence from the study of age differences. *European Journal of Cognitive Psychology*, *11*, 391-406. doi:10.1080/713752324
- Wallentin, M., Weed, E., Østergaard, L., Mouriden, K. & Roepstorff, A. (2007). Accessing the mental space – Spatial working memory processes for language and vision overlap in precuneus. *Human Brain Mapping*, *29*, 524-532. doi:10.1002/hbm.20413
- Weidenmann, B. (1995). *Information und Lernen mit Multimedia*. Weinheim, Germany: Psychologie-Verlags-Union.
- Weidenmann, B. & Strittmatter, S. (1997). "Multimedia": Mehrere Medien, mehrere Codes, mehrere Sinneskanäle? *Unterrichtswissenschaften*, *3*, 197-206.
- Wermke, M., Kunkel-Razum, K. & Scholze-Stubenrecht, W. (2004). *Duden. Das Fremdwörterbuch* (9. aktualisierte Auflage). Mannheim: Dudenverlag.

-
- Woodman, G. F. & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin and Review*, 11, 269-274.
- Zimmer, H. D. (2010). Visuell-räumliches Arbeitsgedächtnis. Eine emergente Eigenschaft der Repräsentation von Reizen im Netzwerk visueller Informationsverarbeitung. *Psychologische Rundschau*, 61, 25-32.

Anhang

- Anhang A: Instruktion Versuchsablauf
- Anhang B: Schematische Abbildungen der Lernobjekte
- Anhang C: Experiment 1 – Prüfung auf Normalverteilung
- Anhang D: Experiment 1 – Prüfung der Fehlervarianzen
- Anhang E: Experiment 1 – Testung auf Homogenität der Regressionsgeraden der *Digit Span*
- Anhang F: Experiment 1 – Gleichheit der Eingangsbedingungen
- Anhang G: Experiment 2 – Instruktion für die Lernphase
- Anhang H: Experiment 2 – Prüfung auf Normalverteilung
- Anhang I: Experiment 2 – Prüfung der Fehlervarianzen
- Anhang J: Experiment 2 – Testung auf Homogenität der Regressionsgeraden der *Digit Span*
- Anhang K: Experiment 2 – Gleichheit der Eingangsbedingungen
- Anhang L: Experiment 3 – Prüfung auf Normalverteilung
- Anhang M: Experiment 3 – Prüfung der Fehlervarianzen
- Anhang N: Experiment 3 – Testung auf Homogenität der Regressionsgeraden der *Digit Span*
- Anhang O: Experiment 3 – Gleichheit der Eingangsbedingungen

Anhang A Instruktion Versuchsablauf

Lieber Teilnehmer, liebe Teilnehmerin!

Herzlich willkommen und vielen Dank, dass Du Dich entschlossen hast, an dieser Untersuchung teilzunehmen. Sie wird circa 1,5 Stunden dauern.

In dieser Untersuchung werden Dir sechs verschiedene Fische präsentiert. Dabei werden Dir pro Fisch jeweils ein Bild und ein Text dargeboten. Auf dem Bild ist der Fisch abgebildet. Die spezifischen Charakteristika des Fisches sind jeweils in dem Text aufgeführt. Diese Charakteristika unterscheiden den jeweiligen Fisch von anderen, verwandten Arten und sind daher zur Kategorisierung des Fisches von hoher Bedeutsamkeit.

Deine Aufgabe ist es, Dir die Bilder der Fische und alle im Text genannten Charakteristika gut einzuprägen.

Im Anschluss an die Lernphase werden Dir dann Aufgaben zu den in den Bildern und den Texten enthaltenen Informationen gestellt.

Alle Daten werden selbstverständlich anonym gespeichert, nur im Kontext dieser Untersuchung verwendet und nicht an Dritte weitergegeben.

Abbildung A1. Die allgemeine Instruktion zu Beginn der Untersuchung in den ersten beiden Experimenten.

Anhang B

Schematische Abbildungen des Lernobjekts

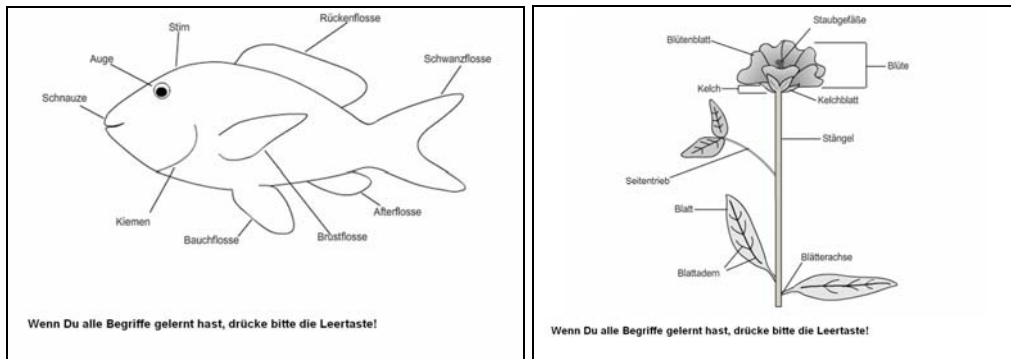


Abbildung B1. Darbietung des Fisches (links, Experiment 1 und 2) und der Pflanze (rechts, Experiment 3) mit relevanten Benennungen.

Anhang C
Experiment 1 – Prüfung auf Normalverteilung

Tabelle C1

Ergebnisse der Prüfung auf Normalverteilung für die kontinuierlichen unabhängigen Variablen in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	gesprochen	geschrieben	gesprochen	geschrieben
Corsi Block Test	$D(15) = .30^*$	$D(16) = .29^*$	$D(14) = .25^*$	$D(14) = .38^*$
<i>Visual Pattern Test</i>	$D(15) = .11$	$D(16) = .11$	$D(14) = .12$	$D(14) = .16$
<i>Digit Span Test</i>	$D(15) = .26^*$	$D(14) = .20$	$D(14) = .26^*$	$D(16) = .26^*$

Anmerkung. $*p < .05$.

Tabelle C2

Ergebnisse der Prüfung auf Normalverteilung für die Verifikationsitems in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	gesprochen	geschrieben	gesprochen	geschrieben
<i>textbezogene Erinnerungsleistung</i>				
Charakteristika	$D(15) = .14$	$D(16) = .13$	$D(14) = .12$	$D(14) = .18$
abstrakte Textinhalte	$D(15) = .11$	$D(16) = .12$	$D(14) = .17$	$D(14) = .24^*$
<i>bildbezogene Erinnerungsleistung</i>				
visuelle Bildinhalte	$D(15) = .13$	$D(16) = .21$	$D(14) = .23$	$D(14) = .15$
räumliche Bildinhalte	$D(15) = .12$	$D(16) = .21$	$D(14) = .19$	$D(14) = .18$

Anmerkung. $*p < .05$.

Anhang C
Experiment 1 – Prüfung auf Normalverteilung

Tabelle C3

Ergebnisse der Prüfung auf Normalverteilung für die Bewertungssitems in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	gesprochen	geschrieben	gesprochen	geschrieben
Textmodalität				
Wie anstrengend war die Lernaufgabe für Dich?	$D(15) = .28^*$	$D(16) = .25^*$	$D(14) = .24^*$	$D(14) = .20$
Wie schwierig war es für Dich, die Inhalte zu verstehen?	$D(15) = .27^*$	$D(16) = .27^*$	$D(14) = .29^*$	$D(14) = .24^*$
Wie sehr hast Du Dich beim Lernen auf das Bild konzentriert?	$D(15) = .31^*$	$D(16) = .25^*$	$D(14) = .22$	$D(14) = .15$
Wie sehr hast Du Dich beim Lernen auf den Text konzentriert?	$D(15) = .23^*$	$D(16) = .20$	$D(14) = .26^*$	$D(14) = .24^*$
Wie sehr hast Du Dich während der Lernphase konzentriert?	$D(15) = .22^*$	$D(16) = .24^*$	$D(14) = .22$	$D(14) = .27^*$
Wie sehr hast Du Dich angestrengt, um das Lernmaterial zu verstehen?	$D(15) = .23^*$	$D(14) = .24^*$	$D(14) = .29^*$	$D(16) = .20$
Wie viel Mühe hast Du in die Lernphase investiert?	$D(15) = .31^*$	$D(14) = .25^*$	$D(14) = .25^*$	$D(16) = .24^*$

Anmerkung. $*p < .05$.

Tabelle D1

Ergebnisse der Levene-Tests für die kontinuierlichen unabhängigen Variablen

	Ergebnisse des Levene Tests
Corsi Block Test	$F(3, 55) = 1.18, p = .33$
<i>Visual Pattern</i> Test	$F < 1$
<i>Digit Span</i> Test	$F < 1$

Tabelle D2

Ergebnisse der Levene-Tests für die Verifikationsitems

	Ergebnisse des Levene-Tests
<i>textbezogene Erinnerungsleistung</i>	
Charakteristika	$F(3, 55) = 2.06, p = .12$
abstrakte Textinhalte	$F(3, 55) = 1.52, p = .22$
<i>bildbezogene Erinnerungsleistung</i>	
visuelle Bildinhalte	$F < 1$
räumliche Bildinhalte	$F(3, 55) = 2.20, p = .10$

Tabelle D3

Ergebnisse der Levene-Tests für die Bewertungsitems

	Ergebnisse des Levene Tests
Wie anstrengend war die Lernaufgabe für Dich?	$F(3, 55) = 1.17, p = .33$
Wie schwierig war es für Dich, die Inhalte zu verstehen?	$F < 1$
Wie sehr hast Du Dich beim Lernen auf das Bild konzentriert?	$F(3, 55) = 2.76, p = .051$
Wie sehr hast Du Dich beim Lernen auf den Text konzentriert?	$F(3, 55) = 1.79, p = .16$
Wie sehr hast Du Dich während der Lernphase konzentriert?	$F(3, 55) = 1.12, p = .35$
Wie sehr hast Du Dich angestrengt, um das Lernmaterial zu verstehen?	$F < 1$
Wie viel Mühe hast Du in die Lernphase investiert?	$F < 1$

Anhang E

Experiment 1 – Testung auf Homogenität der Regressionsgeraden der *Digit Span*

Tabelle E1

Ergebnisse der MANCOVAs für die Verifikationsitems mit den Faktoren *Digit Span*, *VPT*, *Corsi Block (Corsi)*, *Textmodalität (TM)* und *Textinhalt (TI)*

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
räumliche Bildinhalte Charakteristika	<i>Digit Span</i>	$V = .09, F(2, 42) = 2.03, p = .14$
	VPT	$V = .01, F < 1$
	Corsi	$V = .05, F(2, 42) = 1.15, p = .33$
	TM	$V < .001, F < 1$
	TI	$V = .27, F(2, 42) = 7.93, p = .001$
	TM × TI	$V = .03, F < 1$
	TM × Corsi	$V = .06, F(2, 42) = 1.30, p = .28$
	TM × VPT	$V = .01, F < 1$
	TI × Corsi	$V = .03, F < 1$
	TI × VPT	$V = .003, F < 1$
	TM × TI × Corsi	$V = .07, F(2, 42) = 1.68, p = .20$
	TM × TI × VPT	$V = .07, F(2, 42) = 1.50, p = .24$
	TM × TI × <i>Digit Span</i>	$V = .19, F(6, 86) = 1.51, p = .18$
visuelle Bildinhalte abstrakte Textinhalte	<i>Digit Span</i>	$V = .13, F(2, 42) = 3.02, p = .06$
	VPT	$V = .01, F < 1$
	Corsi	$V = .01, F < 1$
	TM	$V = .12, F(2, 42) = 2.90, p = .07$
	TI	$V = .28, F(2, 42) = 8.05, p = .01$
	TM × TI	$V = .14, F(2, 42) = 3.31, p = .05$
	TM × Corsi	$V = .06, F(2, 42) = 1.40, p = .26$
	TM × VPT	$V = .07, F(2, 42) = 1.51, p = .23$
	TI × Corsi	$V = .01, F < 1$
	TI × VPT	$V = .01, F < 1$
	TM × TI × Corsi	$V = .08, F(2, 42) = 1.89, p = .16$
	TM × TI × VPT	$V = .18, F(2, 42) = 4.54, p = .02$
	TM × TI × <i>Digit Span</i>	$V = .26, F(6, 86) = 2.14, p = .06$

Tabelle E2

Ergebnisse der MANCOVAs für die Bewertungsitems mit den Faktoren *Digit Span*, *VPT*, *Corsi Block (Corsi)*, *Textmodalität (TM)* und *Textinhalt (TI)*

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	<i>Digit Span</i>	$V = .16, F < 1$
	VPT	$V = .09, F < 1$
	Corsi	$V = .23, F(7, 37) = 1.35, p = .25$
	TM	$V = .33, F(7, 37) = 2.23, p = .05$
	TI	$V = .13, F < 1$
	TM × TI	$V = .20, F(7, 37) = 1.09, p = .39$
	TM × Corsi	$V = .24, F(7, 37) = 1.42, p = .22$
	TM × VPT	$V = .22, F(7, 37) = 1.28, p = .28$
	TI × Corsi	$V = .09, F < 1$
	TI × VPT	$V = .05, F < 1$
	TM × TI × Corsi	$V = .01, F < 1$
	TM × TI × VPT	$V = .18, F < 1$
	TM × TI × <i>Digit Span</i>	$V = .50, F < 1$

Anhang F

Experiment 1 – Gleichheit der Eingangsbedingungen

Tabelle F1

Ergebnisse der Varianzanalysen mit den Faktoren Textmodalität (TM) und Textinhalt (TI) für die Kontrollvariablen

		Ergebnisse der ANOVA
Alter ^a (in Jahren)	TM	$F(1, 55) = 3.06, MSE = 14.37, p = .09, \eta_p^2 = .05$
	TI	$F < 1$
	TM × TI	$F(1, 55) = 1.20, MSE = 14.37, p = .28, \eta_p^2 = .02$
Semesteranzahl ^a	TM	$F < 1$
	TI	$F < 1$
	TM × TI	$F < 1$
Letzte Bionote ^a	TM	$F < 1$
	TI	$F(1, 54) = 2.91, MSE = 8.05, p = .09, \eta_p^2 = .05$
	TM × TI	$F(1, 54) = 2.05, MSE = 8.05, p = .16, \eta_p^2 = .04$
Vorkenntnisse bzgl. Fischen ^a	TM	$F < 1$
	TI	$F < 1$
	TM × TI	$F < 1$

Anmerkung. ^aEine Absicherung der Analyse mit dem Kruskal-Wallis Test bestätigte das Ergebnis.

Tabelle F2

Ergebnisse der Loglinearen Analysen mit den Faktoren Textmodalität (TM) und Textinhalt (TI) für die nominalen Kontrollvariablen

	Schritt	Effekte	χ^2	df	Sig.	Anzahl Iterationen
Geschlecht	6	HE Geschlecht	12.83	1	<.001	0
Biologie 4- stündig/Leistungskurs	6	HE Bionote	22.19	1	<.001	0
Studienfach	Berechnung nicht möglich, da zu wenige Teilnehmer mit biologischen Vertiefungsfach					

Anhang G

Experiment 2 – Instruktion für die Lernphase

Im folgenden Teil der Untersuchung werden Dir nun sechs verschiedene Fische präsentiert. Pro Fisch werden Dir ein Bild und ein Text dargeboten. Auf dem Bild ist der Fisch abgebildet. Im Text werden zusätzlich spezifische Charakteristika des Fisches genannt. Diese im Text genannten Merkmale machen den jeweiligen Fisch von verwandten Fischen unterscheidbar. Beachte dabei bitte, dass ALLE im Text genannten Informationen spezifische Charakteristika sind und somit für die eindeutige Kategorisierung des Fisches eine hohe Bedeutung haben!

DEINE BBFGABE IST ES, DIR DAS BILD, ALLE IM TEXT ENTHALTENEN INFORMATIONEN UND DIE NAMEN DER FISCH EINZUPRÄGEN.

Die Darbietung läuft von selbst ab. Dabei werden Dir die Inhalte nur einmal und in einer begrenzten Zeit dargeboten. Bitte verfolge die Darbietung daher aufmerksam und merke Dir die Inhalte gut! Weiter mit Leertaste!

Während der Darbietung der Fische sollst Du nebenbei die Nebenaufgabe durchführen, die Dir vom Versuchsleiter gezeigt wurde. Beachte dabei bitte, dass Deine Hauptaufgabe im Lernen der dargebotenen Informationen besteht. [1]

Während der Lernphase werden Dir einige Fragen zu Deiner Wahrnehmung der Lernphase gestellt. Es gibt dabei keine richtigen oder falschen Antworten. [2]

Während der Darbietung der Fragen kannst Du die Nebenaufgabe unterbrechen. Nehme sie aber bitte sofort wieder auf, sobald der nächste Fisch dargeboten wird. [3]

Im Anschluss an die Lernphase werden Dir dann Fragen und Aufgaben zu den Bildern und den im Text genannten Charakteristika gestellt, die Du mit Hilfe der gelernten Informationen beantworten sollst.

Setze nun bitte noch die Kopfhörer auf, da die Texte teilweise gesprochen dargeboten werden. [4]

Wenn Du auf die Leertaste drückst, beginnt die Darbietung.

Nehme jetzt bitte die Nebenaufgabe auf! [5]

Abbildung G1. Instruktion der Lerner mit Zweitaufgabe vor Beginn der Lernphase. Neuerungen im Vergleich zum ersten Experiment sind durchnummeriert. Die Sätze [2] und [4] wurden in allen Bedingungen eingefügt, die Sätze [1], [3] und [5] nur in Bedingungen mit Zweitaufgabe.

Anhang H

Experiment 2 – Prüfung auf Normalverteilung

Tabelle H1

Ergebnisse der Prüfung auf Normalverteilung für die kontinuierlichen unabhängigen Variablen in Abhängigkeit des Textinhalts und der Zweitaufgabe (ZA)

Textinhalte	visuell		räumlich	
	ohne ZA	mit ZA	ohne ZA	mit ZA
Zweitaufgabe				
Corsi Block Test	$D(18) = .25^*$	$D(19) = .24^*$	$D(18) = .32^*$	$D(19) = .36^*$
Visual Pattern Test	$D(18) = .17^*$	$D(19) = .13$	$D(18) = .18$	$D(19) = .11$
Digit Span Test	$D(18) = .22^*$	$D(19) = .27^*$	$D(18) = .27^*$	$D(19) = .25^*$

Anmerkung. $*p < .05$.

Tabelle H2

Ergebnisse der Prüfung auf Normalverteilung für die Verifikationsitems in Abhängigkeit des Textinhalts, der Zweitaufgabe (ZA) und der Textmodalität

Textinhalt	visuell				räumlich			
	ohne ZA		mit ZA		ohne ZA		mit ZA	
Zweitaufgabe								
Textmodalität	gespr.	geschr.	gespr.	geschr.	gespr.	geschr.	gespr.	geschr.
VERIFIKATIONSSITEMS								
<i>textbezogene Erinnerungsleistung</i>								
Charakteristika	$D(18) = .13$	$D(18) = .11$	$D(19) = .14$	$D(19) = .13$	$D(18) = .12$	$D(18) = .19$	$D(19) = .17$	$D(19) = .13$
abstrakte Textinhalte	$D(18) = .15$	$D(18) = .16$	$D(19) = .12$	$D(19) = .15$	$D(18) = .15$	$D(18) = .22^*$	$D(19) = .18$	$D(19) = .17$
<i>bildbezogene Erinnerungsleistung</i>								
visuelle Bildinhalte	$D(18) = .18$	$D(18) = .17$	$D(19) = .14$	$D(19) = .18$	$D(18) = .14$	$D(18) = .11$	$D(19) = .13$	$D(19) = .16$
räumliche Bildinhalte	$D(18) = .10$	$D(18) = .09$	$D(19) = .09$	$D(19) = .09$	$D(18) = .16$	$D(18) = .09$	$D(19) = .12$	$D(19) = .24^*$
OFFENE FRAGEN								
<i>textbezogene Erinnerungsleistung</i>								
Charakteristika	$D(18) = .24^{**}$	$D(18) = .16$	$D(19) = .24^{**}$	$D(19) = .19$	$D(18) = .32^{**}$	$D(18) = .30^{**}$	$D(19) = .37^{**}$	$D(19) = .30^{**}$
abstrakte Textinhalte	$D(18) = .23^*$	$D(18) = .18$	$D(19) = .20^*$	$D(19) = .27^{**}$	$D(18) = .23^*$	$D(18) = .21^*$	$D(19) = .23^*$	$D(19) = .21^*$
<i>bildbezogene Erinnerungsleistung</i>								
visuelle Bildinhalte	$D(18) = .11$	$D(18) = .20$	$D(19) = .12$	$D(19) = .14$	$D(18) = .17$	$D(18) = .12$	$D(19) = .16$	$D(19) = .12$
räumliche Bildinhalte	$D(18) = .15$	$D(18) = .12$	$D(19) = .12$	$D(19) = .14$	$D(18) = .21^*$	$D(18) = .17$	$D(19) = .15$	$D(19) = .21^*$

Anmerkung. gespr. = gesprochen, geschr. = geschrieben

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

Anhang H

Experiment 2 – Prüfung auf Normalverteilung

Tabelle H3

Ergebnisse der Prüfung auf Normalverteilung für die Dual Task Leistung in Abhängigkeit des Textinhalts und der Textmodalität

	visuell			räumlich		
	Baseline	gespr.	geschr.	Baseline	gespr.	geschr.
Fehlerrate (%)	$D(18) = .35^*$	$D(18) = .21^*$	$D(18) = .13$	$D(16) = .30^*$	$D(16) = .25^*$	$D(16) = .14$
Tipprate (ms)	$D(18) = .15$	$D(18) = .09$	$D(18) = .10$	$D(16) = .22^*$	$D(16) = .18$	$D(16) = .18$

Anmerkung. gespr. = gesprochen, geschr. = geschrieben

* $p < .05$.

Tabelle H4

Ergebnisse der Prüfung auf Normalverteilung für die Bewertungsitems in Abhängigkeit des Textinhalts, der Zweitaufgabe (ZA) und der Textmodalität

Textinhalt	visuell				räumlich			
	ohne ZA		mit ZA		ohne ZA		mit ZA	
Zweitaufgabe	gespr.	geschr.	gespr.	geschr.	gespr.	geschr.	gespr.	geschr.
Textmodalität	gespr.	geschr.	gespr.	geschr.	gespr.	geschr.	gespr.	geschr.
Wie sehr hast Du Dich während der Lernphase konzentriert?	$D(18) = .22^*$	$D(18) = .17$	$D(19) = .17$	$D(19) = .23^{**}$	$D(18) = .27^{**}$	$D(18) = .20^*$	$D(19) = .28^{**}$	$D(19) = .27^{**}$
Wie viel Mühe hast Du in die Lernphase investiert?	$D(18) = .25^{**}$	$D(18) = .23^*$	$D(19) = .32^{**}$	$D(19) = .32^{**}$	$D(18) = .28^{**}$	$D(18) = .27^{**}$	$D(19) = .24^{**}$	$D(19) = .21^*$

Anmerkung. gespr. = gesprochen, geschr. = geschrieben

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

Anhang I

Experiment 2 – Prüfung der Fehlervarianzen

Tabelle I1

Ergebnisse der Levene-Tests für die kontinuierlichen unabhängigen Variablen

Ergebnisse des Levene Tests	
Corsi Block Test	$F(3, 70) = 2.16, p = .10$
Visual Pattern Test	$F(3, 70) = 1.43, p = .24$
Digit Span Test	$F < 1$

Tabelle I2

Ergebnisse der Levene-Tests für die Verifikationsitems und die offenen Fragen

Ergebnisse des Levene Test		
	gesprochen	geschrieben
VERIFIKATIONSITEMS		
<i>textbezogene Erinnerungsleistung</i>		
Charakteristika	$F(3, 70) = 1.62, p = .19$	$F(3, 70) = 1.30, p = .28$
abstrakte Textinhalte	$F < 1$	$F < 1$
<i>bildbezogene Erinnerungsleistung</i>		
visuelle Bildinhalte	$F(3, 70) = 10.14, p < .001$	$F(3, 70) = 2.12, p = .11$
räumliche Bildinhalte	$F(3, 70) = 3.09, p = .03$	$F(3, 70) = 3.04, p = .04$
OFFENE FRAGEN		
<i>textbezogene Erinnerungsleistung</i>		
Charakteristika	$F(3, 70) = 1.46, p = .23$	$F < 1$
abstrakte Textinhalte	$F(3, 70) = 1.07, p = .37$	$F < 1$
<i>bildbezogene Erinnerungsleistung</i>		
visuelle Bildinhalte	$F(3, 70) = 1.73, p = .17$	$F < 1$
räumliche Bildinhalte	$F(3, 70) = 6.91, p < .001$	$F(3, 70) = 2.50, p = .07$

Tabelle I3

Ergebnisse der Levene-Tests für die Zweitaufgabenleistung

Ergebnisse des Levene Tests			
	Baseline	gesprochen	geschrieben
Fehlerrate (%)	$F < 1$	$F < 1$	$F(1, 32) = 1.71, p = .20$
Tipprate (ms)	$F(1, 32) = 1.38, p = .25$	$F < 1$	$F(1, 32) = 1.27, p = .27$

Anhang I
Experiment 2 – Prüfung der Fehlervarianzen

Tabelle I4

Ergebnisse der Levene-Tests für die Bewertungsitems

Ergebnisse des Levene Tests		
	gesprochen	geschrieben
Wie sehr hast Du Dich während der Lernphase konzentriert?	$F(3, 70) = 2.21, p = .10$	$F < 1$
Wie viel Mühe hast Du in die Lernphase investiert?	$F(3, 70) = 3.33, p = .02$	$F < 1$

Anhang J

Experiment 2 – Testung auf Homogenität der Regressionsgeraden der *Digit Span*

Tabelle J1

Ergebnisse der (M)ANCOVAs für die Verifikationsitems mit den Faktoren Digit Span, VPT, Corsi Block (Corsi), Textmodalität (TM), Textinhalt (TI) und Zweitaufgabe (ZA)

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA	
Charakteristika	<i>Digit Span</i>	$F(1, 60) = 7.42, MSE = 0.32, p = .01, \eta_p^2 = .11$	
	VPT	$F < 1$	
	Corsi	$F < 1$	
	TI	$F(1, 60) = 12.88, MSE = 0.32, p = .01, \eta_p^2 = .18$	
	ZA	$F(1, 60) = 11.62, MSE = 0.32, p = .01, \eta_p^2 = .16$	
	ZA × Corsi	$F < 1$	
	TI × Corsi	$F(1, 60) = 1.07, MSE = 0.32, p = .31, \eta_p^2 = .02$	
	TI × VPT	$F(1, 60) = 5.20, MSE = 0.32, p = .03, \eta_p^2 = .08$	
	TI × ZA	$F < 1$	
	TI × ZA × Corsi	$F < 1$	
	TI × ZA × <i>Digit Span</i>	$F(1, 60) = 2.25, MSE = 0.32, p = .09, \eta_p^2 = .10$	
	TM	$F(1, 60) = 7.00, MSE = 0.23, p = .01, \eta_p^2 = .10$	
	TM × <i>Digit Span</i>	$F < 1$	
	TM × VPT	$F(1, 60) = 7.00, MSE = 0.23, p = .01, \eta_p^2 = .10$	
	TM × Corsi	$F < 1$	
	TM × TI	$F < 1$	
	TM × ZA	$F(1, 60) = 3.41, MSE = 0.23, p = .07, \eta_p^2 = .05$	
	TM × TI × Corsi	$F < 1$	
	TM × TI × VPT	$F(1, 60) = 2.38, MSE = 0.23, p = .13, \eta_p^2 = .04$	
	TM × TI × ZA	$F < 1$	
	TM × ZA × Corsi	$F < 1$	
	TM × TI × ZA × Corsi	$F < 1$	
	TM × TI × ZA × <i>Digit Span</i>	$F < 1$	
	abstrakte Textinhalte	<i>Digit Span</i>	$F(1, 60) = 9.01, MSE = 1.39, p = .01, \eta_p^2 = .13$
		VPT	$F(1, 60) = 1.14, MSE = 1.39, p = .29, \eta_p^2 = .02$
		Corsi	$F < 1$
TI		$F < 1$	
ZA		$F < 1$	
ZA × Corsi		$F(1, 60) = 1.01, MSE = 1.39, p = .32, \eta_p^2 = .02$	
TI × Corsi		$F < 1$	
TI × VPT		$F < 1$	
TI × ZA		$F < 1$	
TI × ZA × Corsi		$F < 1$	
TI × ZA × <i>Digit Span</i>		$F(1, 60) = 1.50, MSE = 1.39, p = .23, \eta_p^2 = .07$	
TM		$F < 1$	
TM × <i>Digit Span</i>		$F < 1$	
TM × VPT		$F(1, 60) = 3.00, MSE = 1.36, p = .09, \eta_p^2 = .05$	
TM × Corsi		$F < 1$	
TM × TI		$F < 1$	
TM × ZA		$F(1, 60) = 5.01, MSE = 1.36, p = .03, \eta_p^2 = .08$	
TM × TI × Corsi		$F < 1$	
TM × TI × VPT		$F < 1$	
TM × TI × ZA		$F < 1$	
TM × ZA × Corsi		$F < 1$	
TM × TI × ZA × Corsi		$F < 1$	
TM × TI × ZA × <i>Digit Span</i>		$F(1, 60) = 1.27, MSE = 1.36, p = .29, \eta_p^2 = .06$	

Fortsetzung Tabelle J1

		Ergebnisse der mixed (M)ANCOVA
räumliche Bildinhalte	<i>Digit Span</i>	$V = .01, F < 1$
visuelle Bildinhalte	VPT	$V = .004, F < 1$
	Corsi	$V = .01, F < 1$
	TI	$V = .01, F < 1$
	ZA	$V = .10, F(2, 59) = 3.41, p = .04$
	ZA × Corsi	$V = .004, F < 1$
	TI × Corsi	$V = .004, F < 1$
	TI × VPT	$V = .01, F < 1$
	TI × ZA	$V = .01, F < 1$
	TI × ZA × Corsi	$V = .02, F < 1$
	TI × ZA × <i>Digit Span</i>	$V = .11, F(2, 59) = 1.16, p = .33$
	TM	$V = .02, F < 1$
	TM × <i>Digit Span</i>	$V = .01, F < 1$
	TM × VPT	$V = .02, F < 1$
	TM × Corsi	$V = .01, F < 1$
	TM × TI	$V = .002, F < 1$
	TM × ZA	$V = .01, F < 1$
	TM × TI × Corsi	$V = .004, F < 1$
	TM × TI × VPT	$V = .04, F < 1$
	TM × TI × ZA	$V = .01, F < 1$
	TM × ZA × Corsi	$V = .01, F < 1$
	TM × TI × ZA × Corsi	$V = .03, F < 1$
	TM × TI × ZA × <i>Digit Span</i>	$V = .09, F < 1$

Anhang J

Experiment 2 – Testung auf Homogenität der Regressionsgeraden der *Digit Span*

Tabelle J2

Ergebnisse der (M)ANCOVAs für die offenen Fragen mit den Faktoren Digit Span, VPT, Corsi Block (Corsi), Textmodalität (TM), Textinhalt (TI) und Zweitaufgabe (ZA)

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA	
Charakteristika	<i>Digit Span</i>	$F(1, 60) = 3.38, MSE = 753.93, p = .07, \eta_p^2 = .05$	
	VPT	$F < 1$	
	Corsi	$F < 1$	
	TI	$F(1, 60) = 7.09, MSE = 753.93, p = .01, \eta_p^2 = .11$	
	ZA	$F(1, 60) = 2.18, MSE = 753.93, p = .15, \eta_p^2 = .04$	
	ZA × Corsi	$F < 1$	
	TI × Corsi	$F < 1$	
	TI × VPT	$F(1, 60) = 5.25, MSE = 753.93, p = .03, \eta_p^2 = .08$	
	TI × ZA	$F < 1$	
	TI × ZA × Corsi	$F < 1$	
	TI × ZA × <i>Digit Span</i>	$F(1, 60) = 1.29, MSE = 753.93, p = .29, \eta_p^2 = .06$	
	TM	$F < 1$	
	TM × <i>Digit Span</i>	$F < 1$	
	TM × VPT	$F < 1$	
	TM × Corsi	$F < 1$	
	TM × TI	$F < 1$	
	TM × ZA	$F < 1$	
	TM × TI × Corsi	$F < 1$	
	TM × TI × VPT	$F < 1$	
	TM × TI × ZA	$F < 1$	
	TM × ZA × Corsi	$F < 1$	
	TM × TI × ZA × Corsi	$F < 1$	
	TM × TI × ZA × <i>Digit Span</i>	$F < 1$	
	abstrakte Textinhalte	<i>Digit Span</i>	$F < 1$
		VPT	$F(1, 60) = 4.87, MSE = 195.71, p = .03, \eta_p^2 = .08$
		Corsi	$F(1, 60) = 3.68, MSE = 195.71, p = .06, \eta_p^2 = .06$
TI		$F < 1$	
ZA		$F(1, 60) = 2.48, MSE = 195.71, p = .12, \eta_p^2 = .04$	
ZA × Corsi		$F < 1$	
TI × Corsi		$F < 1$	
TI × VPT		$F(1, 60) = 3.89, MSE = 195.71, p = .05, \eta_p^2 = .06$	
TI × ZA		$F < 1$	
TI × ZA × Corsi		$F(1, 60) = 5.24, MSE = 195.71, p = .02, \eta_p^2 = .08$	
TI × ZA × <i>Digit Span</i>		$F(1, 60) = 3.42, MSE = 195.71, p = .02, \eta_p^2 = .15$	
TM		$F(1, 60) = 2.02, MSE = 327.77, p = .16, \eta_p^2 = .03$	
TM × <i>Digit Span</i>		$F(1, 60) = 2.08, MSE = 327.77, p = .15, \eta_p^2 = .03$	
TM × VPT		$F(1, 60) = 4.44, MSE = 327.77, p = .04, \eta_p^2 = .07$	
TM × Corsi		$F < 1$	
TM × TI		$F < 1$	
TM × ZA		$F < 1$	
TM × TI × Corsi		$F < 1$	
TM × TI × VPT		$F < 1$	
TM × TI × ZA		$F(1, 60) = 2.58, MSE = 327.77, p = .11, \eta_p^2 = .04$	
TM × ZA × Corsi		$F < 1$	
TM × TI × ZA × Corsi		$F < 1$	
TM × TI × ZA × <i>Digit Span</i>		$F < 1$	

Fortsetzung Tabelle J2

		Ergebnisse der mixed (M)ANCOVA
räumliche Bildinhalte	<i>Digit Span</i>	$V = .01, F < 1$
visuelle Bildinhalte	VPT	$V = .01, F < 1$
	Corsi	$V = .04, F(2, 59) = 1.20, p = .31$
	TI	$V = .49, F(2, 59) = 27.76, p = .01$
	ZA	$V = .12, F(2, 59) = 3.93, p = .03$
	ZA × Corsi	$V = .03, F < 1$
	TI × Corsi	$V = .04, F(2, 59) = 1.21, p = .31$
	TI × VPT	$V = .01, F < 1$
	TI × ZA	$V = .03, F(2, 59) = 1.00, p = .37$
	TI × ZA × Corsi	$V = .04, F(2, 59) = 1.07, p = .35$
	TI × ZA × <i>Digit Span</i>	$V = .03, F < 1$
	TM	$V = .04, F(2, 59) = 1.33, p = .27$
	TM × <i>Digit Span</i>	$V = .04, F(2, 59) = 1.08, p = .35$
	TM × VPT	$V = .04, F(2, 59) = 1.11, p = .34$
	TM × Corsi	$V = .07, F(2, 59) = 2.30, p = .11$
	TM × TI	$V = .01, F < 1$
	TM × ZA	$V = .01, F < 1$
	TM × TI × Corsi	$V = .01, F < 1$
	TM × TI × VPT	$V = .01, F < 1$
	TM × TI × ZA	$V = .08, F(2, 59) = 2.43, p = .10$
	TM × ZA × Corsi	$V = .02, F < 1$
	TM × TI × ZA × Corsi	$V = .04, F(2, 59) = 1.12, p = .33$
	TM × TI × ZA × <i>Digit Span</i>	$V = .07, F(6, 120) = 1.01, p = .42$

Tabelle J3

Ergebnisse der (M)ANCOVAs für die Bewertungsitems mit den Faktoren *Digit Span*, *VPT*, *Corsi Block (Corsi)*, *Textmodalität (TM)*, *Textinhalt (TI)* und *Zweitaufgabe (ZA)*

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	<i>Digit Span</i>	$V = .10, F(2, 59) = 2.17, p = .10$
	VPT	$V = .06, F(2, 59) = 1.19, p = .32$
	Corsi	$V = .01, F < 1$
	TI	$V = .10, F(2, 59) = 2.19, p = .01$
	ZA	$V = .18, F(2, 59) = 4.15, p = .01$
	ZA × Corsi	$V = .20, F(2, 59) = 4.81, p = .01$
	TI × Corsi	$V = .03, F < 1$
	TI × VPT	$V = .06, F(2, 59) = 1.23, p = .31$
	TI × ZA	$V = .04, F < 1$
	TI × ZA × Corsi	$V = .17, F(2, 59) = 4.02, p = .01$
	TI × ZA × <i>Digit Span</i>	$V = .02, F < 1$
	TM	$V = .23, F(2, 59) = 5.83, p = .01$
	TM × <i>Digit Span</i>	$V = .09, F(2, 59) = 1.80, p = .16$
	TM × VPT	$V = .11, F(2, 59) = 2.40, p = .08$
	TM × Corsi	$V = .04, F < 1$
	TM × TI	$V = .03, F < 1$
	TM × ZA	$V = .04, F < 1$
	TM × TI × Corsi	$V = .03, F < 1$
	TM × TI × VPT	$V = .04, F < 1$
	TM × TI × ZA	$V = .02, F < 1$
	TM × ZA × Corsi	$V = .03, F < 1$
	TM × TI × ZA × Corsi	$V = .08, F(2, 59) = 1.63, p = .19$
	TM × TI × ZA × <i>Digit Span</i>	$V = .04, F < 1$

Anhang J

Experiment 2 – Testung auf Homogenität der Regressionsgeraden der *Digit Span*

Tabelle J4

Ergebnisse der ANCOVAs für die Leistung in der Zweitaufgabe mit den Faktoren Digit Span, VPT, Corsi Block (Corsi), Baseline in der Zweitaufgabe, Textmodalität (TM) und Textinhalt (TI)

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
Fehlerrate	<i>Digit Span</i>	$F < 1$
	VPT	$F < 1$
	Corsi	$F < 1$
	Baseline	$F(1, 24) = 13.63, MSE = 1.84, p = .01, \eta_p^2 = .36$
	TI	$F < 1$
	TI \times Corsi	$F(1, 24) = 1.61, MSE = 1.84, p = .22, \eta_p^2 = .06$
	TI \times VPT	$F(1, 24) = 2.26, MSE = 1.84, p = .15, \eta_p^2 = .09$
	TI \times <i>Digit Span</i>	$F(1, 24) = 1.01, MSE = 1.84, p = .31, \eta_p^2 = .04$
	TI \times Baseline	$F < 1$
	TM	$F(1, 24) = 3.46, MSE = .10, p = .08, \eta_p^2 = .13$
	TM \times <i>Digit Span</i>	$F(1, 24) = 3.04, MSE = .10, p = .09, \eta_p^2 = .11$
	TM \times VPT	$F < 1$
	TM \times Corsi	$F < 1$
	TM \times Baseline	$F(1, 24) = 1.28, MSE = .10, p = .27, \eta_p^2 = .05$
	TM \times TI	$F < 1$
	TM \times TI \times Corsi	$F < 1$
	TM \times TI \times VPT	$F(1, 24) = 1.53, MSE = .10, p = .23, \eta_p^2 = .06$
	TM \times TI \times Corsi	$F < 1$
	TM \times TI \times <i>Digit Span</i>	$F < 1$
	TM \times TI \times Baseline	$F < 1$
Tipp- geschwindigkeit	<i>Digit Span</i>	$F < 1$
	VPT	$F(1, 24) = 3.88, MSE = 119784.40, p = .06, \eta_p^2 = .14$
	Corsi	$F < 1$
	Baseline	$F < 1$
	TI	$F(1, 24) = 2.41, MSE = 119784.40, p = .13, \eta_p^2 = .09$
	TI \times Corsi	$F(1, 24) = 2.29, MSE = 119784.40, p = .14, \eta_p^2 = .09$
	TI \times VPT	$F < 1$
	TI \times <i>Digit Span</i>	$F < 1$
	TI \times Baseline	$F(1, 24) = 1.10, MSE = 119784.40, p = .30, \eta_p^2 = .04$
	TM	$F(1, 24) = 5.38, MSE = 6873.80, p = .03, \eta_p^2 = .18$
	TM \times <i>Digit Span</i>	$F < 1$
	TM \times VPT	$F(1, 24) = 2.18, MSE = 6873.80, p = .15, \eta_p^2 = .08$
	TM \times Corsi	$F(1, 24) = 1.90, MSE = 6873.80, p = .18, \eta_p^2 = .07$
	TM \times Baseline	$F < 1$
	TM \times TI	$F(1, 24) = 1.61, MSE = 6873.80, p = .22, \eta_p^2 = .06$
	TM \times TI \times Corsi	$F < 1$
	TM \times TI \times VPT	$F < 1$
	TM \times TI \times <i>Digit Span</i>	$F < 1$
	TM \times TI \times Baseline	$F(1, 24) = 1.71, MSE = 6873.80, p = .20, \eta_p^2 = .07$

Anhang K

Experiment 2 – Gleichheit der Eingangsbedingungen

Tabelle K1

Ergebnisse der Varianzanalysen mit den Faktoren Textinhalt (TI) und Zweitaufgabe (ZA) für die Kontrollvariablen

		Ergebnisse der ANOVA
Alter ^a	TI	$F < 1$
	ZA	$F(1, 70) = 11.97, MSE = 9.62, p = .27, \eta_p^2 = .02$
	TI \times ZA	$F < 1$
Semesteranzahl ^a	TI	$F(1, 70) = 1.33, MSE = 6.92, p = .25, \eta_p^2 = .02$
	ZA	$F(1, 70) = 2.68, MSE = 6.92, p = .11, \eta_p^2 = .04$
	TI \times ZA	$F < 1$
Letzte Bionote ^a	TI	$F < 1$
	ZA	$F < 1$
	TI \times ZA	$F(1, 67) = 1.48, MSE = 3.05, p = .23, \eta_p^2 = .02$
Vorkenntnisse bzgl. Fischen ^a	TI	$F < 1$
	ZA	$F < 1$
	TI \times ZA	$F < 1$

Anmerkung. ^aEine Absicherung der Analyse mit dem Kruskal-Wallis Test bestätigte das Ergebnis.

Tabelle K2

Ergebnisse der Loglinearen Analysen mit den Faktoren Textmodalität (TM) und Textinhalt (TI) für die nominalen Kontrollvariablen

	Schritt	Effekte	χ^2	df	Sig.	Anzahl Iterationen
Geschlecht	6	HE Geschlecht	36.99	1	<.001	0
Biologie 4-stündig/Leistungskurs	6	HE Bionote	14.31	1	<.001	0
Studienfach	Berechnung nicht möglich, da zu wenige Teilnehmer mit biologischen Vertiefungsfach					

Anhang L
Experiment 3 – Prüfung auf Normalverteilung

Tabelle L1

Ergebnisse der Prüfung auf Normalverteilung für die kontinuierlichen unabhängigen Variablen in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	gesprochen	geschrieben	gesprochen	geschrieben
Corsi Block Test	$D(20) = .37^*$	$D(20) = .34^*$	$D(20) = .25^*$	$D(20) = .35^*$
<i>Visual Pattern</i> Test	$D(20) = .12$	$D(20) = .18$	$D(20) = .10$	$D(20) = .10$
<i>Digit Span</i> Test	$D(20) = .22^*$	$D(20) = .21^*$	$D(20) = .19^*$	$D(20) = .23^*$

Anmerkung. ^aein Missing Data

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

Tabelle L2

Ergebnisse der Prüfung auf Normalverteilung für die Verifikationsitems und offenen Fragen in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	gesprochen	geschrieben	gesprochen	geschrieben
VERIFIKATIONSITEMS				
<i>textbezogene Erinnerungsleistung</i>				
Charakteristika	$D(20) = .09$	$D(20) = .15$	$D(20) = .11$	$D(20) = .11$
abstrakte Textinhalte	$D(20) = .15$	$D(20) = .15$	$D(20) = .10$	$D(20) = .13$
<i>bildbezogene Erinnerungsleistung</i>				
visuelle Bildinhalte	$D(20) = .18$	$D(20) = .12$	$D(20) = .17$	$D(20) = .13$
räumliche Bildinhalte	$D(20) = .12$	$D(20) = .13$	$D(20) = .17$	$D(20) = .15$
OFFENE FRAGEN				
<i>textbezogene Erinnerungsleistung</i>				
Charakteristika	$D(20) = .32^{**}$	$D(20) = .16$	$D(20) = .21^*$	$D(20) = .17$
abstrakte Textinhalte	$D(20) = .20^*$	$D(20) = .26^*$	$D(20) = .16$	$D(20) = .16$
<i>bildbezogene Erinnerungsleistung</i>				
visuelle Bildinhalte	$D(20) = .29^{**}$	$D(20) = .18$	$D(20) = .16$	$D(20) = .11$
räumliche Bildinhalte	$D(20) = .11$	$D(20) = .13$	$D(20) = .19$	$D(20) = .21^*$

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

Anhang L
Experiment 3 – Prüfung auf Normalverteilung

Tabelle L3

Ergebnisse der Prüfung auf Normalverteilung für die Blickbewegungsindizes in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	gesprochen	geschrieben	gesprochen	geschrieben
Ausmaß an Augenbewegungen	$D(17) = .17$	$D(18) = .22^*$	$D(17) = .19$	$D(19) = .13$
Fixationszeit auf Bild	$D(17) = .15$	$D(18) = .15$	$D(17) = .17$	$D(19) = .17$
Fixationszeit auf Text	—	$D(18) = .12$	—	$D(19) = .15$

Anmerkung. * $p < .05$.

Tabelle L4

Ergebnisse der Prüfung auf Normalverteilung für die Bewertungsitems in Abhängigkeit des Textinhalts und der Textmodalität

Textinhalt	visuell		räumlich	
	gesprochen	geschrieben	gesprochen	geschrieben
Wie anstrengend war die Lernaufgabe für Dich?	$D(20) = .15$	$D(20) = .21^*$	$D(20) = .23^*$	$D(20) = .33^*$
Wie schwierig war es für Dich, die Inhalte zu verstehen?	$D(20) = .27^*$	$D(20) = .26^*$	$D(20) = .16$	$D(20) = .27^*$
Wie sehr hast Du Dich beim Lernen auf das Bild konzentriert?	$D(20) = .34^*$	$D(20) = .16$	$D(20) = .26^*$	$D(20) = .20^*$
Wie sehr hast Du Dich beim Lernen auf den Text konzentriert?	$D(20) = .17$	$D(20) = .19^*$	$D(20) = .24^*$	$D(20) = .21^*$
Wie sehr hast Du Dich während der Lernphase konzentriert?	$D(20) = .20^*$	$D(20) = .35^*$	$D(20) = .29^*$	$D(20) = .27^*$
Wie sehr hast Du Dich angestrengt, um das Lernmaterial zu verstehen?	$D(20) = .22^*$	$D(20) = .20^*$	$D(20) = .24^*$	$D(20) = .20^*$
Wie viel Mühe hast Du in die Lernphase investiert?	$D(20) = .21^*$	$D(20) = .23^*$	$D(20) = .24^*$	$D(20) = .24^*$

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

Tabelle M1

Ergebnisse der Levene-Tests für die kontinuierlichen unabhängigen Variablen

Ergebnisse des Levene Tests	
Corsi Block Test	$F < 1$
Visual Pattern Test	$F < 1$
Digit Span Test	$F < 1$

Tabelle M2

Ergebnisse der Levene-Tests für die Verifikationsitems und offenen Fragen

Ergebnisse des Levene-Tests	
VERIFIKATIONSITEMS	
<i>textbezogene Erinnerungsleistung</i>	
Charakteristika	$F < 1$
abstrakte Textinhalte	$F < 1$
<i>bildbezogene Erinnerungsleistung</i>	
visuelle Bildinhalte	$F(3, 76) = 2.00, p = .12$
räumliche Bildinhalte	$F < 1$
OFFENE FRAGEN	
<i>textbezogene Erinnerungsleistung</i>	
Charakteristika	$F < 1$
abstrakte Textinhalte	$F(3, 76) = 3.77, p = .01$
<i>bildbezogene Erinnerungsleistung</i>	
visuelle Bildinhalte	$F(3, 76) = 1.65, p = .18$
räumliche Bildinhalte	$F < 1$

Tabelle M3

Ergebnisse der Levene-Tests für die Blickbewegungsindizes

Ergebnisse des Levene Tests	
Sakkaden	$F < 1$
Fixationszeit auf Bild	$F(3, 67) = 3.43, p = .02$
Fixationszeit auf Text	$F < 1$

Anhang M
Experiment 3 – Prüfung der Fehlervarianzen

Tabelle M4

Ergebnisse der Levene-Tests für die Bewertungsitems

	Ergebnisse des Levene Tests
Wie anstrengend war die Lernaufgabe für Dich?	$F < 1$
Wie schwierig war es für Dich, die Inhalte zu verstehen?	$F < 1$
Wie sehr hast Du Dich beim Lernen auf das Bild konzentriert?	$F(3, 76) = 1.81, p = .15$
Wie sehr hast Du Dich beim Lernen auf den Text konzentriert?	$F(3, 76) = 1.47, p = .23$
Wie sehr hast Du Dich während der Lernphase konzentriert?	$F < 1$
Wie sehr hast Du Dich angestrengt, um das Lernmaterial zu verstehen?	$F < 1$
Wie viel Mühe hast Du in die Lernphase investiert?	$F < 1$

Tabelle N1

Ergebnisse der MANCOVAs für die Verifikationsitems mit den Faktoren *Digit Span*, *VPT*, *Corsi Block (Corsi)*, *Textmodalität (TM)* und *Textinhalt (TI)*

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
räumliche Bildinhalte visuelle Bildinhalte Charakteristika	<i>Digit Span</i>	$V = .02, F < 1$
	VPT	$V = .05, F < 1$
	Corsi	$V = .02, F < 1$
	TI	$V = .36, F(3, 53) = 9.92, p < .001$
	TM	$V = .02, F < 1$
	TM × TI	$V = .06, F(3, 53) = 1.19, p = .32$
	TM × Corsi	$V = .07, F(3, 53) = 1.27, p = .30$
	TM × VPT	$V = .15, F(3, 53) = 3.15, p = .03$
	TI × Corsi	$V = .04, F < 1$
	TI × VPT	$V = .12, F(3, 53) = 2.34, p = .09$
	TM × TI × Corsi	$V = .02, F < 1$
	TM × TI × VPT	$V = .05, F < 1$
	TM × TI × <i>Digit Span</i>	$V = .21, F(9, 165) = 1.36, p = .21$
abstrakte Textinhalte	<i>Digit Span</i>	$F(1, 79) = 1.50, MSE = 42.72, p = .22, \eta_p^2 = .02$
	VPT	$F < 1$
	Corsi	$F(1, 79) = 2.33, MSE = 42.72, p = .13, \eta_p^2 = .04$
	TM	$F < 1$
	TI	$F(1, 79) = 1.60, MSE = 42.72, p = .21, \eta_p^2 = .02$
	TM × TI	$F < 1$
	TM × Corsi	$F(1, 79) = 1.58, MSE = 42.72, p = .21, \eta_p^2 = .02$
	TM × VPT	$F < 1$
	TI × Corsi	$F(1, 79) = 1.49, MSE = 42.72, p = .23, \eta_p^2 = .02$
	TI × VPT	$F < 1$
	TM × TI × Corsi	$F(1, 79) = 1.61, MSE = 42.72, p = .09, \eta_p^2 = .04$
	TM × TI × VPT	$F < 1$
	TM × TI × <i>Digit Span</i>	$F < 1$

Tabelle N2

Ergebnisse der MANCOVAs für die offenen Fragen mit den Faktoren *Digit Span*, *VPT*, *Corsi Block (Corsi)*, *Textmodalität (TM)* und *Textinhalt (TI)*

AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
räumliche Bildinhalte visuelle Bildinhalte Charakteristika	<i>Digit Span</i>	$V = .05, F < 1$
	VPT	$V = .03, F < 1$
	Corsi	$V = .13, F(3, 53) = 2.64, p = .06$
	TM	$V = .11, F(3, 53) = 2.25, p = .09$
	TI	$V = .14, F(3, 53) = 2.93, p = .04$
	TM × TI	$V = .05, F < 1$
	TM × Corsi	$V = .04, F < 1$
	TM × VPT	$V = .02, F < 1$
	TI × Corsi	$V = .06, F(3, 53) = 1.06, p = .37$
	TI × VPT	$V = .02, F < 1$
	TM × TI × Corsi	$V = .04, F < 1$
	TM × TI × VPT	$V = .06, F(3, 53) = 1.09, p = .36$
	TM × TI × <i>Digit Span</i>	$V = .15, F < 1$
abstrakte Textinhalte	<i>Digit Span</i>	$F(1, 79) = 1.12, MSE = 252.28, p = .29, \eta_p^2 = .02$
	VPT	$F < 1$
	Corsi	$F < 1$
	TM	$F < 1$
	TI	$F(1, 79) = 5.95, MSE = 252.28, p = .02, \eta_p^2 = .09$
	TM × TI	$F < 1$
	TM × Corsi	$F < 1$
	TM × VPT	$F < 1$
	TI × Corsi	$F < 1$
	TI × VPT	$F < 1$
	TM × TI × Corsi	$F(1, 79) = 2.33, MSE = 252.28, p = .13, \eta_p^2 = .04$
	TM × TI × VPT	$F(1, 79) = 1.66, MSE = 252.28, p = .20, \eta_p^2 = .03$
	TM × TI × <i>Digit Span</i>	$F < 1$

Tabelle N3

Ergebnisse der ANCOVAs für die Blickverhaltensindizes mit den Faktoren *Digit Span*, *VPT*, *Corsi Block (Corsi)*, *Textmodalität (TM)* und *Textinhalt (TI)*

Ergebnisse der ANCOVA		
AV: Fixationen auf Bild	<i>Digit Span</i>	$F < 1$
	VPT	$F < 1$
	Corsi	$F < 1$
	TM	$F(1, 70) = 1127.76, MSE = 713579423760.93, p = .01, \eta_p^2 = .95$
	TI	$F < 1$
	TM \times TI	$F < 1$
	TM \times Corsi	$F < 1$
	TM \times VPT	$F < 1$
	TI \times Corsi	$F < 1$
	TI \times VPT	$F < 1$
	TM \times TI \times Corsi	$F < 1$
	TM \times TI \times VPT	$F(1, 70) = 1.43, MSE = 713579423760.93, p = .24, \eta_p^2 = .03$
	TM \times TI \times <i>Digit Span</i>	$F(1, 70) = 1.34, MSE = 713579423760.93, p = .27, \eta_p^2 = .07$
	AV: Fixationen auf Text	<i>Digit Span</i>
VPT		$F < 1$
Corsi		$F < 1$
TI		$F < 1$
TI \times Corsi		$F < 1$
TI \times VPT		$F < 1$
TI \times <i>Digit Span</i>		$F < 1$
AV: Ausmaß an Blickbewegungen		<i>Digit Span</i>
	VPT	$F(1, 70) = 1.17, MSE = 6602324.93, p = .29, \eta_p^2 = .02$
	Corsi	$F < 1$
	TM	$F(1, 70) = 221.10, MSE = 6602324.93, p = .01, \eta_p^2 = .80$
	TI	$F(1, 70) = 2.22, MSE = 6602324.93, p = .14, \eta_p^2 = .04$
	TM \times TI	$F < 1$
	TM \times Corsi	$F(1, 70) = 4.88, MSE = 6602324.93, p = .03, \eta_p^2 = .08$
	TM \times VPT	$F < 1$
	TI \times Corsi	$F < 1$
	TI \times VPT	$F < 1$
	TM \times TI \times Corsi	$F < 1$
	TM \times TI \times VPT	$F < 1$
	TM \times TI \times <i>Digit Span</i>	$F(1, 70) = 1.24, MSE = 6602324.93, p = .30, \eta_p^2 = .06$

Tabelle N4

Ergebnisse der MANCOVAs für die Bewertungsitems mit den Faktoren Digit Span, VPT, Corsi Block (Corsi), Textmodalität (TM) und Textinhalt (TI)

AVn der MANCOVA	Effekte der MANCOVA	Ergebnisse der MANCOVA
Bewertungsitems	<i>Digit Span</i>	$V = .14, F(7, 58) = 1.17, p = .34$
	VPT	$V = .22, F(7, 58) = 2.00, p = .06$
	Corsi	$V = .10, F < 1$
	TM	$V = .29, F(7, 58) = 2.97, p = .01$
	TI	$V = .07, F < 1$
	TM × TI	$V = .16, F(7, 58) = 1.31, p = .26$
	TM × Corsi	$V = .09, F < 1$
	TM × VPT	$V = .04, F < 1$
	TI × Corsi	$V = .16, F(7, 58) = 1.36, p = .24$
	TI × VPT	$V = .13, F(7, 58) = 1.07, p = .40$
	TM × TI × Corsi	$V = .09, F < 1$
	TM × TI × VPT	$V = .12, F < 1$
	TM × TI × <i>Digit Span</i>	$V = .48, F(21, 180) = 1.39, p = .12$

Tabelle N5

Ergebnisse der MANCOVAs für die Verifikationsitems mit den Faktoren *Digit Span*, *VPT*, *Corsi Block (Corsi)*, *Anzahl an Sakkaden (AS)* und *Textinhalt (TI)*

	AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
Gesprochene Textdarbietung	räumliche Bildinhalte	<i>Digit Span</i>	$V = .15, F(3, 19) = 1.16, p = .35$
	visuelle Bildinhalte	VPT	$V = .18, F(3, 19) = 1.37, p = .28$
	Charakteristika	Corsi	$V = .08, F < 1$
		AS	$V = .08, F < 1$
		TI	$V = .25, F(3, 19) = 2.15, p = .13$
		AS × TI	$V = .08, F < 1$
		AS × TI × Corsi	$V = .25, F < 1$
		AS × TI × VPT	$V = .14, F(6, 40) = 1.47, p = .21$
		AS × TI × <i>Digit Span</i>	$V = .14, F < 1$
		abstrakte Textinhalte	<i>Digit Span</i>
	VPT	$F(1, 21) = 2.50, MSE = 0.53, p = .13, \eta_p^2 = .11$	
	Corsi	$F < 1$	
	AS	$F < 1$	
	TI	$F(1, 21) = 4.18, MSE = 0.53, p = .054, \eta_p^2 = .17$	
	AS × TI	$F(1, 21) = 1.59, MSE = 0.53, p = .22, \eta_p^2 = .07$	
	AS × TI × Corsi	$F < 1$	
	AS × TI × VPT	$F(2, 21) = 1.40, MSE = 0.53, p = .27, \eta_p^2 = .12$	
	AS × TI × <i>Digit Span</i>	$F < 1$	
Geschriebene Textdarbietung	räumliche Bildinhalte	<i>Digit Span</i>	$V = .19, F(3, 22) = 1.71, p = .19$
	visuelle Bildinhalte	VPT	$V = .21, F(3, 22) = 1.96, p = .15$
	Charakteristika	Corsi	$V = .03, F < 1$
		AS	$V = .20, F(3, 22) = 1.86, p = .17$
		TI	$V = .10, F < 1$
		AS × TI	$V = .04, F < 1$
		AS × TI × Corsi	$V = .31, F(6, 46) = 1.41, p = .23$
		AS × TI × VPT	$V = .43, F(6, 46) = 2.12, p = .07$
		AS × TI × <i>Digit Span</i>	$V = .25, F(6, 46) = 1.08, p = .39$
		abstrakte Textinhalte	<i>Digit Span</i>
	VPT	$F < 1$	
	Corsi	$F(1, 24) = 1.12, MSE = 0.53, p = .30, \eta_p^2 = .05$	
	AS	$F(1, 24) = 2.55, MSE = 0.53, p = .12, \eta_p^2 = .10$	
	TI	$F < 1$	
	AS × TI	$F < 1$	
	AS × TI × Corsi	$F(2, 24) = 1.02, MSE = 37.82, p = .37, \eta_p^2 = .08$	
	AS × TI × VPT	$F < 1$	
	AS × TI × <i>Digit Span</i>	$F < 1$	

Tabelle N6

Ergebnisse der (M)ANCOVAs für die offenen Fragen mit den Faktoren *Digit Span*, *VPT*, *Corsi Block (Corsi)*, *Anzahl an Sakkaden (AS)* und *Textinhalt (TI)*

	AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
Gesprochene Textdarbietung	räumliche Bildinhalte	<i>Digit Span</i>	$V = .03, F < 1$
	visuelle Bildinhalte	VPT	$V = .25, F(3, 19) = 2.13, p = .13$
	Charakteristika	Corsi	$V = .11, F < 1$
		AS	$V = .03, F < 1$
		TI	$V = .40, F(3, 19) = 4.29, p = .02$
		AS × TI	$V = .31, F(3, 19) = 2.79, p = .07$
		AS × TI × Corsi	$V = .29, F(6, 40) = 1.16, p = .35$
		AS × TI × VPT	$V = .29, F(6, 40) = 1.12, p = .37$
		AS × TI × <i>Digit Span</i>	$V = .11, F < 1$
Gesprochene Textdarbietung	abstrakte Textinhalte	<i>Digit Span</i>	$F < 1$
		VPT	$F < 1$
		Corsi	$F < 1$
		AS	$F < 1$
		TI	$F < 1$
		AS × TI	$F(1, 21) = 1.45, MSE = 4.32, p = .24, \eta^2 = .07$
		AS × TI × Corsi	$F < 1$
		AS × TI × <i>Digit Span</i>	$F < 1$
Geschriebene Textdarbietung	räumliche Bildinhalte	<i>Digit Span</i>	$V = .17, F(3, 22) = 1.49, p = .24$
	visuelle Bildinhalte	VPT	$V = .03, F < 1$
	Charakteristika	Corsi	$V = .08, F < 1$
		AS	$V = .21, F(3, 22) = 1.93, p = .15$
		TI	$V = .23, F(3, 22) = 2.15, p = .12$
		AS × TI	$V = .24, F(3, 22) = 2.26, p = .11$
		AS × TI × Corsi	$V = .39, F(6, 46) = 1.86, p = .11$
		AS × TI × VPT	$V = .14, F < 1$
		AS × TI × <i>Digit Span</i>	$V = .13, F < 1$
		Geschriebene Textdarbietung	abstrakte Textinhalte
VPT	$F(1, 24) = 2.44, MSE = 1.36, p = .13, \eta_p^2 = .09$		
Corsi	$F(1, 24) = 8.77, MSE = 1.36, p = .01, \eta_p^2 = .26$		
AS	$F(1, 24) = 4.66, MSE = 1.36, p = .04, \eta_p^2 = .16$		
TI	$F(1, 24) = 12.37, MSE = 1.36, p = .002, \eta_p^2 = .34$		
AS × TI	$F(1, 24) = 7.41, MSE = 1.36, p = .03, \eta_p^2 = .19$		
AS × TI × Corsi	$F(1, 24) = 8.24, MSE = 1.36, p = .002, \eta_p^2 = .41$		
AS × TI × <i>Digit Span</i>	$F(1, 24) = 7.44, MSE = 1.36, p = .09, \eta_p^2 = .18$		

Tabelle N7

Ergebnisse der MANCOVAs für die Bewertungsitems mit den Faktoren Digit Span, VPT, Corsi Block (Corsi), Anzahl an Sakkaden (AS) und Textinhalt (TI)

	AVn der (M)ANCOVA	Effekte der (M)ANCOVA	Ergebnisse der (M)ANCOVA
Gesprochene Textdarbietung	Bewertungsitems	<i>Digit Span</i>	$V = .27, F < 1$
		VPT	$V = .41, F(7, 15) = 1.22, p = .35$
		Corsi	$V = .18, F < 1$
		AS	$V = .35, F < 1$
		TI	$V = .41, F(7, 15) = 1.20, p = .36$
		AS \times TI	$V = .53, F(7, 15) = 1.96, p = .13$
		AS \times TI \times Corsi	$V = .33, F < 1$
		AS \times TI \times VPT	$V = .87, F(7, 15) = 1.43, p = .19$
		AS \times TI \times <i>Digit Span</i>	$V = .87, F(14, 32) = 1.43, p = .19$
Geschriebene Textdarbietung	Bewertungsitems	<i>Digit Span</i>	$V = .25, F < 1$
		VPT	$V = .28, F < 1$
		Corsi	$V = .61, F(7, 18) = 3.41, p = .02$
		AS	$V = .14, F < 1$
		TI	$V = .23, F < 1$
		AS \times TI	$V = .16, F < 1$
		AS \times TI \times Corsi	$V = .84, F(14, 38) = 1.62, p = .11$
		AS \times TI \times VPT	$V = .64, F(14, 38) = 1.06, p = .42$
		AS \times TI \times <i>Digit Span</i>	$V = .39, F < 1$

Anhang O
Experiment 3 – Gleichheit der Eingangsbedingungen

Tabelle O1

Ergebnisse der Varianzanalysen mit den Faktoren Textmodalität (TM) und Textinhalt (TI) für die Kontrollvariablen

Ergebnisse der ANOVA		
Alter ^a (in Jahren)	TM	$F(1, 76) = 1.79, MSE = 7.26, p = .19, \eta_p^2 = .02$
	TI	$F < 1$
	TM × TI	$F < 1$
Semesteranzahl ^a	TM	$F(1, 76) = 1.21, MSE = 16.53, p = .28, \eta_p^2 = .02$
	TI	$F < 1$
	TM × TI	$F < 1$
Letzte Bionote ^a	TM	$F < 1$
	TI	$F < 1$
	TM × TI	$F(1, 75) = 1.77, MSE = 10.07, p = .19, \eta_p^2 = .02$
Vorkenntnisse bzgl. Pflanzen ^a	TM	$F(1, 76) = 3.30, MSE = 0.46, p = .07, \eta_p^2 = .04$
	TI	$F(1, 76) = 1.34, MSE = 0.46, p = .25, \eta_p^2 = .02$
	TM × TI	$F < 1$

Tabelle O2

Ergebnisse der Loglinearen Analysen mit den Faktoren Textmodalität (TM) und Textinhalt (TI) für die nominalen Kontrollvariablen

	Schritt	Effekte	χ^2	df	Sig.	Anzahl Iterationen
Geschlecht	6	HE Geschlecht	27.93	1	<.001	0
Biologie 4- stündig/Leistungskurs	6	HE Bionote	25.60	1	<.001	0
Studienfach	Berechnung nicht möglich, da zu wenige Teilnehmer mit biologischen Vertiefungsfach					