

**Lernen mit Animationen:
Effekte dynamischer und statischer
Visualisierungen auf die Bildung perzeptueller und
kognitiver Repräsentationen beim Erwerb von
Wissen über dynamische Sachverhalte**

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

vorgelegt von

Inga Wagner

Vorsitzender des Promotionsausschusses: Prof. Dr. M. Schmitt

Erster Berichterstatter: Prof. Dr. W. Schnotz

Zweite Berichterstatterin: Prof. Dr. M. Bannert

Datum der wissenschaftlichen Aussprache: 15.01.2013

Danksagung

Bei der Erstellung dieser Arbeit wurde ich beruflich und privat von einigen Menschen unterstützt, bei denen ich mich gerne an dieser Stelle bedanken würde.

Als erstes möchte ich mich bei meinem Betreuer Herrn Prof. Dr. Schnotz für die Möglichkeit bedanken, in seiner Abteilung mit einem Stipendium promovieren zu können, wodurch ich mich auf das Gelingen der Dissertation konzentrieren konnte. Seine fachlichen Ratschläge sowie seine Geduld, Ruhe und Zuversicht waren mir während des längeren Entstehungsprozesses der Dissertation eine große Unterstützung.

Desweiteren möchte ich Frau Prof. Dr. Bannert für die sofortige Zusage danken, das Zweitgutachten zu übernehmen.

Ein weiterer großer Dank gebührt allen meinen ehemaligen und aktuellen Kollegen und Kolleginnen der Arbeitseinheit „Allgemeine und Pädagogische Psychologie“ für die vielfältige fachliche und motivationale Hilfe sowie für die sehr angenehme Arbeitsatmosphäre. Hervorheben möchte ich an der Stelle Herrn Dr. Christoph Mengelkamp, der mir vor allem in statistischen Fragen mit Rat und Tat und mit großer Geduld zur Seite stand.

Bei der Vorbereitung und Durchführung der Studien wurde ich von studentischen Hilfskräften unterstützt, bei denen ich mich ebenfalls bedanken möchte. Zu nennen sind hier Radu Georghiu, der das qualitativ hochwertige Lernmaterial technisch erstellt hat sowie Katharina Allgaier und Lena Büscher, die einige der Erhebungen selbständig übernommen haben. Eine große Hilfe bei den Durchführungen war mir Sabine Boysen, deren zuverlässige und kompetente Arbeit mir die Datensammlung sehr erleichtert hat.

Zuletzt möchte ich meinen Eltern, Walter und Naemi Wagner, sowie meiner Schwester und meinem Schwager, Haike und Jörg Blinn, für den jahrelangen Rückhalt und für die motivationale Unterstützung während meiner Promotion danken.

Landau, im September 2012

Inga Wagner

Inhaltsverzeichnis

1	Einleitung.....	13
2	Theoretischer Hintergrund	17
2.1	Lernen mit statischen Bildern.....	17
2.1.1	Allgemeine Merkmale von statischen Bildern	17
2.1.2	Grundlegende Prozesse beim Bildverstehen	22
2.1.2.1	Mentale Repräsentationen beim Bildverstehen	23
2.1.2.2	Aspekte des Gedächtnisses beim Bildverstehen.....	25
2.1.2.3	Visuelles Bildverstehen nach dem Modell von Schnotz und Bannert (2003).....	26
2.1.3	Wahrnehmungs- und Verstehensprozesse dynamischer Sachverhalte anhand statischer Bilder.....	28
2.1.3.1	Allgemeine Befunde	28
2.1.3.2	Vorteile beim Lernen dynamischer Sachverhalte anhand statischer Bilder .	29
2.1.3.3	Anforderungen beim Lernen dynamischer Sachverhalte anhand statischer Bilder	31
2.2	Lernen mit Animationen.....	35
2.2.1	Allgemeine Merkmale von Animationen	35
2.2.2	Grundlegende Prozesse beim Verstehen von Animationen	38
2.2.3	Wahrnehmungs- und Verstehensprozesse dynamischer Sachverhalte anhand von Animationen.....	39
2.2.3.1	Allgemeine Befunde	39
2.2.3.2	Vorteile beim Lernen dynamischer Sachverhalte anhand von Animationen	40

2.2.3.3	Anforderungen beim Lernen dynamischer Sachverhalte anhand von Animationen	42
2.2.3.4	Gestaltungsempfehlungen zur Verbesserung des Designs von Animationen	45
2.3	Studien zum Vergleich von Animationen mit statischen Bildern	49
3	Zielsetzungen der vorliegenden Arbeit	55
3.1	Theoretische Herleitung der Grundannahmen von Studie 1: Vergleich von Animationen und statischen Bildern hinsichtlich ihrer Wirksamkeit auf die Lösung von Inferenzaufgaben	56
3.1.1	Teilstudie 1	56
3.1.2	Teilstudie 2	73
3.2	Theoretische Herleitung der Grundannahmen von Studie 2: Vergleich von Animationen und statischen Bildern hinsichtlich ihrer Wirksamkeit auf die Lösung von bildlichen Kategorisierungsaufgaben	79
4	Empirischer Teil von Studie 1: Vergleich von Animationen und statischen Bildern hinsichtlich ihrer Wirksamkeit auf die Lösung von Inferenzaufgaben.....	91
4.1	Vorpilotstudie	91
4.1.1	Fragestellungen und Zielsetzung	91
4.1.2	Methoden.....	93
4.1.2.1	Stichprobe	93
4.1.2.2	Design	93
4.1.2.3	Lernmaterialien.....	95
4.1.2.4	Messinstrumente	97
4.1.2.5	Durchführung.....	101
4.1.3	Ergebnisse	103

4.1.4	Diskussion	118
4.2	Pilotstudie	122
4.2.1	Fragestellung und Hypothesen	122
4.2.2	Methoden.....	126
4.2.2.1	Stichprobe	126
4.2.2.2	Design	126
4.2.2.3	Lernmaterialien.....	127
4.2.2.4	Messinstrumente	134
4.2.2.5	Durchführung.....	140
4.2.3	Ergebnisse	143
4.2.4	Diskussion	161
4.3	Hauptstudie	167
4.3.1	Teilstudie 1	167
4.3.1.1	Fragestellung und Hypothesen	167
4.3.1.2	Methoden	173
4.3.1.2.1	Stichprobe	173
4.3.1.2.2	Design	174
4.3.1.2.3	Lernmaterialien.....	174
4.3.1.2.4	Messinstrumente	175
4.3.1.2.5	Durchführung.....	179
4.3.1.3	Ergebnisse.....	181
4.3.1.4	Diskussion	209
4.3.2	Teilstudie 2	223

4.3.2.1	Fragestellung und Hypothesen	223
4.3.2.2	Methoden	226
4.3.2.3	Ergebnisse.....	227
4.3.2.4	Diskussion	240
5 Empirischer Teil von Studie 2: Vergleich von Animationen und statischen Bildern hinsichtlich ihrer Wirksamkeit auf die Lösung von bildlichen Kategorisierungsaufgaben		
247		
5.1	Pilotstudie	247
5.1.1	Fragestellung und Hypothesen	247
5.1.2	Methoden.....	249
5.1.2.1	Stichprobe	249
5.1.2.2	Design	249
5.1.2.3	Lernmaterialien.....	251
5.1.2.4	Messinstrumente	259
5.1.2.5	Durchführung.....	268
5.1.3	Ergebnisse	271
5.1.4	Diskussion	287
5.2	Hauptstudie	293
5.2.1	Fragestellung und Hypothesen	293
5.2.2	Methoden.....	297
5.2.2.1	Stichprobe	297
5.2.2.2	Design	297
5.2.2.3	Lernmaterialien.....	298

5.2.2.4	Messinstrumente	299
5.2.2.5	Durchführung.....	302
5.2.3	Ergebnisse	304
5.2.4	Diskussion	324
6	Integration und Gesamtdiskussion der Befunde.....	335
7	Methodische Kritik und Ausblick	343
8	Literaturverzeichnis.....	349
9	Anhang	357

Zusammenfassung

Das Lernen und Verstehen dynamischer Sachverhalte kann sowohl anhand von statischen Bildern als auch von Animationen erfolgen. Da die bisher vorliegenden Befunde keine eindeutige Überlegenheit der einen oder der anderen Repräsentationsform erkennen lassen, sind differenziertere Wirkungsanalysen erforderlich. Grundannahme der vorliegenden Dissertation war, dass beide Darstellungsformen unterschiedlich gut für den Aufbau dynamischer mentaler Modelle einerseits und den Aufbau dynamischer perzeptueller Repräsentationen andererseits geeignet sind.

Es wurde erstens angenommen, dass dynamische mentale Modelle keine kontinuierlichen mentalen Simulationen ermöglichen, sondern aus einer Sequenz von diskreten prototypischen Prozess-Zuständen (sog. „key states“) bestehen und deshalb besser anhand statischer Bilder dieser „key states“ als anhand einer Animation konstruiert werden können. Zweitens wurde angenommen, dass perzeptuelle Repräsentationen dynamischer Sachverhalte aus dynamischen Schemata bestehen, die besser anhand einer Animation als anhand von statischen Bildern gebildet werden können. Als Indikator für die Qualität mentaler Modelle wurden Inferenzaufgaben verwendet, bei denen durch Manipulation dieser Modelle neue Informationen zu erschließen sind. Als Indikator für die Qualität dynamischer perzeptueller Repräsentationen wurden Leistungen bei der Kategorisierung von Bewegungsabläufen verwendet.

Die erste Annahme wurde am Beispiel der synaptischen Informationsübertragung (nach einer Vorpilot- und Pilotstudie) in einer ersten Hauptstudie mit 61 Probanden überprüft. Die Ergebnisse weisen darauf hin, dass die Hypothese eher bei Probanden mit höherem Vorwissen als bei solchen mit niedrigerem Vorwissen zutrifft. Die zweite Annahme wurde am Beispiel des Erkennens der Gangart von Pferden (nach einer Pilotstudie) in einer zweiten Hauptstudie mit 52 Probanden überprüft. Die Ergebnisse zeigten durchwegs Unterschiede in der erwarteten Richtung, die jedoch nicht oder nur marginal signifikant sind. Allerdings führten die Animationen zu einer hochsignifikant geringeren kognitiven Belastung als statische Bilder. Darüber hinaus zeigten sich signifikante Interaktionen mit den räumlichen Fähigkeiten: Probanden mit geringeren Fähigkeiten profitieren bei der Kategorisierung von Bewegungsabläufen bedeutsam mehr von Animationen als von statischen Bildern.

1 Einleitung

Animationen sind ein moderner Bestandteil von Lernmaterialien geworden und werden auch gerne auf Homepages oder im E-learning-Bereich eingesetzt. Sie gelten als innovativ und dekorativ und vor allem für junge Leute, die mit Video- und Computerspielen aufgewachsen sind, als motivierend und optisch ansprechend.

Allerdings sind Animationen aus pragmatischer Sicht sehr kostspielig und zeitaufwändig in der Herstellung. Und auch aus lernpsychologischer Perspektive bergen sie nicht nur Vorteile. Die bisherige Forschung hat nicht ergeben, dass in der Vermittlung von dynamischen Lerninhalten Animationen gegenüber statischen Bildern generell überlegen sind. Eher stellen sich die Befunde diesbezüglich als sehr inkonsistent dar (z.B. Lowe, 2001; Rieber, 1990).

Daher empfiehlt es sich, zu untersuchen, unter welchen Umständen Animationen sinnvoll und gewinnbringend eingesetzt werden können und unter welchen Bedingungen sie statischen Bildern gegenüber in der Repräsentation von Dynamik vorzuziehen sind.

In der vorliegenden Dissertation soll ein solcher differenzierterer Vergleich zwischen Animationen und statischen Bildern angestellt werden. Der Grundgedanke ist, dass das Lernen dynamischer Sachverhalte sowohl auf der kognitiven als auch auf der perzeptuellen Ebene ablaufen kann, wobei kognitive oder perzeptuelle dynamische mentale Repräsentationen aufgebaut werden. Möglicherweise sind Animationen und statische Bilder unterschiedlich gut für den Aufbau dieser mentalen Repräsentationen geeignet, was sich auf empirischer Ebene in einer Aufgabenabhängigkeit der Effektivität von Animationen und statischen Bildern äußern würde. Der Grund hierfür liegt darin, dass für die Lösung bestimmter Aufgabenarten entweder kognitive oder perzeptuelle mentale Repräsentationen herangezogen werden, wie im Folgenden näher erläutert werden wird.

Wenn ein linearer Prozess, wie beispielsweise die „Signalübertragung an einer Synapse“, die in der ersten Studie der vorliegenden Dissertation auch Thema des Lernmaterials sein wird, verstanden und gelernt wird, wird vom Betrachter auf kognitiver Ebene ein dynamisches mentales Modell dieses dynamischen Vorganges aufgebaut. Die Konstruktion des dynamischen mentalen Modells erfolgt vermutlich aus den „key states“, also den prototypischen Zuständen eines Ereignisses, heraus, da mit Hilfe dieser Bilder der Prozess „kognitiv ökonomischer“ repräsentiert wird (Schnotz & Lowe, 2008). Daher erleichtert vermutlich eine externe Präsentation eines dynamischen Sachverhaltes in Form statischer

Bilder, die die „key states“ des Vorganges zeigen, den Aufbau eines dynamischen mentalen Modells, da der menschliche kognitive Apparat dann nicht wie bei einer Animation den dargestellten Prozess erst strukturieren und die Schlüsselzustände extrahieren muss. Sollen nach einem Lernprozess Inferenzaufgaben gelöst werden, bei denen neue Informationen durch eine Manipulation des dynamischen mentalen Modells erschlossen werden, ist es aufgrund dieser Argumentation vermutlich von Vorteil, wenn zum Lernen statische Bilder herangezogen werden. Daher wird in der vorliegenden Arbeit davon ausgegangen, dass für den Aufbau dynamischer mentaler Modelle und somit für die Lösung von Inferenzaufgaben statische „key frames“ (= statische Bilder, die „key states“ zeigen), geeigneter sind als Animationen. Die Fragestellung bezüglich der Lösung von Inferenzaufgaben durch statische Bilder oder durch Animationen soll in der ersten Studie der Arbeit untersucht werden.

Wenn ein Mensch hingegen biologische, repetitive Bewegungsabläufe, wie beispielsweise das Galoppieren eines Pferdes, beobachtet und versucht, die invarianten Bewegungsmuster zu analysieren, findet Lernen auf perzeptueller Ebene statt. Der kognitive und perzeptuelle Apparat des Menschen ist evolutionär dazu ausgestattet, die räumlich-zeitlichen Bewegungsmuster von biologischen Bewegungen schnell und einfach zu extrahieren und in dynamische perzeptuelle Schemata zu enkodieren. Ein solches dynamisches perzeptuelles Schema hilft in der Folge dabei, die entsprechende Bewegung wiedererkennen zu können (Schnotz & Lowe, 2008). Sollen verschiedene Bewegungsmuster, wie die von Pferdegangarten, wie es in der vorliegenden Arbeit der Fall sein wird, erlernt werden, bietet sich vermutlich eher ein Lernvorgang mit Animationen an, da die dynamischen perzeptuellen Schemata dann analog zum alltäglichen Wahrnehmungsprozess aufgebaut werden können. Die Lösung von bildlichen Kategorisierungsaufgaben, bei denen verschiedene Bewegungsarten mit Hilfe der dynamischen perzeptuellen Schemata erkannt und voneinander unterschieden werden müssen, dürfte nach einem Lernprozess mit Animationen dann auch besser gelingen als wenn mit statischen Bildern gelernt wurde. Daher ist die Annahme in der zweiten Studie der Dissertation, dass Animationen für den Aufbau dynamischer perzeptueller Schemata und somit für die Lösung bildlicher Kategorisierungsaufgaben besser geeignet sind, als statische Bilder, die die „key states“ der Bewegungen zeigen.

In beiden Studien werden zudem mögliche Wechselwirkungen zwischen den Lernbedingungen und individuellen Merkmalen, wie dem Vorwissen, den räumlichen Fähigkeiten oder der Merkfähigkeit einer Person mit berücksichtigt, so dass Aptitude-Treatment-Interaktionen erfasst werden können.

Der Aufbau der Dissertation ist wie folgt: Zunächst wird ein theoretischer Hintergrund zum Lernen mit stehenden Bildern und mit Animationen gegeben. Ein kurzes Review veranschaulicht nochmals die Inkonsistenz der Befunde zum Vergleich von Animationen und statischen Bildern, wodurch auf die Herleitung der aktuellen Hypothesen der Arbeit übergeleitet wird.

Es folgt der empirische Teil zu der ersten Studie der Arbeit, die in zwei Teilstudien untergliedert ist.

In der ersten Teilstudie werden in einem einfaktoriellen Design Animationen mit statischen Key-Frames verglichen, die seriell hintereinander gezeigt werden, wobei die Abfolgegeschwindigkeit der einzelnen Frames in einem Fall schneller und im anderen Fall langsamer ist. Zudem werden auch noch die beiden seriellen Bildbedingungen miteinander kontrastiert. Es werden deshalb zwei verschiedene Präsentationsgeschwindigkeiten der seriellen Bilder verwendet, da davon ausgegangen wird, dass bei der seriell langsamen Variante eher strukturelle und bei der seriell schnellen Form eher dynamische Aspekte des Prozesses betont werden, wodurch unterschiedliche Kontraste im Vergleich zur Animation entstehen dürften. Wie bereits erwähnt, hat das Lernmaterial die „Signalübertragung an Synapsen“ (= Verbindungen zwischen zwei Nervenzellen) zum Thema. Abhängige Variablen sind die Wissens- und die Inferenzskala eines selbst konstruierten Nachtests sowie die mentale Belastung („cognitive load“, Sweller, 1988), die von den Probanden während des Lernens und während des Nachtests empfunden wird.

In der zweiten Teilstudie werden als erweiternde Fragestellung die Effekte unterschiedlicher Präsentationsformen von statischen Bildern auf die Lösung von Inferenzaufgaben untersucht und miteinander verglichen. Die beiden seriellen Bildbedingungen der ersten Teilstudie werden in einem einfaktoriellen Design mit einer Gruppe kontrastiert, die die Key-Frames parallel und gleichzeitig auf dem Bildschirm angeordnet betrachtet. Das Lernmaterial und die abhängigen Variablen sind somit die gleichen wie in der ersten Teilstudie.

Die beiden Teilstudien werden aus pragmatischen Gründen gemeinsam an einer studentischen Stichprobe durchgeführt.

Vor der Darstellung der beiden Hauptstudien werden in der Arbeit auch die Vorpilotstudie und die Pilotstudie der ersten Studie berichtet, um den Verlauf der aufwändigeren Materialentwicklung und die diesbezüglichen Überlegungen zu verdeutlichen.

Danach folgt der empirische Teil der zweiten Studie der Dissertation. Hier wird in einem zweifaktoriellen Design mit Messwiederholung eine Animationsbedingung mit einer Bedingung verglichen, die das Lernmaterial mit Hilfe von simultan auf dem Screen angeordneter Key-Frames betrachtet. Das Thema des Lernmaterials sind „Pferdegangarten“. Kriterien sind die Leistungen bei der Kategorisierung von Bewegungen und wiederum der Cognitive Load, der während des Lernens und während des Nachtests empfunden wird. Die Studie wird ebenfalls an Studenten durchgeführt.

Auch bei der zweiten Studie wird der Hauptstudie im Bericht eine Beschreibung der Pilotstudie vorangestellt.

Die Ergebnisse der Studien werden in den empirischen Teilen zunächst einzeln diskutiert. Am Ende der Arbeit werden die Befunde miteinander integriert und in ein Gesamtbild geordnet. Zuletzt wird noch ein Ausblick für weitere Forschungen gegeben.

2 Theoretischer Hintergrund

2.1 Lernen mit statischen Bildern

2.1.1 Allgemeine Merkmale von statischen Bildern

Statische Bilder sind depiktive Repräsentationen. Depiktive Repräsentationen bestehen aus sog. „icons“ bzw. Bildzeichen, die mit ihrem Bezeichneten durch Ähnlichkeit oder strukturelle Gemeinsamkeiten verbunden sind. Bei realistischen Bildern besteht eher eine Ähnlichkeit zwischen der Darstellung und dem Dargestellten, während bei Diagrammen eine Analogierelation zwischen diesen beiden Entitäten vorliegt. In depiktiven Repräsentationen können nur spezifische Dinge dargestellt werden, wobei sie den Vorteil haben, komplette Informationen über einen Sachverhalt oder einen Gegenstand zu liefern, beispielsweise kann eine Karte vollständige geometrische Informationen über ein dargestelltes Gebiet beinhalten. Depiktive Repräsentationen sind besonders nützlich, um Inferenzen zu ziehen, da neue Informationen direkt von der Abbildung abgelesen werden können (Schnotz, Böckheler & Grzondziel, 1999; Kosslyn, 1994).

Generell gibt es drei Arten von statischen Bildern: Realistische Bilder, wie Fotos oder Zeichnungen, die einer dargestellten Realität ähneln. Analogiebilder zeigen hingegen nicht den gemeinten Sachverhalt, sondern einen anderen, der zum eigentlich Gemeinten in einer Analogiebeziehung steht. Logische Bilder dienen zur Veranschaulichung von abstrakten Sachverhalten, die teilweise nicht unmittelbar wahrnehmbar sind. Linien- oder Streudiagramme zählen beispielsweise zu den logischen Bildern (Schnotz, 2006).

Bilder können kognitive, motivationale oder dekorative Funktionen erfüllen. Zu den kognitiven Funktionen gehören beispielsweise die Konkretisierungsfunktion, bei der ein verbal beschriebener Sachverhalt veranschaulicht wird oder die Organisationsfunktion, bei der Bilder einen Überblick über einen schwierigeren Sachverhalt geben und somit einen Bezugsrahmen für detailliertere Informationen bilden. Die Motivationsfunktion besagt, dass Bilder auf den Lerner interessant wirken und eine höhere Motivation zur kognitiven Verarbeitung des Lerninhalts schaffen können. Bilder erfüllen dann eine Dekorationsfunktion, wenn sie ästhetisch ansprechend wirken und das Bedürfnis des Lerners nach Wohlgeformtheit befriedigen (Schnotz, 2006).

Im Vergleich zu sprachlichen Darstellungen werden durch Bilder mehr Informationen pro Zeiteinheit präsentiert, die in Texten anhand vieler Wörter umständlich umschrieben werden müssten. Zudem kann die Bedeutung einer Abbildung auch dann gut erschlossen werden, wenn einzelne Teile von ihr fehlen, was bei Texten schwieriger möglich ist (Engelkamp, 1991; Seufert, 2003). Statische Bilder haben in Kombination mit Texten vor allem die Funktion, Informationen zu visualisieren oder eine verbale Beschreibung leichter zu machen. Insbesondere die Enkodierung von Informationen wird bei Texten durch das Hinzufügen von Bildern verbessert, da die Daten dann sowohl im Sprach- als auch im Bildsystem enkodiert werden („dual-encoding“, Paivio, 1978). Bilder sind im Zusammenhang mit Texten dann hilfreich, wenn sie mit dem sprachlichen Material kongruent sind, sie können aber auch ablenkend wirken, wenn sie keinen Bezug zu dem Text haben oder wenn sie zu komplex sind. Vor allem Kinder verlassen sich bis zu einem Alter von 9 bis 10 Jahren auf äußere Bilder, wobei diese Abhängigkeit von externen Darstellungen im Laufe der Entwicklung abnimmt, da Jugendliche und Erwachsene besser in der Lage sind, ihre eigenen mentalen Bilder zu erzeugen (Willows, 1978; Guttman, Levin & Pressley, 1977).

Bilder sollten syntaktischen, semantischen und pragmatischen Gestaltungskriterien genügen. Syntaktische Gestaltungskriterien bedeuten, dass die Elemente gut voneinander unterscheidbar und so angeordnet sein sollten, dass eindeutige Konfigurationen entstehen, da zusammengehörige Komponenten vom perzeptuellen Apparat aufgrund von Ähnlichkeit oder räumlicher Nähe als Einheit wahrgenommen werden. Semantische Kriterien stehen dafür, dass die wahrgenommene visuelle Konfiguration gut mit der Struktur des darzustellenden Sachverhaltes übereinstimmen sollte. Visuelle Merkmale der Bilder haben hierbei bestimmte semantische Funktionen. Bei den pragmatischen Kriterien spielt eine Rolle, für welchen Zweck ein Bild verwendet werden soll. Insbesondere soll ein Bild einen Betrachter nicht zu falschen Schlüssen verleiten (Schnotz, 2006).

Ein weiteres wichtiges Gestaltungskriterium statischer Bilder betrifft ihren Grad an Realismus. In einem Review von über 100 Studien berichtet Dwyer (1978), dass sehr detaillierte, realistische Bilder dann von Vorteil sind, wenn genügend Zeit für ihre Inspektion vorhanden ist, da der Lerner bei einem hohen Detailliertheitsgrad nach relevanten Lernhinweisen innerhalb der Abbildung suchen muss. Unter Zeitdruck werden zu realistische Bilder von Lernern ignoriert, so dass es unter restriktiveren Bedingungen ratsam

ist, dass die Bilder nur geringe Mengen an Details enthalten. Bei Lewalter (1997) werden ebenfalls Illustrationen, die einen mittleren Grad an Komplexität aufweisen, als hilfreich angesehen. Eine größere Schematisierung von Abbildungen, bei der wichtige Informationen akzentuiert und irrelevante Informationen hingegen weggelassen werden, könnte dem „Apprehension Principle“ (Tversky, Morrison & Bétrancourt, 2002) von effektiven Bildern eher Rechnung tragen, das aussagt, dass die Struktur und der Inhalt einer externen Repräsentation so beschaffen sein sollten, dass sie unmittelbar und akkurat wahrgenommen und verstanden werden können. Oft sind statische Bilder gerade dann von Vorteil, wenn sie Zeit und Raum nicht kontinuierlich abbilden, wie es beispielsweise bei Straßenkarten der Fall ist, die bei einer Wegbeschreibung unwesentliche Informationen weglassen oder verzerren (Tversky, Heiser, Mackenzie, Lozano & Morrison, 2008). Wie detailliert ein Bild ist, wird auch durch seine räumliche Granularität beschrieben. Bei hoher räumlicher Granularität liegt ein hoher räumlicher Realismus vor. Der Realismus einer Abbildung steigt jedoch nur bis zu einem gewissen Grad an Granularität mit an, liegen die Items des Bildes allerdings zu eng beieinander, werden sie von dem menschlichen Auge nicht mehr erkannt. Bei einer Serie von statischen Bildern bedeutet Granularität auch die Anzahl der verwendeten Frames. Bei einer geringen Granularität muss der Lerner große Inferenzen zwischen den Bildern machen, bei einer hohen Granularität hingegen können die Veränderungsinformationen eher direkt abgelesen werden. Allerdings wird bei einer hohen Anzahl von Bildern die räumliche Realität eines Vorganges verletzt. Beispielsweise würde der Sprung eines Känguruhs dann weiter erscheinen, als er eigentlich ist. Zudem ist der Vergleich zwischen den Frames schwieriger, da sie teilweise weiter voneinander entfernt sind, was auf eine geringe intrapiktoriale räumliche Kontiguität dieser Darstellungsform hindeutet (Schnotz & Lowe, 2008; Mayer, 2001).

In gestaltungstechnischer Hinsicht ist weiterhin zu sagen, dass statische Bilder durch Namen, Symbole, Kästen oder Pfeile angereichert werden können. Jedoch können statische Bilder Objekte nur aus einer festgelegten Perspektive heraus darstellen.

Generell sollten die relevanten Gestaltungsaspekte von statischen Bildern für die konkrete Lernsituation jeweils einzeln festgelegt werden, wobei Lernermerkmale und Lernziele eine wichtige Rolle spielen sollten (Tversky et al., 2008; Lewalter, 1997).

Statische Bilder sind ideal, um Strukturen darzustellen, da sie die wichtigen Teile eines Objektes in den richtigen räumlichen Beziehungen zueinander zeigen. Die häufige Annahme, die analog zu dem „Congruence Principle“ (Tversky et al., 2002) ist, dass eine

externe visuelle Repräsentation so sein sollte wie ihr dargestellter Inhalt, würde für statische Bilder bedeuten, dass sie statischen Inhalt präsentieren sollen, damit ein statisches mentales Modell vom kognitiven Apparat aufgebaut werden kann. Ein dynamischer Vorgang kann aber auch durch mehrere statische Einzelbilder dargestellt werden, die wichtige Ausschnitte des Prozesses zeigen. Die Einzelbilder unterscheiden sich nur durch Details, wodurch der perzeptuelle und kognitive Apparat des Menschen die Lücken zwischen den Frames ausfüllen kann, so dass der Eindruck einer kontinuierlichen Bewegung entsteht. Durch „top down“-Prozesse werden die Bilder somit zu einer organisierten, bedeutungsvollen Gesamtinformation zusammengefügt (Lewalter, 1997). Durch diese menschliche Fähigkeit, bei der Wahrnehmung von Ereignissen, Inferenzen über die vorangegangenen oder über die nachfolgenden Ereignisse zu ziehen, können dynamische mentale Modelle auch aus statischen Bildern heraus aufgebaut werden. Dieses Phänomen ist auch in der Wahrnehmungspsychologie als „repräsentationaler Impuls“ bekannt (Freyd, 1983a). Die Darstellung von Bewegungen anhand statischer Bilder hat zudem den Vorteil, dass Details der Vorgänge genauer inspiziert werden können, wie das historische Beispiel der Fotos von Edward Muybridge zeigt. Bis zu seinen Fotografien, bei denen Bewegungen Schritt für Schritt präsentiert wurden, wurde der Galopp von Pferden von Malern falsch gezeichnet, da das exakte Bewegungsmuster der vier Beine zu komplex war, um in Echtzeit verstanden zu werden. Werden die Bilder bei einer Bewegungsdarstellung zudem simultan präsentiert, können die einzelnen Frames auch direkt miteinander verglichen werden, was einen weiteren Vorzug dieser Präsentationsform darstellt (Schnotz & Lowe, 2008). Tversky et al. (2008) haben in einer Durchsicht von Colledgelehrbüchern herausgefunden, dass Sequenzen von statischen Bildern u. a. für die Darstellung von Veränderungen über die Zeit hinweg bzw. für die Darstellung von Zyklen (z.B. der Lebenszyklus eines Schmetterlings), für die Präsentation von Partonomien (z.B. der Aufbau eines Baums oder eines Computers) oder für die Darstellung eines Objektes aus verschiedenen Perspektiven heraus (z.B. die Außen- oder die Innenansicht eines Gebäudes) verwendet werden.

Um Dynamik auszudrücken und um eine mentale Animation anzuleiten, brauchen statische Bilder häufig zudem „extra-pictorial features“ („zusätzliche Hinweise“), wie beispielsweise Pfeile, die relativ gut Veränderungen über die Zeit, Abfolgen, Kausalitäten und Bewegungsrichtungen vermitteln können, wobei ihre Effektivität und die von bildlichen Bewegungsindikatoren im Allgemeinen nicht absolut bestimmbar ist, da sie vom Verlauf der darzustellenden Bewegung jeweils abhängt (Tversky et al., 2008; Lewalter, 1997). Bei

Pfeilen wird die Richtung der Bewegung durch die Richtung des Pfeils und die Geschwindigkeit der Bewegung durch die Länge des Pfeils ausgedrückt. Eine ähnliche Methode ist das Zeichnen von Bewegungspunkten hinter einem Objekt, deren Platzierung die Bewegungsrichtung des Objektes und deren Abstände zueinander seine Geschwindigkeit angeben.

Weitere Möglichkeiten, Bewegungsabläufe mit stehenden Bildern darzustellen, wären beispielsweise, ein Objekt an mehreren Orten gleichzeitig zu präsentieren und die Bilder dabei zu numerieren bzw. frühere Bilder verblasst darzustellen oder eine besonders prägnante Bewegungsphase als Repräsentant des gesamten Vorganges auszuwählen. Man kann auch Bewegungslinien verwenden, die angeben, in welche Richtung ein Objekt wandert oder von woher es gekommen ist. Hohe Geschwindigkeiten innerhalb eines Prozesses können durch mehrere nebeneinander liegende Bewegungslinien oder durch symbolische Ergänzungen ausgedrückt werden, während Indikatoren für Geschwindigkeitsveränderungen nur schwierig zu finden sind (Lewalter, 1997).

Ein Spezialfall von dynamischer Darstellung durch statische Bilder ist ein „exploded diagram“, das dem Lerner helfen soll, beispielsweise das Zusammenbauen eines technischen Gerätes zu verstehen. „Exploded diagrams“ zeigen die Zerlegung eines Ganzen in seine Bestandteile und deren räumliche Beziehungen zueinander in einer Weise, dass dennoch der gesamte Gegenstand erkennbar bleibt. Allerdings können diese Bilder nicht die Abfolge und die Methode des Zusammenbaus deutlich machen, und sie liefern diesbezüglich auch keine zeitlichen Informationen (Tversky et al., 2008). Eine weitere Möglichkeit, temporäre Daten durch statische Bilder auszudrücken, ist die „temporal collage“, bei der verschiedene Ereignisse eines Prozesses in einem Frame plziert werden, wie es beispielsweise in Comicialogon häufig der Fall ist. Bei einer „temporal collage“ ist ein sequentielles Lesen des Lerners erforderlich (McCloud, 1994). In wissenschaftlichen Visualisierungen werden auch realistische Bilder mit einer Zeitachse verwendet, um Verläufe darzustellen. Auch logische Bilder können Veränderungen repräsentieren, indem die Veränderung eines visuellen Aspektes des Bildes die Veränderung eines abstrakteren Attributes des Repräsentanden anzeigt (Schnotz & Lowe, 2008).

Um einen Prozess anhand statischer Bilder zu verstehen, muss die Bildinformation korrekt dekodiert werden, d.h., aus einer zweidimensionalen Darstellung muss ein dreidimensionales Vorstellungsbild generiert werden, was eine richtige Interpretation der

Bewegungsindikatoren voraussetzt. Nur wenn ein Lerner hierzu in der Lage ist, können statische Bilder ausreichen, um ein Ereignis zu präsentieren, ansonsten muss dafür auf Animationen zurückgegriffen werden (Rieber & Kini, 1991). Zudem tritt bei statischen Bildern häufiger das Problem auf, dass es keine Anweisungen gibt, wie sie analysiert werden sollen. Der Lerner ist für die Art und die Reihenfolge der Dekomposition selbst verantwortlich, während eine Animation, beispielsweise bei Zerlegungsprozessen, die Aufmerksamkeit besser leiten kann. Die Grenzen der Präsentation von Dynamik durch statische Bilder liegen vor allem bei komplexen oder ungleichmäßigen Bewegungsabläufen oder bei simultanen Bewegungen mehrerer Objekte (Schnotz & Lowe, 2008; Lewalter, 1997).

2.1.2 Grundlegende Prozesse beim Bildverstehen

Die folgende Beschreibung des Bildverstehens anhand des Modells von Schnotz und Bannert (2003) verläuft größtenteils analog zu dem Buchkapitel von Schnotz (2005). Zitate aus anderen Quellen sind gesondert gekennzeichnet.

Schnotz und Bannert (2003) haben ein Modell formuliert, um Text- und Bildverstehen zu analysieren. Das Modell beinhaltet einen verbalen und einen bildlichen Kanal für die Verarbeitung und Speicherung von Informationen. Im verbalen Kanal wird externer Text zu einer mentalen Textoberflächenrepräsentation sowie zu einer propositionalen Repräsentation transformiert, wobei die Verarbeitung symbolisch abläuft. Der bildliche Kanal umfasst hingegen den Aufbau eines Images und eines mentalen Modells aus der Wahrnehmung eines externen Bildes heraus. Die Verarbeitung geht hier durch „structure mapping“ vonstatten. Das Modell entspricht dem „Dual-Coding-Konzept“ von Paivio (1986), der auch von einem verbalen und einem bildlichen System ausgeht, wobei im Modell von Schnotz und Bannert angenommen wird, dass mehrere Repräsentationen im Text- und Bildverstehen geformt werden.

Da in der vorliegenden Arbeit das Verstehen von Texten eine untergeordnete Rolle spielt, sollen in den folgenden Abschnitten vor allem die Aspekte des Bildverstehens innerhalb des Modells von Schnotz und Bannert thematisiert werden. Abbildung 2.1 zeigt zunächst das vollständige integrierte Modell zum Verstehen von Texten und Bildern von Schnotz und Bannert (2003).

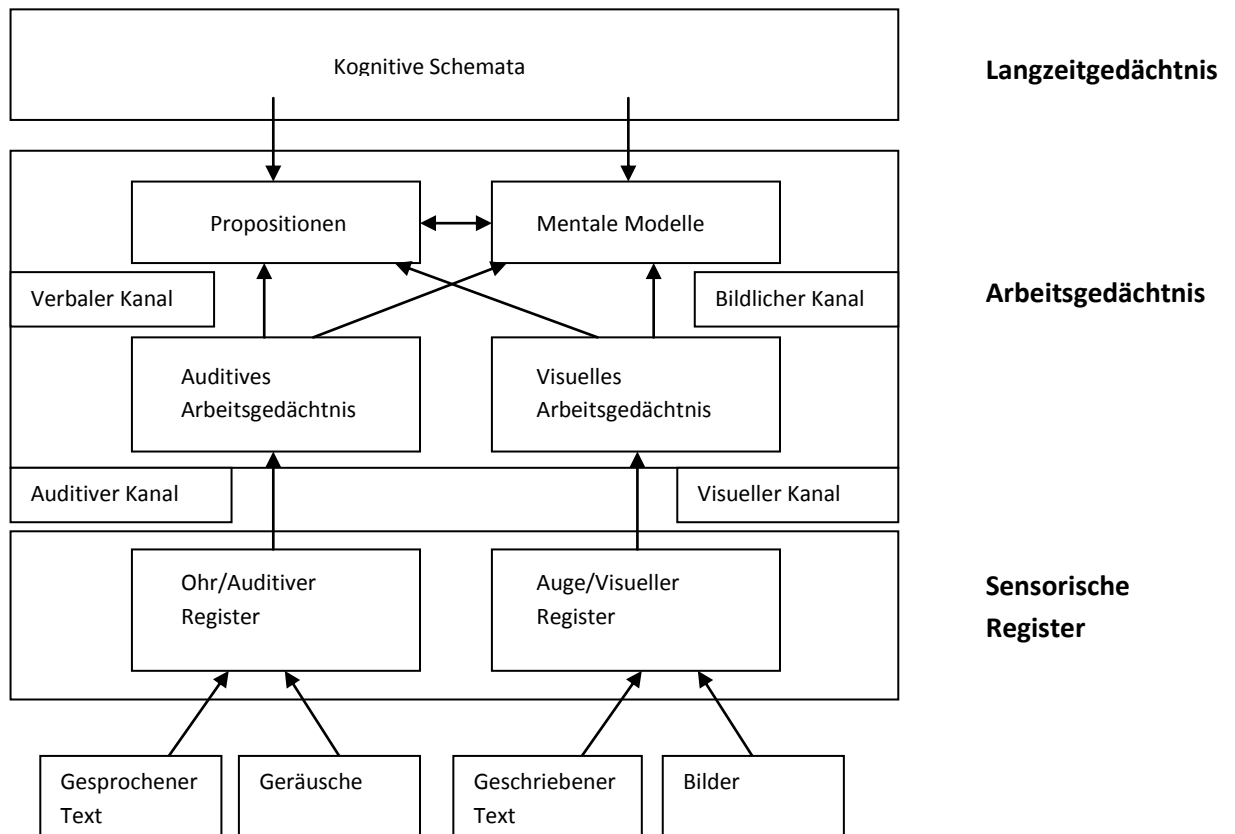


Abbildung 2.1: Das integrierte Modell des Text- und Bildverstehens (angelehnt an die Darstellung in Schnotz, 2005)

2.1.2.1 Mentale Repräsentationen beim Bildverstehen

Während des Bildverstehensvorganges, kreiert der Lerner zunächst eine perzeptuelle Repräsentation des Bildes, ein „visual image“, auf dem aufbauend dann ein mentales Modell des Bildinhaltes konstruiert wird. Das mentale Modell kann genutzt werden, um spezifische Informationen abzulesen, wobei die so erhaltene Information als Proposition enkodiert wird. Images und mentale Modelle sind depiktive Repräsentationen, von denen angenommen wird, dass sie eine inhärente Struktur haben, die der Struktur des jeweiligen dargestellten Gegenstandes entspricht (Johnson-Laird, 1983). Ein „visual image“ ist sinnesspezifisch, ein mentales Modell hingegen nicht, da es in der Lage ist, Informationen aus verschiedenen Sinnesmodalitäten miteinander zu integrieren. Ein mentales Modell ist abstrakter als ein Image und kann daher weniger als eine bildhafte Vorstellung bezeichnet werden, wobei es aber oft einen gewissen Imaginationsgehalt beinhaltet, weswegen der Aufbau eines mentalen

Modells auch „envisioning“ genannt wird (de Kleer, 1975). Irrelevante Details eines Bildes, die im Image noch enthalten sind, werden im mentalen Modell eventuell ignoriert. Jedoch beinhaltet das mentale Modell häufig auch zusätzliche Informationen aus dem Vorwissen des Lerners zu einem bestimmten Gegenstand sowie Informationen über das Original, die der Wahrnehmung nicht direkt zugänglich sind. Mentale Modelle sind hypothetische Quasi-Objekte, die dazu dienen, bestimmte Aufgaben und Probleme stellvertretend mental an ihnen statt am realen Objekt zu lösen. Sie stellen eine Basis dar, um über einen Lernstoff nachzudenken (Schnotz, 1994; Lowe, 1999a). Die Eigenschaften des mentalen Modells und die des repräsentierten Objektes sind funktional zueinander analog, so dass eine natürliche Isomorphie zwischen dem Modell und dem Original herrscht. Mentale Modelle werden somit auch als „intrinsische Repräsentationen“ bezeichnet, da sie lediglich strukturelle Übereinstimmungen mit dem repräsentierten System aufweisen. Die Struktur eines mentalen Modelles folgt hierbei dem Prinzip der „kognitiven Ökonomie“, was bedeutet, dass sie nicht über das hinausgeht, was für eine bestimmte Aufgabe gerade benötigt wird. Dadurch fallen konzeptuelle Defizite innerhalb des Modells nur dann auf, wenn geeignete Aufgaben gestellt werden, die die Konstruktion eines ausreichenden mentalen Modelles erfordern. Bei einem effektiven Lernvorgang muss der Lerner die gezeigte Information zum Aufbau eines solchen „high quality“-Modells nutzen können (Schnotz & Preuß, 1999; Lowe, 1999a). Mentale Modelle lassen sich in drei Unterarten unterscheiden: Strukturmodelle stellen den topologischen Aufbau eines Sachverhaltes dar; Attributmodelle repräsentieren zusätzlich verschiedene Zustände bzw. Attribute der einzelnen Systemkomponenten, und Kausalmodelle spiegeln auch die Kausalbeziehungen zwischen den Zuständen wider. Auch dynamische Sachverhalte werden in mentalen Modellen dargestellt. Mentale Modelle können Systeme aus unterschiedlichen Perspektiven und mit unterschiedlicher Granularität repräsentieren. Novizen konstruieren zunächst eher mentale Modelle, die die Oberflächeneigenschaften eines Gegenstandes beinhalten (Larkin, 1983), während in den Modellen von Experten strukturell und funktional wesentliche Eigenschaften eines Objektes vorhanden sind. Bei mentalen Modellen können die von ihnen repräsentierten Informationen direkt abgelesen werden. Sie eignen sich zudem gut, um Inferenzen auf neue Informationen zu ziehen. Hierfür werden die mentalen Modelle manipuliert, und die gesuchte Information wird durch das Ablesen der betreffenden Modellmerkmale bestimmt. Diese Informationen werden dann propositional abgespeichert, da mentale Modelle intern durch Propositionen beschrieben werden, so dass eine intensive Interaktion zwischen diesen beiden

Repräsentationsformen besteht. Der empirische Nachweis analoger mentaler Repräsentationen ist generell schwierig. Für ihre Existenz spricht beispielsweise der Befund von Kosslyn, Ball und Reiser (1978): In ihrer Studie führten Versuchspersonen im Geiste bestimmte Bewegungen, wie z.B. Drehungen, von Objekten durch. Die Zeit, die sie für die mentalen Bewegungen benötigten, war proportional zu der Zeit, die für die entsprechende äußere Bewegung gebraucht wird. Zudem wurde der „semantische Distanzeffekt“ durch Bryant und Trabasso (1971) empirisch belegt. Bei diesem Effekt wird davon ausgegangen, dass der mentale Vergleich von Objekten analog zu dem Vergleich von realen Objekten abläuft, so dass der Vergleich umso mehr Zeit benötigt, je ähnlicher die Objekte hinsichtlich des Vergleichsmerkmals zueinander sind. Das liegt daran, dass Personen für den Vergleich mentale Modelle konstruieren, bei denen die Objekte entsprechend ihrer Merkmalsausprägung auf einem Kontinuum unterschiedlich lokalisiert sind. Der Vergleich zwischen Objekten wird somit mit zunehmender Ähnlichkeit schwieriger, da sie dann schwerer auf diesem Kontinuum voneinander zu unterscheiden sind.

Die Beschreibung der mentalen Modelle in diesem Abschnitt erfolgte, soweit nicht anders gekennzeichnet, analog zu der Literatur von Schnotz (1994, 2005).

2.1.2.2 Aspekte des Gedächtnisses beim Bildverstehen

Der deskriptive und der depiktive Kanal im Modell von Schnotz und Bannert (2003) werden durch die begrenzte Kapazität des Arbeitsgedächtnisses beschränkt. Das visuelle Arbeitsgedächtnis wird bei Baddeley (1986) als ein „visuo-spatial sketchpad“ bezeichnet. Dieser „visuo-spatial sketchpad“ ist auf die Verarbeitung räumlicher Informationen spezialisiert, die visuell präsentiert werden, was auch die Verarbeitung von visuell dargestellten Bildern miteinschließt. Forschungsergebnisse zeigen zudem, dass es im Arbeitsgedächtnis ein eigenes Subsystem für den Aufbau mentaler Modelle gibt. Der Aufbau mentaler Modelle wird zwar durch die Kapazität des „visuo-spatial sketchpads“ beeinflusst und ist sehr mit der räumlich-kognitiven Verarbeitung verbunden (Sims & Hegarty, 1997), allerdings fanden Knauff und Johnson-Laird (2002) Hinweise dafür, dass visuelles Imaginieren und räumliches Schlussfolgern auf verschiedenen kognitiven Subsystemen basieren. Somit erscheint die Unterscheidung zwischen einem visuellen Arbeitsgedächtnis für visuelle Images und einem räumlichen Arbeitsgedächtnis für den Aufbau mentaler Modelle als plausibel.

Dem Arbeitsgedächtnis sind zwei sensorische Kanäle vorangestellt: der visuelle und der auditive Kanal. Der visuelle Kanal, der für das Bildverstehen relevant ist, transportiert Informationen vom Auge zum visuellen Arbeitsgedächtnis. Visuelle Information, die auf das Auge trifft, wird sehr kurz im „visual register“ gespeichert. Wenn die Aufmerksamkeit des Betrachters auf diese Information im „visual register“ gerichtet wird, wird sie über den visuellen Kanal ins visuelle Arbeitsgedächtnis zur weiteren Verarbeitung geleitet.

Beim Bildverstehen beeinflusst auch das Vorwissen des Lernalers aus seinem Langzeitgedächtnis, wie leicht er bildliche Informationen kategorisieren kann. Objekte werden dabei von ihm schneller und einfacher erkannt, wenn sie aus einer für sie typischen Perspektive präsentiert werden. Begriffliches domänenspezifisches Wissen ist für den Aufbau eines mentalen Modells beim Bildverstehen essentiell. Vorwissen kann einen Mangel an externer Information teilweise kompensieren, wobei Bilder intensiver analysiert werden, wenn ihr Inhalt schwierig und das Vorwissen des Lernalers niedrig ist (Carney & Levin, 2002).

2.1.2.3 Visuelles Bildverstehen nach dem Modell von Schnotz und Bannert (2003)

Text- und Bildverstehen finden in einer kognitiven Architektur statt, in der es ein Arbeitsgedächtnis mit begrenzter Kapazität, modalitätsspezifische sensorische Register und ein Langzeitgedächtnis gibt. Verbale und bildliche Informationen werden zum Arbeitsgedächtnis durch den visuellen und den auditiven Kanal transportiert. Die Kanäle haben eine begrenzte Kapazität, um Informationen zu verarbeiten und zu speichern. Die weitere Informationsverarbeitung im Arbeitsgedächtnis findet in zwei verschiedenen repräsentationalen Kanälen statt: Informationen von geschriebenem oder gesprochenem Text werden im verbalen Kanal verarbeitet, Informationen von Bildern oder Geräuschen werden hingegen im bildlichen Kanal verarbeitet. Auch diese beiden Kanäle verfügen nur über eine begrenzte Kapazität. Beim Verstehensprozess baut ein Lerner kohärente Wissensstrukturen aus der externen verbalen und bildlichen Information und aus seinem Vorwissen heraus auf. Das integrierte Modell unterscheidet somit sensorische Kanäle auf der perzeptuellen Ebene von repräsentationalen Kanälen auf der kognitiven Ebene. Am Ende des Verstehensprozesses von Texten und Bildern steht im integrierten Modell ein mentales Modell, das Informationen aus verschiedenen Quellen miteinander verbindet. Es wird davon

ausgegangen, dass ein mentales Modell länger und leichter im Gedächtnis gespeichert wird als eine Proposition und dass mentale Modelle enger mit Bildern als mit Texten verknüpft sind.

Beim Verstehen von Bildern im Speziellen kommt visuelle bildliche Information durch das menschliche Auge in den „visual register“. Die Wahrnehmung eines Bildes basiert hierbei auf präattentiven Prozessen, die automatisierte visuelle Routinen beinhalten. Diese Routinen verlaufen beispielsweise entsprechend den Gestaltgesetzen der Wahrnehmung und sind primär datengeleitet. Graphische Entitäten werden so identifiziert und voneinander diskriminiert (Schnotz, 2006; Schnotz, Böckheler & Grzondziel, 1999). Die wahrgenommene Information wird dann durch den visuellen Kanal zum visuellen Arbeitsgedächtnis geleitet, wo sie in eine visuelle perzeptuelle Repräsentation des Bildes umgewandelt wird. Ein visueller bildlicher Filter wählt im visuellen Arbeitsgedächtnis bildliche Informationen aus, die relevant sind, und leitet sie durch den piktorialen Kanal, wo sie zum Auf- oder Ausbau eines mentalen Modells führen. Das mentale Modell wird hierbei so konstruiert, dass visuelle Strukturen des wahrgenommenen Bildes mit semantisch bedeutungsvollen Strukturen im mentalen Modell korrespondieren. Die Konstruktion eines mentalen Modells aufgrund eines realistischen Bildes ist somit ein schemageleiteter Prozess des Struktur-Mappings von visuellen in semantische Relationen. Bei realistischen Bildern kann der Lerner hierfür auf kognitive Schemata der Alltagswahrnehmung und des generellen Weltwissens zurückgreifen (Schnotz, Böckheler & Grzondziel, 1999). Das mentale Modell kann dann genutzt werden, um neue Informationen abzulesen, die in das propositionale Arbeitsgedächtnis enkodiert werden. Dies führt zum zusätzlichen Aufbau von Propositionen. Beim Verstehen visueller Bilder wählt der Lerner somit relevante bildliche Informationen aus, organisiert sie, aktiviert entsprechendes Vorwissen und konstruiert ein kohärentes mentales Modell, das durch Propositionen ergänzt wird. Kehrt man diesen Vorgang um, wenn beispielsweise ein Graphikdesigner einen Gegenstand zeichnet, geht er zunächst von einem mentalen Modell dieses Gegenstandes aus, „mapt“ es in ein visuelles Image und transformiert dieses dann in ein externes Bild.

Beim Bildverstehen beeinflusst die Form der Visualisierung die Struktur des mentalen Modells. Die Effizienz eines mentalen Modells für eine bestimmte Aufgabe entspricht hierbei der Effizienz des dazugehörigen Bildes für die Aufgabe was auch als „Structure-Mapping Effect“ bezeichnet wird (Larkin & Simon, 1987). Schnotz und Bannert (2003)

merken zudem an, dass Bilder das Verstehen nur dann verbessern, wenn der Lerninhalt in aufgabenangemessener Weise visualisiert wird, ansonsten interferiert das Bild eher mit dem Aufbau eines adäquaten mentalen Modells. Gerade für Lerner mit höherem Vorwissen sind gut designte Bilder von Wichtigkeit, da unangemessene Visualisierungen bei ihnen zu Interferenzen beim Aufbau mentaler Modelle führen können.

Unklar beim integrierten Modell von Schnotz und Bannert (2003) bleibt noch die Antwort auf die Frage, ob es eventuell mehrere Ebenen von mentalen Modellen im bildlichen Kanal gibt, die von groben Übersichtsmodellen hin zu detaillierten Modellen mit hoher Granularität reichen können. Zudem wäre es noch interessant, zu klären, ob Propositionen direkt aus einer perzeptuellen Repräsentation eines Bildes heraus kreiert werden können, ohne dass der Zwischenschritt des mentalen Modells notwendig ist.

2.1.3 Wahrnehmungs- und Verstehensprozesse dynamischer Sachverhalte anhand statischer Bilder

2.1.3.1 Allgemeine Befunde

Mit statischen Bildern kann man, wie oben bereits angedeutet, indirekt auch dynamische Informationen darstellen. Beispielsweise wird die Fotografie einer Bewegungsphase mental nicht als statisch, sondern als Bewegungsablauf repräsentiert. Dies geschieht allerdings nur bei vertrauten Bewegungsabläufen, bei nicht vertrauten Prozessen fehlt zu einer solchen dynamischen Erweiterung die Vorerfahrung (Freyd, 1983b).

Reichert man statische Bilder mit Pfeilen an, können Veränderungen über die Zeit oder Kausalitäten besser vermittelt werden. Durch solche optischen Bewegungshinweise wird der Lerner dazu veranlasst, eine mentale Vorstellung des Bewegungsablaufs zu generieren, wobei er die relevanten Lerninhalte selbständig interpretieren und verarbeiten muss. Dabei besteht auch die Gefahr einer falschen Interpretation der Bewegungsindikatoren (Lewalter, 1997). Dass sich beispielsweise Pfeile gut für die Vermittlung von Dynamik eignen, zeigte sich in einer Studie von Heiser und Tversky (2002). Hier wurden statische Bilder einer Fahrradpumpe in einem Fall mit Pfeilen versehen, im anderen Fall jedoch nicht. Beinhalteten die Bilder keine Pfeile, beschrieben die Versuchspersonen ihren Inhalt nur strukturell, bei den Bildern mit Pfeilen hingegen machten die Probanden funktionelle Beschreibungen zu

der Fahrradpumpe. Ebenso wie Pfeile können auch sprachliche Ergänzungen die Aussagekraft von Bildern verbessern. Gerade Novizen profitieren von einer verbalen Anreicherung von Bildern, da auf diese Weise sequentielle Informationen gut vermittelt werden können und auch über nicht sichtbare Sachverhalte geredet werden kann (Hegarty & Kriz, 2008).

Statische Bilder können zur Darstellung von Dynamik auf zwei Arten präsentiert werden: seriell und parallel. Bei der seriellen Präsentationsform erscheinen die Bilder unabhängig voneinander, so dass ältere Informationen durch neue Informationen im Arbeitsgedächtnis überschrieben werden. Bei dieser Darstellungsart besteht zudem eine Interferenz mit der Wahrnehmung der Kontinuität von Bewegungen. Eine weitere kognitive Anforderung serieller Bilder besteht darin, dass der Lerner jedes Frame im Gedächtnis halten muss, während er das jeweils nachfolgende verarbeitet. Die Information kann über die einzelnen Schritte hinweg schlechter integriert werden. Somit kann eine sequentielle Präsentation von Bildern mit inferentieller Aktivität interferieren, da sie den „extraneous load“ erhöht (Ayres & Paas, 2007; Paas, van Gerven & Wouters, 2007). Als zweite Möglichkeit der Darstellung können die Bilder parallel angeordnet werden, d.h., die statischen Frames sind alle gleichzeitig auf dem Bildschirm sichtbar. Diese Präsentationsform kann den Aufbau einer präzisen mentalen Repräsentation unterstützen, da jeder Schritt immer zugänglich ist, so dass der Lerner in der Wahrnehmung des Prozesses Kontinuität aufrechterhalten kann. Es sind direkte Vergleiche zwischen den einzelnen Ereignisschritten möglich, wodurch der „germane load“ erhöht wird. Diese Vergleiche erlauben zudem, zwischen den breiteren Schritten Inferenzen auf die feineren Schritte zu ziehen, was sehr effektiv für die Informationsverarbeitung ist. Allerdings weisen sequentielle Bilder auch Vorteile auf, wie auch parallele Bilder Nachteile in sich bergen, da dort beispielsweise der Betrachter seine Lernsequenz selbst zusammenstellt, was auch zu Defiziten führen kann (Ayres & Paas, 2007; Boucheix & Schneider, 2009).

2.1.3.2 Vorteile beim Lernen dynamischer Sachverhalte anhand statischer Bilder

Personen denken über ein komplexes System als eine gerichtete Sequenz von Schritten vom Anfangs- bis zum Endzustand. Ereignisse in Form von statischen Bildern darzustellen, bietet sich daher aus kognitionspsychologischer Perspektive an, denn wenn über Ereignisse als eine

Sequenz diskreter Schritte gedacht wird, sollten sie auf diese Weise präsentiert werden. Lässt man mehrere Betrachter eine Animation segmentieren, liefert dies wertvolle Hinweise, was die wichtigen Schritte innerhalb eines Ereignisses sind. Wenn zudem eine interne Animation schrittweise abläuft und die Bewegung einer Komponente die nächste Bewegung bewirkt, könnte ein „mismatch“ zwischen diesem internen Vorgang und einer kontinuierlich ablaufenden Animation bestehen, so dass statische Bilder für die Darstellung von Dynamik günstig sein dürften (Tversky et al., 2008; Hegarty & Kriz, 2008).

Dass sich statische Bilder, entgegen alltäglicher Vermutungen, gut für die Präsentation von Dynamik eignen könnten, ist auch an manchen Gestaltungsrichtlinien für Animationen zu erkennen. Es wird häufiger empfohlen, die Handlung einer Animation in klare Schritte zu zerlegen und die Handlung jedes Schrittes aus der Perspektive des Betrachters zu zeigen. Dieses Vorgehen ähnelt der „Segmenting Method“, bei der Animationen in Segmenten gezeigt werden sollten, die wichtige Teile des Prozesses darstellen (Mayer & Chandler, 2001). Eine segmentierte Präsentation, die die Hauptschritte eines Prozesses zeigt, hilft vor allem Personen mit weniger räumlichem Vorstellungsvermögen, relevante Inferenzen zu ziehen. Hierbei sind parallele Bilddarstellungen oder teils kontrollierbare Animationen als Präsentationsformen gut geeignet. Bei einer „Teil-zu-Teil-Kontrolle“ einer Animation, die eher einen statisch diskreten Charakter hat, kann der Lerner erst einmal einen Teil des Prozesses verstehen, bevor er zum nächsten Teil übergeht, was weniger mentale Anstrengung verursacht. Ebenso besteht eine höhere Wahrscheinlichkeit, die einzelnen Teile zu verarbeiten und miteinander zu verbinden, wenn eine schrittweise Darstellung präferiert wird. Lerner können zuerst Komponentenmodelle der Teile bilden und sie in ein gesamtes kausales Modell während einer weiteren Gesamtpräsentation integrieren („Parts-First-Hypothesis“). Würde umgekehrt zunächst der ganze Prozess und dann erst seine Teile gezeigt werden, hätte der Lerner einen „overload“, und die Gesamtpräsentation würde keinen effektiven Kontext liefern, um die Teile zu integrieren (Mayer & Chandler, 2001). Werden hingegen zuerst statische Bilder gezeigt, kann der Lerner ein mentales Modell konstruieren, das er während des Betrachtens einer Animation nur noch revidieren muss. Die Bilder zeigen wichtige Aspekte der Animation vorab und lenken so die Aufmerksamkeit des Lerners auf relevante Inhalte. Animationen sollten weiterhin so ausgestattet sein, dass sie verlangsamt oder pausiert werden können, damit der Lerner einzelne Frames genauer studieren kann. Es können dabei im Vorherein solche Frames ausgewählt werden, die informativ sind und die die „breakpoints“ in der Sequenz darstellen (Hegarty & Kriz, 2008;

Tversky et al., 2008). An diesen Empfehlungen ist zu erkennen, dass die Verwendung von statischen Elementen in der Darstellung von Ereignissen hilfreich sein kann. Auch subjektiv scheint in einigen Fällen die Verwendung statischer Bilder für die Präsentation von Dynamik angebracht zu sein: In einer Studie von Lowe (2008) verlangsamten Probanden eine interaktive Animation zur Untersuchung ihres Lerninhalts oft derart stark, dass sie die einzelnen Bilder der Animation Frame für Frame hintereinander sahen. Wenn die Versuchspersonen somit die Wahl hatten, lernten sie lieber mit statischen Bildern der Animation als sie dynamisch laufen zu lassen. Die Probanden wollten in der Studie anscheinend den dynamischen Charakter der Darstellung minimieren. Diese Tendenz zum Lernen mit statischen Bildern hilft bei der visuellen Analyse, um die einzelnen Frames zu gliedern und um die unmittelbaren räumlichen Beziehungen der Teile zueinander zu erkennen. Dies geht allerdings zu Lasten der höheren Analyseebenen, auf denen man weitläufigere Muster sieht.

Statische Bilder haben bei der Darstellung von Dynamik weiterhin den Vorteil, dass die visuelle Aufmerksamkeit bei ihrer Betrachtung wandern kann und dass somit immer mehr bildliche Information verarbeitet wird, bis das ganze Display abgesehen ist (Kosslyn, 1994). Der Lerner kann wiederholt schauen, bis er alles verstanden hat. Letztendlich können statische Bilder Handlungen in wichtige Schritte einteilen und sie mit Pfeilen sowie Linien miteinander verbinden.

2.1.3.3 Anforderungen beim Lernen dynamischer Sachverhalte anhand statischer Bilder

Bei einer Animation muss ein Lerner die zeitlichen Veränderungen nur wahrnehmen, bei statischen Bildern muss er sie hingegen inferieren, um ein mentales Modell aufzubauen, wobei diese Inferenzen anstrengender sind als die bloße Wahrnehmung bei Animationen (Hegarty, Kriz & Cate, 2003). Die Inferenzen von Bewegungen bei statischen Bildern hängen auch vom Vorwissen oder den Orientierungen der lernenden Personen ab. Animationen sind für Personen mit geringerem Vorwissen besser geeignet, da alle Informationen gezeigt werden und nicht das Risiko von Fehlinferenzen besteht, wie es bei statischen Bildern der Fall ist. Die Konstruktion einer mentalen Repräsentation mit statischen Bildern ist für Novizen hingegen anstrengend und fehleranfällig. Da Personen mit geringerem Vorwissen nicht sichtbare Informationen aus Bildern nicht extrahieren können,

bietet sich für sie die Verwendung zusätzlicher sprachlicher Beschreibungen an (Narayanan & Hegarty, 1998; Lowe, 1999b).

Statische Bilder erfordern zudem die Durchführung mentaler Animationen, wenn sie Bewegungen darstellen. Narayanan und Hegarty (1998) beschreiben hierzu das Verstehen einer Maschine aus statischen Bildern und Text heraus folgendermaßen: Die Maschine wird vom Lerner mental in einfachere Komponenten zerlegt, es wird Vorwissen zu diesen Teilen herangezogen, und die Relationen zwischen den Teilen werden enkodiert, so dass ein statisches mentales Modell der Konfiguration der Maschine entsteht. Das dynamische mentale Modell wird durch mentale Animation gebildet. Das bedeutet, dass die Bewegungen der verschiedenen Teile durch räumliche Visualisierungsprozesse oder durch regelgeleitetes Schlussfolgern schrittweise inferiert werden. Eine mentale Animation läuft nach Hegarty (1992) „piecemeal“ ab: Das heißt, dass Personen beispielsweise ein mechanisches System in kleinere Einheiten zergliedern und die Bewegungen stückweise vom Anfang an in der Richtung der kausalen Kette der Abläufe innerhalb des Systems inferieren. Personen können kein mechanisches System auf einmal animieren, da dies wahrscheinlich einen „overload“ für ihr Arbeitsgedächtnis darstellt. Zudem können Personen nur Bewegungen aus benachbarten Teilen heraus inferieren. Die Produktionsregeln in einem Modell beinhalten nur Beziehungen zwischen räumlich zueinander benachbarten Teilen, da dies eine effiziente Methode ist, um mechanisches Wissen zu speichern. So kann die Arbeitsweise eines neuen Systems erschlossen werden, wenn man Wissen über seine lokalen Interaktionen hat. Würde man ein System hingegen vollständig mental animieren, bräuchte man verschiedene Wissensstrukturen für verschiedene Systeme. Lediglich bei Wissenschaftlern und Ingenieuren geht man davon aus, dass sie ein mechanisches System in seiner Gesamtheit animieren können. Bei einer mentalen Animation müssen mehr Inferenzen gemacht werden, wenn die Bewegung einer späteren Komponente in der kausalen Kette des Systems erschlossen werden soll. In Experimenten machten die Probanden bei solchen Schlussfolgerungen mehr Fehler. Die Inferenzen sind ebenfalls anstrengender, wenn sie gegen die Kausalitätsrichtung des Systems gemacht werden. Aus Fixationsdaten ließ sich zudem ableiten, dass wenn Personen die Bewegung einer Komponente inferieren sollten, sie auf diese Komponente und auf frühere Komponenten schauten. Wenn während der mentalen Animation das Display angeschaut werden kann, muss der Lerner keine statische Repräsentation des Systems speichern, so dass sein Arbeitsgedächtnis entlastet wird. Durch eine äußere Fixation des Prozesses werden außerdem Unterziele besser fokussiert, und der

Problemlösevorgang wird aufrechterhalten. Eigene Körper- und Augenbewegungen könnten weiterhin die mentale Animation unterstützen (Hegarty, 1992).

Empirisch zeigten sich Hinweise, dass eine mentale Animation eher ein Visualisierungsprozess ist, durch den das Bild der Komponenten mental transformiert wird, als dass sie dadurch abläuft, dass Produktionsregeln auf eine propositionale Repräsentation des Systems angewandt werden. Beispielsweise sollten die Probanden in zwei Studien von Schwartz und Black (1996a,b) die Bewegungen von Zahnrädern in einer Kette inferieren. Die Versuchspersonen zeigten die Bewegungen der Zahnräder mit den Händen und formulierten darauf aufbauend mechanische Prozeduren. Die Handbewegungen wurden als Externalisierung eines räumlich-bildlichen Transformationsprozesses interpretiert. Zudem war in den Studien die Zeit, um die Rotation von zwei Zahnrädern zu inferieren, proportional zu ihrem Rotationswinkel, was auch dafür spricht, dass die Inferenz der Zahnradbewegungen ein analog bildlicher Prozess ist. In einer von Sims und Hegarty (1997) durchgeführten „dual-task“-Studie machte eine räumliche Interferenzgruppe mehr Fehler als eine verbale Interferenzgruppe in sog. „motion verification tasks“, bei denen mentale Animationen geleistet werden mussten, um die Bewegung einer Komponente in einem mechanischen System zu erschließen. Ohne einen zusätzlichen Gedächtnisload unterschieden sich die beiden Gruppen in ihren Leistungen hingegen nicht voneinander. Umgekehrt nahm die Leistung in einem Punktgedächtnistest stärker ab als die in einem Buchstabentest, wenn die beiden Tests mit „motion verification tasks“ gepaart wurden, so dass anscheinend Aufgaben, bei denen mentale Animationen notwendig sind, die gleichen Ressourcen des „visuospatial sketchpads“ beanspruchen wie Punktgedächtnisaufgaben. Die Versuchspersonen machten zudem mehr Fehler in den „motion verification tasks“, wenn sie die Bewegungen von späteren Komponenten des Systems inferieren mussten, da bei einer mentalen Animation die Systemkomponenten eine nach der anderen vom Anfang der kausalen Ereigniskette an transformiert werden. Es zeigte sich in einer weiteren Studie, dass die Reaktionszeit, um die Bewegung eines Elementes in einem Flaschenzug vorherzusagen, monoton mit der Entfernung dieses Elementes zum Anfang der Kausalkette hin anstieg. Zusammengefasst konnte man aus diesen Studien schließen, dass mentale Animationen mehr mit dem visuellen als mit dem verbalen Gedächtnis und dass ein visueller Gedächtnisload mehr mit mentalen Animationen interferierte als ein verbaler Gedächtnisload, so dass mentale Animationen offensichtlich abhängig vom „visuospatial sketchpad“ sind. Eine mentale Animation

erfordert letztendlich von Seiten des Lerners ein gewisses Vorwissen und räumliche Fähigkeiten, um sie adäquat durchführen zu können (Hegarty, 2004).

Enthält eine Präsentation keine explizite Dynamik, wirft dies gerade für Novizen Schwierigkeiten auf, da sie nicht das Vorwissen haben, um aus einem statischen Bild ein mentales Modell aufzubauen. Bei alltäglichen Inhalten ist dies hingegen für sie weniger ein Problem. Bei statischen Bildern, wie auch bei Animationen, ist es häufig der Fall, dass Novizen sich auf die Regionen konzentrieren, die perzeptuell auffällig, konzeptuell aber nicht wichtig sind (Lowe, 1999a). Lowe (1993) fand in einer seiner Studien heraus, dass unerfahrene Meteorologiestudenten bei der Verarbeitung einer statischen Wetterkarte vorwiegend solche relationalen Informationen extrahierten, die auf perzeptuellen, allgemeinen Merkmalen, wie Nähe, Ähnlichkeit oder der Beteiligung der Elemente an einem größeren visuellen Muster, basierten. Die dynamische Information wurde von den Studenten ebenfalls zu einfach und zu verallgemeinert verstanden. In einer weiteren Studie von Lowe (1996) konstruierten Nicht-Meteorologen „low quality“-Modelle aus einer statischen Wetterkarte heraus, die ebenfalls die Dynamik zu sehr vereinfachten. Die Modelle basierten auf einer graphischen Charakterisierung des Displays, und die Probanden behandelten zeitliche Veränderungen derart, als ob die graphischen Entitäten physikalische Objekte wären.

Wenn letztendlich für die Bewältigung einer Aufgabe der dynamische Charakter eines Vorganges wichtig und ein Detailvergleich hingegen weniger essentiell ist, sind Animationen als Präsentationsmedium besser geeignet, da dynamische Informationen in statischen Bildern verzerrt gezeigt werden (Schnotz & Lowe, 2008).

2.2 Lernen mit Animationen

2.2.1 Allgemeine Merkmale von Animationen

Animationen haben einige allgemeine Merkmale gemeinsam mit statischen Bildern. Beispielsweise sind sie auch depiktive Repräsentationen und können ebenfalls kognitive, motivationale sowie dekorative Funktionen übernehmen. Allerdings haben sie auch einige Eigenschaften, die sie von stationären Bildern unterscheiden, die im Folgenden näher beschrieben werden sollen.

Animationen sind nach Schnotz und Lowe (2008) definiert als „bildliche Darstellungen, die ihre Struktur über die Zeit verändern und die die Wahrnehmung kontinuierlicher Veränderungen erleichtern.“ Ähnlich wird es bei Tversky et al. (2008) ausgedrückt, hier wird eine Animation als „ein sich veränderndes graphisches Display“ bezeichnet, das „die Wirkweise eines Systems von Anfang bis Ende zeigt“.

Animationen stellen eine direkte Art dar, zeitliche Änderungen eines Systems zu präsentieren und bereichern die zweidimensionale räumliche Darstellung eines Vorganges durch statische Bilder um die zeitliche Dimension. Allerdings können Animationen und statische Bilder nicht als qualitativ voneinander verschiedene Designs angesehen werden (Schnotz & Lowe, 2008). Nach dem „Congruence Principle“ (Tversky et al., 2002), bei dem Struktur und Inhalt einer externen Repräsentation zur Struktur und dem Inhalt der gewünschten mentalen Repräsentation passen sollten, sollten Animationen natürlicherweise Veränderungen darstellen, da sie realistischer und lebensnaher sind und räumliche sowie zeitliche Kontinuität in Echtzeit beinhalten. Animationen nutzen Veränderungen in der Zeit, um Veränderungen in der Zeit zu vermitteln. Zudem handelt es sich bei einer Animation um eine vollständige Visualisierung eines Bewegungsvorganges (Tversky et al., 2008; Lewalter, 1997).

Animierte Displays zeigen beispielsweise den Aufbau komplexer Objekte oder sie visualisieren Gegenstände aus verschiedenen Perspektiven heraus. Sie können Ansichten und Handlungen selektiv zeigen sowie die räumliche und zeitliche Skalierung von Ereignissen ändern. Weiterhin kann die Präsentation gezoomt werden, so dass Details genauer angeschaut werden können. In Animationen finden sowohl Translationen als auch Transformationen von Elementen statt. Bei Translationen verschieben sich die Positionen

von Entitäten, bei Transformationen hingegen ändern sich die Attribute des Dargestellten in ihrer Farbe oder Größe (Tversky et al., 2008; Lewalter, 1997; Schnotz & Lowe, 2008).

Animationen sollten generell möglichst realistisch sein, jedoch können auch bestimmte Aspekte betont, und die Darstellung kann schematisiert werden, wenn dies sinnvoll erscheint. Wie detailgetreu eine Animation sein sollte, hängt vom Instruktionsziel und der Zielgruppe ab. Wenn Lerner ein geringes Vorwissen haben und vor allem kausale Zusammenhänge erkannt werden sollen, kann eine weniger realistische Zeichnung hilfreich sein. Soll der Lerner hingegen mittels der Animation konkrete, interaktive Aufgaben in seiner Umwelt bewältigen können, sollte das Display realistischer gestaltet sein. Der Grad des „behavioral realism“ (Lowe, 2006) entspricht hierbei nicht immer zwangsläufig dem des „visuospatial realism“ der Darstellung. Manche Animationen haben einen eher diagrammartigen Charakter und weisen somit einen eher niedrigen „visuospatial realism“ auf, verfügen aber dennoch über einen höheren „behavioral realism“, wenn Veränderungen in der dargestellten Situation auch in der Animation stattfinden (Schnotz & Lowe, 2008). Wie realistisch das Verhalten in einer Animation wahrgenommen wird, hängt auch von ihrer zeitlichen Granularität ab. Eine hohe zeitliche Granularität führt zu einem höheren Realismus im Verhalten, allerdings kann das menschliche Auge ab einem gewissen Grad der Granularität die aufeinanderfolgenden Frames nicht mehr voneinander unterscheiden. Gewöhnlich liegt die Bildfrequenz von Animationen bei ungefähr 24 Frames pro Sekunde. Bei dieser Bildzahl erkennt der Betrachter eine kontinuierliche Bewegung, ohne dass die einzelnen Frames wahrgenommen werden, wie es bei einer Frequenz von unter 16 Bildern pro Sekunde der Fall ist. Weiterhin hängt eine realistische Darstellung eines Vorganges auch davon ab, inwiefern die zeitlichen Merkmale des Displays mit denen der Realität übereinstimmen (Rieber & Kini, 1991; Schnotz & Lowe, 2008).

Animationen können mit unterschiedlichen Geschwindigkeiten abgespielt werden. Ist die Geschwindigkeit hoch, werden die Makrostrukturen des Prozesses betont, ist die Geschwindigkeit hingegen niedrig, werden die Mikrostrukturen des Vorganges salient. Daher empfiehlt es sich häufig, Animationen mit verschiedenen Geschwindigkeiten abzuspielen (Lowe, 2006). Dynamische Darstellungen müssen keine realistischen Geschwindigkeiten aufweisen, wenn dies den Aufbau guter mentaler Modelle erleichtert. Eine Animation muss demzufolge kein „one-to-one mapping“ der externen Repräsentation sein, und sie muss auch nicht zwangsläufig die räumliche und zeitliche Kontinuität eines Vorganges einhalten (Schnotz & Lowe, 2008).

Animationen erfüllen in der Instruktion häufig drei Funktionen: Erregung von Aufmerksamkeit, Präsentation von Inhalten und Einübung von Gelerntem (Rieber, 1990). Das Hervorrufen von Aufmerksamkeit ist ein wichtiges Anfangsereignis in der Instruktion (Gagné, 1985) und markiert somit einen praktischen sowie rationalen Gebrauch einer Animation. Animierte Figuren weisen einen Kontrast zu ihrem statischen Hintergrund auf, wodurch ihre Prominenz verstärkt wird. Diese „kosmetische“ Anwendung von Animationen kann das Interesse eines Lerners erhöhen, ihn aber auch von den eigentlichen Inhalten ablenken. Werden Animationen verwendet, um dynamische Informationen zu präsentieren, wird dies als „Learning-by-Viewing-Approach“ beschrieben (Reed, 1985). Animationen finden weiterhin als Teil von Übungsstrategien Verwendung. In strukturierten „Frage und Antwort“-Übungen sind Animationen häufig Bestandteil des Feedbacks für die Antworten des Lerners. Der neuartige Charakter einer Animation verstärkt hierbei noch die Intensität des Lernens, wobei dieser Effekt meist nur vorübergehend ist (Clark, 1983).

Animationen sind besonders effektiv, um Prozesse, wie Wetterverläufe oder Kreislaufsysteme, zu präsentieren und um die zeitliche Komponente von Bewegungsbahnen darzustellen. Sie sind auch von Vorteil beim inzidentellen Lernen sowie für prozedurale Aufgaben, wie Videostudien gezeigt haben (Spangenberg, 1973). Eine weitere häufige Verwendung von Animationen sind beispielsweise Balken, die den Stand eines Computerdownloads anzeigen oder auch interaktive Programme, wie die Echtzeitsimulation, ein Flugzeug zu lenken, die durch den Input des Lerners manipuliert werden können („Learning-by-Doing“-Beispiel einer Animation, Brown, 1983). Im Allgemeinen haben Animationen dann eine positive Wirkung, wenn Veränderungen und Bewegungen wichtige Aspekte des Lerninhalts sind (Lewalter, 1997; Tversky et al., 2002). Befürworter von Animationen argumentieren weiterhin, dass sie attraktiv und motivierend wirken und dass sie deshalb gegenüber statischen Bildern bevorzugt werden sollten (Perez & White, 1985). Allerdings birgt dieser ästhetische Vorzug von dynamischen Darstellungen die Gefahr, dass Animationen eher mit der Absicht verwendet werden, um zu beeindrucken als um zu lehren. Jedoch sollten die Bedürfnisse des Lerners und die Anforderungen einer Aufgabe maßgeblich für die Entscheidung sein, ob eine Animation in der Instruktion notwendig ist oder nicht (Rieber, 1990).

2.2.2 Grundlegende Prozesse beim Verstehen von Animationen

Laut Schnotz und Lowe (2008) werden animierte und statische Bilder vom gleichen perzeptuellen und kognitiven System verarbeitet, weshalb man sie hinsichtlich ihrer Wahrnehmung und ihres Verstehens als nicht so unterschiedlich ansehen sollte. Weiterhin wird bei Lowe (1999a) angemerkt, dass statische und animierte Bilder auf der visuell-räumlichen Ebene gleiche Verarbeitungsanforderungen an den Lerner stellen.

Da somit Animationen und statische Bilder prinzipiell die gleichen Verarbeitungswege im kognitiven System nutzen, soll für die Darstellung des Verstehens von Animationen nochmals das integrierte Modell des Text- und Bildverstehens von Schnotz und Bannert (2003) herangezogen werden. Da zwischen der Verarbeitung von Animationen und der von statischen Bildern große Analogien bestehen, wird der Verarbeitungsvorgang bei Animationen kurz, analog zu dem Buchkapitel von Schnotz (2005) (soweit nicht anders gekennzeichnet), skizziert, und auf Besonderheiten im Vergleich zu dem Verarbeitungsweg von statischen Bildern wird explizit hingewiesen. Für ausführlichere Beschreibungen des Modells von Schnotz und Bannert sowie der in ihm enthaltenen mentalen Repräsentationen und Gedächtnisprozesse sei auf Kapitel 2.1.2 zum grundlegenden Verstehen von statischen Bildern verwiesen.

Animationen erreichen als visuelle bildliche Informationen das menschliche Auge und gelangen von dort aus kurz in den „visual register“. Wenn die Aufmerksamkeit des Betrachters auf Informationen im „visual register“ gerichtet wird, gelangen diese durch den visuellen Kanal ins visuelle Arbeitsgedächtnis. Da eine Animation auch zeitliche Merkmale zeigt, wird bei ihrer Betrachtung vom Lerner mehr Verarbeitungskapazität im Vergleich zu statischen Bildern verlangt (Lowe, 1999a), zumal die Kapazität des visuellen Kanals begrenzt ist. Im visuellen Arbeitsgedächtnis wird aus der ankommenden Information ein visuelles Image aufgebaut. Ein visueller bildlicher Filter wählt aus dieser noch perzeptuellen Repräsentation der Animation relevante Informationen aus, die im bildlichen Kanal weitergeleitet werden, wo sie zum Auf- oder Ausbau eines mentalen Modells führen. Da Animationen dynamische Systeme präsentieren, muss ein Lerner entsprechende kausale, dynamische mentale Modelle konstruieren, mit denen er im Gedächtnis Verhaltensweisen simulieren kann (Schnotz, Böckheler & Grzondziel, 1999). Der Aufbau eines mentalen Modells findet durch den Prozess des „structure mappings“ statt, d.h., dass visuelle Strukturen des wahrgenommenen Bildes mit semantisch bedeutungsvollen Strukturen im

mentalen Modell korrespondieren. Bei animierten Bildern werden neben räumlichen auch zeitliche Strukturen des Prozesses dargestellt, so dass diese auf die Dynamik des mentalen Modells gemappt werden können. Dies liefert eine Unterstützung für eine spätere mentale Simulation des Vorganges. Die kognitive Verarbeitung ist hierbei auch abhängig von „top down“-Prozessen, die durch kognitive Schemata geleitet werden, in denen das Vorwissen des Lerners enthalten ist (Schnotz, Böckheler & Grzondziel, 1999; Kintsch, 1998). Komplexe Aktivitäten, die durch Animationen dargestellt werden können, werden mit Hilfe eines kognitiven Schemas verstanden, das aus einem Set von hierarchischen Unterschemata besteht.

Zusammengefasst kann gesagt werden, dass beim Verstehen eines Prozesses ein dynamisches mentales Modell konstruiert wird, wobei dieser Konstruktionsvorgang ein schemageleiteter Prozess ist, in dem zeitlich-räumliche Muster des Gesehenen auf die Muster des mentalen Modells gemappt werden (Schnotz & Lowe, 2008). Ebenso wie bei statischen Bildern, kann das dynamische mentale Modell, das durch eine Animation aufgebaut wurde, dazu genutzt werden, um neue Informationen abzulesen, die in Form von Propositionen enkodiert werden. Somit wählt der Lerner auch beim Verstehen von Animationen relevante bildliche Informationen aus, organisiert sie, aktiviert entsprechendes Vorwissen und konstruiert ein kohärentes mentales Modell, das durch Propositionen ergänzt wird.

2.2.3 Wahrnehmungs- und Verstehensprozesse dynamischer Sachverhalte anhand von Animationen

2.2.3.1 Allgemeine Befunde

Generell können durch Animationen sowohl die Makro-, als auch die Mikroereignisse eines Prozesses explizit gezeigt werden, wobei die Mikroereignisse in den Makroereignissen eingebettet sind. Makroereignisse sind beispielsweise Translationen, bei denen ein Objekt seine räumliche Position ändert. Mikroereignisse können dabei wiederholte Transformationen der internen räumlichen Relationen des Objektes oder Veränderungen seiner visuellen Attribute, wie Größe oder Farbe, sein. Eine Animation kann die Veränderungen innerhalb eines Prozesses auf verschiedenen Ebenen darstellen, wobei je nach Vermittlungsabsicht die Mikro- oder die Makrostrukturen eines Vorganges wichtiger sind (Schnotz & Lowe, 2008).

Allgemein ist beim Verständnis von Animationen zu sagen, dass sie kausale Relationen innerhalb eines Ereignisses durch die gleichzeitige Kovariation von Attributen repräsentieren. Die Veränderung eines Attributes bewirkt eine systematische Veränderung eines anderen Attributes, wodurch vom Lerner ein kausales mentales Modell aufgebaut wird. In vielen erklärenden Animationen stellen somit zeitliche Veränderungen Ursache-Wirkungs-Verhältnisse zwischen Entitäten dar. Man muss dabei unter den Veränderungen die Beziehungen feststellen, die eine Erklärung für das jeweilige Verhalten in einem Prozess liefern. Verfügt ein Lerner über kein Vorwissen, interpretiert er eine Animation jedoch oft deskriptiv oder er zieht inadäquates Alltagswissen heran, um ein System zu erklären (Mayer, 1997; Lowe, 1999a).

In gestaltungstechnischer Hinsicht können Animationen verschiedene Blickwinkel und Skalierungen beinhalten. Sie können Sprache und „extra-pictorial features“ für die Darstellung von Verbindungen nutzen. Zudem können durch Animationen Informationen verzerrt oder auch hervorgehoben werden (Tversky et al., 2008).

Es bleibt anzumerken, dass Animationen eine sensible Domäne sind, um Forschung zu betreiben, da ihre Effektivität stark durch Interaktivität, Expertise, räumliche Fähigkeiten und Aufgabenarten beeinflusst wird (Ayres & Paas, 2007).

2.2.3.2 Vorteile beim Lernen dynamischer Sachverhalte anhand von Animationen

Animationen zeigen neben räumlicher auch zeitliche Information und sind in der Darstellung von Ereignissen authentischer und natürlicher als statische Bilder. Animationen sind weiterhin effektiver als statische Bilder, um Prozesse wie biologische Regelkreise zu zeigen (Bétrancourt & Realini, 2005). Zudem haben Animationen bei der Darstellung von einfachen physikalischen Systemen mit wenigen Entitäten und Beziehungen positive Effekte auf die Instruktion, während das Betrachten einer Animation bei komplexeren Systemen wahrscheinlich schwieriger ist. (Mayer & Anderson, 1992). Durch Animationen werden Veränderungen oder Bewegungen vollständig und weniger abstrakt präsentiert (Lewalter, 1997). Wenn für eine Aufgabe der dynamische Charakter eines Vorganges wichtig und Detailvergleiche hingegen weniger essentiell sind, sind Animationen besser zum

Lernen geeignet als statische Bilder, da dynamische Informationen in statischen Bildern verzerrt gezeigt werden.

Bei Animationen wird im Vergleich zu statischen Bildern zudem die Notwendigkeit reduziert, Inferenzen zu ziehen. Dadurch entlastet eine Animation das Kurzzeitgedächtnis des Menschen, da es Bewegungen bei dieser Darstellungsform weniger selbst generieren muss. Der Verlauf einer Bewegung kann direkt ins visuelle Gedächtnis aufgenommen werden, und der Transfer ins Langzeitgedächtnis wird vereinfacht (Rieber & Kini, 1991).

Animationen können die Aufmerksamkeit des Lernalers besser leiten. Dies kann durch Perspektivenwechsel und Bewegung von Bildelementen geschehen. Animationen können zeigen, wie Gegenstände aufgebaut werden, so dass deren Struktur jeweils klar wird und dem Lerner eine direkte Anleitung für die Konstruktion eines mentalen Modells gegeben wird. Auch umgekehrt können die Dekompositionen von Objekten für analytische Zwecke animiert werden (Lewalter, 1997; Schnotz & Lowe, 2008).

Emotional gesehen, denken Mediendesigner häufig, dass Animationen motivierend und unterhaltend sind. Auch bei Lowe (2001) wird angemerkt, dass animierte Bilder einen positiveren Effekt als statische Bilder haben können, weil Animationen positive Auswirkungen auf die Motivation des Lernalers zeigen. Sie können aber auch negative Effekte aufs Lernen haben, wenn vom Betrachter inadäquate Verarbeitungsstrategien eingesetzt werden (Schnotz & Lowe, 2008).

Befürwortend ist bei Animationen weiterhin anzumerken, dass sie auf das Lernen sowohl einen „Enabling Effect“ als auch einen „Facilitating Effect“ haben können (Schnotz & Lowe, 2008). Bei dem „Enabling Effect“ wird die mentale Belastung beim Lernen derart reduziert, dass die Lösung einer Aufgabe möglich wird. Beispielsweise kann eine Animation einen Sachverhalt aus verschiedenen Perspektiven heraus zeigen, so dass ein Lerner mehr kognitive Verarbeitung leisten kann. Bei dem „Facilitating Effect“ wird die mentale Belastung ebenfalls reduziert, so dass eine Aufgabe einfacher wird. Animationen weisen zum Beispiel dann einen solchen erleichternden Effekt auf, wenn sie eine externe Unterstützung für mentale Simulationen sind und wenn sie somit eine „Supplantationsfunktion“ ausüben (Salomon, 1979). Bei dieser Funktion übernehmen die Animationen für den Lerner die Aufgabe, Vorstellungsbilder der räumlichen Bewegungsabläufe zu generieren und zu transformieren. Auf diese Weise wird der Lerner

bei der mentalen Simulation des Vorganges unterstützt. Dadurch, dass Animationen die Grenzen von Transformationen innerhalb des Prozesses zeigen, wird weniger „cognitive load“ („mentale Belastung“, Sweller, 1988) beim Betrachter erzeugt (Schnotz, Böckheler & Grzondziel, 1999). Gerade bei Kindern ist diese unterstützende Eigenschaft von Animationen von Bedeutung, da sich die Fähigkeit, räumliche Vorstellungsbilder selbständig zu generieren, erst mit dem Alter entwickelt. Aus demselben Grund können Animationen für Personen mit geringeren visuellen Vorstellungsfähigkeiten eine instruktionale Hilfestellung sein. Manchmal ist jedoch eine externe Unterstützung nicht notwendig für eine mentale Animation, und gerade Personen mit mehr Vorwissen werden durch eine Animation eher entmutigt, die mentale Simulation selbst auszuführen, wodurch sie sogar Nachteile in Simulationsaufgaben haben können (Lewalter, 1997; Schnotz, Böckheler & Grzondziel, 1999).

2.2.3.3 Anforderungen beim Lernen dynamischer Sachverhalte anhand von Animationen

Bétrancourt, Morrison und Tversky (2001) formulieren drei Quellen für kognitive Schwierigkeiten bei Animationen. Zum einen kann es bei dynamischen Darstellungen Aufmerksamkeitsschwierigkeiten geben, da zunächst vom Betrachter stets erkannt werden muss, welche Elemente sich bewegen und was ihre Bewegung jeweils ausgelöst hat. Wenn eine Animation nicht angehalten oder kontrolliert werden kann, muss der Lerner sie daher sehr aufmerksam betrachten und diesbezüglich Ressourcen investieren (Lewalter, 1997). Zum anderen kann es Schwierigkeiten durch die mentalen Operationen geben, die man bei einer Animation anwenden muss. Desweiteren sind Gedächtnisschwierigkeiten zu nennen, da ein Lerner, um die kausale Kette eines Systems zu verstehen, sich an die Positionen jedes Elementes zu verschiedenen Phasen der Animation erinnern können muss.

Ein weiterer Grund, weshalb das Lernen mit Animationen anstrengender als das mit statischen Bildern ist, liegt darin, dass mehr Informationen verarbeitet werden müssen, wodurch ein höherer Cognitive Load entsteht (Sweller, van Merriënboer & Paas, 1998). Wenn das zu lernende Thema komplex ist, ist es für den Lerner schwierig, bei einer Animation bestimmten Elementen zu folgen, weil eine hohe Informationsmenge in einer Zeiteinheit präsentiert wird. Vor allem für Novizen ist es sehr anstrengend, bestimmte Teile in einem sich kontinuierlich ändernden Display zu verfolgen, da sie keine Ordnung in ein

vielschichtiges Bild bringen können (Lowe, 2001). An Animationen wird daher häufiger kritisiert, dass sie nicht das „Apprehension Principle“ guter Bilder erfüllen, das besagt, dass eine externe Repräsentation sofort und akkurat wahrnehmbar sein sollte (Tversky et al., 2002).

Animationen erzeugen kognitive Kosten, weil sie flüchtig sind (Lowe, 1999a). Das bedeutet, dass die Informationen auf dem Display schon verschwunden sind, bevor sie ins Langzeitgedächtnis des Lerners aufgenommen werden können. Versucht der Lerner, die Inhalte im Arbeitsgedächtnis zu halten, ist dies schwierig, da ständig neue Informationen hinzukommen (Paas et al., 2007). Durch die begrenzte Darbietungszeit ist somit eine zügige Informationsverarbeitung notwendig. Zudem kann sich der Lerner durch die Flüchtigkeit der Animation nicht so intensiv mit den Inhalten beschäftigen. Bei Animationen laufen häufig zu viele Vorgänge zu schnell ab, was ihre Wahrnehmung erschwert, während statische Bilder genauer inspiziert und reinspiziert werden können (Lewalter, 1997; Tversky et al., 2008). Bei Animationen ist eine differenzierte Analyse der visuellen Strukturen schwieriger. Durch ihre inhärente Flüchtigkeit muss die visuelle Aufmerksamkeit des Betrachters bei Animationen zur richtigen Zeit am richtigen Ort sein, da aufgrund der „Immediacy Assumption“ und der „Eye-Mind Assumption“ (Schnotz & Lowe, 2008) angenommen wird, dass perzeptuelle Informationen sofort verarbeitet werden und dass Informationen, die vom Auge erfasst werden, direkt mit den Informationen verbunden werden, die in der kognitiven Verarbeitung sind.

Wie bei statischen Bildern gibt es bei Animationen auch das Problem, dass der Lerner eventuell nicht auf die konzeptuell wichtigen Regionen schaut, vor allem dann, wenn diese perzeptuell nicht salient sind. Animationen, die sehr realistisch und mit vielen Details gestaltet sind, bergen gerade für Lerner mit weniger Vorwissen perzeptuell die Schwierigkeit, dass sie oft nicht die relevanten Informationen extrahieren können. Sie führen weniger eine „top down“-, sondern mehr eine „bottom up“-Verarbeitung durch, da ihre Aufmerksamkeit mehr von perzeptuell salienten als von thematisch relevanten Aspekten gelenkt wird. Bei Experten läuft die Verarbeitung hingegen strategischer ab, und sie werden nicht so stark von perzeptuell Auffälligem abgelenkt (Schnotz & Lowe, 2008). Diese Schwäche von Animationen für Novizen kann man durch „highlighting-techniques“, wie Zoomen oder Übertreiben, beheben, die zu einer Manipulation des Displays und somit zu einer Veränderung der „bottom-up“-Effekte führen. Wenn die perzeptuellen Eigenschaften einer Animation auf relevante Informationen verweisen und wenn die perzeptuelle

Verarbeitung des Lerners eine räumlich-zeitliche Struktur generiert, die man für ein mentales Modell benötigt, wird dessen Aufbau erleichtert.

Während der Betrachtung einer Animation werden aus ihr vom Lerner Aufmerksamkeitszentren ausgeschnitten und mental zusammengefügt, so dass ein „attentional core“ entsteht, der das Rohmaterial für die weitere kognitive Verarbeitung liefert. Diese hohe Selektivität beim Betrachten einer Animation kann durch eine Minderung der Animationsgeschwindigkeit reduziert werden, so dass eine größere Informationsmenge für die weitere Verarbeitung und den Aufbau eines mentalen Modells zur Verfügung steht. Bei komplexeren Animationen verändern sich zudem mehrere visuelle Aspekte oft gleichzeitig, so dass ein „intra-representational split of visual attention“ entstehen kann, da der Lerner sich auf bestimmte Teile konzentriert, so dass er andere Aspekte nicht beachten kann. Daher sind externe Anleitungen und interne Strategien des Lerners zur Lenkung seiner Aufmerksamkeit wichtig für das Lernen mit Animationen. Eine Möglichkeit, die Aufmerksamkeit zu leiten, besteht in verbalen Anreicherungen der Animation, wobei hier gesprochener Text („narration“) besser geeignet ist als geschriebener Text, da dann zwischen Animation und Text temporale Kontiguität und kein „interrepresentational split of visual attention“ vorhanden ist (Schnotz & Lowe, 2008). Die Anreicherung einer Animation mit einer Erzählung erhöht ihren Effekt, da der Lerner sich dann auf die wichtigen Teile und Veränderungen in der Animation konzentriert und da durch den Text eine Erklärung für die Vorgänge gegeben wird. Umgekehrt lernen Personen besser von Animationen mit Erzählungen als von einer Erzählung allein, was bei Mayer (2008) auch als „Multimedia-effect“ bezeichnet wird. Da die visuelle Wahrnehmung präattentiv abläuft und es für einen Lerner daher schwierig ist, ein auffallendes Objekt zu ignorieren, ist es auch ein aufmerksamkeitslenkendes Mittel, wichtige Aspekte der Animation perzeptuell auffällig, mit Hilfe von Pfeilen oder Kreisen, zu gestalten. Sich bewegende Pfeile sind hierbei für das Problemlösen besser geeignet als statische Pfeile. Auch durch dynamische Kontraste kann die Aufmerksamkeit des Betrachters geleitet werden (Tversky et al., 2008; Schnotz & Lowe, 2008).

Ein weiteres Problem beim Lernen mit Animationen ist der Umstand, dass ein Vergleich zwischen Zuständen schwierig ist, da ein voriger Zustand, mit dem ein jetziger Zustand verglichen werden soll, im Arbeitsgedächtnis behalten oder aus dem Langzeitgedächtnis geholt werden muss (Lowe, 1999a). Durch diesen Umstand kann eine Animation das Lernen

dann schwieriger machen, wenn genaue Vergleiche zwischen verschiedenen Zuständen wichtig für das Verständnis sind. Durch die Verwendung einer Animation wird eine solche vergleichende Verarbeitung erschwert (Lowe, 2001).

Weiterhin muss bei statischen Bildern über zeitliche und räumliche Einschränkungen innerhalb des Prozesses reflektiert werden, um ein mentales Modell aufzubauen, während der Lerner dies bei einer Animation an das Medium delegieren kann, wodurch die Animation konzeptuell weniger genau analysiert wird. Die Kontrolle des eigenen Verständnisses und des Annäherungsgrades an ein Lernziel ist daher vor allem bei Animationen wichtig, da scheinbar leichtverständliche Informationen präsentiert werden (Schnotz, Böckheler & Grzondziel, 1999; Lewalter, 1997).

2.2.3.4 Gestaltungsempfehlungen zur Verbesserung des Designs von Animationen

Um die Anforderungen, die Animationen an den kognitiven Apparat des Menschen stellen, zu mindern, schlagen Paas et al. (2007) folgende Methoden zur Animationsgestaltung vor:

- „Segmenting Method“: Eine Animation sollte in Segmenten gezeigt werden, die die wichtigen Teile des Prozesses darstellen. Die Segmentstrategie scheint ziemlich robust für individuelle Eigenheiten im Design zu sein (Mayer & Chandler, 2001; Ayres & Paas, 2007).
- „Cueing Method“: Man kann perzeptuelle Hinweise geben, die die Aufmerksamkeit des Lerners auf relevante Aspekte der Animation lenken (Bétrancourt, 2005). Beim Cueing werden Kernaspekte eines Vorganges betont und andere Aspekte verdunkelt, so dass dem Lerner unnötige Suchprozesse erspart werden, was seine mentale Belastung reduziert. Bei dieser Methode zeigen Lerner auch bessere Leistungen in dem nicht betonten Inhalt, da durch das Cueing kognitive Ressourcen freigesetzt werden, die zum intensiven Betrachten des verdunkelten Materials genutzt werden können. Bei komplexen Animationen brauchen Lerner häufig nur eine kleine Hilfe in Form von Cueing, um eine Hürde zu überschreiten und das Verständnis zu verbessern (Ayres & Paas, 2007).

- „Tracing Method“: Man kann innerhalb der Animation Spuren hinterlassen, in denen Schlüsselinformationen noch vorhanden sind (Marcus, KhengJoo, Beng - Fei & Ayres, 2006).

Zudem ist Interaktivität ein Mittel, um den „germane load“ (Sweller, 2005) bei einer Animation zu erhöhen, da so bedeutsame Lernaktivitäten erzeugt werden können. Bei interaktiven Animationen manipuliert ein Lerner eine Animation explorativ und sieht den Lernstoff so aus verschiedenen Perspektiven. Dieser Lernvorgang ist sehr informativ und aktiv. Im Normalfall ist bei einer Animation die Lernerkontrolle geringer ausgeprägt, und die Sequenz sowie die Geschwindigkeit des Ablaufs sind vorgegeben („Technology-Centered Approach“ vs. „Learner-Centered Approach“, Mayer, 2001). Bei benutzerkontrollierten Animationen hat der Lerner jedoch mehr Freiheiten, wodurch diese Darstellungsform weniger herausfordernd wird (Narayanan & Hegarty, 1998). Bei der Betrachtung von Animationen muss der Lerner selektiv Informationen zur Verarbeitung auswählen, da er nur eine begrenzte Verarbeitungskapazität hat (Lowe, 2001). In manchen Phasen einer Animation passt die gezeigte Informationsmenge in die Verarbeitungskapazität des Lerners, in anderen Phasen ist dies hingegen nicht der Fall. Ist die Animation mit Benutzerkontrolle ausgestattet, kann ein Lerner selbst bestimmen, wie intensiv er die Animation anschaut. Er kann weniger wichtige Phasen überspringen, und er hat die Möglichkeit, wichtige Phasen einer Animation mehrmals anzusehen, so dass er einen Verarbeitungsvorteil wie bei statischen Bildern hat. Auch eine interaktive Manipulation der Animationsgeschwindigkeit bietet sich an, um den Lerneffekt zu erhöhen. Bei höheren Geschwindigkeiten werden die Makroinformationen, bei geringeren Geschwindigkeiten werden die Mikroinformationen eines Vorganges salienter (Lowe, 2006). Spielt man eine Animation mehrmals mit verschiedenen Geschwindigkeiten ab, entstehen beim Lerner unterschiedliche perzeptuelle Profile des Vorganges, die in einem mehrschichtigen mentalen Modell münden.

Damit die Benutzerkontrolle bei einer Animation jedoch effektiv ist, muss ein Lerner sie strategisch nutzen. Beispielsweise muss er bei einer hohen Informationsdichte der Animation ihre Geschwindigkeit verlangsamen und sollte sie umgekehrt bei einer geringeren Informationsdichte schneller laufen lassen, was er aber nur dann sinnvoll tun kann, wenn er die Bedeutungen der einzelnen Phasen kennt und voneinander unterscheiden kann. Zudem stellen interaktive Animationen auch höhere kognitive Anforderungen an einen Lerner, da ein Teil seines Arbeitsgedächtnisses dazu benutzt wird, um das Medium zu bedienen

(Schnotz, Böckheler & Grzondziel, 1999). Beim Lernen mit interaktiven Animationen kann das Entscheiden über Untersuchungsstrategien und das Überwachen der Lernerfolge gerade für Novizen eine hohe Anforderung sein. Da Novizen wahrscheinlich keine geeigneten Strategien für die Nutzerkontrolle haben, sollte man sie bei deren Verwendung unterstützen und die Interaktivität an das Vorwissen und die Verarbeitungskapazität des Lerners jeweils anpassen. Man kann hierbei die Interaktivitätsmöglichkeiten sinnvoll begrenzen oder Anleitungen geben, wie man die Nutzerkontrolle gewinnbringend bedient. Generell empfehlen sich Lernhilfen, die den Lerner dabei unterstützen, mit der Animation adäquat zu interagieren und die Informationen tiefer zu verarbeiten (Schnotz & Lowe, 2008).

Weiterhin gibt Rieber (1990) folgende Designempfehlungen für animierte Bilder:

Empfehlung 1: Animationen sollten nur verwendet werden, wenn ihre Attribute zur Lernaufgabe kongruent sind. Animationen bringen in ein instruktionales Setting die Elemente der Visualisierung, der Bewegung und der Bewegungsrichtung mit ein. Die Effektivität einer Animation hängt von dem Bedürfnis des Lerners ab, ob er eines dieser Attribute für die erfolgreiche Bewältigung einer Aufgabe braucht.

Empfehlung 2: Wenn Lerner Novizen sind, wissen sie oft nicht, welche Hinweise oder Details der Animation sie beachten müssen. Daher empfiehlt es sich, relevante Inhalte perceptuell zu betonen. Beispielsweise zeigten sich in einer Studie von Baek und Layne (1988) signifikante Effekte bei einer Animation, wenn die Materialien sorgfältig designt wurden, um die Aufmerksamkeit der Lerner auf die sachdienlichen Details zu richten.

Empfehlung 3: Der größte Beitrag von Animationen dürfte in ihren interaktiven Anwendungen liegen, die oft als Teil von Übungen eingesetzt werden. Die Stärke einer Animation besteht in dem Kontext darin, dass sie eine breitere Vielfalt von Übungsmöglichkeiten schafft. Die Interaktionsmöglichkeiten des Lerners mit der Instruktionaufgabe werden stark erweitert.

Zuletzt formuliert Mayer (2008) mehrere Prinzipien, um „extraneous processing“ („überflüssige Verarbeitung“) zu vermindern und um „essential processing“ („konzeptuell wesentliche Verarbeitung“) sowie „generative processing“ („Verarbeitung, die zu einem

tieferen Verständnis führt“) zu fördern. Um an der Stelle nicht konzeptuell zu weit auszuufern, werden nur ein paar dieser Prinzipien genannt und beschrieben.

Um „extraneous overload“ bei Animationen zu vermeiden, sollte man laut Mayer (2008) beispielsweise das „Coherence Principle“ berücksichtigen. Animationen können belanglose Bilder oder Geräusche beinhalten, die ablenken, oder sie können von Erzählungen begleitet sein, die unwichtige Informationen enthalten. Solche unwesentlichen Elemente sollten in einer Animation oder in einer Erzählung vermieden werden. Reichert man eine Animation mit einer gesprochenen narrativen Erzählung an, sollten die Animation und die Erzählung gleichzeitig präsentiert werden („Temporal Contiguity Principle“). Bei einer sequentiellen Präsentation der Animation und der Erzählung entsteht eine zusätzliche Belastung, da die Lerner dann die Wörter bzw. die Animation im Arbeitsgedächtnis halten müssen, wodurch weniger Kapazität für eine tiefere Verarbeitung zur Verfügung steht.

Um das essentielle Verstehen des Lernmaterials zu fördern, bietet sich das oben schon genannte „Segmenting Principle“ an, bei dem Personen besser lernen, wenn die Animationen in bedeutsame Teile zergliedert werden, die vom Lerner kontrolliert werden können. Das „Pre-training Principle“ besagt zudem, dass Personen besser von einer Animation lernen, wenn sie vorher ein Training in den Namen und den Merkmalen der Hauptkonzepte des dargestellten Vorganges hatten. Wenn ein Lerner schon die Namen und Zustände der einzelnen Komponenten kennt und somit Komponentenmodelle aufgebaut hat, kann er während der Animation mehr kognitive Ressourcen verwenden, um die Schritte der kausalen Kette zu verarbeiten.

Beim „generative processing“ geht es darum, relevante Informationen auszuwählen, sie mental zu organisieren und zu integrieren, um ein tiefes Verstehen zu gewährleisten. Um diese Verarbeitung zu unterstützen, empfiehlt Mayer (2008), eine „soziale Verbindung“ zwischen dem Lerner und dem Computerprogramm aufzubauen. Das bedeutet, dass wenn eine Animation mit gesprochenem Text angereichert ist, sollte dieser eher in Konversations- als in formalem Stil gehalten sein („Personalization Principle“). Zudem sollte der Text mit Standard-Akzent und menschlicher Stimme gesprochen werden und eher nicht mit einer menschlichen akzentbehafteten oder einer Maschinenstimme, damit der Lerner den Text als soziale Konversation akzeptiert („Voice Principle“).

2.3 Studien zum Vergleich von Animationen mit statischen Bildern

Die bisherige Forschung hat nicht ergeben, dass animierte Bilder generell eine Überlegenheit zu statischen Bildern haben. Es ist eher unwahrscheinlich, dass es eine einfache Dichotomie bezüglich des Vergleichs dieser beiden Darstellungsformen gibt (Lowe, 2001), da sowohl statische Bilder als auch Animationen Vor- und Nachteile hinsichtlich ihrer Wahrnehmung und ihrer kognitiven Verarbeitung aufweisen. Boucheix und Schneider (2009) resümieren, dass eine geringe Anzahl von Studien positive Effekte der Animationen beim Verständnis komplexer Systeme gezeigt hat (Hidrio & Jamet, 2002), wohingegen bei anderen Studien kein oder sogar ein negativer Effekt der animierten Darstellungen gefunden wurde (Lowe, 1999a). In manchen Studien erbrachte eine sequentielle Präsentation eines Prozesses mehr Vorteile als eine Animation, um eine mentale Repräsentation aufzubauen (Hegarty, 1992). In anderen Untersuchungen war eine kontinuierliche animierte Darstellung einer statischen Präsentation gegenüber überlegen (Catrambone & Fleming Seay, 2002).

Auch Tversky et al. (2002) berichten, dass Animationen in Evaluationsstudien nicht die erwarteten positiven Effekte gezeigt haben, ebenso wie Hegarty und Kriz (2008), die in ihren Untersuchungen keine empirische Basis dafür gefunden haben, dass Personen von Animationen mehr lernen als von statischen Bildern. Beispielsweise werden Animationen häufig verwendet, um zu demonstrieren, wie ein Computer funktioniert, wobei sie hier aber auch nicht über die Effektivität verfügen, die von ihnen erwartet worden war. Letztendlich merkt auch Rieber (1990) an, dass die Daten zur Forschung mit Animationen inkonsistent sind.

Für diese generelle Inkonsistenz der Befunde beim Vergleich von Animationen und statischen Bildern könnten mehrere Gründe verantwortlich sein:

Um Animationen an sich zu evaluieren, werden sie mit statischen Bildern verglichen, wobei die Animationen nicht mit eigenen Features, wie beispielsweise Interaktivität, angereichert sein dürfen, da diese einen zusätzlichen Effekt auf den Lerner haben können. In manchen Studien, bei denen Animationen mit statischen Bildern verglichen wurden, waren die Animationen jedoch mit Interaktivität angereichert, die statischen Bilder hingegen nicht, wodurch die genaue Vergleichbarkeit der beiden Darstellungsformen herabgesetzt war. In anderen Untersuchungen war der Effekt der Animation mit der lernerleichternden Maßnahme der „Vorhersage“ konfundiert (Tversky et al. 2002).

Vergleichsstudien zwischen Animationen und statischen Bildern müssen sicherstellen, dass diese beiden Präsentationsformen informationsmäßig zueinander äquivalent sind. Oft haben Animationen und statische Bilder keinen vergleichbaren Inhalt, wie beispielsweise in der Studie von Large, Beheshti, Bruleux und Renaud (1996), in der Animationen, die das menschliche Blutkreislaufsystem darstellen sollten, Blutbahnen enthielten, die bei den statischen Bildern nicht gezeigt wurden. Gerade wenn die Animationen komplexe Bewegungen zeigen, bei denen das Timing und die räumlichen Positionen der Elemente gleichzeitig wichtig sind, ist es schwierig, informationsäquivalente statische Bilder zu konstruieren. Im Allgemeinen ist es häufiger so, dass Animationen Informationen präsentieren, die nicht in statischen Bildern enthalten sind, da sie von dem Beobachter in gröbere und feinere Einheiten zerlegt werden können und somit auch die Mikroschritte zwischen den größeren Schritten präsentieren (Zacks, Tversky & Iyer, 2001). Statische Bilder stellen oft nur die groben Einheiten eines Vorganges dar, wodurch beim Lernen mit Animationen Vorteile durch zusätzliche Informationen und nicht durch ein unterschiedliches Präsentationsformat entstehen (Tversky et al., 2002). Animationen und statische Bilder weisen somit häufiger eine unterschiedliche Granularität auf, durch die die widersprüchlichen Ergebnisse der Vergleichsstudien teilweise mit erklärt werden können. Letztendlich muss bei Vergleichsstudien mit Erwachsenen zudem das Bedürfnis der Teilnehmer nach visueller Unterstützung validiert werden. Erwachsene haben bessere persönliche Lernstrategien und sind weniger abhängig von äußeren Bildern, was ein potentiell konfundierender Effekt in Animationsstudien sein kann (Rieber, 1990).

Im Folgenden soll ein kurzer Überblick über Studien gegeben werden, die in verschiedener Form Animationen mit statischen Bildern verglichen haben, wodurch nochmals anhand konkreter Fälle deutlich wird, dass bei keiner der beiden Darstellungsformen eine eindeutige Überlegenheit konstatiert werden kann.

In der Studie von Morrison und Tversky (2001) wurde eine Textbedingung mit zwei weiteren Bedingungen verglichen, in denen der Text mit animierten oder mit statischen Bildern angereichert war, wobei das Kriterium war, Aufgaben zu sozialen Pfaden zwischen Personen zu lösen. Mit den Bildbedingungen wurden bessere Ergebnisse erzielt als mit der Textbedingung, allerdings galt dies nur für Personen mit geringeren räumlichen Fähigkeiten. Das Lernen mit animierten Bildern erzeugte jedoch keine besseren Resultate als das mit statischen Bildern.

Mayer, Hegarty, Mayer und Campbell (2005) verglichen verschiedene Inhalte, wie beispielsweise die Funktionsweise eines Bremssystems, die durch Sequenzen von textbegleiteten statischen Bildern dargestellt wurden, mit einer Präsentation dieser Inhalte durch computerbasierte narrative Animationen. Es wurde keine Überlegenheit der Animationen gefunden.

Studien, in denen das Lernen meteorologischer Informationen von Nöten war, zeigten sogar, dass Versuchspersonen eher statische Bilder hierfür nutzten als kontinuierliche Animationen (Bogacz & Trafton, 2005).

Jedoch zeigten die Probanden in einer Studie von Thompson und Riding (1990) signifikant bessere Lernleistungen zum „Satz des Pythagoras“, wenn sie darüber mit Animationen gelernt hatten als wenn sie mit mehreren statischen Bildern oder nur mit einem Bild gelernt hatten. Die Autoren führten dies auf die vollständigere Informationsdarbietung durch Animationen zurück.

Um Computeralgorithmen zu lernen, erwiesen sich Animationen ebenfalls als eine bessere Hilfe als diskrete statische Bilder (Catrambone & Fleming Seay, 2002).

Rieber (1989) führte eine Studie an Grundschulern durch, bei der das Lernmaterial „Newtons Bewegungsgesetze“ thematisierte und das entweder nicht oder mit stehenden Bildern oder mit animierten Bildern illustriert war. Die Animationsbedingung wurde in den Hypothesen favorisiert, da durch dynamische Darstellungen Bewegungsbahnen gut präsentiert werden können. Es zeigte sich allerdings, dass die Animationsgruppe weder beim Faktenlernen, noch bei den Transferaufgaben signifikant bessere Ergebnisse hatte als die statische Bildgruppe. Zurückgeführt wurde dies auf einen zu hohen Schwierigkeitsgrad der Lerneinheit und darauf, dass die Probanden signifikant weniger Lernzeit beim Betrachten der Animationen nutzten. Eventuell haben die Teilnehmer der animierten Bedingung die Zeit für andere mentale Aufgaben genutzt.

In einer Replikation der Studie (Rieber, 1990) reduzierte Rieber den Schwierigkeitsgrad des Lernstoffs und unterteilte die Animation in kleinere Sequenzen, die in direkter Beziehung zum Lerntext standen. In dieser Untersuchung erreichte die Animationsbedingung signifikant bessere Ergebnisse als die anderen Bedingungen, da die Animation inhaltlich besser in die restliche Lerneinheit eingebunden war. Diese positiven Effekte verschwanden allerdings, wenn die Schüler keine Übung im Umgang mit der Animation hatten, wodurch ihre Effektivität von der vollen Unterrichtsunterstützung abhing.

Rieber (1991) nutzte das gleiche Lernmaterial noch einmal und überprüfte die Lernleistung von Viertklässlern bezüglich expliziter und impliziter Lerninhalte. Die Animationsgruppe erzielte signifikant bessere Ergebnisse als die Gruppe, die mit statischen Bildern lernte, wobei dieser Vorteil vorwiegend auf ein höheres implizites Wissen zurückging. Durch die Animationen konnten aber schon bestehende Fehlvorstellungen zum Lernstoff nicht korrigiert werden, was mit statischen Bildern eher der Fall war.

Bei Verwendung des Lernmaterials zu den Newtonschen Bewegungsgesetzen bei Erwachsenen (Rieber, Boyce & Assad, 1990) zeigte sich, dass durch die Animationen keine besseren Lernleistungen erzielt werden konnten, dass aber der nachfolgende Test schneller beantwortet wurde. Animationen können somit für die Enkodierung und für den Abruf von Informationen eine sinnvolle Hilfestellung sein. Sie erleichtern die Organisation und die Rekonstruktion des Inhalts während des Abrufs.

Ein Spezialfall des Vergleichs von Animationen mit statischen Bildern stellen solche Studien dar, bei denen die Wechselwirkungen des Präsentationsformats mit individuellen Eigenschaften, wie beispielsweise den räumlichen Fähigkeiten einer Person, mit untersucht werden.

Zu den Interaktionen von Animationen mit den räumlichen Fähigkeiten einer Person gibt es ebenfalls widersprüchliche Befunde: Es ist gut möglich, dass Animationen für Personen mit niedrigeren räumlichen Fähigkeiten effektiver sind als statische Bilder, da diesen Personen die mentale Animation mit statischen Bildern schwerer fallen dürfte. Es gibt allerdings auch Untersuchungen, bei denen keine Interaktion zwischen den räumlichen Fähigkeiten einer Person und dem Präsentationsformat vorlag, in anderen Studien stellte sich sogar heraus, dass Personen mit höheren räumlichen Fähigkeiten besser mit den Animationen lernten (Garg, Norman & Sperotable, 2001).

In einer Studie von Blake (1977) wurden die Bewegungsabfolgen von fünf Schachzügen drei unterschiedlichen Lernergruppen präsentiert. Die erste Gruppe lernte mit einer vollständig animierten Darstellung, die zweite mit einer teilweise animierten Darstellung, und die dritte Gruppe betrachtete statische Bilder. Im Anschluss sollten die Probanden die Bewegungsabfolgen zur Lösung von Schachproblemen aufzeichnen. Durch die animierten Präsentationen wurden signifikant bessere Lernleistungen erreicht als durch die statische Informationsdarbietung. Die animierten Versionen führten bei Personen mit niedrigerem räumlichem Vorstellungsvermögen zu signifikant besseren Ergebnissen als die statischen Bilder. Animationen können dazu dienen, Lerner mit geringerem räumlichem

Vorstellungsvermögen beim Erlernen von dreidimensionalen Inhalten und Abläufen zu unterstützen. Die Effektivität einer Animation könnte somit auch durch das räumliche Vorstellungsvermögen von Probanden beeinflusst werden.

Dieser Befund konnte jedoch durch die Studie von Hegarty, Kriz und Cate (2003) nicht unterstützt werden. Die Autoren verglichen eine Gruppe, die die Funktionsweise eines Spülkastens mit einem statischen Bild lernte, mit einer Gruppe, die den Vorgang mit mehreren statischen Bildern lernte und mit einer Gruppe, die den Prozess mit Hilfe einer Animation sah. Es zeigte sich, dass die Personen, die mit der Animation lernten, keine besseren Ergebnisse erzielten als die Personen, die mit mehreren statischen Bildern lernten. Jedoch waren diese beiden Gruppen besser als diejenige, die mit einem statischen Bild lernte. Weiterhin fanden sich keine signifikanten Interaktionen zwischen dem Instruktionsformat und den räumlichen Fähigkeiten, dem theoretischen Wissen oder der praktischen Vorerfahrung der Teilnehmer. Generell konnten Hegarty und Kriz (2008) in mehreren, von ihnen durchgeführten, Experimenten keine Hinweise dafür finden, dass Animationen für bestimmte Personen besser geeignet waren als statische Bilder, auch wenn sie mit Text kombiniert waren.

Somit kann auch bezüglich der Frage nach möglichen Wechselwirkungen von Animationen und statischen Bildern mit individuellen Eigenschaften von Personen keine klare Antwort gegeben werden.

Trotz der widersprüchlichen empirischen Befunde lassen sich dennoch aus der bestehenden Literatur Hinweise dazu extrahieren, unter welchen Umständen Animationen eine positive Wirkung haben können, was als Anhaltspunkt für ihren sinnvollen Einsatz dienen kann.

Baek und Layne (1988) weisen darauf hin, dass Animationen nicht per se eine positive Wirkung haben, sondern vor allem dann hilfreich sind, wenn sie mit verbalen Erklärungen ergänzt sind, die ihre Verständlichkeit erhöhen.

Nach Lewalter (1997) wirken sich Animationen dann positiv auf die Lernleistung aus, wenn

- a) Veränderungen und Bewegungen im zeitlichen Verlauf dargestellt werden sollen und einen relevanten Aspekt des Lerninhalts ausmachen.

- b) Lerner mit geringem Vorwissen vorher auf die relevanten Aspekte der Animation aufmerksam gemacht werden. Falsches Vorwissen führt dazu, dass Informationen nicht wahrgenommen oder falsch interpretiert werden.

- c) Lerner über ein geringes räumliches Vorstellungsvermögen verfügen und mit einer dynamischen Informationspräsentation bei der Generierung eines dreidimensionalen mentalen Vorstellungsbildes unterstützt werden. Inwieweit die räumlichen Fähigkeiten durch Animationen grundlegend gefördert werden können, ist noch unklar.

Positive Effekte einer Animation können jedoch durch zu schwierige Aufgaben oder durch ein allgemein schwaches Instruktionsdesign untergraben werden.

Generell hängt die Effektivität von Animationen auch davon ab, inwiefern der Lerner ihrer spezifischen Eigenschaften bedarf, um einen bestimmten Inhalt zu verstehen und inwiefern er dazu fähig ist, mit den Anforderungen dieser Darstellungsform umzugehen.

3 Zielsetzungen der vorliegenden Arbeit

Der Überblick über die bereits vorhandene Literatur hat gezeigt, dass es keine eindeutige Überlegenheit von Animationen oder statischen Bildern bezüglich der Frage gab, welche der beiden Präsentationsformen für einen Lernvorgang im Allgemeinen geeigneter ist. Da sich somit die Befundlage als inkonsistent darstellt, soll in der vorliegenden Arbeit ein differenzierterer Vergleich zwischen statischen Bildern und Animationen angestrebt werden. Der grundlegende Gedanke ist, dass es auch an der Aufgabenart liegen könnte, die nach einem Lernprozess gelöst werden soll, ob man besser Animationen oder statische Bilder zum Lernen heranzieht. Konkret zielen die beiden Studien der aktuellen Arbeit darauf ab, die Wirksamkeit von Animationen und Bildern auf die Lösung von Inferenz- und bildlichen Kategorisierungsaufgaben zu untersuchen und jeweils miteinander zu vergleichen. Zusätzlich sind als Teil der ersten Studie auch noch die Einflüsse verschiedener Präsentationsformen von statischen Bildern auf die Bearbeitung von Inferenzaufgaben von Interesse. Ein aufgabenabhängiger Vergleich von Animationen und statischen Bildern birgt auf einer weiteren Ebene die Abklärung der Frage in sich, ob sich diese beiden Darstellungsformen für den Aufbau perzeptueller und kognitiver Repräsentationen dynamischer Sachverhalte unterschiedlich gut eignen, da für die Lösung von Inferenzaufgaben kognitive mentale Modelle, für die Bearbeitung bildlicher Kategorisierungsaufgaben jedoch perzeptuelle dynamische Schemata benötigt werden, wie unten näher erläutert werden wird. Daher tritt innerhalb der ersten Studie der Arbeit mehr die kognitive Komponente des Lernens in den Vordergrund, während bei der zweiten Studie der Akzent auf perzeptuellen Lernaspekten liegt.

Im Folgenden sollen zunächst die Grundannahmen der beiden Studien getrennt voneinander theoretisch hergeleitet und konkret genannt werden. Bei den zugehörigen empirischen Teilen der Studien werden zu Beginn jeweils die entsprechenden Hypothesen nochmals aufgezählt und, sofern nötig, mit Erläuterungen versehen.

3.1 Theoretische Herleitung der Grundannahmen von Studie 1: Vergleich von Animationen und statischen Bildern hinsichtlich ihrer Wirksamkeit auf die Lösung von Inferenzaufgaben

3.1.1 Teilstudie 1

Grundannahmen zu den Haupteffekten zwischen den Lernbedingungen

Haupteffekte zwischen Animationen und seriell schnellen sowie seriell langsamen Key-Frames

Im ersten Teil der ersten Studie sollen zunächst Animationen und statische Bilder hinsichtlich ihrer Wirksamkeit auf die Lösung von Inferenzaufgaben untersucht werden. Die statischen Bilder werden hierbei in serieller Abfolge präsentiert, d.h., der Lerner sieht zu einem Zeitpunkt immer nur ein Bild. Die Präsentationsgeschwindigkeit der Bilder wird zudem variiert, so dass sie in einmal Fall eher schnell wechseln und die einzelnen Bilder jeweils nur kurz zu sehen sind, während in einer weiteren Bedingung die Darstellungsdauer der einzelnen Bilder erhöht wird und die Wechsel zwischen den Bildern langsamer erfolgen. Aufgrund dieser verschiedenen Präsentationsgeschwindigkeiten der statischen Bilder könnten sich differentielle Effekte auf das Lernen ergeben, die in dieser Studie auch untersucht werden.

Bei dem Vergleich von Animationen und statischen Bildern ist die Annahme, dass statische Bilder besser für einen Lernvorgang geeignet sind, wenn nach ihm Inferenzaufgaben gelöst werden sollen. Der Grund für diese Aussage liegt darin, dass für die Bewältigung von Inferenzaufgaben zu Prozessen, bei denen neue, im Lernmaterial nicht explizit gesehene Informationen neu hergeleitet werden müssen, ein dynamisches mentales Modell benötigt wird. Die Strukturen des mentalen Modells, die zu denen des gelernten Prozesses analog sind, werden für die Lösung von Inferenzaufgaben in angemessener Form manipuliert, und die neue, inferierte Information wird abgelesen (Schnotz, 1994). Eine weitere Annahme besteht nunmehr darin, dass dynamische mentale Modelle besser durch statische Bilder und besonders durch die sog. „key states“ (Schnotz & Lowe, 2008) eines Vorganges aufgebaut werden als durch Animationen. Um diese Annahme zu erläutern, soll zunächst auf den Term

„key states“ eingegangen werden. Key-States sind prototypische Bilder, die ein Ereignis in seiner Gesamtheit am besten repräsentieren. Daher liegen sie an ungefähr jeweils mittleren Positionen in den Teilprozessen eines Ereignisses. Sie sind sehr informationsreich, da die anderen Zustände der Unterereignisse aus ihnen einfach inferiert werden können. Zudem befinden sich die Key-States auch an den kausalen Schnittstellen eines Ereignisses, da kausale Informationen essentiell für den Aufbau eines funktionsfähigen mentalen Modells sind. Diese kausalen Schnittstellen sind oft dort positioniert, wo die Teilprozesse eines Vorganges beginnen bzw. enden. Aufgrund ihrer Informationsdichte und ihres prototypischen Charakters sind Key-States somit vorteilhaft für einen Lernvorgang. Die Key-States sind deshalb wichtig für den Aufbau eines mentalen Modells, da davon ausgegangen wird, dass eine Person beim Lernen eines dynamischen Vorganges diesen visuell in Unterereignisse zergliedert. Die Grenzen zwischen den Unterereignissen liegen häufig dort, wo innerhalb des Prozesses perzeptuell „maximale physikalische Veränderungen“ im Bewegungsablauf stattfinden. An diesen Stellen befinden sich oft zudem kausale Momente innerhalb des Vorganges (Zacks & Tversky, 2001). Die Zustände innerhalb eines Unterereignisses sind sich hinsichtlich verschiedener Merkmale ähnlich, so dass die Gesamtheit dieser Zustände vom menschlichen kognitiven Apparat in einem Bild, einem Key-State, zusammengefasst werden kann („chunking“). So wird die Gesamtheit eines Prozesses durch wenige Key-States repräsentiert, und diese werden wegen ihrer prototypischen Eigenschaften gut enkodiert. Ein Ereignis wird somit in Form seiner Key-States als ein kognitives Schema im Langzeitgedächtnis gespeichert, da dies einen „kognitiv ökonomischen“ Weg darstellt, den Prozess mental zu repräsentieren. Wird nun ein dynamisches mentales Modell eines Vorganges benötigt, um beispielsweise Inferenzaufgaben zu lösen, werden die Key-States abgerufen, und es werden aus ihnen im Arbeitsgedächtnis die anderen Zustände des Ereignisses inferiert, so dass das Gesamtmodell konstruiert werden kann. Daher rührt die Annahme, dass die Darstellung eines Vorganges in Form seiner Key-States die menschliche kognitive Architektur beim Aufbau eines dynamischen mentalen Modells und bei der anschließenden Lösung von Inferenzaufgaben unterstützt. Dem kognitiven und perzeptuellen Apparat werden Gliederungs- und Chunkingprozesse beim Lernen von Ereignissen erspart, die er beim Betrachten einer Animation leisten muss (Schnotz & Lowe, 2008).

Zudem hat der Lerner bei statischen Bildern den Vorteil, dass sie nicht so flüchtig sind wie Animationen, er hat die Möglichkeit, die Stellen, die er nicht ganz verstanden hat, genauer

zu inspizieren (Lewalter, 1997). Sein Arbeitsgedächtnis wird nicht so sehr mit immer neuen Eindrücken überflutet (Paas et al., 2007). Gerade bei seriellen Bildern werden die Veränderungen zwischen den Frames besonders salient, was lernförderlich ist. Obwohl Animationen vollständige Informationen über einen Prozess liefern und der Lerner hier somit keine Zustände selbst inferieren und den Vorgang nicht eigenständig mental animieren muss, was einen gewichtigen Vorzug dieser Präsentationsform darstellt (Hegarty et al., 2003), wird in der aktuellen Studie davon ausgegangen, dass Inferenzaufgaben nach dem Lernen mit Animationen schlechter gelöst werden bzw. dass der Lernvorgang als anstrengender erlebt wird als wenn mit Key-States gelernt wurde.

Somit sollen in der ersten Teilstudie folgende Grundannahmen untersucht werden:

1. Durch das Lernen mit statischen Key-Frames (= Bilder, die Key-States zeigen), die in seriell *schneller* Abfolge präsentiert werden, werden Wissens- und Inferenzaufgaben besser gelöst als wenn mit Animationen gelernt wird. Zudem werden der Lernprozess selbst und die Bearbeitung des Nachtests bei Key-Frames als weniger anstrengend erlebt als bei Animationen.
2. Durch das Lernen mit statischen Key-Frames, die in seriell *langsamer* Abfolge präsentiert werden, werden Wissens- und Inferenzaufgaben besser gelöst als wenn mit Animationen gelernt wird. Zudem werden der Lernprozess selbst und die Bearbeitung des Nachtests bei Key-Frames als weniger anstrengend erlebt als bei Animationen.

Zwischen Animationen und seriell schnellen bzw. seriell langsamen Bildern werden bewusst eigenständige Grundannahmen aufgestellt. In beiden Fällen wird zwar jeweils von einer Überlegenheit der Key-Frames ausgegangen, durch deren unterschiedliche Präsentationsgeschwindigkeit und die damit einhergehende unterschiedliche Betonung von Aspekten des Lernmaterials, könnten sich aber verschieden große Effekte ergeben.

Bei diesen und den nachfolgenden Grundannahmen ist nicht nur der Einfluss von Animationen und statischen Bildern auf verschiedene Leistungsmaße von Belang, sondern es wird auch das subjektive Maß des „cognitive loads“ (Sweller, 1988), also die beim Lernen empfundene mentale Belastung der Probanden, mit berücksichtigt. Dies dient dazu, die „instruktionale Effizienz“ der Präsentationsformen jeweils zu erfassen, die ein Indikator dafür ist, zu welchen „Kosten“ eine bestimmte Leistung erbracht wird (Paas et al., 2007).

Die Thematik des Lernmaterials bei den beiden Teilstudien der ersten Studie dieser Arbeit wird die „Signalübertragung an Synapsen“ (Synapse = Verbindung zwischen zwei Nervenzellen) sein, da dieser Vorgang eine Kausalkette darstellt, die eindeutig in Unterereignisse zerlegt werden kann und die gut durch entsprechende Key-Frames repräsentiert werden kann.

Haupteffekte zwischen seriell schnellen und seriell langsamen Key-Frames

Wie oben bereits angedeutet wurde, werden in der ersten Teilstudie auch Vergleiche zwischen seriell schnellen und seriell langsamen Key-Frames hinsichtlich ihrer Wirkungen auf die Lösung von Inferenzaufgaben und die mentale Belastung von Personen angestellt. Der Unterschied zwischen diesen beiden Präsentationsarten besteht darin, dass bei seriell schnellen Bildern die einzelnen Key-Frames eher kurz gezeigt werden, wodurch die Wechsel zwischen den Bildern schneller sind und der Gesamtvorgang öfters wiederholt wird. Bei seriell langsamen Bildern ist die Darstellungsweise hingegen umgekehrt: Die einzelnen Frames werden länger gezeigt, dadurch wechseln die Bilder langsamer, und die Zahl der Wiederholungen des Gesamtvorganges ist niedriger. Welche Präsentationsgeschwindigkeit lernförderlicher ist, ist unklar, wodurch zu diesem Haupteffekt im empirischen Teil ungerichtete Hypothesen formuliert werden.

Für die seriell schnellen Bilder spricht, dass bei ihrer Darstellung der Vorgang öfters gezeigt wird, der Lerner hat somit mehrere Möglichkeiten, die Prozesse zu verstehen. Die Inferenzen zwischen den Bildern, die eine Person beim Lernen mit Key-Frames leisten muss, sind leichter zu ziehen, da die Frames rasch aufeinander folgen. Der Zusammenhang zwischen den Bildern und somit der Gesamtkontext wird einfacher verstanden, so dass Inferenzaufgaben, die sich auf mehrere Key-Frames beziehen, besser gelöst werden können. Durch die höhere Präsentationsgeschwindigkeit werden die Makrostrukturen des Vorganges betont (Lowe, 2006), so dass der Lerner abstraktere Informationen erfährt. Zudem wäre eine weniger detailbetonte Darstellung, wie sie bei seriell schnellen Bildern der Fall ist, analog zum Prinzip der „elaborativen Sequenzierung“ nach Reigeluth (Reigeluth & Stein, 1983), bei der auch zunächst die Vorgänge eher grob und dann erst detailliert präsentiert werden („zooming“), was als verstehensförderlich angesehen wird.

Seriell langsame Bilder haben hingegen den Vorteil, dass die einzelnen Frames relativ lange gezeigt werden. Der Lerner hat genügend Zeit, sich in den Bildern zu orientieren, die einzelnen Abläufe zu verstehen und sie mental zu verbalisieren. Neue Inhalte können so besser organisiert und mit bereits vorhandenem Wissen integriert werden. Ein „overload“ wird vermieden, da auf den Lerner nicht so viele Eindrücke gleichzeitig einströmen (Sweller, van Merriënboer & Paas, 1998). Detail- und „intrastate“-Informationen werden gut verstanden, wodurch präzisere Inferenzaufgaben leichter beantwortet werden können. Zuletzt ähnelt eine seriell langsame Darstellung von Bildern einer „parts first“-Präsentation, wie sie bei Mayer und Chandler (2001) beschrieben wird. Bei einer solchen Darstellungsweise, bei der zuerst die einzelnen Teile eines Prozesses und dann erst der Gesamtvorgang gezeigt werden, kann der Lerner besser Komponentenmodelle der einzelnen Teilereignisse erstellen und diese dann miteinander integrieren. Eine Betonung der einzelnen Teile eines Ereignisses, wie sie bei der Präsentation von seriell langsamen Bildern der Fall ist, erlaubt es somit einer Person, ein mentales Modell schrittweise zu erstellen.

Aus diesen Überlegungen heraus ergibt sich die folgende Grundannahme zum Vergleich von seriell schnellen mit seriell langsamen Key-Frames:

3. Ein Lernvorgang mit seriell schnellen Key-Frames führt zu unterschiedlichen Leistungen in Wissens- und Inferenzaufgaben sowie zu einem anderen Cognitive Load als ein Lernvorgang mit seriell langsamen Key-Frames. Welche der beiden Präsentationsformen sich als überlegen erweist, bleibt zunächst unklar, da für beide Darstellungsgeschwindigkeiten Vor- und Nachteile gefunden werden können.

Grundannahmen zu den Interaktionen zwischen den Lernbedingungen und dem Vorwissen von Personen

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit dem Vorwissen von Personen

Innerhalb der ersten Teilstudie wird von bedeutsamen Wechselwirkungen zwischen der Animations- und der seriell schnellen bzw. seriell langsamen Bildbedingung mit dem Vorwissen von Personen ausgegangen. Bei den Haupteffekten zwischen den Gruppen waren

gerichtete Hypothesen aufgestellt worden, die zu Gunsten der seriellen Key-Frames formuliert waren. Aus der bereits vorhandenen Literatur wird jedoch ersichtlich, dass das Vorwissen sowohl das Lernen mit statischen Bildern als auch das mit Animationen moderierend beeinflussen kann, wobei die Richtung des Interaktionsmusters nicht eindeutig ist (Hegarty & Kriz, 2008).

Eine mögliche differentielle Wirkung der statischen Bilder wäre, dass der Lernvorteil durch Key-Frames bei Personen mit weniger Vorwissen stärker ausgeprägt ist als bei Personen mit höherem Vorwissen.

Für Personen mit weniger Vorwissen könnten statische Bilder zum Lernen und anschließenden Lösen von Inferenzaufgaben in der aktuellen Studie deshalb geeigneter sein, da Key-Frames, wie oben bereits ausgeführt wurde, generell einen förderlichen Einfluss auf den Aufbau eines dynamischen mentalen Modelles haben. Der zu lernende Vorgang wird external durch die Key-Frames gegliedert und gehunkt, so dass Personen mit geringerem Vorwissen diese schwierigeren Anforderungen nicht mehr leisten müssen. Statische Bilder, die die Key-States eines Prozesses zeigen, sind somit kongruent zur menschlichen kognitiven Architektur, was für Novizen besonders hilfreich sein könnte (Schnotz & Lowe, 2008).

Bei statischen Bildern wird durch ihren stationären Charakter das Arbeitsgedächtnis des Lerners zudem nicht so stark mit immer neuen Reizen überflutet, wie es bei einer Animation der Fall ist, obwohl zumindest bei den seriell schnellen Key-Frames auch eine gewisse Flüchtigkeit vorhanden ist. Dennoch können bei seriellen Bildern die einzelnen Zustände mehr mit externer Unterstützung miteinander verglichen und integriert werden. Der Lerner kann eher bedürfnisgerecht die Details betrachten, die er nicht verstanden hat (Kosslyn, 1994). Zudem werden gerade bei seriellen Key-Frames die Veränderungen zwischen den Bildern besonders salient. Zumindest seriell langsame Key-Frames haben für Personen mit weniger Vorwissen den Vorteil, dass das mentale Modell schrittweise aufgebaut werden kann (Mayer & Chandler, 2001). Hingegen könnten Novizen beim Lernen mit einer Animation mit deren Informationsdichte überfordert sein. Zu viele Eindrücke strömen auf den Lerner ein, die er aufgrund ihrer Flüchtigkeit nicht alle beachten kann. Dadurch kann es sein, dass er konzeptuell irrelevanten, aber perzeptuell salienten Details seine Aufmerksamkeit schenkt, wodurch sein Verständnis des Gesamtprozesses beeinträchtigt wird (Lowe, 1999a). Nur wenn man über ein gewisses Vorwissen verfügt,

kann man eine dynamische Darstellung adäquat strukturieren. Daher könnten serielle Bilder für Personen mit niedrigerem Vorwissen geeigneter sein als Animationen.

Für Personen mit höherem Vorwissen hingegen könnte eine Animation eher lernförderlich sein, da sie mit ihrer Informationsdichte gut umgehen können. Sie sind in der Lage, das Material zu strukturieren und wissen, auf welche konzeptuell relevanten Details sie ihre Aufmerksamkeit lenken sollen. Fortgeschrittene Lerner können die einzelnen Zustände des Prozesses durch ihre höhere Verarbeitungskapazität auch besser internal miteinander vergleichen. Die Gliederungs- und Chunkingprozesse bei einer dynamischen Darstellung sind eine angemessene Anforderung für sie. Gerade seriell langsame Bilder könnten durch ihren monotonen Charakter eher demotivierend für fortgeschrittene Lerner wirken, und eventuell schweifen sie bei einer solchen Präsentationsform mehr mit ihrer Konzentration ab als bei Animationen, bei denen durch ihre permanente Dynamik die Aufmerksamkeit des Betrachters stets auf den Screen gelenkt wird. Daher ist es möglich, dass der Lernvorteil durch Key-Frames bei Personen mit höherem Vorwissen nicht so stark ausgeprägt ist wie bei Personen, die mit weniger Vorwissen ausgestattet sind, und eventuell erzielen fortgeschrittene Lerner bessere Ergebnisse mit Animationen.

Allerdings ist auch das gegenteilige Interaktionsmuster plausibel, bei dem der Lernvorteil durch Key-Frames bei Personen mit höherem Vorwissen größer ist als bei Personen, die über ein geringeres Vorwissen verfügen.

Bei statischen Bildern müssen generell die Zustände zwischen den Frames inferiert, und der gesamte Prozess muss mental animiert werden (Hegarty et al., 2003; Hegarty, 1992). Diese Notwendigkeit könnte besonders bei seriell langsamen Bildern schwieriger sein, da dort vor allem die „intrastate“-Information im Fokus steht. Für Personen mit höherem Vorwissen könnten diese Konstruktionsvorgänge jedoch eine angemessene Anforderung sein. Eventuell fördert und erweitert das Inferieren der Zwischenzustände ihr Verständnis und ihren Umgang mit dem Lernmaterial. Eine Animation liefert hingegen vollständige Informationen über einen Prozess (Lewalter, 1997), was aber für fortgeschrittene Lerner bedeuten kann, dass viele dieser Informationen für sie redundant sind und ihr Arbeitsgedächtnis unnötig belasten. Daher erzielen Personen mit höherem Vorwissen möglicherweise bessere Leistungen mit den statischen Key-Frames als mit den Animationen.

Bei Personen mit geringerem Vorwissen besteht in einem Lernvorgang mit statischen Bildern hingegen die Gefahr, dass sie Fehlinferenzen ziehen und aufgrund ihrer mangelnden Vorkenntnisse die „Lücken“ zwischen den Frames nicht korrekt füllen können (Narayanan

& Hegarty, 1998). Die mentale Animation des Vorganges ist zudem fehleranfällig und bedeutet einen zusätzlichen internalen Konstruktionsaufwand, der Novizen überfordern könnte (Hegarty, 2004). Bei Animationen müssen sie den Prozess zunächst nur wahrnehmen, da durch diese Darstellungsform vollständige Informationen geliefert werden und keine Inferenzen geleistet werden müssen. Eine Animation kann dem Lerner sogar dabei helfen, den Vorgang mental zu animieren und zu transformieren, da sie diese Prozesse extern präsentiert. Diese „Supplantationsfunktion“ (Salomon, 1979) von Animationen wird vor allem für den Erwerb von neuen Fähigkeiten, wie es bei Novizen der Fall ist, als hilfreich angesehen. Daher mag es sein, dass sich der Lernvorteil durch Key-Frames sich bei Personen mit weniger Vorwissen verringert und dass sie eventuell bessere Ergebnisse mit der Animation erreichen.

Aus diesen Überlegungen erfolgen als weitere zu untersuchende Grundannahmen:

4. Es besteht eine Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *schnell*“ mit dem Vorwissen von Personen sowohl hinsichtlich der Lösung von Wissens- und Inferenzaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Welche Richtungen die Interaktionen jeweils annehmen, bleibt zunächst unklar, da mehrere Interaktionsmuster plausibel sind.

5. Es besteht eine Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *langsam*“ mit dem Vorwissen von Personen sowohl hinsichtlich der Lösung von Wissens- und Inferenzaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Welche Richtungen die Interaktionen jeweils annehmen, bleibt zunächst unklar, da mehrere Interaktionsmuster plausibel sind.

Für den Vergleich von Animationen mit seriell schnellen und seriell langsamen Key-Frames werden getrennte Annahmen zu den Interaktionen mit der Variable „Vorwissen“ formuliert. Das ist dadurch begründet, dass bei seriell langsamen Bildern die Frames länger gezeigt werden, wodurch unterschiedliche Effekte in der Interaktion mit dem Vorwissen von Personen erzeugt werden können als es bei seriell schnellen Bildern der Fall ist, da bei seriell langsamen Key-Frames mehr die strukturellen Aspekte der Darstellung betont werden, während bei seriell schnellen Key-Frames die dynamischen Aspekte eines Vorganges im Vordergrund stehen.

Interaktionen zwischen den Bedingungen „Key-Frames seriell schnell“ und „Key-Frames seriell langsam“ mit dem Vorwissen von Personen

Als eine weitere Annahme der ersten Teilstudie wird von signifikanten Interaktionen zwischen den Bedingungen „Bild seriell schnell“ und „Bild seriell langsam“ mit dem Vorwissen von Personen ausgegangen. Da es denkbar ist, dass die diesbezüglichen Interaktionsmuster unterschiedliche Formen annehmen können, wird zusammengefasst eine ungerichtete Grundannahme aufgestellt werden. Zuvor waren bei den Haupteffekten zwischen seriell schnellen und seriell langsamen Key-Frames ebenfalls ungerichtete Hypothesen formuliert worden.

Ein mögliches Interaktionsmuster könnte sein, dass Personen mit einem geringeren Vorwissen mehr von seriell langsamen als von seriell schnellen Bildern profitieren im Vergleich zu Personen, die über ein höheres Vorwissen verfügen. Seriell langsame Key-Frames haben für Novizen den Vorteil, dass die einzelnen Bilder lange gezeigt werden, was zu einem guten Verständnis der Teilprozesse führt. Der Betrachter hat genügend Zeit, die einzelnen Vorgänge zu verbalisieren und die Teile der Abbildung zu benennen. Die Darstellung durch seriell langsame Bilder ähnelt einer „parts first“-Präsentation, bei der der Lerner zunächst die Komponentenmodelle der Teilvorgänge erstellen und darauf aufbauend schrittweise das gesamte mentale Modell des Prozesses konstruieren kann. Eine solche Darstellungsform wird in der Studie von Mayer und Chandler (2001) als besonders förderlich für Personen mit weniger Vorwissen beschrieben. Zudem wird der Betrachter durch die höhere Langsamkeit der Bilder nicht so stark mit immer neuen Eindrücken überflutet. Die Präsentation ist nicht so flüchtig, weshalb der Lerner die einzelnen Zustände mehr externalisiert miteinander vergleichen und integrieren kann. Das Risiko, in einem Frame ausschließlich auf die weniger relevanten Details zu achten und im weiteren Lernprozess dadurch den Anschluss zu verlieren, ist geringer (Lowe, 1999a). Bei den seriell schnellen Key-Frames kann es hingegen durch deren höhere Flüchtigkeit, die allerdings in der vorliegenden Studie durch die mehrfachen Wiederholungen des Gesamtprozesses etwas abgemildert ist, eher zu diesen eben genannten Defiziten kommen. Zudem ist bei seriell schnellen Bildern die Präsentationsdauer der einzelnen Frames zu kurz, um ein gutes Verständnis der Teilprozesse zu gewährleisten. Der Lerner kann kein Grundlagenwissen über die einzelnen Zustände erwerben, was für einen Novizen für das Gesamtverstehen

essentiell ist (Mayer & Chandler, 2001). Daher sind möglicherweise seriell langsame Key-Frames für Personen mit weniger Vorwissen besser geeignet als seriell schnelle Key-Frames. Lerner mit höherem Vorwissen können hingegen mit dem flüchtigeren Charakter der seriell schnellen Bilder besser umgehen. Sie wissen eher, auf welche Aspekte sie achten müssen und haben mehr Verarbeitungskapazität, um die einzelnen Frames teilweise auch internal miteinander integrieren zu können (Schnotz & Lowe, 2008; Lowe, 1999a). Personen mit höherem Vorwissen brauchen zudem auch weniger Detailwissen über die einzelnen Teilvorgänge, da sie wahrscheinlich über solche Kenntnisse vorab schon verfügen. Deshalb wirken seriell langsame Key-Frames vielleicht eher demotivierend für sie und liefern ihnen redundante Informationen, die ihr Arbeitsgedächtnis unnötig belasten. Bei seriell schnellen Key-Frames werden hingegen mehr die größeren Zusammenhänge innerhalb des Prozesses betont (Lowe, 2006), was den Kenntnisstand von fortgeschrittenen Lernern erweitern kann. Aus diesen Gründen könnte sich der Lernvorteil durch seriell langsame Bilder bei Personen mit höherem Vorwissen verringern, und eventuell erzielen sie auch bessere Lernergebnisse mit den seriell schnellen Key-Frames. Für den Fall, dass ein größerer oder signifikanter Haupteffekt der seriell schnellen Bilder vorliegt, wäre ein weiteres Interaktionsmuster, dass dann der Lernnachteil durch seriell langsame Key-Frames bei Personen mit geringerem Vorwissen weniger stark ausgeprägt ist als bei Personen mit höherem Vorwissen.

Andererseits wäre auch die Variante möglich, dass Personen mit weniger Vorwissen mehr von den seriell schnellen als von den seriell langsamen Key-Frames profitieren im Vergleich zu Personen, die mit einem höheren Vorwissen ausgestattet sind. Als Grund für diese Annahme ist zum einen zu nennen, dass seriell schnelle Key-Frames für Personen mit geringerem Vorwissen den Vorzug haben, dass die Zustände zwischen den Bildern mehr betont werden und der Gesamtkontext besser überschaubar ist. Daher sind möglicherweise die Inferenzen auf die fehlenden Zustände zwischen den Frames leichter zu ziehen. Bei den seriell schnellen Bildern werden eher die Makrostrukturen des Prozesses betont, wodurch ein größerer Überblick über die Vorgänge erst einmal gegeben wird. Diese Darstellungsform könnte für Novizen wie ein „advance organizer“ wirken und würde dem Prinzip der „elaborativen Sequenzierung“ entsprechen, bei dem für ein besseres Verständnis die Lerninhalte auch zunächst aus einem größeren Weitwinkel heraus betrachtet werden (Ausubel, 1968; Reigeluth & Stein, 1983). Der Gesamtprozess wird zudem in der festgelegten Lernzeit bei den seriell schnellen Bildern öfters wiederholt werden als bei den seriell langsamen Bildern. Daher hat auch der unerfahrene Lerner bei den seriell schnellen

Key-Frames mehrere Möglichkeiten, den Vorgang zu verstehen und zu strukturieren. Seriell langsame Key-Frames haben für Personen mit weniger Vorwissen den Nachteil, dass durch die längere Präsentationsdauer der einzelnen Bilder die Inferenzen zwischen den Frames schwieriger zu ziehen sind und der Gesamtkontext leichter aus dem Blick gerät. Daher erzielen Personen mit geringerem Vorwissen möglicherweise bessere Lernergebnisse mit seriell schnellen als mit seriell langsamen Key-Frames. Lerner, die über ein höheres Vorwissen verfügen, können hingegen die Inferenzen auch dann gut durchführen und somit den Gesamtzusammenhang begreifen, wenn mehr die „intrastate“-Information innerhalb der einzelnen Frames betont wird, wie es bei seriell langsamen Bildern der Fall ist. Eventuell werden sie durch diese höhere Anforderung an ihre Inferierfähigkeiten mehr gefördert und können dadurch ihre Kenntnisse erweitern. Zudem könnten ebenfalls die seriell schnellen Bilder durch die höhere Zahl an Wiederholungen des Vorganges redundante Informationen für erfahrene Lerner liefern, wodurch ihr Arbeitsgedächtnis unnötig belastet wird. Daher mag es sein, dass der Lernvorteil durch seriell schnelle Bilder bei Personen mit höherem Vorwissen im Vergleich zu Personen mit einem geringeren Vorwissen absinkt oder dass die fortgeschrittenen Lerner mehr von den seriell langsamen Key-Frames profitieren. Im Falle eines größeren bis signifikanten Haupteffektes der seriell langsamen Bilder, wäre ein weiterhin mögliches Interaktionsmuster, dass der Nachteil durch seriell schnelle Key-Frames bei Personen mit einem niedrigeren Vorwissensstand kleiner ist als bei solchen Personen, die über ein höheres Vorwissen verfügen.

Aus dieser Argumentation heraus ergibt sich die folgende Grundannahme:

6. Es besteht eine Interaktion zwischen den Bedingungen „Bild seriell schnell“ und „Bild seriell langsam“ mit dem Vorwissen von Personen sowohl hinsichtlich der Lösung von Wissens- und Inferenzaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Welche Richtungen die Interaktionen jeweils annehmen, bleibt zunächst unklar, da mehrere Interaktionsmuster plausibel sind.

Grundannahmen zu den Interaktionen zwischen den Lernbedingungen und den räumlichen Fähigkeiten von Personen

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit den räumlichen Fähigkeiten von Personen

Eine weitere Annahme innerhalb der ersten Teilstudie betrifft die Interaktionen zwischen der Animations- und der seriell schnellen bzw. der seriell langsamen Bildbedingung mit den räumlichen Fähigkeiten von Personen, bei denen auch davon ausgegangen wird, dass sie ein bedeutsames Maß annehmen können. Das räumliche Vorstellungsvermögen einer Person umfasst ihre Fähigkeit, gute interne räumliche Repräsentationen zu konstruieren, aufrechtzuerhalten und zu transformieren (Hegarty & Waller, 2005). Bezüglich dieser Interaktionen werden ungerichtete Hypothesen aufgestellt, da für mehrere Interaktionsmuster unterstützende Argumente gefunden werden können.

Eine mögliche Variante ist, dass der Lernvorteil durch Key-Frames bei Personen mit geringeren räumlichen Fähigkeiten stärker ausgeprägt ist als bei Personen mit höheren räumlichen Fähigkeiten.

Statische Key-Frames, vor allem in der seriell langsamen Darstellung, könnten für Personen mit geringerem räumlichem Vorstellungsvermögen deshalb von Vorteil sein, da bei ihnen figurale Aspekte und Details des Prozesses länger und bedürfnisgerecht betrachtet werden können (Kosslyn, 1994). Die räumlichen Zustände und Zusammenhänge können mehr mit externer Unterstützung des Displays miteinander verglichen werden (Lowe, 1999a). Zudem ist es bei den Key-Frames eher möglich, das mentale Modell schrittweise aufzubauen, wodurch auch räumliche Inferenzen einzeln und nacheinander gezogen werden können (Mayer & Chandler, 2001). Bei einer Animation braucht der Lerner aufgrund ihrer inhärenten Flüchtigkeit eine gewisse räumliche Verarbeitungsgeschwindigkeit sowie eine gewisse Speicherkapazität des räumlichen Arbeitsgedächtnisses, um die Rotations- und Translationskomponenten der Bewegung verstehen zu können. Eine höhere räumliche Verarbeitungsgeschwindigkeit ist bei einer dynamischen Darstellung deshalb von Nöten, damit der Betrachter die vielen auf ihn einströmenden figuralen Details angemessen miteinander integrieren kann, um der Animation weiterhin folgen zu können. Das räumliche Arbeitsgedächtnis muss zudem ein bestimmtes Maß an Speicherkapazität haben, da darin figurale Zustände des Prozesses internal miteinander verglichen werden (Salthouse, 1996; Shah & Miyake, 1996). Daher ist es möglich, dass Personen, die über ein geringeres

räumliches Vorstellungsvermögen verfügen, besser mit statischen Key-Frames als mit einer Animation lernen.

Personen mit höherem räumlichem Vorstellungsvermögen besitzen hingegen eine genügend hohe figurale Verarbeitungsgeschwindigkeit und figurale Speicherkapazität, um eine dynamische Darstellung verstehen zu können. Diese Lerner sind in der Lage, räumliche Details auch internal miteinander zu integrieren und eventuell auch mehrere räumliche Verbindungen gleichzeitig herzustellen, wie es bei einer Animation notwendig sein kann, wodurch diese Präsentationsform angemessene Anforderungen an ihre Kompetenzen stellt. Aus diesen Argumenten wird ersichtlich, dass es sein mag, dass der Lernvorteil durch Key-Frames bei Personen mit höheren räumlichen Fähigkeiten absinkt und dass sie möglicherweise bessere Lernergebnisse mit der Animation erzielen.

Als eine weitere Variante, wie die Interaktion verlaufen könnte, wäre auch denkbar, dass der Lernvorteil durch Key-Frames bei Personen mit höherem räumlichem Vorstellungsvermögen stärker ausgeprägt ist als bei Personen, die mit niedrigeren räumlichen Fähigkeiten ausgestattet sind.

Für diese Annahme spricht, dass bei Key-Frames generell die Zustände zwischen den Bildern erschlossen werden müssen und damit müssen auch räumliche Inferenzen vom Lerner eigenständig gezogen werden (Hegarty et al., 2003). Der zu lernende Vorgang muss mental animiert werden, wobei der Prozess der mentalen Animation mit den räumlichen Fähigkeiten einer Person korreliert. Eine mentale Animation hat eher einen räumlichen Charakter, da in „dual task“-Studien gezeigt wurde, dass sie mehr durch eine zusätzliche visuelle als durch eine verbale Belastung in ihrer Durchführung behindert wird und dass sie daher wahrscheinlich Ressourcen des „visuospatial sketchpads“ mit beansprucht (Sims & Hegarty, 1997). Bei seriell langsamen Bildern sind die räumlichen Inferenzen eventuell schwieriger zu leisten, da durch die längere Präsentationsdauer der einzelnen Bilder die Abstände zwischen den Frames relativ groß werden. Für Personen mit höheren räumlichen Fähigkeiten könnte dieser teilweise erschwerte Lernvorgang eine angemessene Anforderung sein, bei der sie ihre Kompetenzen üben und erweitern können. Beim Betrachten einer Animation hingegen strömt viel räumliche Information in ihr Arbeitsgedächtnis ein, die für sie redundant ist und die sie nicht adäquat auslastet. Daraus folgt, dass für Personen mit höherem räumlichem Vorstellungsvermögen Key-Frames besser geeignet sein könnten als Animationen.

Für Lerner mit geringeren figuralen Fähigkeiten sind die räumlichen Inferenzen bei statischen Bildern schwer zu ziehen. Eine mentale Animation dürfte ebenfalls für sie schwieriger sein, da sie einen Teil ihrer Arbeitsgedächtniskapazität verbraucht, wodurch sie vom Inhalt des Lernmaterials abgelenkt werden (Sims & Hegarty, 1997). Allgemein ist ein Charakteristikum statischer Bilder, dass sie nur durch visuelle Raumindikatoren Dreidimensionalität und ganze Bewegungsabläufe widerspiegeln können, was für Personen mit geringeren räumlichen Fähigkeiten hinderlich ist (Lewalter, 1997). Eine Animation hingegen liefert vollständige Informationen über einen Prozess und unterstützt den Lerner dabei, die mentale Animation und die damit einhergehenden Rotationen und Transformationen durchzuführen. Diese „Supplantationsfunktion“ einer dynamischen Darstellung kann einen Mangel an räumlichen Fähigkeiten kompensieren (Salomon, 1979). Dadurch ist es möglich, dass der Lernvorteil durch Key-Frames bei Personen mit einem geringeren räumlichen Vorstellungsvermögen kleiner wird und dass sie bessere Resultate mit einer Animation erreichen.

Aus diesen Überlegungen heraus folgen die ungerichteten Grundannahmen:

7. Es besteht eine Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *schnell*“ mit den räumlichen Fähigkeiten von Personen sowohl hinsichtlich der Lösung von Wissens- und Inferenzaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Welche Richtungen die Interaktionen jeweils annehmen, bleibt zunächst unklar, da mehrere Interaktionsmuster plausibel sind.

8. Es besteht eine Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *langsam*“ mit den räumlichen Fähigkeiten von Personen sowohl hinsichtlich der Lösung von Wissens- und Inferenzaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Welche Richtungen die Interaktionen jeweils annehmen, bleibt zunächst unklar, da mehrere Interaktionsmuster plausibel sind.

Wie in den obigen Abschnitten werden auch hier jeweils eigene Grundannahmen für die Interaktionen mit den seriell schnellen und den seriell langsamen Bildern aufgestellt, da durch ihre verschiedenen Präsentationsgeschwindigkeiten auch Unterschiede in ihren räumlichen Lernmerkmalen entstehen können, wie sich in der Herleitung schon angedeutet hat.

Grundannahmen zu den Interaktionen zwischen den Lernbedingungen und der Merkfähigkeit von Personen

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit der Merkfähigkeit von Personen

Bei den letzten Annahmen der ersten Teilstudie wird vermutet, dass bedeutsame Interaktionen zwischen der Animations- und der seriell schnellen bzw. der seriell langsamen Bildbedingung mit der Merkfähigkeit von Personen vorhanden sind. Die Hypothesen werden dabei ungerichtet formuliert, da sowohl dynamische als auch statische Darstellungen eine höhere Merkfähigkeit beanspruchen können.

Eine erste Möglichkeit eines plausiblen Interaktionsmusters wäre, dass der Lernvorteil durch Key-Frames bei Personen mit geringerer Merkfähigkeit stärker ausgeprägt ist als bei Personen mit höherer Merkfähigkeit.

Key-Frames könnten für Personen mit geringerer Gedächtniskapazität den Vorzug haben, dass durch ihren stationären Charakter nicht so viele Reize gleichzeitig auf den Lerner einströmen und sein Arbeitsgedächtnis überlasten (Sweller, van Merriënboer & Paas, 1998). Er hat mehr Zeit, sich Detailinformationen zu betrachten und einzuprägen, und er muss die einzelnen Prozesszustände weniger internal im Arbeitsgedächtnis miteinander vergleichen, da er bei der Integration einzelner Aspekte mehr durch die externe Präsentation unterstützt wird (Lowe, 1999a). Gerade bei seriell langsamen Key-Frames kann das mentale Modell eher schrittweise aufgebaut werden, was ebenfalls entlastend für Personen mit geringerer Merkfähigkeit sein kann (Mayer & Chandler, 2001). Beim Lernen mit Animationen wird ein Vorgang vollständig präsentiert, wodurch relativ viele Informationen dicht nacheinander auf den Lerner einströmen, die er filtern muss. Auf diese Weise gelangen erst einmal auch weniger relevante Details in das Arbeitsgedächtnis des Betrachters, die redundant sind und es somit unnötig belasten. Der Vergleich einzelner Prozessabschnitte muss bei einer Animation durch ihre Flüchtigkeit eher internal und teilweise mit Hilfe von Informationen aus dem Langzeitgedächtnis durchgeführt werden (Lowe, 1999a). Daher könnte es sein, dass Personen mit geringerer Merkfähigkeit bessere Lernergebnisse mit Key-Frames als mit Animationen erzielen.

Personen, die über eine höhere Merkfähigkeit verfügen, haben hingegen genügend Gedächtniskapazität, um die Vielzahl an Eindrücken bei einer Animation bis zu einem gewissen Grad aufnehmen zu können. Sie können die einzelnen Zustände auch gut ohne

externe Unterstützung miteinander integrieren und sind in der Lage, die dafür notwendigen Informationen im Arbeitsgedächtnis zu halten. Eine dynamische Darstellung schöpft ihre Fähigkeiten somit optimal aus und kann sie auch erweitern. Daher verringert sich der Lernvorteil durch Key-Frames bei Personen mit höherer Merkfähigkeit, und möglicherweise profitieren sie mehr von einem Lernvorgang mit Animationen.

Es könnte allerdings auch der Fall sein, dass der Lernvorteil durch Key-Frames bei Personen mit höherer Merkfähigkeit größer ist als bei Personen, die mit geringeren Gedächtnisfähigkeiten ausgestattet sind.

Statische Bilder könnten deshalb vorteilhaft für Personen mit mehr Merkfähigkeit sein, da bei ihnen die Zustände zwischen den Frames inferiert und der Prozess mental animiert werden muss (Hegarty et al., 2003; Hegarty, 1992). Die mentale Animation beansprucht das „visuospatial sketchpad“ und belastet das Arbeitsgedächtnis, so dass andere Prozesse, wie das Einprägen von Details oder von verbalen Benennungen von Teilen der Abbildung, weniger gut stattfinden können (Sims & Hegarty, 1997). Bei seriell langsamen Key-Frames dürften die Inferenzen zwischen den Bildern zudem schwieriger zu ziehen sein, da die einzelnen Frames relativ lange gezeigt werden. Personen mit höherer Merkfähigkeit haben genügend Gedächtniskapazität, um diese Anforderungen zu meistern und werden dadurch gut ausgelastet. Bei einer Animation müssen die einströmenden Reize nur wahrgenommen werden (Hegarty et al., 2003), was Personen mit guten Gedächtnisfähigkeiten weniger gut fordern könnte als es die aktiven Konstruktionsvorgänge bei statischen Bildern tun. Daher ist es möglich, dass statische Key-Frames für Personen mit höherer Merkfähigkeit besser geeignet sind als Animationen.

Für Personen, die mit weniger guten Gedächtnisfähigkeiten ausgestattet sind, könnten die mentale Animation und das Inferieren der Zwischenzustände bei statischen Bildern zu anspruchsvoll sein und könnten ihr Arbeitsgedächtnis zu sehr belasten, so dass die übrigen lern- und merkrelevanten Prozesse bei ihnen nur eingeschränkt ablaufen können. Bei einer Animation bleiben diesen Lernern hingegen aktive interne Konstruktionsvorgänge stärker erspart (Rieber & Kini, 1991). Aus diesen Überlegungen folgt, dass der Lernvorteil durch Key-Frames bei Personen mit geringerer Merkfähigkeit absinken könnte und dass sie eventuell bessere Leistungen mit einer Animation erzielen.

Die ungerichteten Grundannahmen bezüglich der Interaktionen zwischen der Animations- und der seriell schnellen bzw. der seriell langsamen Bildbedingung mit der Merkfähigkeit von Personen lauten somit:

9. Es besteht eine Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *schnell*“ mit der Merkfähigkeit von Personen sowohl hinsichtlich der Lösung von Wissens- und Inferenzaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Welche Richtungen die Interaktionen jeweils annehmen, bleibt zunächst unklar, da mehrere Interaktionsmuster plausibel sind.

10. Es besteht eine Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *langsam*“ mit der Merkfähigkeit von Personen sowohl hinsichtlich der Lösung von Wissens- und Inferenzaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Welche Richtungen die Interaktionen jeweils annehmen, bleibt zunächst unklar, da mehrere Interaktionsmuster plausibel sind.

Auch bei diesen Interaktionen werden unabhängige Annahmen für seriell schnelle und seriell langsame Key-Frames aufgestellt, da bei seriell langsamen Bildern beispielsweise das Einprägen von Details erleichtert, die Inferenz der Zustände zwischen den Frames jedoch gegenüber seriell schnellen Bildern erschwert ist, so dass sich verschiedene Interaktionsmuster ergeben könnten.

3.1.2 Teilstudie 2

Grundannahmen zu den Haupteffekten zwischen den Lernbedingungen

Haupteffekte zwischen parallelen Key-Frames und seriell schnellen sowie seriell langsamen Key-Frames

Im zweiten Teil der ersten Studie geht es darum, zusätzlich zu vergleichen, wie sich verschiedene Anordnungen von Key-Frames auf die Lösung von Wissens- und Inferenzaufgaben und auf die mentale Belastung von Personen auswirken. Kontrastiert werden sollen hierbei die oben beschriebenen seriell schnellen und seriell langsamen Key-Frames mit parallel angeordneten Key-Frames. „Parallel“ oder „simultan angeordnet“ bedeutet, dass alle Bilder gleichzeitig auf dem Bildschirm erscheinen und dass jedes Bild für den Lerner jederzeit beim Betrachten des Materials verfügbar ist. Bei den seriellen Bildern kann man hingegen immer nur ein Bild zu einem Zeitpunkt sehen, und die Reihenfolge der Key-Frames wird extern vorgegeben.

Welche Präsentationsform zu besseren Lernergebnissen führt, bleibt zu untersuchen, da für beide Optionen positive und negative Argumente gefunden werden können.

Bei parallel angeordneten Bildern ist als ein großer Vorzug zu nennen, dass der Lerner hier mehr Zeit hat, ein einzelnes Bild zu betrachten, wenn er es benötigt. Alle Key-Frames stehen ihm die ganze Lernzeit über zur Verfügung, er kann selbst und bedürfnisgerecht entscheiden, welche Bilder er länger betrachten und inspizieren möchte. Durch diese höhere Selbststeuerung des Lernprozesses hat der Lerner die Möglichkeit, auch für ihn schwierige Details gut zu verstehen. Er kann zudem Vergleiche zwischen den verschiedenen Key-Frames besser anstellen, da sie alle auf dem Screen die volle Lernzeit über zu sehen sind (Boucheix & Schneider, 2009). Der Lerner ist in der Lage, ein mentales Modell des Vorganges schrittweise und nach seinen Bedürfnissen aufzubauen, er hat aber auch die Möglichkeit, den Gesamtprozess als Ganzes zu betrachten und größeren Zusammenhängen seine Aufmerksamkeit zu schenken.

Seriell angeordnete Key-Frames könnten allerdings ebenfalls von Vorteil sein, da manche der Vorzüge von parallelen Bildern unter bestimmten Umständen auch Lerndefizite verursachen können. Das hervorstechende Merkmal eines Lernprozesses mit parallelen Bildern besteht darin, dass dem Lerner die Möglichkeit, aber auch die Anforderung zur

Selbststeuerung gegeben wird. Im Gegensatz zu seriellen Bildern muss der Lerner bei parallelen Bildern selbst entscheiden, welches Frame er wie lange betrachten möchte. Dies stellt zusätzliche Anforderungen an die kognitiven und motivationalen Voraussetzungen des Lerners. Er muss seinen Lernprozess stärker selbst überwachen und sich selbst navigieren, wodurch er ein höheres Maß an Metakognition benötigt, die sein Arbeitsgedächtnis zusätzlich belastet. Er muss zudem über ein gewisses Maß an intrinsischer Motivation verfügen, um von diesem aktiv ausgerichteten Lernvorgang profitieren zu können. Durch die höhere Anstrengung beim Lernen mit parallelen Bildern könnte die Lernzeit nicht voll ausgenutzt werden, was bei seriellen Bildern wahrscheinlicher ist, da durch ihre Dynamik die Aufmerksamkeit des Lerners eher auf den Screen gelenkt wird. Zudem könnte, ähnlich wie beim Arbeiten mit Hypermedien, die Lernsequenz, die sich eine Person beim Lernen mit parallelen Bildern selbst kreiert, eine qualitativ schlechtere sein als die, die bei seriellen Bildern extern und am Lehrbuch orientiert vorgegeben wird (Schnotz, 2006). Letztlich werden gerade bei seriellen Bildern Unterschiede zwischen den einzelnen Frames besonders salient.

Insgesamt gesehen resultieren einige der Vorteile von seriellen Bildern aus möglichen Nachteilen von parallelen Bildern.

Da dennoch theoretisch nicht eindeutig geklärt werden kann, welche der beiden Präsentationsformen überlegen ist, ergeben sich folgende Grundannahmen:

1. Ein Lernvorgang mit parallelen Key-Frames führt zu unterschiedlichen Leistungen in Wissens- und Inferenzaufgaben sowie zu einem anderen Cognitive Load als ein Lernvorgang mit seriell *schnellen* Key-Frames. Welche der beiden Präsentationsformen sich als überlegen erweist, bleibt zunächst unklar, da für beide Anordnungsmöglichkeiten Vor- und Nachteile gefunden werden können.
2. Ein Lernvorgang mit parallelen Key-Frames führt zu unterschiedlichen Leistungen in Wissens- und Inferenzaufgaben sowie zu einem anderen Cognitive Load als ein Lernvorgang mit seriell *langsamen* Key-Frames. Welche der beiden Präsentationsformen sich als überlegen erweist, bleibt zunächst unklar, da für beide Anordnungsmöglichkeiten Vor- und Nachteile gefunden werden können.

Für den Vergleich zwischen parallelen und seriell schnellen bzw. seriell langsamen Key-Frames werden explizit verschiedene Grundannahmen formuliert. Zwar ist die Argumentation für die Ungerichtetheit der Annahmen bei beiden Vergleichen als analog zu sehen, es könnten sich aber durch die unterschiedlichen Präsentationsgeschwindigkeiten der seriellen Bilder verschieden große Effekte ergeben, die man differentiell betrachten sollte. Seriell langsame Key-Frames ähneln durch die längere Dauer der einzelnen Bilder in manchen Aspekten mehr den parallelen Bildern, beispielsweise kann bei beiden Präsentationsformen ein mentales Modell schrittweise aufgebaut werden (Mayer & Chandler, 2001), so dass sich ein unterschiedlicher Kontrast im Vergleich zu dem mit seriell schnellen Bildern ergeben könnte.

Grundannahmen zu den Interaktionen zwischen den Lernbedingungen und dem Vorwissen von Personen

Interaktionen zwischen den Bedingungen „Key-Frames parallel“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit dem Vorwissen von Personen

Es wird in der zweiten Teilstudie der ersten Studie davon ausgegangen, dass signifikante Wechselwirkungen zwischen den Bedingungen „Bild parallel“ und „Bild seriell schnell“ bzw. „Bild seriell langsam“ mit dem Vorwissen von Personen bestehen. Die Grundannahmen hierzu werden ungerichtet formuliert werden, da mehrere Interaktionsmuster theoretisch sinnvoll sind. Die Haupteffektshypothesen zwischen den Bedingungen wiesen ebenfalls keine eindeutige Richtung auf.

Unabhängig davon, ob sich die parallele oder die seriellen Bildbedingungen generell als überlegen erweisen werden, wäre ein mögliches Interaktionsmuster, dass Personen mit einem geringeren Vorwissen mehr von parallelen als von sequentiellen Key-Frames profitieren im Vergleich zu Personen, die über ein höheres Vorwissen verfügen. Die Gründe hierfür liegen darin, dass parallele Bilder für Novizen von daher besonders gut geeignet sind, da sie bei dieser Darstellungsform die einzelnen Zustände bedürfnisgerecht und in Ruhe inspizieren und ihr Augenmerk auf die Details richten können, die sie nicht verstanden haben (Kosslyn, 1994). Die einzelnen Prozessschritte können external miteinander verglichen werden und müssen nicht im Arbeitsgedächtnis gespeichert werden. Das mentale Modell des Vorganges kann schrittweise aufgebaut werden, was vor allem für Personen ohne größeres

Vorwissen hilfreich ist (Mayer & Chandler, 2001). Bei den seriellen Key-Frames hingegen ist das individuell gesteuerte Betrachten der einzelnen Bilder nicht möglich. Diese Präsentationsform ist flüchtiger und kann daher eher das Arbeitsgedächtnis überlasten. Novizen richten ihre Aufmerksamkeit eventuell auf Aspekte, die konzeptuell nicht relevant sind (Lowe, 1999a). Zumindest bei der seriell schnellen Variante ist es schwieriger, die Teilprozesse zu verstehen und das gesamte mentale Modell darauf aufbauend zu konstruieren. Daher ist es möglich, dass Personen mit einem geringeren Vorwissen mit parallelen Bildern bessere Lernergebnisse erzielen als mit sequentiellen Bildern. Personen, die über ein höheres Vorwissen verfügen, können mit der Flüchtigkeit der seriellen Key-Frames besser umgehen, es ist für sie weniger ein Problem, Prozesszustände teilweise im Arbeitsgedächtnis miteinander zu integrieren (Lowe 1999a). Bei seriell schnellen Key-Frames werden besonders die größeren Zusammenhänge innerhalb des Gesamtprozesses betont (Lowe, 2006), was eher den Kenntnisstand von Lernern mit höherem Vorwissen erweitert. Zudem könnten die parallelen Key-Frames für sie den Nachteil haben, dass sie die ständige Verfügbarkeit der Bilder nicht brauchen und daher die Möglichkeit des selbstgesteuerten Lernens für sie redundant ist und eher eine Form von „extraneous load“ darstellt (Sweller, 1988). Fortgeschrittene Lerner könnten auch eher die Tendenz haben, ihren Wissensstand zu überschätzen und die Lernzeit bei parallelen Key-Frames nicht vollständig zu nutzen, was bei seriellen Bildern eher weniger vorkommen dürfte, da die Aufmerksamkeit des Betrachters durch die Flüchtigkeit der Vorgänge mehr auf den Bildschirm gelenkt wird.

Daher könnte bei Personen mit höherem Vorwissen der Lernvorteil durch parallele Bilder geringer ausgeprägt sein als bei Personen mit niedrigerem Vorwissen, und eventuell lernen sie besser mit den seriellen Key-Frames. Eine weitere Möglichkeit dieses Interaktionsmusters könnte auch darin bestehen, dass bei einem größeren bis signifikanten Haupteffekt der seriell schnellen bzw. der seriell langsamen Bilder der Lernnachteil durch parallele Key-Frames bei Personen mit geringerem Vorwissen kleiner ist als bei Personen mit höherem Vorwissen.

Auf der anderen Seite wäre auch das entgegengesetzte Interaktionsmuster plausibel: Personen, die mit einem geringeren Vorwissen ausgestattet sind, profitieren mehr von den seriellen Key-Frames als von den parallelen Key-Frames im Vergleich zu Personen, die ein größeres Vorwissen haben. Dies könnte darin begründet sein, dass serielle Bilder für Novizen den Vorzug haben, dass ihnen eine gute Sequenz der einzelnen Bilder extern

vorgegeben wird. Bei den parallelen Key-Frames wählt sich der Betrachter seine Lernreihenfolge der Zustände selbst aus, was eventuell zu einem qualitativ schlechteren Vorgehen führt. Bei den seriellen Key-Frames werden die Veränderungen zwischen den Bildern perzeptuell besonders salient, und zumindest bei den seriell langsamen Key-Frames ist es zudem eher möglich, das mentale Modell wie bei den parallelen Key-Frames schrittweise aufzubauen (Mayer & Chandler, 2001). Der Lernvorgang mit simultanen Bildern ist aktiver und selbstgesteuerter als der mit seriellen Key-Frames, was aber auch höhere Anforderungen an die Metakognition und an das Arbeitsgedächtnis des Betrachters stellt. Gerade Personen mit geringerem Vorwissen könnten kognitiv und motivational mit einem solchen, in gewisser Hinsicht anspruchsvolleren, Lernprozess überfordert sein. Daher könnte es der Fall sein, dass Lerner mit geringerem Vorwissen eine bessere Leistung mit seriellen als mit parallelen Key-Frames erzielen. Personen, die ein höheres Vorwissen haben, können hingegen mit einem mehr selbstgesteuerten Lernvorgang gut umgehen, da sie inhaltlich mit dem Lernstoff nicht so stark gefordert sind. Dadurch werden sie durch die zusätzlichen kognitiven Prozesse bei den parallelen Bildern nicht zu sehr belastet, ihre Arbeitsgedächtniskapazität wird eventuell besser ausgeschöpft. Personen mit mehr Vorwissen haben bei parallelen Key-Frames die Möglichkeit, sich gezielt die Details und Zusammenhänge anzuschauen, die ihren Kenntnisstand erweitern. Daher sinkt bei höherem Vorwissen der Lernvorteil durch serielle Bilder, und vielleicht erzielen fortgeschrittene Lerner mit parallelen Key-Frames auch bessere Ergebnisse. Sollte bei der Studie ein größerer oder signifikanter Haupteffekt der parallelen Key-Frames vorliegen, könnte es in dem Fall sein, dass der Lernnachteil durch serielle Bilder bei Personen mit geringerem Vorwissen nicht so groß ist wie bei Personen mit höherem Vorwissen.

Aus diesen Überlegungen heraus wird deutlich, dass mehrere Interaktionsmuster plausibel sind, so dass ungerichtete Grundannahmen aufgestellt werden:

3. Es besteht eine Interaktion zwischen den Bedingungen „Bild parallel“ und „Bild seriell *schnell*“ mit dem Vorwissen von Personen sowohl hinsichtlich der Lösung von Wissens- und Inferenzaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Welche Richtungen die Interaktionen jeweils annehmen, bleibt zunächst unklar, da mehrere Interaktionsmuster plausibel sind.

4. Es besteht eine Interaktion zwischen den Bedingungen „Bild parallel“ und „Bild seriell langsam“ mit dem Vorwissen von Personen sowohl hinsichtlich der Lösung von Wissens- und Inferenzaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Welche Richtungen die Interaktionen jeweils annehmen, bleibt zunächst unklar, da mehrere Interaktionsmuster plausibel sind.

Für den Kontrast zwischen parallelen und seriell schnellen sowie für den zwischen parallelen und seriell langsamen Key-Frames wird von jeweils unabhängigen Grundannahmen ausgegangen. Die Argumentation für die verschiedenen Interaktionsmuster ist zwar größtenteils analog, aber wie sich auch in manchen Aspekten der Herleitung andeutete, könnte es zu unterschiedlichen Effekten kommen, da in den seriellen Bildbedingungen durch ihre verschiedenen Präsentationsgeschwindigkeiten andere Informationen jeweils stärker betont werden (Lowe, 2006). Bei der seriell langsamen Variante liegt der Schwerpunkt auf den einzelnen Teilprozessen, während bei der seriell schnellen Form mehr die Zusammenhänge zwischen den Zuständen hervorgehoben werden.

3.2 Theoretische Herleitung der Grundannahmen von Studie 2: Vergleich von Animationen und statischen Bildern hinsichtlich ihrer Wirksamkeit auf die Lösung von bildlichen Kategorisierungsaufgaben

Grundannahmen zu den Haupteffekten zwischen den Lernbedingungen

Haupteffekte zwischen Animationen und parallelen Key-Frames

In der zweiten Studie der vorliegenden Dissertation geht es darum, Animationen und statische Bilder hinsichtlich ihrer Wirksamkeit auf die Lösung bildlicher Kategorisierungsaufgaben zu untersuchen. Bildliche Kategorisierungsaufgaben beinhalten die Anforderung, unterschiedliche Bewegungen anhand von bildlichem Material, das sowohl animiert als auch statisch sein kann, erkennen und voneinander unterscheiden zu können. Im vorliegenden Fall wird das Thema des Lernmaterials „Die Gangarten des Pferdes“ sein, weshalb vier Gangarten erlernt und voneinander diskriminiert werden müssen (s. auch Abschnitt 5.1.2.3). In der Lernphase werden die Bilder parallel bzw. simultan auf dem Bildschirm angeordnet sein. Die Annahme ist, dass ein Lernvorgang mit Animationen zu besseren Leistungen in bildlichen Kategorisierungsaufgaben führt als ein Lernprozess, der mit statischen Bildern abläuft. Bei den statischen Bildern handelt es sich wiederum um Key-Frames, denen die gleichen theoretischen Überlegungen wie in der ersten Studie zugrunde liegen. Key-Frames sind zum einen mittlere, prototypische Bilder der Unterereignisse eines Vorganges (Schnotz & Lowe, 2008), und zum anderen liegen sie an kausalen Schnittstellen, die in der aktuellen Studie mit biologischen Bewegungen eher temporal zu sehen sind und sich an den Anfangs- und Endpunkten der Teilereignisse befinden.

Die theoretische Begründung für die Annahme, dass sich Animationen für das spätere Lösen von Kategorisierungsaufgaben besser eignen als statische Bilder liegt darin, dass in der Umwelt des Menschen häufig biologische Bewegungen auftreten, wie das Gehen eines Menschen oder das Laufen eines Hundes, die sich durch invariante und sich ständig wiederholende räumlich-zeitliche Bewegungsmuster auszeichnen. Der Mensch ist evolutionär dazu ausgestattet, diese Bewegungsmuster zu erkennen und in perzeptuellen dynamischen Schemata zu enkodieren. Diese Schemata repräsentieren typische, oft erlebte zeitlich-räumliche Muster, die uns dabei helfen, dynamische Information effizient zu

verarbeiten. Wird eine natürliche Bewegung im Umfeld wahrgenommen, wird das entsprechende dynamische Schema aktiviert, das dann zur Erkennung der Bewegung beiträgt. Die Erkennung funktioniert umso besser und schneller bei Bewegungen, mit denen man häufig konfrontiert ist, da das entsprechende dynamische Schema dann elaboriert und leicht abrufbar ist. Dynamische Schemata helfen durch ihren Informationsgehalt über zeitlich-räumliche Strukturen ebenfalls dabei, Bewegungen zu antizipieren oder Anomalien einer Bewegung (z.B. das Hinken als Anomalie des Gehens) zu erfassen (Schnotz & Lowe, 2008).

Generell läuft das Erkennen von biologischen Bewegungen recht schnell ab, und die einzelnen Bewegungen können auch fein voneinander diskriminiert werden (Troje, 2002; Wade & Swanston, 2001). So wird menschliche Bewegung innerhalb von 0,1 s erkannt, wie sich in den „Lichtpunktläufer-Experimenten“ („point-light walker stimulus“) von Johansson (1975) zeigte, in denen an eine Person vor einer dunklen Wand Lichter befestigt wurden, so dass von der Person sonst nichts zu sehen war. Erst als das Modell sich in Bewegung setzte, wurde rasch von den Betrachtern erkannt, dass es sich um einen gehenden Menschen handelte. Das Geschlecht oder die Emotionen können ebenfalls aus dem Gang einer Person heraus erschlossen werden, wobei bei der Analyse des Geschlechts zwei Gangzyklen zu seiner Erkennung ausreichen (Barclay, Cutting & Kozlowski, 1978; Dittrich, Troscianko, Lea & Morgan, 1996). Diese Befunde legen die Vermutung nahe, dass es für die Verarbeitung biologischer Bewegungen ein eigenes Hirnareal gibt. Grossman und Blake (2001) lieferten Befunde dafür, dass der Sulcus temporalis superior bei der Wahrnehmung biologischer Bewegung stärker aktiv ist als bei der Perzeption einer durcheinandergewürfelten Bewegung, was die Idee eines eigenen Hirnareals stützt.

Animationen stellen nun biologische Bewegungen dynamisch und in der Form dar, wie sie im natürlichen Umfeld auftreten. Daher können das Erkennen der Bewegungsmuster und der Aufbau der dynamischen Schemata analog zu den entsprechenden Prozessen im Alltag und damit sehr schnell und evolutionär prädisponiert ablaufen. Die Bewegungsmuster neuer Gangarten können daher leicht erlernt werden, wobei der Lernprozess eher implizit und auf perzeptueller Ebene abläuft. Bei Animationen werden vor allem die Prozessinformationen einer Bewegung betont. Die Hirnareale, die bei der Wahrnehmung tatsächlicher Bewegung aktiv sind, sprechen teilweise auch beim Betrachten stationärer Bilder an, die Bewegung implizieren (Senior, Barnes, Giampietro, Simmons, Bulhore, Brammer & David, 2000), weshalb eventuell auch für einen statischen Lernvorgang eine gewisse Prädisposition

vorliegen mag. Dennoch ist das Lernen biologischer Bewegungen mit statischen Bildern unnatürlicher als mit Animationen. Die Zustände zwischen den einzelnen Frames müssen inferiert, und der Prozess muss mental animiert werden (Hegarty et al., 2003; Hegarty, 1992). Dies belastet das menschliche Arbeitsgedächtnis. Im Vergleich zu einem Lernprozess mit Animationen geht das Lernen mit statischen Bildern eher explizit und mit größerer Beteiligung der Kognition vonstatten. Strukturelle Informationen über die jeweilige Bewegung treten in den Vordergrund. Zudem liegen die typischen Nachteile von Animationen nur in verminderter Form vor, wenn diese biologische Bewegungen präsentieren: Da natürliche Vorgänge repetitiv sind, sich bestimmte Bewegungszyklen also ständig wiederholen, hat der Lerner mehrere Möglichkeiten, das entsprechende Bewegungsmuster zu extrahieren. Durch die permanenten Wiederholungen hat die Animation einen weniger flüchtigen Charakter (Lowe, 1999a), und das Arbeitsgedächtnis des Lerners wird nicht mit immer neuen Eindrücken überlastet (Paas et al., 2007), was ansonsten ein gewichtiger Nachteil dieser Darstellungsform ist. Da weniger Reize auf dem Display vorhanden sind als bei einem linearen Vorgang, der eventuell mehrere Komponenten mit einbezieht, ist die Wahrscheinlichkeit geringer, die Aufmerksamkeit auf konzeptuell irrelevante Inhalte zu lenken (Lowe, 1999a).

Dies könnten weitere Gründe dafür sein, weshalb bei dieser Fragestellung Animationen gegenüber statischen Bildern im Vorteil sein könnten, was zu der folgenden Grundannahme führt:

1. Durch das Lernen mit Animationen werden sowohl animierte als auch statische bildliche Kategorisierungsaufgaben besser gelöst als wenn mit parallel angeordneten statischen Key-Frames gelernt wird. Zudem werden der Lernprozess selbst und die Bearbeitung des Nachtests bei Animationen als weniger anstrengend erlebt als bei Key-Frames.

Auch in dieser Studie wird der Cognitive Load, der nach dem Lernen und nach dem Nachtest gemessen wurde, jeweils herangezogen, um die „instruktionale Effizienz“ (Paas et al., 2007) der Lernbedingungen zu beurteilen. Ein solches Kriterium stellt eine gute Ergänzung zu den Leistungsmaßen dar, um herauszufinden, mit welchen „kognitiven Kosten“ eine Leistung jeweils erbracht wurde.

Grundannahme zu der Interaktion zwischen dem Präsentationsformat der Lernbedingung und dem Präsentationsformat der Testitems

Wie oben bereits angedeutet, werden die bildlichen Kategorisierungsaufgaben den Teilnehmern der Studie sowohl in animierter als auch in statischer Form dargeboten werden. Daraus erwächst die zusätzliche Fragestellung, ob der vermutete Lernvorteil durch Animationen auf einer Skala, die nur animierte Items enthält, größer ist als auf einer Skala, die ausschließlich statische Aufgaben beinhaltet. Eine Begründung für diese Annahme liegt darin, dass bei Verwendung des gleichen Präsentationsformates in der Lern- und Testphase Transfereffekte entstehen können, da in diesem Fall das enkodierte Material dem Testmaterial entspricht. In der Lernphase werden sich bestimmte strukturelle Aspekte der Darstellung und des Displays eingeprägt sowie bestimmte kognitive Prozeduren, die zum Verstehen des Materials nötig sind, eingeübt, so dass, wenn Items mit dem gleichen Präsentationsformat bearbeitet werden sollen, Gedächtniseffekte und die größere Routine mit den erforderlichen mentalen Operationen, die Lösung der Aufgaben begünstigen. Soll man beispielsweise statische Bilditems bearbeiten, bei denen die Inferenz von Zwischenzuständen nötig ist, kann es helfen, wenn man zuvor mit statischen Bildern gelernt und diese Inferenzen schon eingeübt hat. Zudem konnte man sich einige strukturelle Aspekte des Displays beim Lernen einprägen. Hat man hingegen mit Animationen gelernt, bei denen man eher wahrnehmen musste und bei denen das Display anders gestaltet war, kann man bei der Lösung von statischen Aufgaben weniger auf Gedächtnisinhalte zurückgreifen, und man muss auch erst mit dem Inferieren und dem mentalen Animieren vertraut werden (Hegarty et al., 2003; Hegarty, 1992). Auch in der Studie von Goldstein, Chance, Hoisington und Buescher (1982) zeigte sich, dass Personen, die mit Filmen, also einer dynamischen Darstellung, gelernt hatten, darin enthaltene dynamische Videoausschnitte besser wiedererkennen konnten als statische Bildausschnitte, obwohl der Leistungsunterschied hierbei nicht groß war.

Aufgrund dieser Überlegungen ergibt sich als eine weitere Grundannahme:

2. Es besteht eine Interaktion zwischen dem Präsentationsformat des Lernmaterials und dem der Nachtestaufgaben, in der Richtung, dass der Lernvorteil durch Animationen bei animierten Items größer ist als bei statischen Bilditems.

Grundannahmen zu den Interaktionen zwischen den Lernbedingungen und der Wahrnehmungsvorerfahrung von Personen

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames parallel“ mit der Wahrnehmungsvorerfahrung von Personen

In der vorliegenden Studie ist es sinnvoll, davon auszugehen, dass signifikante Interaktionen zwischen den beiden Bedingungen „Animation“ und „Bild parallel“ mit der Wahrnehmungsvorerfahrung von Personen vorhanden sein werden. Diese Vermutung hängt damit zusammen, dass je mehr Erfahrung eine Person mit den Bewegungen von Pferden hat, ihre dynamischen Schemata der entsprechenden Gangarten schon elaborierter und leichter abrufbar sind. Dies wiederum könnte die Qualität ihres Lernprozesses sowie die Bearbeitung des Nachtests beeinflussen, so dass Teilnehmer mit unterschiedlich hoher Wahrnehmungsvorerfahrung eventuell mit den beiden Präsentationsformen in verschiedenem Ausmaß gut klarkommen.

Dabei liegt es näher, zu vermuten, dass der Lernvorteil durch Animationen bei Personen mit geringerer Vorerfahrung stärker ausgeprägt ist als bei Personen mit höherer Vorerfahrung. Für diese Vermutung sprechen mehrere theoretische Gründe, wobei die Argumentation einer solchen ähnelt, wie sie auch mit der Variable „Vorwissen“ verlaufen würde, da sowohl beim Vorwissen einer Person wie auch bei ihrer Wahrnehmungsvorerfahrung an schon bestehende themenspezifische Inhalte und Strukturen angeknüpft wird, was den Lernvorgang in ähnlicher Weise mit bestimmen kann.

Wie oben bereits erwähnt, laufen das Erlernen biologischer Bewegungen sowie der Aufbau von dynamischen Schemata mit Hilfe von Animationen eher natürlich und analog zu den Wahrnehmungsprozessen im Alltag ab. Der Lernvorgang geschieht mehr auf perzeptueller Ebene, ist automatisiert und genetisch prädisponiert (Schnotz & Lowe, 2008; Wade & Swanston, 2001). Daher ist das Lernen mit Animationen in diesem Fall wahrscheinlich auch für Novizen geeignet, weshalb sie mit dieser Darstellungsform vergleichsweise gute Ergebnisse erzielen können. Bei statischen Bildern hingegen sind mehr kognitive Fähigkeiten von Nöten, das Lernen ist explizit, und die Zustände zwischen den Bildern müssen inferiert werden, was immer die Gefahr von Fehlinferenzen in sich birgt (Narayanan & Hegarty, 1998). Zudem müssen die Vorgänge mental animiert werden (Hegarty, 1992). Damit diese internen Konstruktionsprozesse möglichst fehlerfrei ablaufen und damit die „Lücken“ zwischen den Frames korrekt gefüllt werden, können

Wahrnehmungsvorerfahrungen mit den Pferdegangarten und zumindest rudimentär vorhandene dynamische Schemata sehr hilfreich sein. Bei Personen mit geringer Wahrnehmungsvorerfahrung besteht somit eher das Risiko, dass ihre Fähigkeiten nicht ausreichen, um die inneren Konstruktionsprozesse durchzuführen und dass von ihnen Fehlinferenzen gezogen werden (Narayanan & Hegarty, 1998). Daher erzielen sie wahrscheinlich mit statischen Bildern im Vergleich zu Animationen relativ schlechte Lernergebnisse, obwohl es sich in dem vorliegenden Experiment um sehr informationsreiche Key-Frames handeln wird, bei denen die Inferenzen etwas erleichtert sind (Schnotz & Lowe, 2008).

Personen mit höherer Vorerfahrung hingegen können sich die Zustände zwischen den Bildern aufgrund ihrer teilweise schon vorhandenen dynamischen Schemata relativ gut erschließen und den Vorgang somit auch mental animieren. Die inneren Konstruktionsprozesse stellen für sie kein übermäßig großes Problem dar, weshalb sie mit statischen Key-Frames relativ gut zurecht kommen, so dass der Lernvorteil durch die Animationen bei ihnen nicht so stark ausgeprägt ist wie bei den Personen mit geringerer Wahrnehmungsvorerfahrung. Eventuell sind erfahrenere Lerner mit Animationen sogar im Nachteil, da sie vielleicht durch die statischen Bilder optimal gefordert werden, bei denen sie durch gezielte Inspektionen von Details ihren Kenntnisstand besser erweitern können.

Wie ebenfalls bereits erwähnt, haben Animationen bei biologischen Bewegungen weniger die typischen Nachteile, die sie sonst für Novizen haben. Die Gangzyklen der Pferde werden in der Darstellung oft wiederholt, es handelt sich hierbei also um repetitive Vorgänge, weshalb die Animation einen weniger flüchtigen Charakter hat und das Arbeitsgedächtnis des Lerners nicht so stark überflutet wird (Lowe, 1999a; Paas et al., 2007). Der Betrachter hat mehrere Gelegenheiten, das Bewegungsmuster zu verstehen und zu strukturieren, wodurch auch Novizen beim Lernen mit dieser dynamischen Darstellung weniger stark überfordert sind. Da beim Design der Animationen nichts konzeptuell Irrelevantes betont werden wird und die Pferdegangarten nicht so vielschichtig und komplex sind wie einige lineare Vorgänge, haben auch Personen mit geringerer Vorerfahrung weniger das Risiko, Details auf dem Display zu beachten, die perzeptuell auffällig, aber inhaltlich eher unwesentlich sind (Lowe, 1999a).

Zusammengefasst bedeutet dies, dass in der Studie die Nachteile von statischen Bildern für Novizen durchaus gegeben sind, während die Defizite von Animationen nur in abgemilderter Form vorhanden sind, weshalb Personen mit geringerer Wahrnehmungsvorerfahrung hier

vermutlich besser mit Animationen als mit statischen Bildern lernen. Personen mit höherer Vorerfahrung können besser mit den Schwierigkeiten von statischen Key-Frames umgehen, weshalb sich bei ihnen die Unterschiede zwischen den beiden Bedingungen eher ausgleichen oder sie eventuell auch bessere Ergebnisse mit den Key-Frames erzielen könnten.

Aus diesen Überlegungen folgt als Grundannahme:

3. Es besteht eine Interaktion zwischen den Bedingungen „Animation“ und „Bild parallel“ mit der Wahrnehmungsvorerfahrung von Personen sowohl hinsichtlich der Lösung von animierten und statischen bildlichen Kategorisierungsaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Die Interaktionen verlaufen jeweils in der Richtung, dass der Lernvorteil durch Animationen bei Personen mit geringerer Vorerfahrung größer ist als bei Personen mit höherer Vorerfahrung.

Grundannahmen zu den Interaktionen zwischen den Lernbedingungen und den räumlichen Fähigkeiten von Personen

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames parallel“ mit den räumlichen Fähigkeiten von Personen

Es wird in der aktuellen Studie ebenfalls davon ausgegangen, dass bedeutsame Wechselwirkungen zwischen den beiden Lernbedingungen „Animation“ und „Bild parallel“ mit den räumlichen Fähigkeiten von Personen bestehen. Das räumliche Vorstellungsvermögen beinhaltet die Fähigkeit, gute interne räumliche Repräsentationen zu konstruieren, aufrechtzuerhalten und zu transformieren (Hegarty & Waller, 2005). Dieses Vermögen wird auch benötigt, um dynamische Schemata zu erstellen, da in ihnen räumliche Informationen über Bewegungen enthalten sind (Schnotz & Lowe, 2008).

Als gerichtete Vermutung wird hierbei formuliert, dass der Lernvorteil durch Animationen bei Personen mit geringeren räumlichen Fähigkeiten größer ist als bei Personen mit höheren räumlichen Fähigkeiten.

Ein Grund für diese Annahme liegt darin, dass das Erlernen biologischer Bewegungen mit statischen Key-Frames eher explizit und unter größerer Beteiligung der Kognition abläuft. Die Zustände zwischen den einzelnen Bildern müssen inferiert, und die Bewegungsprozesse

müssen mental animiert werden (Hegarty et al., 2003; Hegarty, 1992). Für eine mentale Animation werden räumliche Fähigkeiten benötigt, was durch hohe Interkorrelationen zwischen diesen Variablen belegt wird. Zudem wurde in „dual-task“-Studien festgestellt, dass ein visueller Gedächtnisload mehr mit einer „mentalen Animation“ interferierte als ein verbaler Gedächtnisload. Dies waren Hinweise dafür, dass eine mentale Animation abhängig vom „visuospatial sketchpad“ ist und einen mehr räumlichen Charakter hat (Sims & Hegarty, 1997).

Die Inferenzen bei den Bewegungen der Pferdegangarten dürften wahrscheinlich gerade in räumlicher Hinsicht etwas schwieriger sein, da sich bei mehreren Vorgängen die vier Beine des Pferdes jeweils in ihrer Stellung zueinander und in ihrem Zusammenspiel im Raum vorgestellt werden müssen. Generell ist es bei statischen Bildern so, dass sie nur visuelle Raumindikatoren liefern können, um Dreidimensionalität und ganze Bewegungsabläufe darzustellen (Lewalter, 1997). Daher werden beim Lernen mit statischen parallelen Key-Frames räumliche Fähigkeiten benötigt, um adäquate dynamische Schemata der Pferdegangarten zu erstellen.

Es kann aus diesen Gründen vermutet werden, dass Personen mit geringerem räumlichem Vorstellungsvermögen mit statischen Bildern schlechtere Lernergebnisse erzielen als mit Animationen, bei denen der Lernvorgang in räumlicher Hinsicht einfacher ist, wie weiter unten erklärt werden wird. Personen mit höheren figuralen Fähigkeiten haben hingegen wahrscheinlich die Kompetenz, richtige Inferenzen zu ziehen, weshalb ihre Leistung mit statischen Key-Frames sich mehr der mit Animationen annähert. Eventuell kommen sie mit den Bildern sogar besser zurecht, da sie vielleicht durch diese Darstellungsform optimal gefordert werden und da somit eine bessere Passung zwischen der Aufgabe und ihren Fähigkeiten herrscht, weil ihnen durch statische Bilder keine redundanten räumlichen Informationen vermittelt werden.

In der vorliegenden Studie werden beim Lernen mit Animationen wahrscheinlich weniger räumliche Fähigkeiten benötigt. Generell spielt das räumliche Vorstellungsvermögen auch beim Betrachten einer dynamischen Darstellung eine Rolle, da Animationen einen flüchtigen Charakter haben (Lowe, 1999a). Daher braucht man eine gewisse räumliche Verarbeitungsgeschwindigkeit, und das räumliche Arbeitsgedächtnis muss eine gewisse Kapazität haben, um die Rotations- und Translationskomponenten der Bewegung auf dem Display zu verstehen. Die räumlichen Komponenten der Darstellung müssen zudem teilweise im Arbeitsgedächtnis gespeichert und dort miteinander abgeglichen und integriert

werden (Salthouse, 1996; Shah & Miyake, 1996). Im aktuellen Fall ist die Animation jedoch repetitiv, d.h., die Bewegungsmuster werden oft wiederholt, so dass der Lerner die figuralen Details nicht so schnell verarbeiten und im Gedächtnis speichern muss, da er sie in der externen Präsentation immer wieder betrachten kann. Aus diesen Gründen sind im vorliegenden Fall räumliche Fähigkeiten beim Lernen mit Animationen wahrscheinlich weniger wesentlich, zumal der Aufbau dynamischer Schemata von biologischen Bewegungen aus dynamischen Darstellungen heraus eher implizit ist und auf perzeptueller Ebene abläuft. Zudem könnte die Animation eine „Supplantationsfunktion“ für Personen mit einem geringen figuralen Vorstellungsvermögen haben, da die Rotations- und Transformationsprozesse extern präsentiert werden, was ihnen dabei helfen kann, diese Vorgänge selbst nachzuvollziehen (Salomon, 1979).

Aus dieser Argumentation heraus resultiert die Annahme, dass Personen, die über geringere räumliche Fähigkeiten verfügen, bessere Leistungen mit Animationen als mit parallelen Key-Frames erzielen, während Personen mit höheren räumlichen Fähigkeiten auch gut mit statischen Bildern lernen können, so dass bei ihnen der Lernvorteil durch Animationen nicht so stark ausgeprägt ist oder so dass sie mit Animationen sogar im Nachteil sind.

Somit lautet die Grundannahme zu den Interaktionen zwischen der Lernbedingung und den räumlichen Fähigkeiten von Personen:

4. Es besteht eine Interaktion zwischen den Bedingungen „Animation“ und „Bild parallel“ mit den räumlichen Fähigkeiten von Personen sowohl hinsichtlich der Lösung von animierten und statischen bildlichen Kategorisierungsaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Die Interaktionen verlaufen jeweils in der Richtung, dass der Lernvorteil durch Animationen bei Personen mit geringeren räumlichen Fähigkeiten größer ist als bei Personen mit höheren räumlichen Fähigkeiten.

Grundannahmen zu den Interaktionen zwischen den Lernbedingungen und der Merkfähigkeit von Personen

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames parallel“ mit der Merkfähigkeit von Personen

Zuletzt wird noch davon ausgegangen, dass signifikante Wechselwirkungen zwischen der Animations- und der parallelen Bildgruppe mit der Merkfähigkeit von Personen vorliegen.

Auch hier liegt es näher, anzunehmen, dass der Lernvorteil durch Animationen bei Personen mit geringerer Merkfähigkeit größer ist als bei Personen mit höherer Merkfähigkeit. Die Argumentation verläuft hierbei ähnlich wie bei den vorigen Interaktionen.

Der Einfluss der Merkfähigkeit beim Lernen mit Animationen von biologischen Bewegungen ist insgesamt wahrscheinlich geringer ausgeprägt, was mit an dem perzeptuellen und impliziten Charakter des Lernvorganges liegen könnte. Der Betrachter hat eine genetische Prädisposition, die zeitlich-räumlichen Muster einer Bewegung aus dynamischen Darstellungen, und somit auch aus Animationen, während ihrer Wahrnehmung zu extrahieren und sie als perzeptuelle dynamische Schemata zu enkodieren (Schnotz & Lowe, 2008). Der Lernprozess ist eher automatisiert und belastet das Arbeitsgedächtnis in geringerem Ausmaß, so dass Animationen auch für Personen mit einer niedrigeren Merkfähigkeit geeignet sind, um gute Lernergebnisse erzielen zu können.

Bei statischen Key-Frames hingegen ist es neben dem reinen Betrachten des Lernmaterials notwendig, die Inferenzen auf die Zustände zwischen den Bildern zu ziehen und eine mentale Animation durchzuführen (Hegarty et al., 2003; Hegarty, 1992), was Ressourcen des Arbeitsgedächtnisses bzw. des „visuospatial sketchpads“ beansprucht (Sims & Hegarty, 1997). Der interne Konstruktionsprozess der dynamischen Schemata ist aufwendiger, weshalb weniger Gedächtniskapazität für sonstige merk- und lernrelevante Prozesse zur Verfügung steht. Von Personen mit geringerer Merkfähigkeit kann es somit als anstrengender erlebt werden, mit parallelen Bildern zu lernen, was sich auch in schlechteren Leistungen äußern kann, obwohl in der Studie Key-Frames verwendet werden, bei denen die Inferenzen etwas erleichtert sein könnten (Schnotz & Lowe, 2008). Daher erzielen mit niedrigerer Gedächtnisfähigkeit ausgestattete Personen wahrscheinlich bessere Ergebnisse mit Animationen als mit statischen Bildern. Personen mit höherer Merkfähigkeit haben hingegen genügend Gedächtniskapazität, um die internen Konstruktionsprozesse zu vollziehen, so dass bei ihnen der Unterschied zwischen den Bedingungen kleiner wird.

Eventuell profitieren sie sogar durch die bessere Auslastung ihres Arbeitsgedächtnisses mehr von den parallelen Key-Frames.

Statische Bilder haben zwar durch ihren stationären Charakter für den Lerner den Vorteil, dass sein Arbeitsgedächtnis nicht durch immer neue Eindrücke überflutet wird (Paas et al., 2007) und dass Vergleiche zwischen den einzelnen Zuständen externalisiert durchgeführt werden können (Lowe, 1999a). Allerdings verfügen Animationen im vorliegenden Fall größtenteils auch über diese Vorzüge. Die Pferdegangarten sind repetitive Bewegungen, d.h., das gleiche dynamische Muster wird immer wieder gezeigt, und die einzelnen Komponenten auf dem Display unterliegen keinen so großen Veränderungen. Das bedeutet, dass das Arbeitsgedächtnis nicht so stark mit immer neuen Reizen belastet wird, was im Normalfall ein gewichtiger Nachteil von Animationen ist (Sweller et al., 1998; Paas et al., 2007). Da die Gangzyklen mehrfach wiederholt werden, kann der Vergleich und die Integration zweier Bewegungszustände mehr mit externer Unterstützung durchgeführt werden. Die Informationen müssen hierfür nicht mehr im Arbeitsgedächtnis gehalten oder aus dem Langzeitgedächtnis abgerufen werden (Lowe, 1999a). Insgesamt gesehen, weisen somit in der vorliegenden Studie die Animationen die typischen Nachteile von dynamischen Darstellungen nur in verminderter Form auf, während die Defizite der mentalen Animation und der Inferenz der Zwischenzustände bei den parallelen Key-Frames gegeben sind.

Daher profitieren Personen mit geringerer Merkfähigkeit vermutlich mehr von Animationen als von statischen Bildern. Bei Personen mit höherer Merkfähigkeit ist dieser Vorteil wahrscheinlich geringer ausgeprägt, da sie über genügend Gedächtniskapazität verfügen, um die zusätzlichen internen Konstruktionsprozesse bei den Key-Frames zu vollziehen. Möglicherweise erreichen sie auch bessere Ergebnisse mit den statischen Bildern.

Aus dieser Argumentation heraus resultiert die folgende Grundannahme:

5. Es besteht eine Interaktion zwischen den Bedingungen „Animation“ und „Bild parallel“ mit der Merkfähigkeit von Personen sowohl hinsichtlich der Lösung von animierten und statischen bildlichen Kategorisierungsaufgaben als auch bezüglich des Cognitive Loads, der während des Lernens und während des Nachtests empfunden wird. Die Interaktionen verlaufen jeweils in der Richtung, dass der Lernvorteil durch Animationen bei Personen mit geringerer Merkfähigkeit größer ist als bei Personen mit höherer Merkfähigkeit.

4 Empirischer Teil von Studie 1: Vergleich von Animationen und statischen Bildern hinsichtlich ihrer Wirksamkeit auf die Lösung von Inferenzaufgaben

4.1 Vorpilotstudie

4.1.1 Fragestellungen und Zielsetzung

Die Grundkonzeption der Experimente der Pilot- sowie der beiden Hauptstudien bestand darin, dass die Inhalte zu synaptischen Vorgängen mit Hilfe einer Animation oder mit Hilfe von verschiedenartig präsentierten statischen Bildern von Probanden gelernt werden sollten, die daraufhin Inferenz- und Wissensaufgaben zu den gesehenen Inhalten lösen sollten. In der Vorpilotstudie ging es zunächst darum, die für dieses Forschungsvorhaben bereits bestehenden Materialien zu überprüfen und zu verbessern.

Die Evaluation verschiedener Faktoren des Designs war deshalb notwendig, da es zum einen nicht völlig klar war, ob man die Vorgänge, die an einer Synapse ablaufen, mit Bildern bzw. mit Animationen ausreichend darstellen konnte oder ob man zur besseren Verständlichkeit noch Text hinzufügen musste. Eine solche Überarbeitung hätte allerdings auch die theoretischen Grundannahmen der Studie verändert.

Ebenfalls getestet werden sollten die Items eines Vorwissenstests hinsichtlich ihrer Verständlichkeit, ihrer Schwierigkeit und ihrer Trennschärfe. Es war zudem ein Nachtest (zunächst in zwei Versionen) erstellt worden, dessen Items ebenfalls auf ihre Gütekriterien hin überprüft werden sollten.

Ein weiterer Aspekt, der getestet werden sollte, war der, ob man den Versuchspersonen zur Beantwortung der Nachtestitems ein Standbild vorlegen sollte, auf dem eine Synapse dargestellt war, deren einzelne Teile mit Fachbegriffen beschriftet waren. Das hatte den Sinn, dass in den Fragen des Nachtests viele Fachbegriffe vorkamen, die man kennen musste, um die Aufgaben beantworten zu können. Es war nicht unmittelbar ersichtlich, ob eine Integration eines solchen Bildes in die Lernphase vollkommen ausreichen würde, damit die Teilnehmer die Fragen des Nachtests lösen konnten.

Desweiteren wurde mit den Versuchspersonen jeweils ein Interview geführt, um ihre Meinung und Bewertung von verschiedenen Aspekten des Versuchs zu erfahren.

Eine Erhebung der Bearbeitungszeiten für die einzelnen Tests sollte eine bessere Einschätzung der späteren Versuchsdauer ermöglichen und die Ökonomie der einzelnen Verfahren sicherstellen.

Die Zielsetzung der Vorpilotstudie war somit, zu überprüfen, ob die Grundkonzeption des Versuchs stimmig und praktisch durchführbar war. Es sollte eine angemessene Schwierigkeit der Materialien gefunden und Bearbeitungszeiten erhoben werden, um die spätere Dauer des Versuchs schätzen zu können. Daher hatten Vergleiche zwischen den Gruppen und eine Miteinbeziehung von Kontrollvariablen eine geringere Wichtigkeit zugunsten eines explorativen Charakters der Studie, die verwertbare Hinweise für eine adäquate Durchführung der Pilot- und der beiden Hauptstudien geben sollte.

Für die Vorpilotstudie ergaben sich folgende Fragestellungen:

Forschungsfrage 1: Sind die Items des Vorwissenstests sowie der Wissens- und Inferenzskalen der beiden konstruierten Nachtests verständlich und weisen sie zufriedenstellende Gütekriterien auf?

Forschungsfrage 2: Kann man die Vorgänge an einer Synapse ausreichend verständlich mit Bildern darstellen, so dass Wissens- und Inferenzitems beantwortet werden können?

Forschungsfrage 3: Wie hoch sind die Bearbeitungszeiten des Vorwissens- und des Nachtests?

Forschungsfrage 4: Ist das angefertigte Lernmaterial in seiner Gestaltung für die Probanden handhabbar und verständlich?

Forschungsfrage 5: Sind die angesetzten Lernzeiten für die Lerneinheit ausreichend?

Forschungsfrage 6: Wird für die Bearbeitung des Nachtests ein Standbild mit Fachbegriffen benötigt oder genügt eine Integration eines solchen Bildes in die Lerneinheit?

4.1.2 Methoden

4.1.2.1 Stichprobe

An der Vorpilotstudie nahmen 14 Studenten der Universität Landau teil. Die Fachrichtungen umfassten größtenteils die Magister- und Lehramtsstudiengänge sowie den Studiengang „Sozialwissenschaften“. Es wurden von vornherein Psychologie-, Biologie- sowie Umweltwissenschaftsstudenten von der Untersuchung ausgeschlossen, da sie zu der Thematik „Die Vorgänge an einer Synapse“ durch ihr Studium tendenziell mehr Vorwissen hätten haben können und dadurch ein höheres Risiko von Deckeneffekten in den Analysen gegeben gewesen wäre. Angeworben wurden die Teilnehmer per Aushang an verschiedenen Instituten der Universität sowie durch direkte Ansprache in einer Lehramtsvorlesung. Das Durchschnittsalter innerhalb der Stichprobe lag bei $M = 23.0$ Jahren ($SD = 1.8$), mit einem Altersrange von 21 bis 27 Jahren. Von den 14 Teilnehmern waren 4 männlich und 10 weiblich, was einem Verhältnis von knapp einem Drittel Männern zu ungefähr zwei Dritteln Frauen entsprach (28,6% zu 71,4%).

4.1.2.2 Design

Innerhalb der Vorpilotstudie lagen drei verschiedene Versuchsbedingungen vor. Es wurden mehrere Faktoren variiert, um möglichst vielfältige und vollständige Informationen darüber zu erhalten, welche Bedingungen für eine spätere Durchführung am geeignetsten waren. Von daher rührte aber auch der Umstand, dass die Durchführungen in den Gruppen zwar größtenteils, aber nicht komplett standardisiert und damit miteinander vergleichbar waren.

Variiert wurden die Faktoren:

- Nachtest (leichtere vs. schwierigere Version)

In der schwierigeren Version waren konzeptuell komplexere Items enthalten, bei denen aus dem Itemstamm heraus umfangreichere mentale Modelle aufgebaut und schwierigere Inferenzen gezogen werden mussten.

- Verwendung eines beschrifteten Standbildes zur Beantwortung des Nachtests (wird vorgegeben vs. wird nicht vorgegeben)
- Lernzeit für das beschriftete Standbild (1 min 30 vs. 2 min)

Die Gruppe, die bei der Beantwortung des Nachtests kein Standbild zur Verfügung hatte, bekam für das Lernen der Begriffe die höhere Lernzeit von 2 min.

- Zahl der Prozessbilder (7 vs. 8)

Zum Lernen der Prozesse an einer Synapse wurden zunächst sieben Bilder vorgelegt, auf denen die wichtigsten Abläufe des Gesamtvorganges dargestellt waren. Da dabei der Vorgang „Spaltung des Transmitters“ nicht deutlich erkennbar war, wurde er bei einer Gruppe von Probanden durch ein eigenes Bild präsentiert, was die Gesamtzahl der Bilder auf 8 erhöhte.

Durch Manipulation und Kombination dieser Faktoren entstanden die folgenden drei Bedingungen:

Tabelle 4.1.1

Beschreibung der drei Versuchsbedingungen der Vorpilotstudie

	Version Nachtest	Standbild bei Nachtest	Lernzeit Standbild	Anzahl Prozessbilder	Stichproben- größe (n)
Gruppe 1	Leichter	Ja	1 min 30	7	5
Gruppe 2	Leichter	Nein	2 min 00	7	5
Gruppe 3	Schwieriger	Ja	1 min 30	8	4

Die Gruppe 1 bearbeitete die einfachere Version des Nachtests und hatte dafür auch das beschriftete Standbild zur Verfügung. In der Lernphase konnten sie sich daher das Standbild nur 1 min 30 lang einprägen. Die Prozesse an der Synapse wurden der Gruppe anhand von sieben Bildern präsentiert. Im Vergleich zu der ersten Bedingung löste die zweite Gruppe den Nachtest ohne das Standbild, das dafür in der Lernphase länger gezeigt wurde. Der Gruppe 3 wurde eine schwierigere Version des Nachtests vorgelegt, zu dessen Bearbeitung sie das Standbild zu Hilfe nehmen konnten. Für die Darstellung der Prozesse wurden nun acht Bilder eingesetzt.

Die Versuchspersonen wurden in der Reihenfolge ihrer Anmeldung, aber ansonsten randomisiert, den verschiedenen Bedingungen zugewiesen, so dass in der ersten Gruppe 5, in der zweiten ebenfalls 5 und in der dritten Gruppe 4 Teilnehmer waren.

Abhängige Variablen waren die Wissens- und die Inferenzskala des Nachtests.

Auf eine Erhebung von Kontrollvariablen wurde hingegen innerhalb des Designs der Vorpilotstudie noch verzichtet (mit Ausnahme des Vorwissens), da keine aussagekräftigen Mittelwertvergleiche gemacht werden konnten.

4.1.2.3 Lernmaterialien

Für die Vorpilotstudie wurde eine Vorläuferversion des eigentlichen Lernmaterials von der Doktorandin selbst produziert, das den Zweck hatte, Hinweise zu erhalten, ob Bilder allein ausreichen, um die Inhalte der synaptischen Vorgänge zu vermitteln oder ob es notwendig war, erklärenden Text hinzuzufügen. Deshalb beschränkte sich, auch aus technischen Gründen, das Material auf parallel angeordnete Bilder, die der Form nach der späteren „Bild parallel“-Bedingung in der Pilot- und der zweiten Hauptteilstudie ähnelten. Auf die Erstellung weiterer Variationen in der Präsentation der Prozesse wurde verzichtet, da in der Vorpilotstudie grundlegende Überprüfungen des Materials sowie des Nachtests im Vordergrund standen. Das Lernmaterial wurde mit dem Windows-Programm „Paint“ erstellt, und die so entstandenen Bilder wurden in das Programm „MS Word“ kopiert (Version „Office 2007“). Was die Art der Darstellung der Bilder hinsichtlich ihrer Formen und Farben betraf, wurde sich an dem gut verständlichen Gymnasiallehrbuch von Bayrhuber und Kull (1989) orientiert. Die wesentlichen Bestandteile der Synapse, die eine Verbindung zwischen zwei Nervenzellen ist, wurden farblich gut voneinander abgegrenzt eingezeichnet, und es wurde auf die Präsentation von unwesentlichen Details verzichtet.

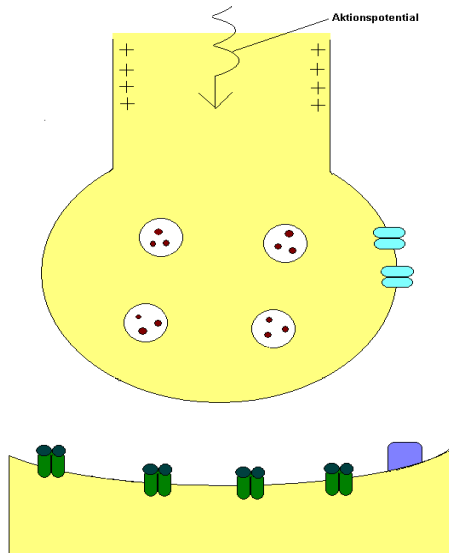


Abbildung 4.1.1: Beispielbild für die Bildgestaltung der Synapse in der Vorpilotstudie

Zur Enkodierung der Fachbegriffe, die die Teile der Synapse benannten und die zur Lösung des Nachttests notwendig waren, wurde ein Standbild produziert, auf dem eine Synapse mit den zehn Namen ihrer wichtigsten Bestandteile dargestellt war. Die Begriffe wurden im Vergleich zum Lehrbuch einfacher gestaltet, da erste Befragungen von Personen während der Materialentwicklung ergeben hatten, dass die Namen sonst zu schwierig gewesen wären, um sie sich in so kurzer Zeit einzuprägen (Beispiel: „präsynaptische Membran“ wurde zu „vorsynaptischer Membran“).

Die Prozesse wurden zunächst in sieben Bildern präsentiert. Die Bilder stellten eine Vorform der Key-Frames dar, da mittlere Bilder und kausale Schnittstellen der Teilprozesse der synaptischen Übertragung verwendet wurden. (Eine genauere Beschreibung der Teilprozesse und der Extraktion der Key-Frames erfolgt in der Materialbeschreibung der Pilotstudie, s. Abschnitt 4.2.2.3). Die Bilder waren auf dem Computerbildschirm so angeordnet, dass jeweils zwei von ihnen gleichzeitig gesehen werden konnten. Zum Aufdecken der weiteren Graphiken war ein Scrollen mit der Maus notwendig, wie es die Teilnehmer aus ihrem Alltag mit Computerprogrammen gewohnt gewesen sein sollten. Die Bilder waren zur besseren Orientierung zusätzlich durchnummeriert, und es wurden drei weitere Begriffe eingebaut, um das „Aktionspotential“ und die beiden Ionenströme zu benennen, da sonst der Ablauf nicht gut verständlich und das Beantworten des Nachttests schwerer möglich gewesen wäre. Dabei wurde darauf geachtet, dass textlich keine explizite Prozessinformation gegeben wurde. Bei den letzten vier Versuchspersonen wurde noch ein weiteres Bild integriert, auf

dem die Spaltung des Transmitters eigens dargestellt und benannt wurde, da dieser Vorgang für die Versuchspersonen sonst nicht ersichtlich war, wie sich in den ersten Durchführungen der Studie gezeigt hatte.

Die Lernzeiten betragen für die Prozesse 4 min und für das Standbild, je nach Bedingung, 1 min 30 bzw. 2 min.

(Für eine vollständige Ansicht des Standbildes und der Prozessbilder, s. Anhang A1 auf der beigefügten CD.)

4.1.2.4 Messinstrumente

Die Instrumente werden in der Reihenfolge ihrer Bearbeitung durch die Probanden vorgestellt und beschrieben.

Soziodemographischer Fragebogen

Mit Hilfe dieses Fragebogens wurden Angaben zu Alter, Geschlecht, Schulabschluss und Beruf der Versuchspersonen erfasst. Bei den letzten drei Variablen wurden jeweils verschiedene Antwortalternativen vorgegeben, bei denen die Teilnehmer das für sie Zutreffende ankreuzen konnten. Weiterhin wurde bei Studenten, die die eigentliche Zielgruppe der Studie waren, mit Hilfe eines offenen Antwortformats nach ihrer Fachrichtung und der Semesterzahl ihres Studiums gefragt (zur Ansicht, s. Anhang A2 – 1).

Vorwissenstest

Der Vorwissenstest sollte in der späteren Pilot- und den Hauptstudien das Vorwissen der Probanden zu Synapsen als Kontrollvariable erfassen. Im Rahmen der Vorpilotstudie sollten die Aufgaben zunächst auf ihre Gütekriterien hin überprüft werden. Der Vorwissenstest bestand aus 11 Items (für die vollständige Version des Vorwissenstests, s. Anhang A2 - 2), wobei sich innerhalb der Testkonstruktion für ein halboffenes Antwortformat entschieden wurde, da auf diese Weise das bereits bestehende mentale Modell der Teilnehmer über Synapsen besser und detaillierter erfasst werden konnte. Für jede Frage wurde anhand von physiologischen Lehrbüchern (z.B. Bayrhuber & Kull, 1989; Rosenzweig, Leiman &

Breedlove, 1999) eine Musterantwort erstellt, die auch für das Rating innerhalb der Datenauswertung verwendet wurde. Auf die Konstruktion umfassenderer Ratingsysteme wurde zunächst verzichtet, da die eingesetzten Items einen noch vorläufigen Charakter hatten. Anhand der Musterantworten wurden die Lösungen der Teilnehmer in „richtig“ (2 Punkte), „zum Teil richtig“ (1 Punkt) und „falsch / nicht beantwortet“ (0 Punkte) eingestuft. Die Summe der Punkte stellte den Vorwissensscore der jeweiligen Person dar.

Der Vorwissenstest umfasste zwei Seiten und wurde als Paper-Pencil-Verfahren vorgegeben. Es gab keine Zeitbegrenzung für die Versuchspersonen hinsichtlich seiner Lösung (Power-Test). In der Instruktion wurde den Teilnehmern der Hinweis gegeben, bei Fragen, die sie nicht beantworten konnten, ein „?“ oder einen „/“ zu machen, mündlich wurde ihnen darüber hinausgehend mitgeteilt, dass es nicht schlimm wäre, wenn sie noch nicht so viel zu der Thematik wüssten. Dieser Zusatz diente dazu, ihre Lernmotivation sowie ihr Selbstvertrauen aufrechtzuerhalten.

Bei der Itemkonstruktion wurde darauf geachtet, alle Wissensbereiche, die auch im Nachtest explizit oder implizit abgefragt wurden, zu berücksichtigen, damit eventuelle Beeinflussungen des Vorwissens auf den Nachtestscore ermittelt werden konnten.

Die ersten vier Items des Vorwissenstests waren konzeptuell einfacher gestaltet und fragten eher grundlegendes Wissen zu Synapsen und Naturwissenschaften ab (*Beispielitem*: „Was findet grundlegend an einer Synapse statt (1 Satz)?“).

In den Items 5 bis 8 ging es um speziellere Aspekte der synaptischen Vorgänge, wie sie auch später in dem Lernmaterial präsentiert wurden. Dabei wurden sowohl die Teilprozesse am Anfang, wie auch die am Ende des Gesamtvorganges erfragt (*Beispielitem*: „Was ist die direkte Folge des Einstroms von Calciumionen im synaptischen Endknopf?“).

Die Items 9 und 10 fragten nach Nervengiften, da im Nachtest viele Fragen toxische Wirkungen thematisierten (*Beispielitem*: „Nennen Sie drei Gifte, die die Funktion einer Synapse beeinträchtigen.“).

Mit dem letzten Item wurde schließlich eher breit nach dem Gesamtwissen der Versuchspersonen zu den synaptischen Vorgängen gefragt, um zu ermitteln, wie viele Teilaspekte sie zu dem Prozess wussten und auf welchem Abstraktionsgrad ein solches Wissen vorhanden war (*Item*: „Nennen Sie die wichtigsten Teilprozesse des Gesamtvorganges an einer erregenden Synapse in der richtigen Reihenfolge (ca. 8 Teilprozesse)“).

Bei manchen der Fragen des Vorwissenstests wurden konkrete Angaben zu der Länge der jeweils zu gebenden Antwort gemacht, um die Items konkreter und einfacher zu gestalten.

Nachtest

Innerhalb der Vorpilotstudie waren zwei Versionen eines Nachtests konzipiert worden. Zehn der Versuchspersonen bearbeiteten Version 1, die restlichen vier Teilnehmer füllten Version 2 aus (s. auch Anhang A2 - 3 und A2 - 4). Beide Nachtests bestanden aus einer Wissens- und einer Inferenzskala, die die abhängigen Variablen von dieser und der nachfolgenden Studien darstellten. Die Wissensskala umfasste 12 Items, die Inferenzskala 18 Items. Der Nachtest war wie der Vorwissenstest ein Power-Verfahren über 5 bzw. 6 Seiten, das in Paper-Pencil-Form vorgelegt wurde. Es wurde wieder ein halboffenes Antwortformat gewählt, was bedeutete, dass die Versuchspersonen ihre Lösungen in schriftlicher Form auf vorgegebenen Antwortlinien festhalten sollten. Dadurch konnte der Komplexitätsgrad ihres erworbenen mentalen Modells besser beurteilt und zwischen den Personen genauer differenziert werden. Für ein anschließendes Rating der Daten wurden wieder Musterantworten mit Hilfe von biologischen Lehrbüchern (z. B. Bayrhuber & Kull, 1989; Rosenzweig et al., 1999) erstellt. Auf die Konstruktion von Ratingsystemen wurde wiederum verzichtet. Die Musterantworten dienten dazu, die Lösungen der Teilnehmer in die Kategorien „richtig“ (2 Punkte), „zum Teil richtig“ (1 Punkt) oder „falsch / nicht beantwortet“ (0 Punkte) einzustufen. Die Summen der erreichten Punkte in den Wissens- und Inferenzaufgaben stellten getrennt voneinander den Nachwissens- und den Inferenzscore der jeweiligen Person dar.

Sollten die Teilnehmer eine Frage nicht beantworten können, wurde die Instruktion gegeben, in einem solchen Fall ein „?“ oder einen „/“ zu machen.

Die Wissensskalen der beiden Versionen des Nachtests waren bis auf ein Item identisch, die Fragen wurden lediglich in unterschiedlicher Reihenfolge dargeboten. Konzeptuell gesehen fragten die Items Wissen ab, das auf dem Lernmaterial explizit zu sehen war und daher einfach reproduziert werden konnte. Zur Beantwortung der meisten Items war jeweils ein Satz ausreichend, bei den schwierigeren Aufgaben mussten zur Lösung zwei bis drei aufeinanderfolgende Prozesse genannt werden (*Beispielitem*: „Welcher Prozess geschieht, bevor die Bläschen auf die Zellmembran zuwandern?“). Bei der Testkonstruktion war von Wichtigkeit, dass die Items alle Teilprozesse der synaptischen Übertragung abdeckten und

dass die Fragen von ihrer Reihenfolge her so angeordnet waren, dass man keine von ihnen aus Informationen einer vorherigen heraus beantworten konnte. Daher wurde den Versuchspersonen auch die Instruktion gegeben, nicht mehr zurückzublättern, wenn sie eine Seite des Fragebogens bearbeitet hatten.

In der Inferenzskala zielten die Items auf Informationen ab, die nicht direkt auf dem Lernmaterial sichtbar waren, aber aus den gezeigten Vorgängen hergeleitet werden konnten. Viele der Fragen thematisierten die Wirkungen von Nervengiften oder Nervenkrankheiten, wodurch die Funktion einer Synapse anormal verändert wird. Die Teilnehmer mussten daher zur Beantwortung der Items ein mentales Modell einer gesunden Synapse aufbauen, dieses Modell durch die Wirkung des Gifts bzw. der Krankheit manipulieren und die entstandene Funktionsänderung ablesen (*Beispielitem*: „Bei der Alzheimer-Erkrankung produzieren die Nervenzellen zu wenig Acetylcholin. Warum verschreibt der Arzt hier häufig Cholinesterasehemmer?“). Durch die Verwendung dieser Themen wurde ebenfalls das Verständnis der Prozesse an verschiedenen Stellen der Synapse abgefragt, je nachdem wo das störende Element angriff.

Die zweite Version des Nachtests hatte zur Hälfte die gleichen Inferenzitems wie die erste Version, allerdings waren die übrigen Items schwieriger gestaltet worden als zuvor. Es wurden zum einen Fragen aus der ersten Form elaboriert, so dass eine höhere Zahl von Inferenzen gezogen werden musste (*Beispielitem: Leichteres Item in Version 1*: „Bei einer Form der Calciumkanalkrankheit sind die Calciumkanäle abnorm häufig geschlossen. Inwiefern beeinflusst dies die Erregungsweiterleitung, wenn ein Aktionspotential ankommt?“ *Schwierigeres Item in Version 2*: „Bei der Calciumkanalkrankheit sind die Calciumkanäle entweder abnorm häufig geschlossen oder abnorm häufig geöffnet. In welche beiden Richtungen können diese beiden Krankheitsfälle jeweils die Erregungsweiterleitung beeinflussen und warum?“)

Zum anderen wurden theoretisch schwierigere Items konstruiert, bei denen komplexere mentale Modelle aus dem Itemstamm heraus aufgebaut und ebenfalls eine höhere Zahl von Inferenzen gezogen werden mussten (*Beispielitem*: „Muscarin setzt sich wie der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle und öffnet sie. Es wird aber im Spalt nicht abgebaut. Atropin dockt ebenfalls an die Rezeptoren der Natriumkanäle an, öffnet sie aber nicht. Würden Sie bei einer Vergiftung durch E 605, das die Cholinesterase hemmt, eher Atropin oder Muscarin als lindernde Maßnahme geben? Begründen Sie.“).

Die Items wurden in den beiden Nachtests in verschiedener Reihenfolge vorgegeben, wobei wieder darauf geachtet wurde, dass keine thematisch ähnlichen Fragen aufeinander folgten, außer in solchen Fällen, in denen Items direkt aufeinander aufbauten.

Interviewleitfaden

Nach der Bearbeitung des Nachtests wurde mit den Versuchspersonen ein Interview geführt, zu dem zuvor ein Leitfaden erstellt worden war, um das Gespräch zu standardisieren und nach wichtigen Aspekten zu strukturieren (zur Ansicht des Leitfadens, s. Anhang A2 - 5). Das Interview diente in erster Linie dazu, Hinweise zu erhalten, wie man die schon bestehenden Materialien und Instrumente noch verbessern konnte. Der Leitfaden bestand aus acht Fragen, wobei auch Raum zum Notieren der Bearbeitungszeiten der Versuchspersonen vom Vorwissens- und vom Nachtest zur Verfügung stand.

Die erste Frage des Interviews zielte auf die Einschätzung der jeweiligen Versuchsperson des eigenen Vorwissens ab, um einen Anhaltspunkt zu erhalten, von welchem Expertenstand aus der Teilnehmer das Experiment beurteilte. Danach wurde nach dem empfundenen Schwierigkeitsgrad des Vorwissenstests gefragt. Anschließend lag der Fokus auf den Lernmaterialien. Die Teilnehmer konnten ihre Meinung zu der Art der Darstellung (Farbe, Form, Zahl der Bilder, etc.) kundtun und Verbesserungsvorschläge machen. Bei der Frage danach, wie die Versuchspersonen die Dauer der Lernzeit jeweils empfanden, wurde auch der Aspekt mit einbezogen, ob diese ebenfalls für eine Person ohne Vorwissen geeignet wäre, da die weiteren Experimente für diese Zielgruppe konzipiert werden sollten. Danach sollten die Teilnehmer noch die Schwierigkeit des Nachtests einschätzen und die Frage beantworten, ob ihrer Meinung nach alle Items des Nachtests aus dem Lernmaterial heraus gelöst werden konnten. Dies diente der Vermeidung von zu schwierigen oder gar unlösbaren Aufgaben innerhalb des Nachtests.

4.1.2.5 Durchführung

Die Erhebungen der Vorpilotstudie fanden im Sommersemester 2009 statt. Als Räumlichkeit diente bei allen Untersuchungen das Labor der Abteilung „Allgemeine und Pädagogische Psychologie“ der Universität Landau, so dass jeweils vergleichbare Bedingungen zwischen

den Versuchspersonen herrschten. Die Studie wurde in Einzelversuchen durchgeführt, wobei die Doktorandin selbst die Versuchsleitung übernahm.

Nachdem den Teilnehmern kurz die Thematik der Studie zu Beginn des Versuchs erklärt und ihnen Anonymität ihrer Daten zugesichert worden war, wurden sie gebeten, den Vorwissenstest zu Synapsen auszufüllen. Bei jedem Teilnehmer wurde die Bearbeitungszeit für diesen Fragebogen mit einer Stoppuhr festgehalten. Danach begaben sich die Versuchspersonen an einen Computer, an dem sie zunächst, je nach Versuchsbedingung, 1 min 30 bzw. 2 min lang das Standbild mit den synaptischen Fachbegriffen zu Synapsen lernten. Es folgte das Verstehen und Einprägen der synaptischen Vorgänge anhand von sieben bzw. acht Bildern, das in jeder Gruppe 4 min lang dauerte. Schließlich bearbeiteten die Teilnehmer den Nachtest, was in der Regel 25 bis 35 Minuten in Anspruch nahm, wie mittels einer Stoppuhr festgestellt wurde. Teils wurde dazu das beschriftete Standbild am Computer wieder gezeigt. Zuletzt wurde mit den Versuchspersonen ein teilstandardisiertes Interview mit Hilfe eines Leitfadens über ihre Meinung zu verschiedenen Aspekten des Experimentes geführt, und ihre Antworten wurden dabei mit protokolliert. Die Entlohnung für die Teilnahme betrug bei jeder Person 8 Euro, und die Experimente dauerten jeweils insgesamt zwischen 45 und 60 Minuten. Auf Erhebungen von Kontrollvariablen wurde im Rahmen der Vorpilotstudie noch verzichtet, da keine genauen Mittelwertsvergleiche, sondern eher explorative Erkenntnisse angestrebt wurden.

4.1.3 Ergebnisse

Die Berechnung der Itemanalysen sowie der deskriptiven Statistiken wurde innerhalb der Vorpilotstudie mit dem Programm „SPSS, Version 17.0“ durchgeführt.

Itemanalyse

Forschungsfrage 1: Sind die Items des Vorwissenstests sowie der Wissens- und Inferenzskalen der beiden konstruierten Nachtests verständlich und weisen sie zufriedenstellende Gütekriterien auf?

Um diese Forschungsfrage zu beantworten, wurde das Cronbachs Alpha des Vorwissenstests sowie die Trennschärfen und Schwierigkeiten seiner Items berechnet. Ebenso wurde für die Wissens- und Inferenzskalen der Nachtests verfahren, wobei die Analysen für die beiden Versionen jeweils getrennt erfolgten und deshalb die Ergebnisse nach den entsprechenden Untersuchungsgruppen getrennt dargestellt werden. Als Kriterium für die Weiterverwendung der Aufgaben für die anschließenden Studien diente bei der Trennschärfe der in der Literatur (Bortz & Döring, 2002) empfohlene Grenzwert von $r_{it} = .30$, sowie bei der Schwierigkeit das mittlere Intervall von $p = .20$ bis $p = .80$. Generell ist zu sagen, dass diese Kriterien aufgrund der sehr geringen Stichprobengrößen milder angelegt wurden, um Items, die mäßige Abweichungen von den Grenzwerten aufwiesen, nochmals bei einer größeren Stichprobe mit stabileren Schätzungen testen zu können. Zudem wurde als Hilfskriterium die von SPSS ausgegebenen Werte „Cronbachs Alpha, wenn Item weggelassen“ herangezogen, die die Reliabilitäten darstellen, die entstünden, wenn man das betreffende Item aus der Gesamtskala eliminieren würde. Für die anschließende Berechnung der deskriptiven Statistiken wurde innerhalb der Vorpilotstudie auf eine Itemelimination verzichtet, da anhand dieser Kennwerte beurteilt werden sollte, ob die Gesamtskalen eine konzeptuell angemessene Schwierigkeit für das Design aufwiesen und ob die Items prinzipiell unter den gegebenen Lernbedingungen lösbar waren.

Zunächst erfolgt die Darstellung der Ergebnisse der Itemanalyse des Vorwissenstests:

a) Vorwissenstest ($n = 14$)

Tabelle 4.1.2

Trennschärpen und Schwierigkeiten der Items des Vorwissenstests in der Vorpilotstudie

Item	<i>rit</i>	<i>p</i>
Item 1	.23	.07
Item 2	.17	.46
Item 3	.62	.39
Item 4	.43	.46
Item 5	.33	.07
Item 6	.41	.04
Item 7	.00	.00
Item 8	-.07	.04
Item 9	.69	.25
Item 10	.19	.11
Item 11	.41	.07

Anmerkungen. Die Itemnummern entsprechen denen des Vorwissenstests in der Vorpilotstudie (s. Anhang A2 - 2).

Das Cronbachs Alpha der Vorwissenskala betrug $\alpha = .64$. Schwächere Trennschärpen waren bei den Items 7 und 8 zu verzeichnen. Sieben der elf Aufgaben hatten niedrige Schwierigkeitswerte, was auf die konzeptuell hohe Komplexität der Items für Novizen zurückzuführen sein dürfte.

b) Wissensskalen der beiden Nachtests (Version 1, $n = 10$; Version 2, $n = 4$)

Tabelle 4.1.3

Trennschärfen und Schwierigkeiten der Items der Wissensskala des Nachtests (Version 1) in der Vorpilotstudie

Item	<i>rit</i>	<i>p</i>
Item 1	.00	1.00
Item 2	.09	.85
Item 6	-.23	.90
Item 9	.06	.85
Item 12	.26	.90
Item 16	.00	1.00
Item 17	-.30	.85
Item 20	.08	.70
Item 23	.58	.45
Item 26	.41	.95
Item 27	-.25	.95
Item 30	.47	.75

Anmerkungen. Die Itemnummern entsprechen denen des Nachtests (Version 1) in der Vorpilotstudie (s. Anhang A2 - 3).

Es konnte bei dieser Skala ein Cronbachs Alpha von $\alpha = .27$ festgestellt werden. Acht der zwölf Items wiesen schwächere Trennschärfen auf ($rit < .10$), wobei die Schwierigkeitswerte generell auf einem hohen Niveau lagen. Bei den Items 2, 9, 12, 23, 26 und 30 ermittelte SPSS eine Verringerung des Alpha, wenn man diese Aufgaben aus der Skala eliminieren würde.

Tabelle 4.1.4

Trennschärfen und Schwierigkeiten der Items der Wissensskala des Nachttests (Version 2) in der Vorpilotstudie

Item	<i>rit</i>	<i>p</i>
Item 1 (1)	.26	.75
Item 2 (2)	-.32	.63
Item 6 (6)	.00	1.00
Item 7 (12)	.00	1.00
Item 8 (27)	.26	.75
Item 16 (30)	-.71	.50
Item 17 (16)	-.19	.75
Item 18 (--)	-.32	.50
Item 23 (26)	-.83	.88
Item 24 (23)	-.57	.38
Item 27 (9)	.56	.88
Item 30 (17)	.00	1.00

Anmerkungen. Die Itemnummern entsprechen denen des Nachttests (Version 2) in der Vorpilotstudie (s. Anhang A2 - 4). Bei in beiden Nachttestversionen vorkommenden Items sind in Klammern die Itemnummern aus Version 1 hinzugefügt.

Bei der Wissensskala des zweiten Nachttests belief sich der Reliabilitätswert nur auf $\alpha = -1.31$, was auch an der sehr geringen Stichprobengröße der Untersuchungsgruppe liegen könnte. Drei der zwölf Items zeigten zufriedenstellende Trennschärfen, während bei sieben Aufgaben mittlere Schwierigkeiten errechnet werden konnten. Durch Ausschließen der Items 1, 6, 7, 8, 27 und 30 wäre die Reliabilität der Skala geringer ausgefallen.

b) Inferenzskalen der beiden Nachttests (Version 1, $n = 10$; Version 2, $n = 4$)

Tabelle 4.1.5

Trennschärfen und Schwierigkeiten der Items der Inferenzskala des Nachttests (Version 1) in der Vorpilotstudie

Item	<i>rit</i>	<i>p</i>
Item 3	-.02	.60
Item 4	.11	.70
Item 5	.28	.85
Item 7	.44	.90
Item 8	.20	.65
Item 10	.72	.80
Item 11	.15	.75
Item 13	.82	.45
Item 14	.72	.80
Item 15	.43	.25
Item 18	.30	.40
Item 19	.76	.65
Item 21	.73	.55
Item 22	.00	1.00
Item 24	.52	.40
Item 25	.73	.30
Item 28	.42	.80
Item 29	.81	.55

Anmerkungen. Die Itemnummern entsprechen denen des Nachttests (Version 1) in der Vorpilotstudie (s. Anhang A2 - 3)

Die Inferenzskala dieses Nachttests hatte ein sehr gutes Cronbachs Alpha von $\alpha = .84$. Bis auf die von Item 3 und Item 22 konnten die Trennschärfen der Aufgaben als zufriedenstellend bis gut bezeichnet werden. Die Itemschwierigkeiten lagen ebenfalls bis auf drei Ausnahmen im mittleren Bereich. Der Miteinbezug der Items 5, 7, 10, 13, 14, 15, 19, 21, 24, 25, 28 und 29 führte zu einer Erhöhung der Reliabilität der Skala.

Tabelle 4.1.6

Trennschärfen und Schwierigkeiten der Items der Inferenzskala des Nachttests (Version 2) in der Vorpilotstudie

Item	<i>rit</i>	<i>p</i>
Item 3 (3)	-.76	.38
Item 4 (--)	.61	.25
Item 5 (4)	-.37	.13
Item 9 (5)	.00	1.00
Item 10 (--)	-.49	.75
Item 11(28)	.00	1.00
Item 12 (--)	-.16	.75
Item 13 (13)	.64	.38
Item 14 (14)	.28	.25
Item 15 (15)	.00	.00
Item 19 (--)	.74	.38
Item 20 (19)	.00	.50
Item 21 (--)	.17	.25
Item 22 (--)	.88	.50
Item 25 (24)	.95	.63
Item 26 (25)	.88	.50

Tabelle wird auf der nächsten Seite fortgesetzt

Fortsetzung Tabelle 4.1.6

Item	<i>rit</i>	<i>p</i>
Item 28 (--)	.74	.38
Item 29 (--)	.67	.50

Anmerkungen. Die Itemnummern entsprechen denen des Nachttests (Version 2) in der Vorpilotstudie (s. Anhang A2 - 4). Bei in beiden Nachttestversionen vorkommenden Items sind in Klammern die Itemnummern aus Version 1 hinzugefügt

Das Cronbachs Alpha der Skala lag bei einem zufriedenstellenden Wert von $\alpha = .71$. Bei den Trennschärfen wiesen acht der achtzehn Items schwächere Werte auf ($rit < .10$). Die Schwierigkeiten befanden sich hingegen vorwiegend im mittleren Bereich. Ohne die Aufgaben 4, 13, 14, 19, 22, 25, 26, 28 und 29 wäre die Reliabilität der Skala geringer ausgefallen.

Deskriptive Statistiken der Testverfahren

Forschungsfrage 2: Kann man die Vorgänge an einer Synapse ausreichend verständlich mit Bildern darstellen, so dass Wissens- und Inferenzitems beantwortet werden können?

Forschungsfrage 6: Wird für die Bearbeitung des Nachttests ein Standbild mit Fachbegriffen benötigt oder genügt eine Integration eines solchen Bildes in die Lerneinheit?

Um diese Fragen zu beantworten, gingen bei der Berechnung der deskriptiven Statistiken alle Items der verwendeten Tests mit ein, um deren prinzipielle Lösbarkeit und Schwierigkeit zu bestimmen. Hierzu wurden die Mittelwerte der Gruppen mit den maximal möglichen Punktzahlen verglichen, um einen Anhaltspunkt dafür zu erhalten, wie gut die Tests unter den gegebenen Bedingungen gelöst wurden. Die maximal mögliche Punktzahl betrug beim Vorwissenstest 22, bei den Nachwissensskalen jeweils 24 und bei den Inferenzskalen jeweils 36 Punkte. Die Mittelwerte der Gruppe 2, die den Nachttest ohne beschriftetes Standbild bearbeitete, dienten als Indikator dafür, ob auch eine derartige Erhebungsform für das Design geeignet war.

a) Vorwissenstest

Tabelle 4.1.7

Deskriptive Statistiken der Variable „Vorwissenstest“ bei den Gruppen 1, 2 und 3 der Vorpilotstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Gruppe 1	5	0.00	7.00	3.60	2.70
Gruppe 2	5	0.00	9.00	5.80	3.49
Gruppe 3	4	0.00	4.00	1.75	1.71

Von 22 möglichen Punkten wurden im Schnitt von der Gruppe 1 nur 3.60, von der Gruppe 2 5.80 und von der Gruppe 3 nur 1.75 Punkte erreicht. Der Vorwissenstest war für die Probanden konzeptuell sehr schwierig, was auch in der Itemanalyse an den geringen Schwierigkeitswerten deutlich wurde. Die Ergebnisse entsprachen den Erwartungen, da bewusst Personen ohne hohes Vorwissen an der Studie teilnehmen sollten, die die teilweise spezifischeren Fragen nur zu einem geringeren Maß beantworten konnten.

b) Nachwissensskala (Version 1 und Version 2 zusammengefasst dargestellt)

Tabelle 4.1.8

Deskriptive Statistiken der Variable „Nachwissensskala“ bei den Gruppen 1, 2 und 3 der Vorpilotstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Gruppe 1	5	17.00	22.00	20.40	2.07
Gruppe 2	5	17.00	24.00	20.20	2.59
Gruppe 3	4	16.00	20.00	18.00	1.83

Die Nachwissensitems wurden von allen drei Gruppen gut gelöst. Gruppe 1 erreichte im Durchschnitt 20.40, Gruppe 2 20.20 und Gruppe 3 18.00 von 24 möglichen Punkten. Dieses Ergebnis konnte als Hinweis gewertet werden, dass die Nachwissensskalen auch nach einem

Lernprozess mit Bildern ohne Text gut bewerkstelligt werden konnten. Da die Gruppe 2, die das Standbild bei der Bearbeitung des Nachttests nicht zur Verfügung hatte, in der Wissensskala genauso gute bzw. bessere Leistungen erbrachte als die anderen Gruppen, schien eine Integration des Standbildes in die Lerneinheit zumindest in diesem Fall ausreichend zu sein.

c) Inferenzskala (Version 1 und Version 2 zusammengefasst dargestellt)

Tabelle 4.1.9

Deskriptive Statistiken der Variable „Inferenzskala“ bei den Gruppen 1, 2 und 3 der Vorpilotstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Gruppe 1	5	12.00	25.00	19.60	5.55
Gruppe 2	5	16.00	35.00	26.00	8.40
Gruppe 3	4	11.00	23.00	17.25	5.32

Generell gesprochen, waren die Inferenzskalen der Nachttests für alle drei Bedingungen lösbar. Gruppe 1 erreichte mit durchschnittlich 19.60 von 36 Punkten einen mittelmäßigen Wert, ebenso wie Gruppe 3, die das geringste Vorwissen hatte und die schwierigeren Items bearbeiten musste ($M = 17.25$). Auffällig war das hohe Ergebnis der Gruppe 2 ($M = 26.00$), bei der die Lösung des Nachttests ohne das Standbild anscheinend zu keinen Defiziten führte. Allerdings lag bei den Probanden dieser Gruppe auch das höchste Vorwissen vor ($M = 5.80$). Zusammengefasst konnte gesagt werden, dass ein Lernprozess, in dem ausschließlich Bilder verwendet wurden, ausreichend war, um die Aufgaben zufriedenstellend lösen zu können.

Deskriptive Statistiken der Bearbeitungszeiten der Testverfahren

Forschungsfrage 3: Wie hoch sind die Bearbeitungszeiten des Vorwissens- und des Nachttests?

a) Vorwissenstest

Tabelle 4.1.10

Deskriptive Statistiken der Variable „Bearbeitungszeit des Vorwissenstests“ bei den Gruppen 1, 2 und 3 der Vorpilotstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Gruppe 1	5	2 min 05	11 min 30	4 min 41	3 min 53
Gruppe 2	5	2 min 45	10 min 13	5 min 14	2 min 55
Gruppe 3	4	2 min 18	3 min 10	2 min 43	0 min 21

Durchschnittlich benötigten die Gruppen für die Bearbeitung des Vorwissenstests ca. 3 bis 5 min. Dieses Ergebnis sprach für eine zeitökonomische Durchführbarkeit dieses Verfahrens.

b) Nachtest (Version 1 und Version 2 zusammengefasst dargestellt)

Tabelle 4.1.11

Deskriptive Statistiken der Variable „Bearbeitungszeit des Nachtests“ bei den Gruppen 1, 2 und 3 der Vorpilotstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Gruppe 1	5	17 min 20	37 min 17	26 min 55	8 min 10
Gruppe 2	5	28 min 20	44 min 55	36 min 54	6 min 37
Gruppe 3	4	21 min 50	40 min 20	29 min 37	7 min 58

Die Bearbeitungszeiten des Nachtests lagen bei den Gruppen 1 bis 3 durchschnittlich zwischen ungefähr 27 und 37 min. Gruppe 2, die den Nachtest ohne Standbild bearbeitete, benötigte am längsten, was aber auch zusätzlich daran liegen konnte, dass diese Gruppe eine gute Leistung in der Inferenzskala aufwies und den Nachtest offensichtlich zum größten Teil vollständig ausgefüllt hat. Gruppe 3, deren Items etwas schwieriger waren, hatte keine nennenswert höhere Bearbeitungszeit als Gruppe 1, die einen leichteren Nachtest löste.

Interview

Innerhalb der Vorpilotstudie wurde ein Interview durchgeführt, bei dem die Probanden ihre Meinung zu verschiedenen Aspekten der Untersuchung äußern konnten. Um die Forschungsfragen 4 und 5 des Theorieteils zu beantworten und um die Ergebnisse der deskriptiven Statistiken zu den Fragen 2 und 6 noch zu ergänzen, wurden die Aussagen zu den Fragen im Interview untersucht, deren Inhalt mit dem der Forschungsfragen korrespondierte. Da es sich um ein exploratives Interview zur Materialentwicklung handelte, wurde auf qualitative Auswertungsmethoden mit Hilfe von Kategoriensystemen verzichtet. Es wurden bei jeder Untersuchungsgruppe die häufigeren, typischen Antworten zu den einzelnen Fragen zusammengefasst, die im Folgenden in Abstimmung zu den Forschungsfragen dargestellt werden sollen.

a) *Forschungsfrage 4*: Ist das angefertigte Lernmaterial in seiner Gestaltung für die Probanden handhabbar und verständlich?

Zur Beantwortung dieser Forschungsfrage wurden die Fragen 3 und 4 des Interviews herangezogen:

„Wie gut findest Du die Materialien? (Farbe, Art der Darstellung, Formen, etc.)“

Antworten:

- Gruppe 1: Die Bilder wurden durchgängig als „gut“ und als „sehr anschaulich“ beschrieben.
- Gruppe 2: Auch hier wurden die Bilder als „gut“ oder als „sehr anschaulich“ beschrieben.
- Gruppe 3: Die Bilder wurden als „ganz gut“ und „verständlich“ beurteilt, gerade wenn „man ein Laie sei“.

„Was fehlt Deiner Meinung nach an den Materialien oder könnte man an ihnen besser machen?“

Antworten:

- Gruppe 1: Viele Versuchspersonen (Vp) fanden die Bilder „so schon gut“. Bis auf das 7. Bild (Spaltung und Reuptake des Transmitters) wurde kaum etwas bemängelt.
- Gruppe 2: Nach der Meinung der Vp waren die Bilder „gut“, bis auf das 7. Bild. Hier sei die Deaktivierung des Transmitters „nicht erkennbar“, und die Spaltung wurde „nicht verstanden“.
- Gruppe 3: Zwei der vier Vp meinten, man könne „nichts“ besser machen. Sonstige Kritikpunkte waren: „Es war nicht ganz klar, ob es sich um Atome oder Ionen handelt“, „Farbe von den Natriumkanälen und den Rezeptoren etwas unterschiedlicher machen“

Das Lernmaterial wurde von den Probanden zu großen Teilen als „gut“ und „anschaulich“ bewertet. Als Mangel wurde die Darstellung der Spaltung des Transmitters empfunden, weshalb in der dritten Gruppe ein weiteres Bild in das Lernmaterial integriert wurde, auf dem die Deaktivierung explizit dargestellt war.

b) *Forschungsfrage 5*: Sind die angesetzten Lernzeiten für die Lerneinheit ausreichend?

Hinweise zur Beantwortung dieser Forschungsfrage gab Frage 5 des Interviews:

„Wie fandest Du die Länge der Lernzeit (zu lang / zu kurz)? Wie findest Du die Länge der Lernzeit für eine Person ohne Vorwissen?“

Antworten:

- Gruppe 1: Die Vp empfanden die Lernzeit als „OK“ oder sogar als „optimal“. Es war eher die Tendenz zu einer geringeren Lernzeit („3 min für die Prozesse hätten gereicht“) zu erkennen.

- Gruppe 2: Die Lernzeit wurde als „OK“ oder als „passend“ beurteilt. Eine Vp hätte für die Begriffe „ länger“ gebraucht. Eine andere Vp fand die Lernzeiten „eher zu lang“.
- Gruppe 3: Die Lernzeit wurde auch als „OK“ empfunden, eine Vp fand die Zeit „für das Standbild zu kurz“, eine andere die Lernzeit „für die Prozesse zu lang“.

Die Lernzeit (1 min 30 bzw. 2 min für das Standbild, 4 min für die Prozesse) wurde im Großen und Ganzen als „passend“ und als „OK“ beurteilt. Bei dem Standbild war eine Tendenz zu mehr, bei den Prozessen zu eher weniger Lernzeit erkennbar.

c) *Forschungsfrage 2*: Kann man die Vorgänge an einer Synapse ausreichend verständlich mit Bildern darstellen, so dass Wissens- und Inferenzitems beantwortet werden können?

Forschungsfrage 6: Wird für die Bearbeitung des Nachttests ein Standbild mit Fachbegriffen benötigt oder genügt eine Integration eines solchen Bildes in die Lerneinheit?

Um diese Forschungsfragen beantworten zu können, wurden die Antworten der Interviewfragen 6 und 8 einer näheren Betrachtung unterzogen:

„Wie schwierig fandest Du es, den Nachttest zu bearbeiten?“

Antworten:

- Gruppe 1: Die Vp beurteilten den Nachttest von „nicht so schwierig“ und „wesentlich leichter als der Vortest“ bis hin zu „teilweise schwierig“ und „machbar“. Die Fragen zu den Giften wurden als „schwieriger“ angesehen. Die Vp gaben an, dass sie die Fragen ohne das Standbild während des Nachttests „nicht ganz so gut hingekriegt“ hätten. Nur eine Vp war der Meinung, es sei ohne das Standbild „nicht zu schaffen“.
- Gruppe 2: Der Nachttest wurde von „anspruchsvoll, durch Mangel an Fachbegriffen“ bis hin zu „nicht so schwierig“ oder „ganz gut“ bzw. „sehr gut“ beschrieben. Die Fragen seien „gut gestellt“ gewesen. Vier von fünf Vp gaben an, dass sie das „Standbild nicht gebraucht hätten“ und „keine Probleme mit den Begriffen“ gehabt hätten (bei einer Vp war der Begriff „Cholinesterase“ schwierig).

- Gruppe 3: Die Vp bewerteten den Nachtest als „nicht ganz so schwer“ oder als „nicht so schwierig“. Als „schwierig“ wurden die „Transferaufgaben mit den Giften“ empfunden, bei denen man sich mehr „konzentrieren“ musste. Die Vp gaben an, dass das Standbild während des Nachtests „schon eine Erleichterung“ gewesen sei und sie ohne das Bild mit dem Test „weniger hätten anfangen“ können. Eine Meinung war, dass man durch das Bild „weniger überlegt“ und „das Bild anguckt“, es aber „psychologisch gut“ gewesen sei. Eine weitere Vp meinte, dass es „ja nicht so viele Begriffe“ gewesen wären.

„Findest Du, dass man die Items des Nachtests aus dem Lernmaterial heraus beantworten kann? Wenn nein, welche nicht und warum?“

Antworten:

- Gruppe 1: Die Vp meinten, dass die Beantwortung der Fragen aus den Bildern heraus „geht“. Die Fragen zu den Giften und Krankheiten seien „schwer“. Fachbegriffe wie „Verschmelzen“ „hätten gefehlt“.
- Gruppe 2: Die Vp meinten, man könne mit den Bildern „alles beantworten“ oder zumindest, dass „es ging“. Die Fragen mit den Giften seien „schwieriger“, die Fragen, die „direkt auf die Bilder bezogen“ waren, seien „gut“. Beispielsweise seien die Fragen zu Alzheimer, zu der Schwarzen Witwe oder zu der Wechselwirkung von Atropin und Muscarin „etwas schwerer“.
- Gruppe 3: Nach Meinung der Vp konnte man die Fragen „an sich schon“ beantworten, wenn „man ein bisschen überlegt“ hat. Eine Vp meinte, dass sie „manche Fragen nicht beantworten konnte“, da die Antworten für sie „auf der Bilderfolge nicht erkennbar waren (z.B. was zwischen zwei Abläufen passiert)“.

Der Nachtest wurde in allen drei Gruppen oft als „nicht so schwierig“ beurteilt, lediglich die Fragen zu den Nervengiften schienen eine höhere Komplexität für die Probanden aufzuweisen. Die Versuchspersonen gaben ebenfalls an, dass es „ging“, die Fragen aus den Bildern heraus zu beantworten. Eine Darstellung der synaptischen Prozesse anhand von

ausschließlich Bildern schien somit ausreichend zu sein, um Wissens- und Inferenzitems beantworten zu können.

Bezüglich der Frage, ob man für die Lösung des Nachtests ein beschriftetes Standbild benötigte, gaben vier von fünf Versuchspersonen in der Gruppe, die keine solche Hilfestellung hatte, an, dass sie ein Bild „nicht gebraucht“ hätten. In den anderen beiden Gruppen, bei denen das Standbild während des Nachtests gezeigt wurde, war eine häufigere Meinung, dass man es ohne dieses Bild „nicht ganz so gut hingekriegt“ hätte. Nur eine Versuchsperson war der Meinung, es sei ohne das Standbild „nicht zu schaffen“.

4.1.4 Diskussion

Das Ziel der Vorpilotstudie war zunächst, die entwickelten Testverfahren auf ihre Güte hin zu überprüfen. Bei dem Vorwissenstest war das Cronbachs Alpha zufriedenstellend ($\alpha = .64$), und die Mehrheit der Items wies akzeptable bis gute Trennschärfen auf. Daher wurden alle Items in die Pilotstudie übernommen, um die Aufgaben nochmals an einer größeren Stichprobe zu testen. Einschränkend muss gesagt werden, dass die statistischen Schwierigkeitswerte der Aufgaben sehr niedrig waren, was auf eine hohe Komplexität der Items für die Probanden hindeutete. Da die Teilnehmer jedoch bewusst in der Form rekrutiert wurden, dass sie wenig Vorwissen zu Synapsen haben sollten, war dieses Ergebnismuster zu erwarten gewesen. Da zudem die Aufgaben, theoretisch gesehen, Aspekte des späteren Lern- und Nachtestmaterials abfragten, wurde in Hinblick auf die Pilotstudie entschieden, die Items mit kleineren Vereinfachungen (einfachere Fachbegriffe, Vorgabe von Antwortlängen) in die Pilotstudie zu übernehmen.

Bei den Wissensskalen der Nachtests stellten sich die Reliabilitäten als durchaus schwächer dar (Version 1: $\alpha = .27$; Version 2: $\alpha = -1.31$), wobei der negative Wert der zweiten Version der Skala durch den geringen Stichprobenumfang der Untersuchungsgruppe mit erklärt werden könnte. Daher wurden für die Itemelimination eher die Werte der ersten Version herangezogen, da die dortige Stichprobengröße etwas höher war und damit die Schätzungen der Trennschärfen und der Schwierigkeiten größere Stabilitäten aufwiesen. Da die Aufgaben der beiden Skalen bis auf ein Item identisch waren, war ein solches Vorgehen gerechtfertigt. Auffällig an den Items der Nachwissensskala waren die mehrheitlich hohen Schwierigkeitswerte, was als Hinweis darauf gewertet wurde, dass die Aufgaben inhaltlich zu einfach für die Probanden konstruiert worden waren. Es wurde beispielsweise oft nur nach einem Prozess gefragt, den die Probanden in einem Satz oder sogar nur in einem Wort niederschreiben sollten. Die hohen Schwierigkeiten könnten mit ein Grund für die schwächeren Trennschärfen der Items gewesen sein. Deshalb wurde bei der Itemelimination in der Form vorgegangen, dass die Aufgaben aus der Vorpilotstudie übernommen wurden, die etwas bessere Trennschärfen hatten bzw. durch deren Wegnahme man das Alpha der Skala noch weiter verringert hätte. Auf diese Art und Weise gingen sechs der zwölf Items der Nachwissensskala (Version 1) in den Nachtest der Pilotstudie mit ein. Zudem wurden sechs neue Items konstruiert, die konzeptuell schwieriger angesiedelt waren, indem sie beispielsweise nach zwei oder mehr der gelernten synaptischen Vorgänge fragten und die Richtigkeit der Reihenfolge der Prozesse berücksichtigten.

Die Reliabilitäten der Inferenzskalen konnten in der Vorpilotstudie als zufriedenstellend bis gut bezeichnet werden (Version 1: $\alpha = .84$; Version 2: $\alpha = .71$). Da der Wert der ersten Version auf einer größeren Stichprobe beruhte, und da die Items dieser Skala zum größten Teil zufriedenstellende Trennschärfen hatten, wurden 15 der 18 Aufgaben in den Nachtest der Pilotstudie übernommen, wobei in wenigen Fällen sprachliche Umformulierungen oder Neukonstruktionen nötig waren. Aus der konzeptuell schwieriger angelegten zweiten Version der Inferenzskala wurden drei weitere Items in den Nachtest integriert, die sich durch gute bis sehr gute Trennschärfen auszeichneten

Zur Beantwortung der zweiten Forschungsfrage, ob ein Lernprozess mit Bildern allein ausreichend war, um anschließend Wissens- und Inferenzitems zu der Thematik lösen zu können, wurden zunächst die Ergebnisse der deskriptiven Statistiken herangezogen. Bei den Wissensitems konnte gezeigt werden, dass alle Untersuchungsgruppen überdurchschnittlich richtige Lösungshäufigkeiten zeigten (Gruppe 1: $M = 20.40$ von 24.00, Gruppe 2: $M = 20.20$ von 24.00, Gruppe 3: $M = 18.00$ von 24.00) und mit den Aufgaben anscheinend sehr gut zurechtkamen, so dass sogar schwierigere Items für die Pilotstudie nachkonstruiert werden mussten. Bei den Aufgaben, bei denen die Inferenzfähigkeiten der Probanden gefragt waren, wurde bei zwei der Gruppen im Schnitt die Hälfte der Aufgaben richtig gelöst (Gruppe 1: $M = 19.60$ von 36.00, Gruppe 3: $M = 17.25$ von 36.00), die Gruppe 2 erreichte sogar noch einen etwas besseren Score ($M = 26.00$ von 36.00). Aus diesen Ergebnissen konnte zunächst geschlossen werden, dass eine bildliche Darstellung der synaptischen Prozesse genügte, um einen anspruchsvolleren Nachtest bearbeiten zu können. Ergänzend hierzu wurden auch die Antworten der beiden Fragen aus dem Interview der Vorpilotstudie bewertet, die sich mit der von den Probanden empfundenen Schwierigkeit des Nachtests befassten bzw. bei denen die Teilnehmer angeben sollten, ob sie alle Items aus dem Lernmaterial heraus lösen konnten. In jeder Untersuchungsgruppe wurde der Nachtest als „nicht so schwierig“ beurteilt, und die Versuchspersonen waren der Meinung, dass es „ging“, die Aufgaben aus den Bildern heraus zu beantworten. Diese subjektiven Daten bestätigten das obige quantitative Muster, das in den deskriptiven Statistiken gefunden wurde, dass die Wissens- und Inferenzitems des Nachtests nach einem ausschließlich piktoralen Lernvorgang bewältigt werden konnten. Daher wurde auf die Integration von textlicher Prozessinformation in das Lernmaterial der Pilotstudie verzichtet.

Die dritte Forschungsfrage thematisierte die Bearbeitungszeiten des Vorwissens- und des Nachtests. Die Probanden benötigten im Schnitt ca. 3 bis 5 min (bei zwei Ausreißerwerten),

um den Vorwissenstest zu bearbeiten. Der Test stellte somit ein zeitökonomisches Verfahren dar, was unter anderem daran lag, dass die Probanden meist nicht über viel Vorwissen verfügten und daher die meisten Aufgaben nicht beantworten konnten. Der Faktor „Zeit“ stellte bei diesem Verfahren kein Hindernis dafür dar, es in seiner bestehenden Form mit halboffenen Items in die Pilotstudie zu übernehmen.

Die Bearbeitung des Nachttests nahm in den drei Gruppen durchschnittlich ca. 27 bis 37 min in Anspruch. Gruppe 2, die den Test ohne das beschriftete Standbild löste, benötigte die meiste Zeit ($M = 36 \text{ min } 54$), was aber auch daran liegen konnte, dass diese Gruppe gute Leistungen in der Inferenzskala zeigte und die Aufgaben mehrheitlich vollständig ausgefüllt hat. Gruppe 3 brauchte ungefähr gleich lang ($M = 29 \text{ min } 37$) wie Gruppe 1 ($M = 26 \text{ min } 55$) zur Lösung des Tests, obwohl sie mit konzeptuell teilweise schwierigeren Items konfrontiert war. Das Maximum in den Bearbeitungszeiten lag bei ca. 45 min. Insgesamt gesehen, war die Zeitdauer des Nachttests angesichts seiner anspruchsvolleren, schriftlich zu beantwortenden Items für das Gesamtdesign akzeptabel.

Die Verständlichkeit des Lernmaterials, die die Forschungsfrage 4 zum Inhalt hatte, schien für die Probanden insgesamt gegeben zu sein. Die Bilder wurden durchgängig als „gut“ und „anschaulich“ beschrieben, weshalb sie auch zur Orientierung für die Produktion der Animation in den weiteren Studien dienten. Ein öfters genannter Kritikpunkt bestand darin, dass die Spaltung des Transmitters im Lernmaterial nicht erkannt werden konnte, da diese in der Vorpilotstudie zusammen mit dem Reuptake des Acetylcholins auf einem Bild dargestellt wurde. Aus diesem Grund wurde in der Gruppe 3 der Vorpilotstudie und auch in der Materialentwicklung der weiteren Studien angestrebt, die Spaltung des Transmitters auf einem separaten Bild zu zeigen und sie mit einer kurzen textlichen Information zu versehen.

Die fünfte Forschungsfrage fokussierte sich darauf, ob die Lernzeiten für das beschriftete Standbild und für die Prozessbilder angemessen waren. Die Probanden beurteilten die Lernzeiten im Großen und Ganzen als „passend“, wobei sie in der Tendenz angaben, für das Standbild eher mehr und für die Prozesse eher weniger Zeit zu benötigen. Bei den Prozessbildern wurde für die Pilotstudie an einer Lernzeit von 4 min festgehalten, um eine sichere Lösbarkeit des Nachttests zu gewährleisten. Da bei dem Standbild entschieden wurde, es bei der Bearbeitung des Nachttests nicht zu zeigen (Begründung, s.u.), und da die Probanden angaben, für das Bild etwas mehr Lernzeit zu benötigen, wurde diese von 2 min auf 2 min 30 in der Pilotstudie erhöht.

Um die Forschungsfrage 6 zu beantworten, ob ein mit Fachbegriffen beschriftetes Standbild notwendig war, um den Nachtest lösen zu können, wurden zunächst die Ergebnisse der deskriptiven Statistiken der Gruppe 2 herangezogen, die in der Untersuchung eine solche Hilfe nicht zur Verfügung hatte. Es zeigte sich, dass die Gruppe 2 in der Nachwissensskala eine durchschnittlich gleich gute Leistung erbrachte ($M = 20.20$) wie die anderen beiden Gruppen (Gruppe 1: $M = 20.40$; Gruppe 3: $M = 18.00$). In der Inferenzskala lag ihr Gesamtscore sogar höher ($M = 26.00$) als bei Gruppe 1 ($M = 19.60$) und 3 ($M = 17.25$), was aber auch an dem etwas höheren Vorwissen der Gruppe 2 liegen konnte. Eventuell könnte sich das Fehlen eines Standbildes sogar positiv auf die Testleistung ausgewirkt haben, da die Versuchspersonen nicht ständig ihre Aufmerksamkeit zwischen den Items und dem Bild teilen mussten („Split attention-Effect“, Sweller, Chandler, Tierney & Cooper, 1990). In dem Interview gaben zudem vier der fünf Versuchspersonen der Gruppe 2 an, dass sie ein „Standbild nicht gebraucht hätten.“ Die Probanden der anderen Gruppen waren der Meinung, dass sie den Nachtest ohne das Standbild „nicht ganz so gut hingekriegt hätten.“ Es wurde daher entschieden, in der Pilotstudie das Standbild den Probanden nicht mehr während der Lösung des Nachtests zur Verfügung zu stellen, sondern es nur in das Lernmaterial zu integrieren. Die Lernzeit für das Bild wurde etwas höher gewählt als in der Vorpilotstudie (2 min 30), und es wurde eine Beschriftungsaufgabe zur Erfassung des begrifflichen Wissens in das Testheft der Pilotstudie aufgenommen, um zu überprüfen, ob und wie gut die Probanden die Begriffe des Standbildes enkodieren und dann auch abrufen konnten.

4.2 Pilotstudie

4.2.1 Fragestellung und Hypothesen

Im Vergleich zur Vorpilotstudie war der Versuchsablauf der Pilotstudie länger und reichhaltiger, da der Lernphase und dem Ausfüllen des Vorwissens- und des Nachtests noch weitere Messinstrumente hinzugefügt wurden, um beispielsweise die Intelligenz oder die aktuelle Lernmotivation der Probanden zu messen. Diese Veränderungen machten es notwendig, in der Pilotstudie nochmals zu überprüfen, ob der Versuchsablauf stimmig und praktisch umsetzbar war. Aus diesem Grund sollte in der Pilotstudie mittels Beobachtung die folgende Forschungsfrage untersucht werden:

Forschungsfrage 1: Ist der Versuchsablauf stimmig und das Experiment in der geplanten Form durchführbar?

Zudem waren nach der Vorpilotstudie die Items der Wissens- und der Inferenzskala des Nachtests, sowie in geringerem Umfang auch die des Vorwissenstests, überarbeitet bzw. neu konstruiert worden, so dass ihre Gütekriterien in der Pilotstudie an einer größeren Stichprobe getestet werden sollten. Es war auch eine Beschriftungsaufgabe zur Erfassung synaptischer Fachbegriffe in den Versuchsablauf integriert worden, deren Items ebenfalls analysiert werden sollten. Daraus ergab sich als eine weitere zu untersuchende Forschungsfrage:

Forschungsfrage 2: Sind die Gütekriterien des Vorwissenstests, der Beschriftungsaufgabe zum begrifflichen Wissen sowie der Wissens- und der Inferenzskala des Nachtests, die die selbst konstruierten Verfahren des Experimentes darstellen, zufriedenstellend?

Ein weiteres Anliegen der Pilotstudie war, die theoretisch hergeleiteten Grundannahmen zu den Haupteffekten zwischen den Bedingungen von Teilstudie 1 und Teilstudie 2 explorativ zu testen, um erste Hinweise auf die Stärke und Richtung der möglicherweise vorhandenen Effekte zu erhalten. Das bedeutet allerdings auch, dass die Ergebnisse im anschließenden Diskussionsteil nicht zu weitreichend interpretiert werden dürfen, da durch die geringere Stichprobengröße noch keine stabilen Schätzungen der Kennwerte erreicht werden können.

Aufgrund der theoretischen Überlegungen, die in Abschnitt 3.1.1 gemacht wurden, wurden die Hypothesen zwischen der Animations- und der seriell schnellen Bildbedingung der ersten Teilstudie zugunsten der Bildbedingung gerichtet formuliert:

Hypothese 1a: Personen, die mit seriell schnellen Key-Frames lernen, erzielen signifikant bessere Werte in den Wissensaufgaben eines Nachtests, als Personen, die mit Animationen lernen.

Hypothese 1b: Personen, die mit seriell schnellen Key-Frames lernen, erzielen signifikant bessere Werte in den Inferenzaufgaben eines Nachtests, als Personen, die mit Animationen lernen.

Hypothese 1c: Personen, die mit seriell schnellen Key-Frames lernen, empfinden signifikant weniger Cognitive Load beim Lernen, als Personen, die mit Animationen lernen.

Hypothese 1d: Personen, die mit seriell schnellen Key-Frames lernen, empfinden signifikant weniger Cognitive Load beim Lösen des Nachtests, als Personen, die mit Animationen lernen.

Analog zu den theoretischen Grundannahmen wurden auch die Hypothesen zwischen der Animations- und der seriell langsamen Bildbedingung gerichtet formuliert:

Hypothese 2a: Personen, die mit seriell langsamen Key-Frames lernen, erzielen signifikant bessere Werte in den Wissensaufgaben eines Nachtests, als Personen, die mit Animationen lernen.

Hypothese 2b: Personen, die mit seriell langsamen Key-Frames lernen, erzielen signifikant bessere Werte in den Inferenzaufgaben eines Nachtests, als Personen, die mit Animationen lernen.

Hypothese 2c: Personen, die mit seriell langsamen Key-Frames lernen, empfinden signifikant weniger Cognitive Load beim Lernen, als Personen, die mit Animationen lernen.

Hypothese 2d: Personen, die mit seriell langsamen Key-Frames lernen, empfinden signifikant weniger Cognitive Load beim Lösen des Nachtests, als Personen, die mit Animationen lernen.

Beim Vergleich zwischen seriell schnellen und seriell langsamen Key-Frames wurden jeweils ungerichtete Hypothesen aufgestellt, da beide Präsentationsmöglichkeiten sowohl Lernvorteile, als auch -nachteile mit sich bringen können.

Hypothese 3a: Personen, die mit seriell schnellen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Wissensaufgaben eines Nachtests, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 3b: Personen, die mit seriell schnellen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Inferenzaufgaben eines Nachtests, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 3c: Personen, die mit seriell schnellen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lernen, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 3d: Personen, die mit seriell schnellen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lösen des Nachtests, als Personen, die mit seriell langsamen Key-Frames lernen.

In der zweiten Teilstudie der ersten Studie ging es darum, parallel angeordnete mit zum einen seriell schnellen und zum anderen mit seriell langsamen Bildern zu vergleichen. Auch hier wurden zur explorativen Testung in der Pilotstudie jeweils ungerichtete Hypothesen aufgestellt:

Hypothese 4a: Personen, die mit parallelen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Wissensaufgaben eines Nachtests, als Personen, die mit seriell schnellen Key-Frames lernen.

Hypothese 4b: Personen, die mit parallelen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Inferenzaufgaben eines Nachtests, als Personen, die mit seriell schnellen Key-Frames lernen.

Hypothese 4c: Personen, die mit parallelen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lernen, als Personen, die mit seriell schnellen Key-Frames lernen.

Hypothese 4d: Personen, die mit parallelen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lösen des Nachttests, als Personen, die mit seriell schnellen Key-Frames lernen.

Hypothese 5a: Personen, die mit parallelen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Wissensaufgaben eines Nachttests, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 5b: Personen, die mit parallelen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Inferenzaufgaben eines Nachttests, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 5c: Personen, die mit parallelen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lernen, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 5d: Personen, die mit parallelen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lösen des Nachttests, als Personen, die mit seriell langsamen Key-Frames lernen.

4.2.2 Methoden

4.2.2.1 Stichprobe

Es nahmen 24 Studenten, vorwiegend der Universität Landau zugehörig, an der Untersuchung teil. Die von den Teilnehmern studierten Fächer reichten von Lehramts- und diversen Magisterstudiengängen bis hin zu Diplom-Sozialwissenschaften. Ausgeschlossen wurden wiederum Psychologie-, Biologie- und Umweltwissenschaftsstudenten sowie Personen, die schon an der Vorpilotstudie teilgenommen hatten. Die Anwerbung erfolgte durch eine Rundmail, die an alle Studenten der Universität ging. Nach der Durchführung der Pilotstudie musste eine Person wegen ihres auffällig unmotivierten Ausfüllens des Nachtests von der Auswertung ausgeschlossen werden, wodurch 23 Teilnehmer in die Analysen eingingen. Die Probanden waren durchschnittlich 23.2 Jahre alt ($SD = 4.2$), der Altersrange betrug 19 bis 39 Jahre. Das Geschlechterverhältnis lag wieder bei ungefähr einem Drittel Männer zu zwei Dritteln Frauen (30,4% zu 69,6%, entsprach 7 Männern und 16 Frauen), was bei der Gesamtverteilung der Geschlechter innerhalb der Hochschule zu erwarten war. In der Stichprobe waren Frauen somit stärker repräsentiert.

4.2.2.2 Design

Um eine möglichst hohe Stichprobengröße für die Itemanalysen zu erhalten und um die Durchführbarkeit aller Bedingungen zu überprüfen, handelte es sich bei der Pilotstudie um ein einfaktorielles Design (Faktor = Präsentationsform der Lernmaterialien zu den synaptischen Prozessen) mit vier Treatmentstufen. Die Hypothesen der beiden Teilstudien wurden somit explorativ innerhalb einer Pilotstudie getestet. Für die Berechnung der Mittelwertsvergleiche wurde das Gesamtdesign in zwei einfaktorielle Designs mit jeweils drei Treatmentstufen aufgeteilt.

Eine Bedingung erarbeitete sich die Inhalte der synaptischen Vorgänge mit Hilfe einer Animation, die mehrmals hintereinander gezeigt wurde. Die zweite Gruppe lernte die Prozesse anhand von zehn Bildern, die seriell hintereinander geschaltete Key-Frames darstellten, die in höherer Geschwindigkeit wechselten, so dass die Teilnehmer immer nur ein Bild zu einem Zeitpunkt sahen. In einer dritten Gruppe wurden die Prozesse mit Hilfe von ebenfalls seriell hintereinander gezeigten Key-Frames vermittelt, wobei diese aber langsamer wechselten und ein einzelnes Key-Frame länger zu sehen war. Diese Gruppe

stellte einen Kontrast zu der seriell schnellen Bildbedingung dar und sollte einen möglichen Einfluss der Präsentationsgeschwindigkeit der Bilder abklären. In einer vierten Bedingung wurde die Darbietungsform der Bilder variiert, indem die Key-Frames parallel auf dem Bildschirm angeordnet zu sehen waren. Die Probanden wurden den einzelnen Gruppen randomisiert zugewiesen, so dass in der seriell langsamen Bildbedingung 5 und in den übrigen Bedingungen jeweils 6 Teilnehmer waren.

Abhängige Variablen waren die Wissens- und die Inferenzskala eines Nachttests zu synaptischen Inhalten, sowie der Cognitive Load, der zum einen durch das Lernen des Materials und zum anderen durch das Lösen der Nachtestaufgaben verursacht wurde. Dies diente dazu, neben den Leistungsmaßen auch einen Anhaltspunkt für den Grad der instruktionalen Effizienz der Lernmaterialien zu erhalten (Paas et al., 2007). Als Kontrollvariablen, von denen ein höherer Einfluss auf die abhängigen Variablen angenommen wurde, wurden das Vorwissen zu Synapsen, die aktuelle Lernmotivation, das Begriffswissen zu Synapsen sowie räumliche, verbale und Merkfähigkeiten innerhalb des Designs mit erhoben.

4.2.2.3 Lernmaterialien

Extraktion der Key-States der synaptischen Erregungsweiterleitung

Innerhalb der Materialentwicklung der Pilotstudie erfolgte zunächst für die drei Bildbedingungen eine Extraktion der Key-States des synaptischen Übertragungsvorganges analog zu den theoretischen Vorüberlegungen, die hinsichtlich ihrer Charakterisierung gemacht wurden. Die synaptische Übertragung stellt einen Vorgang dar, der in mehrere Teilprozesse gegliedert werden kann. Die Segmentierung fand zum einen anhand des Kriteriums maximaler physikalischer Veränderungen statt, die sich dort ereignen, wo ein jeweils anderer Bestandteil der Synapse seine Aktivität aufnimmt. Dies ist meist dort der Fall, wo sich innerhalb der Übertragung kausale Schnittstellen zwischen den Teilprozessen befinden, die ein weiterer entscheidender Gesichtspunkt für das Setzen einer Ereignisgrenze waren (Zacks & Tversky, 2001). Somit stellt die synaptische Erregungsweiterleitung eine Kausalkette mehrerer Unterereignisse dar, wobei ein Teilprozess auslösendes Moment für den darauffolgenden Vorgang ist.

Wendete man diese Art der Segmentierung an, erhielt man eine Gliederung des Gesamtvorganges in die folgenden Teilprozesse:

- Das Aktionspotential kommt an.
- Die Calciumkanäle öffnen sich, das Calcium strömt ein.
- Die synaptischen Bläschen wandern auf die präsynaptische Membran zu und verschmelzen mit ihr. Das Acetylcholin wird in den Spalt entleert.
- Das Acetylcholin diffundiert im synaptischen Spalt und dockt an die Acetylcholinrezeptoren der Natriumkanäle an.
- Die Natriumkanäle öffnen sich, Natrium strömt ein.
- Ein Aktionspotential entsteht.
- Das Acetylcholin wird von der Cholinesterase gespalten.
- Das Acetylcholin wird in den synaptischen Endknopf wiederaufgenommen.

Diese Darstellung entspricht einer Aufzählung der Teilereignisse, die man von ihrem Abstraktionsgrad her auf einer mittleren bzw. basalen Ebene einordnen könnte. Es wäre sowohl eine feinkörnigere als auch eine gröbere Beschreibung der Vorgänge möglich. Die mittlere Ebene eignet sich zum Lernen besonders gut, da sie, relativ zu anderen Ebenen gesehen, viel Information vermittelt (Rosch, 1978). Auch aus diesem Grund wurde die obige Partonomie für die Extraktion der Key-States favorisiert.

Abgeleitet von den theoretischen Vorüberlegungen zu den Merkmalen von Key-States wurde als Key-Frame (= Bild, das ein Key-State zeigt) von jedem der obigen Teilvorgänge ein mittleres Bild sowie die kausalen Schnittstellen zwischen zwei aufeinanderfolgenden Prozessen gewählt. Diese Schnittstellen kann man sowohl als ein eigenes Bild, aber auch als in die mittleren Bilder integriert darstellen, beispielsweise mit Hilfe von Bewegungspfeilen. Letzteres Vorgehen wurde auch in der aktuellen Materialentwicklung angewandt, da auf diese Art und Weise eine ökonomischere Anzahl von Bildern entstand, was unnötige Such- und Vergleichsprozesse zwischen den Key-Frames ersparte und somit „extraneous load“ (Sweller, 1988) reduzierte.

Der synaptische Übertragungsvorgang wurde daher durch die folgenden zehn Key-Frames in den drei Bildbedingungen der Pilotstudie jeweils repräsentiert (zur Ansicht der Bilder, s. Anhang B1, Dateien: „BildParallel.swf“ und „BildSeriellSchnell.swf“ auf der beigelegten CD):

Bild 1: Mittleres Bild:

Aktionspotential kommt an.

Die kausale Schnittstelle zur Öffnung der Calciumkanäle wird durch Bewegungspfeile integriert.

Bild 2: Mittleres Bild:

Die Calciumkanäle öffnen sich, das Calcium strömt ein.

Bild 3: Kausale Schnittstelle:

Calcium stößt die Bläschen an.

Bild 4: Mittleres Bild:

Die synaptischen Bläschen wandern auf die präsynaptische Membran zu und verschmelzen mit ihr. Das Acetylcholin wird in den Spalt entleert.

Die kausale Schnittstelle zur Diffusion des Acetylcholins ergibt sich durch die Darstellung zwangsläufig.

Bild 5: Mittleres Bild:

Das Acetylcholin diffundiert im synaptischen Spalt und dockt an die Acetylcholinrezeptoren der Natriumkanäle an.

Bild 6: Kausale Schnittstelle:

Das Acetylcholin dockt an die Natriumkanäle an und öffnet sie.

Bild 7: Mittleres Bild:

Die Natriumkanäle öffnen sich, Natrium strömt ein.

Bild 8: Mittleres Bild:

Ein Aktionspotential entsteht.

Die kausale Schnittstelle vom Natriumioneneinstrom her ist noch sichtbar.

Bild 9: Mittleres Bild:

Das Acetylcholin wird von der Cholinesterase gespalten.

Bild 10: Mittleres Bild:

Das Acetylcholin wird in den synaptischen Endknopf wiederaufgenommen.

Zwischen den letzten Teilvorgängen sind kausale Schnittstellen bildlich nicht eindeutig darstellbar, weswegen für die Präsentation nur mittlere Bilder dieser Ereignisse gewählt wurden.

Generell wurden in den Bildern teilweise Bewegungspfeile eingezeichnet, um beispielsweise die Öffnung von Kanälen oder die Fließrichtung der Ionen deutlich zu machen.

Beschreibung der Materialien

Die Animation wurde, ebenso wie die Bildbedingungen, von einem Graphikdesigner mit dem Programm „Macromedia Flash 8“ produziert. Die Dateien lagen im „swf“-Format als „Shockwave Flash-Objekte“ vor (zur kompletten Ansicht aller vier Lerneinheiten, s. Anhang B1 auf der beiliegenden CD).

Bei allen Bedingungen sahen die Teilnehmer zunächst eine Instruktionsseite, auf der je nach ihrer Gruppenzuteilung die anstehende Lernaufgabe kurz beschrieben wurde. Auf der Instruktionsseite war rechts unten ein Pfeil eingebaut, der durch einen Klick mit der Maus aktiviert wurde, so dass die Probanden die Lerneinheit selbständig starten konnten, nachdem sie die Anweisungen gelesen und verstanden hatten.

Danach erschien ein Standbild, das eine Synapse im Ruhezustand zeigte. Die Teile der Synapse waren jeweils mit ihren Namen beschriftet. Es wurde bei der Materialentwicklung darauf geachtet, dass nur die Teile gezeigt und mit Begriffen versehen wurden, die später auch im Nachtest Verwendung fanden und die man daher zur Lösung der Wissens- und Inferenzaufgaben benötigte.

Es wurden somit zehn Teile der Synapse in das Standbild eingezeichnet und mit den folgenden Benennungen versehen:

- Calciumkanal
- Endknopf
- Spalt
- Nachsynaptische Membran
- Cholinesterase

- Natriumkanal
- Rezeptor
- Vorsynaptische Membran
- Transmitter: Acetylcholin
- Bläschen

Wie schon erwähnt, wurden im Vergleich zum Lehrbuch (Bayrhuber & Kull, 1989; Rosenzweig et al., 1999) die Bezeichnungen der Bestandteile vereinfacht, damit das Einprägen der Begriffe innerhalb kurzer Zeit möglich war. Das Standbild wurde den Teilnehmern 2 min 30 lang gezeigt, dann war ein automatischer Wechsel zu den Lerneinheiten über die Prozesse an der Synapse geschaltet. Die Begriffe des Standbildes standen bei der späteren Lösung des Nachttests den Probanden nicht mehr zur Verfügung.

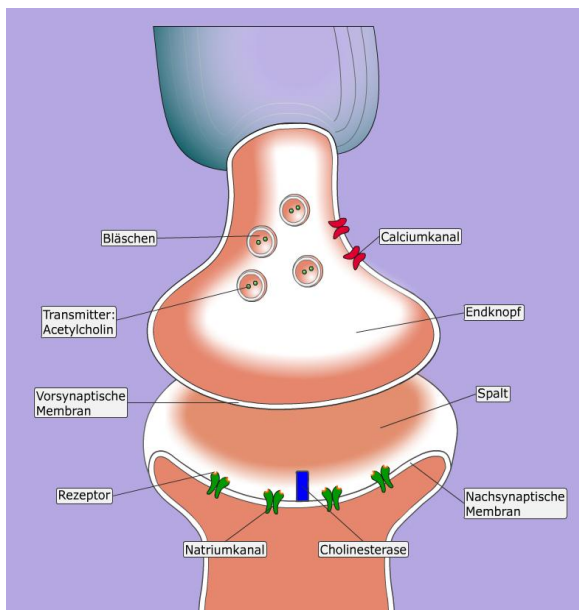


Abbildung 4.2.1: Darstellung des beschrifteten Standbildes

Die Prozesse, die an einer Synapse ablaufen und deren Inhalte die Teilnehmer lernen und verstehen sollten, wurden, wie im Abschnitt „Design“ (s.o.) beschrieben, in den vier Gruppen unterschiedlich dargestellt. Innerhalb der Pilotstudie wurden die Teilnehmer in die Bedingungen „Animation“, „Bild seriell schnell“, „Bild seriell langsam“ und „Bild parallel“ eingeteilt, die im graphischen Design ihrer Lernmaterialien sowohl Gemeinsamkeiten, als auch Unterschiede aufwiesen.

Bei allen Bedingungen lag eine einheitliche Gestaltung der Lerneinheiten hinsichtlich der Farben und Formen der Synapse und ihrer Teile vor. Wie schon bei dem Standbild, wurde bei der Darstellung der Prozesse darauf geachtet, dass alle Bestandteile, die innerhalb der Abläufe einer Synapse eine wesentliche Rolle spielen, deutlich und farblich gut voneinander abgesetzt gezeichnet wurden. Kein Detail stach sonderlich hervor, um nicht durch perzeptuelle Merkmale von den relevanten konzeptuellen Inhalten abzulenken. Die Lernabschnitte zu den dynamischen Vorgängen der Synapse hatten alle eine Dauer von ca. 4 min, wobei zwischen den einzelnen Bedingungen leichte technische Schwankungen von 2 bis 3 s möglich waren. In der Animation, wie auch in den Prozessbildern, wurden zusätzlich zu dem vorherigen Standbild weitere vier Objekte innerhalb des synaptischen Vorganges beschriftet („Aktionspotential“, „Calciumionen“, „Natriumionen“, „gespaltener deaktivierter Transmitter“), da die Zusammenhänge auf diese Art und Weise besser verständlich und die Aufgaben des Nachtests prinzipiell lösbar blieben. Allerdings war es hier von Wichtigkeit, dass bei diesen textlichen Angaben nur statische Informationen durch die Verwendung von Substantiven und Adjektiven gegeben wurden. Prozessinformationen hingegen sollten nur durch die Bilder vermittelt werden.

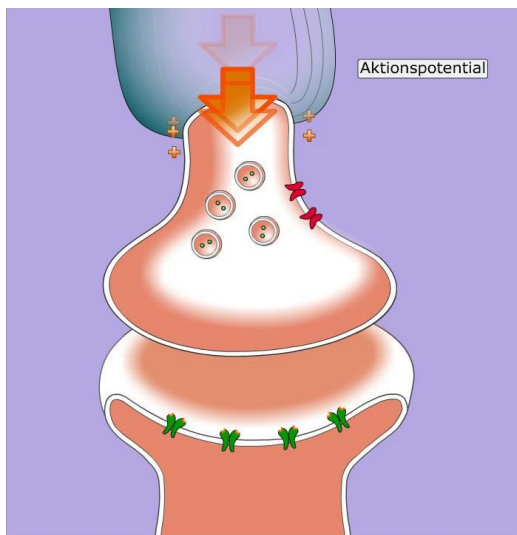


Abbildung 4.2.2: Beispielbild für die graphische Gestaltung der Synapse

Die Animation wurde in selbiger Bedingung sieben Mal hintereinander gezeigt. Ein Durchgang dauerte ca. 34 s lang, wodurch sich eine Gesamtlernzeit von ca. 4 min ergab, die sich in der Vorpilotstudie als für die Thematik geeignet erwiesen hatte. Zwischen den einzelnen Durchgängen war jeweils 1 s Pause geschaltet. Die Geschwindigkeit der

Animation wurde so gewählt, dass die Prozesse perzeptuell gut erkannt und konzeptuell auch von Laien verstanden werden konnten. Erste Hinweise hierzu waren schon innerhalb der Materialentwicklung durch informelle Personenbefragungen erhalten worden, letzten Endes sollte aber auch die Durchführung der Pilotstudie Aufschluss darüber geben, ob die gewählte Geschwindigkeit noch zu verändern war oder belassen werden konnte. Durch die sieben Wiederholungen hatten die Teilnehmer mehrere Lerndurchgänge zu den Prozessen, wodurch eine zu hohe Schwierigkeit der dynamischen, flüchtigen Darstellung vermieden wurde.

Die Lerneinheit mit den seriell schnellen Bildern umfasste wieder sieben Wiederholungen des Gesamtvorganges, der durch die oben genannten zehn Bilder repräsentiert wurde. Die Frames waren in Reihe geschaltet, so dass jeweils nur ein Bild zu einem Zeitpunkt zu sehen war. Ein einzelner Lerndurchgang hatte eine Dauer von ca. 34 s, wodurch jedes Bild 3,4 s lang auf dem Screen erschien. Zwischen den Durchgängen war wieder jeweils 1 s Pause eingebaut. Durch ihre Gestaltungsmerkmale ähnelte diese Bildbedingung sehr der Lerneinheit mit der Animation, wodurch die beiden Bedingungen hinsichtlich ihrer Präsentationsgeschwindigkeit und der Zahl der Lerndurchgänge direkt miteinander vergleichbar waren.

Als Kontrast zu dieser schnelleren Darstellungsform wurden als dritte Bedingung die zehn Frames wieder seriell hintereinander in Reihe geschaltet, allerdings dauerte ein Lerndurchgang nun 80 s lang. Die Präsentationsdauer eines einzelnen Bildes betrug somit 8 s, und die Zahl der Wiederholungen reduzierte sich durch diese Parameter auf drei Lerndurchgänge. Zwischen den Durchgängen wurde wieder jeweils 1 s Pause gelassen.

Die beiden seriellen Bedingungen waren, wie die Animation, vollständig systemgesteuert angelegt.

Bei der parallelen Bildbedingung waren die Frames neben- sowie untereinander auf dem Bildschirm angeordnet, so dass sie alle zu jedem Zeitpunkt sichtbar sein konnten. Da die Bilder eine solche Größe haben sollten, dass die Prozesse in ihren Details gut zu erkennen waren, konnten immer nur zwei Bilder nebeneinander auf das Display angeordnet werden, wobei die nach unten folgenden Bilder noch jeweils zur Hälfte sichtbar waren. Die Bedingung wurde daher mit einer Leiste am rechten Bildschirmrand versehen, damit die Probanden durch ein Scrollen mit der Maus alle Frames erschließen konnten. Dieses technische Feature war sehr intuitiv angelegt, so dass die Probanden mit seiner Nutzung aus dem Computeralltag her vertraut gewesen sein sollten. Durch diese leichte Form der Interaktivität konnten alle Bilder von den Teilnehmern selbständig die vollen 4 min lang

angesteuert werden. Auf diese Weise konnte der Grundvorteil von parallel angeordneten Bildern, dass Personen selbst entscheiden können, welche Bilder sie wie lange inspizieren, weitestgehend erhalten werden.

Die Lerneinheiten aller vier Bedingungen schlossen mit einer Seite ab, auf der der Hinweis „Bitte wenden Sie sich an den Versuchsleiter“ zu lesen war.

4.2.2.4 Messinstrumente

Innerhalb der Untersuchung wurden als abhängige Variablen die Wissens- und die Inferenzskala eines Nachttests sowie der Cognitive Load (= mentale Belastung) der Lerneinheit und der Nachttestaufgaben erhoben.

Zur Erfassung der Kontrollvariablen: Vorwissen, aktuelle Lernmotivation, begriffliches Wissen zu Synapsen, räumliche, verbale und Merkfähigkeiten war zudem die Integration weiterer Verfahren in die Testabfolge notwendig.

Die in dem Experiment zur Anwendung gekommenen Messinstrumente sollen im Folgenden in der Reihenfolge ihrer Bearbeitung durch die Probanden vorgestellt und beschrieben werden.

Soziodemographischer Fragebogen

Der Aufbau des soziodemographischen Fragebogens war an den der Vorpilotstudie angelehnt (s. Abschnitt 4.1.2.4, und zur Ansicht des Fragebogens, s. Anhang B2 - 1). Eine Neuerung fand dahingehend statt, dass die Probanden mit Hilfe eines vorgegebenen Beispiels einen achtstelligen Code kreierte, der die Anonymität ihrer Daten sichern sollte. Der Code setzte sich aus den ersten beiden Buchstaben des Vornamens der Mutter und des Vaters, den ersten beiden Buchstaben des Geburtsortes sowie dem Geburtstag des Probanden zusammen. Somit war mit höherer Wahrscheinlichkeit gewährleistet, dass kein Code zwei Mal vergeben wurde (außer bei Zwillingsgeschwistern).

Vorwissenstest

Der Vorwissenstest der Pilotstudie entsprach in wesentlichen Gestaltungspunkten und -merkmalen der Version, die in der Vorpilotstudie verwendet worden war (s. Abschnitt 4.1.2.4 sowie die Anhänge A2 - 2 und B2 - 2). Bei einigen Items erleichterten sich die Fragestellungen durch eine Vorgabe der Antwortlänge und durch die Nutzung einfacherer Fachbegriffe. Ansonsten wurden die Aufgaben aus der Vorpilotstudie in gleichbleibender Form übernommen. Sie hatten zwar größtenteils niedrige Schwierigkeitsindizes aufgewiesen, was aber durch die differenzierten Fragestellungen der Items und durch die Verwendung einer Stichprobe, die bewusst wenig Vorwissen haben sollte, zu erwarten gewesen war. Zudem hatten bei der Skala ein zufriedenstellendes Cronbachs Alpha sowie bei der Mehrheit der Aufgaben befriedigende Trennschärfen vorgelegen. Angesichts der geringen Stichprobengröße innerhalb der Vorpilotstudie sollten die Items des Vorwissenstests in der Pilotstudie nochmals erprobt.

Hierzu wurde als Weiterentwicklung zur Vorpilotstudie für jedes Item ein Ratingsystem erstellt, das eine jeweils möglichst objektive Bewertung der Antworten ermöglichen sollte. Der Aufbau der Ratingsysteme orientierte sich an den Vorgaben von Mayring (2002). Sie enthielten eine dreistufige Kodierung, mit der die Antworten in „richtig“ (2 Punkte), „zum Teil richtig“ (1 Punkt) und „falsch / nicht beantwortet“ (0 Punkte) eingeteilt werden konnten und deren Skalenniveau in den statistischen Analysen als metrisch angenommen wurde. Bei der Kategorie „richtig“ wurde analog zu physiologischen Lehrbüchern (Bayrhuber & Kull, 1989; Rosenzweig et al., 1999) jeweils eine Musterantwort konstruiert, die als Vorlage für die Einstufung der Lösungen dienen sollte. Die Kategorie „zum Teil richtig“ enthielt meist solche Antworten, die nur Teile oder unpräzisere bzw. fehlerhafte Formulierungen der Musterlösung beinhalteten. Unter die Kategorie „falsch“ fielen alle anderen Antworten oder die Nichtbeantwortung der Aufgabe. Für die Erstellung der Kategorien wurden zudem die Daten der Vorpilotstudie herangezogen, wodurch das theoretisch deduktive durch ein empirisch induktives Vorgehen ergänzt wurde. Um die Beschreibungen der Kategorien anschaulicher und das Rating präziser zu machen, wurden jeweils zusätzlich Ankerbeispiele und Kodierregeln erstellt, die eine trennscharfe Zuordnung der Lösungen erleichtern sollten. Die Ratingsysteme sollten in ihrer Praktikabilität innerhalb der Pilotstudie erprobt und mit den Daten inhaltlich weiterentwickelt werden. Aufgrund der geringen Stichprobengröße wurde auf eine Berechnung der Interraterreliabilitäten der einzelnen Systeme in der Auswertung verzichtet (zur Ansicht der nach der Pilotstudie weiterentwickelten

Ratingsysteme, s. Anhang B3, Datei: „Ratingsysteme_Vorwissenstest_Pilot.pdf“ auf der beigefügten CD.)

Für die statistischen Analysen wurden die Punktwerte der Probanden für die einzelnen Aufgaben summiert, und der Gesamtscore wurde als eine potentielle Kovariate in den Rechnungen berücksichtigt, da dem Vorwissen in der Literatur eine moderierende Funktion beim Lernen mit Animationen und Bildern zugesprochen wird (Hegarty & Kriz, 2008).

Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen (FAM)
(Rheinberg, Vollmeyer & Burns, 2001)

Da Lernleistungen generell nicht nur von kognitiven Komponenten, sondern auch von Motivationsfaktoren determiniert werden (Atkinson, 1974), wurde in der Pilotstudie die aktuelle Lernmotivation der Probanden mit Hilfe einer adaptierten Form des „Fragebogens zur Erfassung aktueller Motivation in Lern- und Leistungssituationen (FAM)“ (Rheinberg, Vollmeyer & Burns, 2001) erhoben (s. auch Anhang B2 - 3). Es handelte sich hierbei um ein reliables, aufgabensensibles und valides Verfahren. Für eine genauere Beschreibung dieses standardisierten Tests sei auf die Literatur von Rheinberg et al., (2001) verwiesen. Der FAM beinhaltete vier Subskalen: Misserfolgsbefürchtung, Erfolgswahrscheinlichkeit, Interesse und Herausforderung, deren Summenscores jeweils als mögliche Kovariaten in die statistischen Analysen miteingingen.

Erfassung des Cognitive Loads

Wegen seiner einfachen Anwendbarkeit und seiner hohen externen Validität (Schnotz & Kürschner, 2007) wurde in der vorliegenden Studie ein Ratingverfahren gewählt, um den „cognitive load“ (Sweller, 1988) zu erfassen. Die mentale Belastung wurde zu zwei Messzeitpunkten während der Untersuchung erhoben. Zum einen wurde von den Probanden beurteilt, wie anstrengend sie die Lerneinheit empfanden, zum anderen sollten sie ihren Cognitive Load bei der Bearbeitung der Nachtestaufgaben bewerten. Für das Rating wurden jeweils zwei Items eingesetzt (s. auch Anhang B2 - 4), wobei das erste Item nach der wahrgenommenen Aufgabenschwierigkeit fragte und sich konzeptuell somit an die Vorlage von Bratfish, Borg und Dornic (1972) anlehnte. Die Probanden konnten auf einer Skala von 1 bis 9 einstufen, wie schwierig sie die Lernaufgabe bzw. die Nachtestaufgaben fanden

(1 = sehr, sehr leicht; 9 = sehr, sehr schwierig). Das jeweils zweite Item war eine übersetzte und adaptierte Form der „Cognitive Load Rating Scale“ von Paas (1992), bei dem die Probanden mit Hilfe eines neunstufigen Antwortformates einschätzen konnten, wie stark sie sich mental bei der Bearbeitung der Lernaufgabe bzw. der Nachtestaufgaben angestrengt haben (1 = sehr, sehr wenig; 9 = sehr, sehr stark).

Für die Gütekriterien der Reliabilität und der Sensibilität hat Paas für die „Cognitive Load Rating Scale“ mit Hilfe von zwei Studien Nachweise erbracht (Paas, 1992; Paas & van Merriënboer, 1994). Das Cronbachs Alpha der Skala lag hier bei $\alpha = .90$ bzw. $\alpha = .82$, zudem erwies sich das Verfahren als sensibel für die Komplexität einer Aufgabe.

Für die Ermittlung des empfundenen Cognitive Loads wurden die von jedem Probanden angekreuzten beiden Werte auf den Items jeweils aufsummiert und stellten als Gesamtscores zwei weitere abhängige Variablen der Studie dar. Dies hatte den Sinn, die Effizienz der Lernbedingungen bestimmen zu können. Leistungsmaße, wie die Scores auf einer Wissens- oder auf einer Inferenzskala, haben das Defizit, keine Auskunft darüber zu geben, mit welchen „kognitiven Kosten“ die Inhalte erworben wurden (Paas et al., 2007). Unterschiede in der Qualität des Lernmaterials können sich demnach nicht nur in Leistungsindizes, sondern auch im empfundenen Cognitive Load der Teilnehmer widerspiegeln, der somit eine wertvolle Ergänzung zur Beurteilung von verschiedenen Lernbedingungen darstellt.

Test zur Erfassung des begrifflichen Wissens zu Synapsen

Dieser Test bezog sich auf das Standbild, das zu Beginn der Lerneinheit 2 min 30 lang gezeigt worden war und wies die Form einer Beschriftungsaufgabe auf (s. auch Anhang B2 - 5). Es wurden die zehn in dem Standbild gezeigten Benennungen der Teile einer Synapse abgefragt, indem die Probanden die richtigen Begriffe in Kästchen eintragen mussten, die auf den jeweils zugehörigen Bestandteil der Synapse zeigten. Für die Bearbeitung des Tests gab es keine Zeitbegrenzung (Power-Test).

Die Teilnehmer erhielten für jede richtige Antwort einen Punkt, wobei bei der Bewertung kleinere Schreibfehler toleriert wurden. Die Summe der richtigen Antworten stellte den Gesamtscore des begrifflichen Wissens des Probanden dar.

Dieses Wissen ging als Kontrollvariable in die Auswertung mit ein, da die Kenntnis der richtigen Fachbegriffe, die die Teile einer Synapse benennen, das Verständnis der

Nachtestitems und somit deren Beantwortung sowie die Präzision der gegebenen Lösungen beeinflusste.

Nachtest

Der Nachtest umfasste, wie die beiden Versionen der Vorpilotstudie, 30 Items, die sich in 12 Wissens- und 18 Inferenzitems gliederten. Die Gestaltungsmerkmale wurden ebenfalls analog zur Vorpilotstudie gewählt (s. Anhang A2 - 3). In der Instruktion erhielten die Probanden den Hinweis, bei Fragen, die sie nicht beantworten konnten, ein „?“ oder ein „/“ zur Kennzeichnung zu machen und nach der Lösung einer Seite im Testheft nicht mehr zurückzublättern.

Was die Zusammensetzung der Wissensskala betraf, wurden sechs der zwölf Items aus der Vorpilotstudie übernommen und teilweise mit Vorgaben zu ihrer Antwortlänge ergänzt (zur Ansicht der Nachtestitems, s. Anhang B2 - 6). Es handelte sich hierbei um solche Aufgaben, die bei der Mehrheit der Probanden bessere Trennschärfen aufgewiesen hatten und durch deren Elimination sich der Messfehler der Skala erhöht hätte. Die Schwierigkeitsindizes der Items waren größtenteils hoch gewesen ($p > .80$), was ein Grund für das insgesamt eher niedrige Niveau der Trennschärfen in der Vorpilotstudie gewesen sein könnte.

Daher wurden für die Wissensskala der Pilotstudie sechs weitere Items konstruiert, die in ihrer konzeptuellen Schwierigkeit höher angesiedelt waren. Inhaltlich fragten die Items nach mehreren statt wie vorher nur nach einem Vorgang der Synapse und berücksichtigten bei der Richtigkeit der Antwort auch die Reihenfolge der Prozesse. (*Beispielitem*: „Nennen Sie die drei letzten Teilprozesse, mit denen der synaptische Übertragungsprozess endet, in der richtigen Reihenfolge.“)

Zudem mussten die Probanden auch die Funktion bzw. die Aufgabe verschiedener synaptischer Bestandteile memorieren (*Beispielitem*: „Was ist die Aufgabe der Calciumionen im Übertragungsprozess?“). Bei der Konstruktion der Aufgaben wurden möglichst alle Teilprozesse, die innerhalb der synaptischen Übertragung ablaufen, berücksichtigt.

Die Inferenzskala des Nachtests umfasste 18 Items, wobei 15 der Aufgaben aus der Version 1 des Tests der Vorpilotstudie übernommen wurden. Teilweise hatten sich die Items ebenfalls in der Version 2 befunden. Die Trennschärfen der Aufgaben waren befriedigend

bis gut (Range (rit) = .20 - .82) gewesen, und ihre Schwierigkeiten hatten vorwiegend im mittleren Bereich gelegen.

Bei drei der Items wurden leichte Änderungen in der Formulierung oder Umkonstruktionen vorgenommen, um die Ratings eindeutiger zu machen. Ansonsten lagen die Aufgaben in der gleichen Form wie in der vorherigen Untersuchung vor.

Zudem wurden in die Inferenzskala drei weitere Items eingebaut, die nur in der zweiten Version des Nachtests der Vorpilotstudie vorhanden gewesen waren und die ebenfalls über gute Trennschärfen ($rit > .60$) und über mittlere Schwierigkeitsindizes verfügt hatten. Bis auf ein Item, bei dem inhaltlich geringfügige Umgestaltungen notwendig waren, wurden die Aufgaben in gleichbleibender Form aus der vorigen Studie übernommen.

Insgesamt gesehen, lagen die Items in einer im Vergleich zur Vorpilotstudie unterschiedlich gemischten Reihenfolge vor, wobei einander ähnliche Aufgaben räumlich voneinander entfernt im Testheft eingesetzt wurden, außer sie bauten inhaltlich direkt aufeinander auf.

Für die Auswertung des Tests wurden für die einzelnen Items Ratingsysteme entwickelt, die in ihrer Gestaltung, wie beim Vorwissenstest (s.o.), den Empfehlungen von Mayring (2002) folgten. Das Rating fand wiederum dreistufig in die Kategorien „richtig“ (2 Punkte), „zum Teil richtig“ (1 Punkt) und „falsch / nicht beantwortet“ (0 Punkte) statt. Das Skalenniveau wurde als metrisch angenommen. Die Kategorie „richtig“ beinhaltete als Definition eine nach Lehrbuch (Bayrhuber & Kull, 1989; Rosenzweig et al., 1999) erstellte Musterantwort, wobei deren unvollständige, fehlerhafte oder vergrößerte Wiedergabe eine Abstufung zur Kategorie „zum Teil richtig“ zur Folge hatte. Die Kategorie „falsch“ berücksichtigte alle anderen Antworten oder den Fall, dass der Proband das Item nicht gelöst hatte. Die Systeme waren zum einen theoretisch deduktiv, zum anderen aber auch mit Hilfe von induktiver Empirie konstruiert worden, da für ihre Entwicklung Daten der Vorpilotstudie herangezogen wurden. In der Pilotstudie sollten die Kategorien in ihrer Handhabbarkeit getestet und mit den Antworten der Probanden inhaltlich weiterentwickelt werden. Wie beim Vorwissenstest enthielten die einzelnen Systeme wieder Ankerbeispiele und Kodierregeln für jede Kategorie (zur Ansicht der nach der Pilotstudie weiterentwickelten Ratingsysteme, s. Anhang B3, Datei: „Ratingsysteme_Nachtest_Pilot.pdf“ auf der beiliegenden CD.)

Die Punkte, die ein Proband bei jedem Item erreicht hatte, wurden, nach Wissens- und Inferenzskala getrennt, aufsummiert, und die so entstandenen Scores stellten, statistisch gesehen, abhängige Variablen innerhalb der Untersuchung dar.

IST-2000-R (Liepmann, Beauducel, Brocke & Amthauer, 2007)

Zur Erfassung von räumlichen, sprachlichen und Merkfähigkeiten, die generell einen Einfluss auf Lernleistungen, vor allem aber auch auf das Lernen mit Bildern und Animationen haben (z.B. Hegarty & Kriz, 2008), wurde in der aktuellen Studie der IST-2000-R von Liepmann, Beauducel, Brocke und Amthauer (2007) verwendet. Für die Messung der räumlichen Fähigkeiten wurden die Aufgabengruppen „Figurenauswahl“ (FA) und „Würfelaufgaben“ (WÜ), für die Bestimmung der verbalen Intelligenz wurden die Aufgabengruppen „Analogien“ (AN) sowie „Gemeinsamkeiten“ (GE) herangezogen. Aufschluss über die Ausprägung der Gedächtnisleistung der Probanden sollten die Aufgabengruppen „Merkfähigkeit verbal“ und „Merkfähigkeit figural“ liefern. Die Reliabilitätsschätzungen der Skalen waren als gut bis sehr gut zu bezeichnen, und es lagen mehrere empirische Hinweise für die Konstrukt- und Kriteriumsvalidität des Verfahrens vor (Liepmann et al., 2007). Für eine genauere Darstellung dieses Intelligenztests und für eine Ansicht der Aufgabengruppen sei auf das Testmanual sowie auf das entsprechende Testheft (Grundmodul, Form A) von Liepmann et al. (2007) verwiesen.

Die Scores der Aufgabengruppen, die sich auf den gleichen Intelligenzbereich bezogen, wurden jeweils aufsummiert, so dass drei Gesamtscores für die verbale und die figurale Intelligenz sowie für die Merkfähigkeit einer Person generiert wurden, die potentielle Kovariaten in den statistischen Analysen darstellten.

4.2.2.5 Durchführung

Die Daten der Pilotstudie wurden in den Semesterferien zwischen dem Sommersemester 2009 und dem Wintersemester 2009 / 2010 innerhalb eines Zeitraumes von drei Wochen erhoben. Die Durchführung fand wieder im Labor der Abteilung „Allgemeine und Pädagogische Psychologie“ der Universität Landau statt, wodurch bei allen Versuchspersonen vergleichbare räumliche Bedingungen herrschten. Die Versuchsleitung wurde wieder von der Doktorandin selbst übernommen, und der Ablauf wurde schriftlich fixiert, was zu einer höheren Durchführungsobjektivität und Standardisierung der einzelnen Experimente führte. Die einzelnen Versuche waren, je nach Zahl der Anmeldungen für die jeweiligen Termine, Einzel- bzw. Gruppentestungen bis zu vier Personen. Vor den Erhebungen wurde sichergestellt, dass die vier verschiedenen Lerneinheiten hinsichtlich

ihrer Bildqualität und Zeitdauer untereinander sowie zwischen den Computern miteinander vergleichbar abliefen.

Den Teilnehmern wurde zu Beginn des Experiments kurz die Thematik der Studie erklärt, und es wurde ihnen die Anonymität ihrer Daten zugesichert. Danach füllten sie die soziodemographischen Angaben und den Vorwissenstest zu Synapsen aus. Da dieser ein Powertest war, gab es für die Bearbeitung keine Zeitbegrenzung, die Probanden wurden gebeten, zu warten, bis die letzte Person mit dem Test fertig war und nicht weiterzublättern. So wurde sichergestellt, dass alle Fragebögen von den Teilnehmern simultan bearbeitet wurden. Nachdem die Probanden innerhalb des Testhefts durch eine schriftliche Instruktion erfahren hatten, welcher Art die anstehende Lernaufgabe war, bearbeiteten sie den FAM. Diese erste Sequenz an Verfahren dauerte ca. 15 min. Daraufhin wurden die Teilnehmer jeweils einem Computer zugewiesen. Die Versuchspersonen waren zuvor den vier Bedingungen randomisiert zugeteilt worden, so dass sie jeweils an einem Computer arbeiteten, an dem ihre Bedingung eingestellt war. Die Instruktion der Lernaufgabe wurde auf dem Bildschirm nochmals gezeigt, und die Teilnehmer konnten die Lernsequenz selbständig mit der Maus starten. Zuerst prägten sie sich 2 min 30 lang das beschriftete Standbild zu Synapsen ein, danach sahen sie, je nach ihrer Bedingung, mit Hilfe von Bildern oder mit Hilfe einer Animation die Prozesse, die an einer Synapse ablaufen. Dieser Lernvorgang dauerte 4 min lang. Nachdem alle Personen die Lerneinheit beendet hatten, beantworteten sie die beiden Items zu dem Cognitive Load, den diese Aufgabe bei ihnen ausgelöst hat. Es folgte die Bearbeitung der Beschriftungsaufgabe und des Nachtests, bei dem die Probanden wieder gebeten wurden, zu warten, bis der letzte Teilnehmer mit dem Ausfüllen fertig war und nicht mehr zurückzublättern, wenn sie eine Seite mit Aufgaben gelöst hatten. Für diese Verfahren benötigten die Versuchspersonen in der Regel 30 bis 60 min. Danach sollten sie noch den Cognitive Load einschätzen, den der Nachtest bei ihnen verursacht hat. Der letzte Teil des Experiments war der Bearbeitung des IST-2000-R gewidmet, die ca. 40 min in Anspruch nahm. Den Versuchspersonen wurde erläutert, dass es sich um einen Test handelte, der ihre „räumlichen und sprachlichen Fähigkeiten“ messen sollte. Die Verwendung des Wortes „Intelligenztest“ wurde bewusst vermieden, um keinen zu großen Leistungsdruck zu erzeugen. Danach wurde ihnen der Aufbau des Tests in Instruktionssseite und Aufgabenseite erklärt, und es wurde gezeigt, wie der Antwortbogen auszufüllen war. Dann lasen sich die Probanden bei jeder Aufgabengruppe zunächst die Instruktion selbst durch, und die Aufgabenstellung wurde durch die Versuchsleiterin

nochmals kurz mündlich erklärt. Es folgte die Bearbeitung der Aufgaben, die zeitmäßig durch standardisierte Vorgaben begrenzt war. Nachdem die Probanden den Intelligenztest beendet hatten, wurde ihnen ihre Belohnung von 10 Euro pro Person ausgehändigt. Die Experimente dauerten insgesamt zwischen eineinviertel und zwei Stunden, wie mittels einer Stoppuhr festgehalten wurde. Je nachdem, wie schnell die Probanden den Vorwissens- und den Nachtest ausfüllten, schwankte die Gesamtdurchführungsdauer, wodurch man aber auch eine genauere Zeitschätzung für die Hauptstudie herleiten konnte. Technische oder sonstige Störungen innerhalb des Versuchs wurden mitnotiert, wobei diese nicht ein solches Ausmaß annahmen, dass Probanden in der Auswertung ausgeschlossen werden mussten. Lediglich bei einer Versuchsperson fiel ein unsorgfältiges Ausfüllen des Nachtests auf, weshalb ihre Daten nicht in die späteren Analysen eingingen. Generell wurde gerade bei dem Nachtest beobachtet, ob die Versuchspersonen ihn ernsthaft bearbeiteten. Insgesamt gesehen, gab es zu den Formulierungen innerhalb der einzelnen Testverfahren kaum Nachfragen, was auf eine gute Verständlichkeit der Aufgaben hindeutete.

4.2.3 Ergebnisse

Die Rechnungen, mit denen die Forschungsfragen und die Hypothesen der Pilotstudie getestet wurden, erfolgten mit dem Statistikprogramm „SPSS, Version 17.0“.

Beobachtung der Durchführbarkeit

Forschungsfrage 1: Ist der Versuchsablauf stimmig und das Experiment in der geplanten Form durchführbar?

Der Versuch konnte in der geplanten Form bei allen Probanden ohne größere technische Störungen durchgeführt werden. Es gab während der einzelnen Testverfahren kaum Nachfragen zu den Formulierungen der Instruktionen oder der Items, was auf deren gute Verständlichkeit hindeutete. Bei der Durchführung fiel lediglich auf, dass es für Personen ohne irgendein Vorwissen zu Synapsen schwierig war, eine Grundorientierung im Lernmaterial zu bekommen. Zudem bestand für die Probanden die Möglichkeit, den Nachtest bei geringer Motivation unvollständig auszufüllen, was bei einer Person auch zu ihrem Ausschluss aus den Daten führte. Die durchschnittliche Versuchszeit betrug $M = 95.86$ min ($SD = 10.86$ min, $Min = 74.00$ min, $Max = 116.00$ min). Informelle Befragungen der Probanden nach dem Experiment ergaben zudem, dass zumindest für ungefähr die Hälfte der Personen die Dauer und der Ablauf des Versuchs stimmig gewesen waren.

Itemanalyse

Forschungsfrage 2: Sind die Gütekriterien des Vorwissenstests, der Beschriftungsaufgabe zum begrifflichen Wissen sowie der Wissens- und der Inferenzskala des Nachtests, die die selbst konstruierten Verfahren des Experimentes darstellen, zufriedenstellend?

Für die Beantwortung dieser Forschungsfrage wurden für die Items der oben genannten Verfahren die Trennschärpen und Schwierigkeiten, sowie das Cronbachs Alpha für die Gesamtskalen jeweils berechnet. Auf die Bestimmung der Intraclass-Korrelationen der für die Auswertung der Items erstellten Ratingsysteme wurde verzichtet, da der Stichprobenumfang hierfür noch zu gering war. Als Grundlage für die Analysen diente die

Gesamtstichprobe der Pilotstudie ($n = 23$), um stabilere Schätzungen der Kennwerte zu erhalten. Bei der Trennschärfe lag der Grenzwert zur Beibehaltung bei $rit = .30$, während bei der Itemschwierigkeit Aufgaben mit Werten zwischen $p = .20$ und $p = .80$ bevorzugt wurden. Generell wurden diese, aus der Literatur (Bortz & Döring, 2002) übernommenen, Kriterien bei der Pilotstudie nicht so streng angelegt, da es sich um eine kleinere Stichprobe mit noch instabilen Größenschätzungen handelte. Es erfolgt zunächst die Darstellung der Ergebnisse der Itemanalyse zum Vorwissenstest.

a) Vorwissenstest

Tabelle 4.2.1

Trennschärpen und Schwierigkeiten der Items des Vorwissenstests in der Pilotstudie

Item	<i>rit</i>	<i>p</i>
Item 1	.29	.26
Item 2	.29	.79
Item 3	.13	.46
Item 4	.44	.35
Item 5	.33	.07
Item 6	.00	.00
Item 7	.42	.05
Item 8	.40	.09
Item 9	.01	.24
Item 10	.09	.07
Item 11	.03	.05

Anmerkungen. Die Itemnummern entsprechen denen des Vorwissenstests in der Pilotstudie (s. Anhang B2 - 2).

Der Vorwissenstest hatte in der Version mit elf Items ein Cronbachs Alpha von $\alpha = .53$. Für die weiteren Analysen der Pilotstudie wurden die Items 6, 9, 10 und 11 aufgrund ihrer geringen Trennschärfen und größtenteils auch geringen Schwierigkeitswerten aus dem Itempool eliminiert. Die Items 5, 7 und 8 wurden trotz ebenfalls niedriger Schwierigkeiten weiterverwendet, da ihre Trennschärfenwerte zufriedenstellend waren und da der Vorwissenstest konzeptuell sehr schwierig für Personen mit wenig Vorwissen war, die die Zielgruppe der Studie darstellten. Ohne die vier eliminierten Items wies der Vorwissenstest ein Alpha von $\alpha = .62$ auf.

b) Beschriftungsaufgabe zum begrifflichen Wissen

Tabelle 4.2.2

Trennschärfen und Schwierigkeiten der Items der Beschriftungsaufgabe zum begrifflichen Wissen in der Pilotstudie

Item	<i>rit</i>	<i>p</i>
Item 1: Calciumkanal	.56	.83
Item 2: Endknopf	.25	.74
Item 3: Spalt	.00	1.00
Item 4: Nachsynaptische Membran	.29	.87
Item 5: Cholinesterase	.07	.57
Item 6: Natriumkanal	.66	.78
Item 7: Rezeptor	.38	.91
Item 8: Vorsynaptische Membran	.29	.87
Item 9: Transmitter	.28	.91
Item 10: Bläschen	.59	.91

Anmerkungen. Die Numerierung der Items erfolgte im Uhrzeigersinn der Anordnung der Begriffe im Testheft der Pilotstudie (s. Anhang B2 - 5).

Das Cronbachs Alpha der Beschriftungsaufgabe betrug $\alpha = .66$. Bei dieser Skala wurde auf eine herkömmliche Itemelimination aus theoretisch konzeptuellen Gründen verzichtet, da das Wissen um jeden einzelnen Begriff den Score des Nachtests sowie den Cognitive Load mit beeinflusst. Daher wurden alle Items in die weiteren Analysen mit aufgenommen.

c) Nachwissensskala

Tabelle 4.2.3

Trennschärfen und Schwierigkeiten der Items der Nachwissensskala in der Pilotstudie

Item	<i>rit</i>	<i>p</i>
Item 1	.11	.78
Item 2	.20	.54
Item 6	.54	.61
Item 9	.68	.70
Item 10	.40	.80
Item 14	.71	.70
Item 15	.71	.70
Item 19	.28	.83
Item 20	.85	.67
Item 26	.29	.85
Item 27	.52	.76
Item 30	.62	.85

Anmerkungen. Die Itemnummern entsprechen denen des Nachtests in der Pilotstudie (s. Anhang B2 - 6).

Die Wissensskala des Nachtests wies ein Cronbachs Alpha von $\alpha = .84$ auf. Es gingen alle Items in die weiteren Rechnungen mit ein. Bei den Items 1 und 2 lagen etwas schwächere

Trennschärpen vor, und bei den Items 19, 26 und 30 waren die Schwierigkeitswerte leicht erhöht. Da es sich aber nur um mäßige Abweichungen von den Grenzwerten handelte und der jeweils andere Kennwert zufriedenstellend war, wurden die Differenzen bei der geringen Stichprobengröße akzeptiert.

d) Inferenzskala

Tabelle 4.2.4

Trennschärpen und Schwierigkeiten der Items der Inferenzskala in der Pilotstudie

Item	<i>rit</i>	<i>p</i>
Item 3	.74	.57
Item 4	.13	.83
Item 5	.37	.67
Item 7	.35	.85
Item 8	.32	.61
Item 11	.56	.43
Item 12	.53	.80
Item 13	.31	.33
Item 16	.26	.78
Item 17	.34	.35
Item 18	.27	.43
Item 21	.25	.70

Tabelle wird auf der nächsten Seite fortgesetzt

Fortsetzung Tabelle 4.2.4

Item	<i>rit</i>	<i>p</i>
Item 22	.37	.67
Item 23	.69	.61
Item 24	.13	.46
Item 25	.46	.80
Item 28	-.04	.76
Item 29	.43	.63

Anmerkungen. Die Itemnummern entsprechen denen des Nachttests in der Pilotstudie (s. Anhang B2 - 6).

Die Reliabilität der Inferenzskala lag bei einem Cronbachs Alpha von $\alpha = .78$. Für die weiteren Analysen der Pilotstudie wurde Item 28 ausgeschlossen, da seine Trennschärfe einen zu niedrigen Wert hatte. Dadurch erhöhte sich das Alpha der Skala auf $\alpha = .80$. Bei den anderen Items lagen für eine Pilotstudie zufriedenstellende bzw. akzeptable Kennwerte vor.

Zusätzlich zu den Itemanalysen der eigens konstruierten Verfahren wurden die Reliabilitäten des Motivationsfragebogens FAM (Rheinberg et al., 2001) bestimmt. Nach der Itemelimination stellten sich die Reliabilitäten dieses Verfahrens wie folgt dar:

Tabelle 4.2.5

Reliabilitäten der Subskalen des FAM in der Pilotstudie

	<i>n</i> (Items)	<i>α</i>
Erfolgswahrscheinlichkeit	2	.86
Misserfolgsbefürchtung	5	.80
Interesse	4	.81
Herausforderung	4	.75

Bei der Skala „Erfolgswahrscheinlichkeit“ waren zwei Items, und bei der Skala „Interesse“ war ein Item eliminiert worden, um ihre etwas schwächeren Kennwerte noch zu erhöhen.

Deskriptive Statistiken der Gruppen „Animation“, „Bild seriell schnell“ und „Bild seriell langsam“

Da im Nachfolgenden für die Mittelwertsvergleiche die Gesamtstichprobe hypothesengemäß in zwei Teilstichproben aufgespalten wurde, erfolgt nun auch die Darstellung der deskriptiven Statistiken für jeweils drei Bedingungen. Auf vorherige Ausreißeranalysen wurde verzichtet, da der Stichprobenumfang der Pilotstudie für derartige Verfahren zu gering war.

Zunächst werden die Kennwerte für die Teilstichprobe mit den Gruppen „Animation“, „Bild seriell schnell“ und „Bild seriell langsam“ angegeben.

Das Durchschnittsalter dieser Stichprobe lag bei 23.65 Jahren ($SD = 4.83$) bei einem Range von 19 bis 39 Jahren. Es waren 6 Männer und 11 Frauen in der Stichprobe, deren Gesamtumfang $n = 17$ betrug.

Es folgt nun die Darstellung der deskriptiven Statistiken der abhängigen und der Kontrollvariablen der Pilotstudie:

Tabelle 4.2.6

Deskriptive Statistiken der abhängigen Variablen für die Teilstichprobe „Animation“ mit „Bild seriell schnell“ und „Bild seriell langsam“

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Nachwissensskala	17	6.00	23.00	18.35	5.01
Inferenzskala	17	6.00	30.00	22.53	6.45
Cognitive Load (Lernaufgabe)	17	5.00	13.00	9.88	2.42
Cognitive Load (Nachttest)	17	6.00	16.00	12.41	2.74

Tabelle 4.2.7

Deskriptive Statistiken der Kontrollvariablen für die Teilstichprobe „Animation“ mit „Bild seriell schnell“ und „Bild seriell langsam“

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Vorwissenstest	17	.00	12.00	4.18	2.98
Begriffswissen	17	3.00	10.00	8.59	1.70
IST verbal	17	15.00	32.00	25.53	4.60
IST räumlich	17	12.00	31.00	23.76	6.50
IST Merkfähigkeit	17	13.00	23.00	18.47	3.02
Erfolgswahrscheinlichkeit	17	6.00	14.00	10.47	2.50
Misserfolgsbefürchtung	17	5.00	20.00	11.47	4.69
Interesse	17	9.00	27.00	19.29	4.70
Herausforderung	16	8.00	25.00	20.13	4.35

Deskriptive Statistiken der Gruppen „Bild parallel“, „Bild seriell schnell“ und „Bild seriell langsam“

Die Teilstichprobe mit den Gruppen „Bild parallel“, „Bild seriell schnell“ und „Bild seriell langsam“ hatte einen Umfang von $n = 17$ mit einer Geschlechterverteilung von 6 Männern und 11 Frauen. Durchschnittlich waren die Probanden 23.41 Jahre alt ($SD = 4.85$). Der Altersrange betrug 19 bis 39 Jahre.

Tabelle 4.2.8

Deskriptive Statistiken der abhängigen Variablen für die Teilstichprobe „Bild parallel“ mit „Bild seriell schnell“ und „Bild seriell langsam“

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Nachwissensskala	17	10.00	23.00	17.41	4.58
Inferenzskala	17	10.00	30.00	21.65	6.40
Cognitive Load (Lernaufgabe)	17	6.00	13.00	10.24	1.99
Cognitive Load (Nachtest)	17	10.00	18.00	13.29	2.20

Tabelle 4.2.9

Deskriptive Statistiken der Kontrollvariablen für die Teilstichprobe „Bild parallel“ mit „Bild seriell schnell“ und „Bild seriell langsam“

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Vorwissenstest	17	2.00	12.00	4.94	2.73
Begriffswissen	17	6.00	10.00	8.47	1.38
IST verbal	17	13.00	32.00	24.18	5.65
IST räumlich	17	14.00	32.00	25.41	5.49
IST Merkfähigkeit	17	10.00	23.00	16.47	4.02
Erfolgswahrscheinlichkeit	17	9.00	14.00	11.06	1.78
Misserfolgsbefürchtung	17	5.00	25.00	12.94	5.43
Interesse	17	14.00	27.00	19.76	3.35
Herausforderung	16	14.00	25.00	21.19	2.74

Mittelwertsvergleiche der Gruppen „Animation“, „Bild seriell schnell“ und „Bild seriell langsam“

Für die Mittelwertsvergleiche wurde die Gesamtstichprobe in zwei Teilstichproben aufgeteilt, um sowohl die Hypothesen der ersten als auch die der zweiten Teilstudie der Hauptstudie explorativ testen zu können. Begonnen wird mit den Vergleichen der ersten Teilstudie mit den Gruppen „Animation“, „Bild seriell schnell“ und „Bild seriell langsam“.

Da die Stichprobengrößen in den einzelnen Zellen der Pilotstudie klein waren ($n = 5$ bzw. $n = 6$), waren zuverlässige Prüfungen von Verteilungsvoraussetzungen wie Normalverteilung oder Varianzhomogenität zwischen den Stichproben nur eingeschränkt möglich. Daher wurden statt Varianz- und Kovarianzanalysen robustere Brown-Forsythe-Tests gerechnet, um die Mittelwerte der drei Gruppen hinsichtlich der Kriterien miteinander zu vergleichen (Diehl & Arbinger, 2001). Um die einzelnen, paarweisen Hypothesen zu testen, wurden w -Tests angewandt. Da es sich hierbei um theoretisch fundierte A-priori-Kontraste handelte, wurde jeweils ein Signifikanzniveau von $\alpha = .05$ gewählt.

Der Einfluss der Kovariaten auf die Mittelwertsunterschiede wurde durch die Berechnung ihrer bivariaten Korrelationen mit den Kriterien einerseits, und durch die Bestimmung ihrer Gruppenunterschiede andererseits, abgeschätzt. Lagen bei einer Kontrollvariable sowohl größere bis signifikante Gruppenunterschiede, als auch signifikante Korrelationen zu abhängigen Variablen vor, wurde ihr Einfluss bei der Interpretation der Ergebnisse post-hoc berücksichtigt, um die Größe der gefundenen Haupteffekte zu relativieren.

Der Aufbau dieses Abschnittes erfolgt in der Form, dass zunächst nochmals die Hypothesen des Theorieteils in statistischer Kurzform genannt und dazu passend jeweils die Rechnungen und Ergebnisse dargestellt werden. Die Reihenfolge der Hypothesen ist dabei im Vergleich zu der im Theorieteil eine andere, was durch die vorgegebene Systematik der statistischen Verfahren zu erklären ist.

Zuerst wurden die Hypothesen zum Kriterium „Nachwissensskala“ getestet:

Hypothese 1a: Bild seriell schnell > Animation; AV = Nachwissensskala

Hypothese 2a: Bild seriell langsam > Animation; AV = Nachwissensskala

Hypothese 3a: Bild seriell schnell ist anders als Bild seriell langsam;

AV = Nachwissensskala

Tabelle 4.2.10

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Nachwissensskala“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild seriell langsam	5	17.20	4.55
Animation	6	18.00	6.33
Bild seriell schnell	6	19.67	4.50

Die Unterschiede zwischen den drei Bedingungen wiesen keine Signifikanz auf ($F(2, 12.96) = .33, p = .72, \eta^2 = .04$). Die Probanden der seriell schnellen Bildbedingung erzielten die im Durchschnitt höchsten Werte. Auch bei der Berechnung der drei Kontraste „Animation vs. Bild seriell schnell“ ($w(9.03) = .53, p(1\text{-seitig}) = .31, d = .30$), „Animation vs. Bild seriell langsam“ ($w(8.87) = .24, p(2\text{-seitig}) = .81, d = .15$) und „Bild seriell schnell vs. Bild seriell langsam“ ($w(8.61) = .90, p(2\text{-seitig}) = .39, d = .55$) konnten keine signifikanten Ergebnisse festgestellt werden, wobei der Unterschied zwischen der Animations- und der seriell langsamen Bildbedingung nicht in die vorhergesagte Richtung ging.

Hypothese 1b: Bild seriell schnell > Animation; AV = Inferenzskala

Hypothese 2b: Bild seriell langsam > Animation; AV = Inferenzskala

Hypothese 3b: Bild seriell schnell ist anders als Bild seriell langsam; AV = Inferenzskala

Tabelle 4.2.11

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Inferenzskala“

	<i>n</i>	<i>M</i>	<i>SD</i>
Animation	6	19.33	6.80
Bild seriell langsam	5	22.60	7.09
Bild seriell schnell	6	25.67	4.76

Der Brown-Forsythe-Test zeigte hinsichtlich der abhängigen Variablen „Inferenzskala“ keine signifikanten Unterschiede ($F(2, 11.94) = 1.50, p = .26, \eta^2 = .18$). Ebenfalls keine bedeutsamen Differenzen erbrachten die Berechnungen der Kontraste „Animation vs. Bild seriell langsam“ ($w(8.49) = .78, p(1\text{-seitig}) = .23, d = .47$) und „Bild seriell schnell vs. Bild seriell langsam“ ($w(6.80) = .82, p(2\text{-seitig}) = .44, d = .51$). Lediglich beim Vergleich zwischen der Gruppe „Animation“ und der Gruppe „Bild seriell schnell“ konnte ein signifikanter Mittelwertsunterschied festgestellt werden ($w(8.95) = 1.87, p(1\text{-seitig}) = .05, d = 1.08$). Die Probanden der Animationsbedingung zeigten die schwächsten, die der seriell schnellen Bildbedingung die besten Inferierleistungen.

Hypothese 1c: Bild seriell schnell < Animation; AV = Cognitive Load beim Lernen

Hypothese 2c: Bild seriell langsam < Animation; AV = Cognitive Load beim Lernen

Hypothese 3c: Bild seriell schnell ist anders als Bild seriell langsam; AV = Cognitive Load beim Lernen

Tabelle 4.2.12

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load bei der Lernaufgabe“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild seriell schnell	6	9.50	1.87
Bild seriell langsam	5	9.80	2.39
Animation	6	10.33	3.20

Die Größe der Mittelwertsunterschiede erreichte bei dem Kriterium „Cognitive Load bei der Lernaufgabe“ keine statistische Signifikanz ($F(2, 12.03) = .17, p = .85, \eta^2 = .02$). Auch die Ergebnisse der Testungen der drei Kontraste stellten sich als statistisch nicht bedeutsam dar („Animation vs. Bild seriell schnell“: $w(8.05) = .55, p(1\text{-seitig}) = .30, d = .32$, „Animation vs. Bild seriell langsam“: $w(8.93) = .32, p(1\text{-seitig}) = .38, d = .19$, „Bild seriell schnell vs. Bild seriell langsam“: $w(7.56) = .23, p(2\text{-seitig}) = .83, d = .14$).

Hypothese 1d: Bild seriell schnell < Animation; AV = Cognitive Load beim Nachttest

Hypothese 2d: Bild seriell langsam < Animation; AV = Cognitive Load beim Nachttest

Hypothese 3d: Bild seriell schnell ist anders als Bild seriell langsam; AV = Cognitive Load beim Nachttest

Tabelle 4.2.13

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load beim Nachttest“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild seriell schnell	6	11.83	1.94
Animation	6	12.33	4.03
Bild seriell langsam	5	13.20	1.92

Bei dem vierten Kriterium „Cognitive Load beim Nachttest“ erbrachte der Brown-Forsythe-Test ebenfalls keine signifikanten Unterschiede ($F(2, 9.68) = .33, p = .73, \eta^2 = .04$). Die

anschließenden, paarweisen w -Tests zeigten ebenfalls keine statistische Bedeutsamkeit, wobei der Kontrast zwischen der Animations- und der seriell langsamen Bildbedingung einen nicht vorhergesagten Effekt ergab. („Animation vs. Bild seriell schnell“: $w(7.20) = .27$, $p(1\text{-seitig}) = .40$, $d = .16$, „Animation vs. Bild seriell langsam“: $w(7.41) = .47$, $p(2\text{-seitig}) = .65$, $d = .28$, „Bild seriell schnell vs. Bild seriell langsam“: $w(8.67) = 1.17$, $p(2\text{-seitig}) = .27$, $d = .71$). Die Probanden der seriell schnellen Bildbedingung verspürten am wenigsten, die der seriell langsamen Bildbedingung am meisten Cognitive Load.

Einfluss der Kovariaten

Größere bis signifikante Gruppenunterschiede wurden bei den Kovariaten „Vorwissenstest“ ($F(2, 9.23) = 5.08$, $p = .03$, $\eta^2 = .41$), „Erfolgswahrscheinlichkeit (FAM)“ ($F(2, 12.17) = 3.06$, $p = .08$, $\eta^2 = .30$) und „Räumliche Fähigkeiten“ ($F(2, 12.59) = 4.74$, $p = .03$, $\eta^2 = .41$) festgestellt.

Bei der Variable „Vorwissenstest“ hatte die Gruppe „Animation“ ($M = 1.67$, $SD = 1.37$) einen geringeren Wert im Vergleich zu der seriell schnellen ($M = 5.50$, $SD = 3.51$) und der seriell langsamen Bildbedingung ($M = 5.60$, $SD = 1.82$). Da das Vorwissen mit der Inferenzskala korrelierte ($r = .44$, $p(1\text{-seitig}) = .04$), könnten sich die Unterschiede in diesem Kriterium unter Einbezug dieser Kontrollvariablen nivellieren. Ebenfalls im Nachteil waren die Probanden der Animationsbedingung bei der Variable „Erfolgswahrscheinlichkeit“ ($M = 8.83$, $SD = 2.79$) im Vergleich zu den seriellen Bildbedingungen (Bild seriell schnell: $M = 12.00$, $SD = 1.67$; Bild seriell langsam: $M = 10.60$, $SD = 2.07$). „Erfolgswahrscheinlichkeit“ korrelierte mit den abhängigen Variablen „Nachwissensskala“ ($r = .49$, $p(1\text{-seitig}) = .03$), und „Inferenzskala“ ($r = .56$, $p(1\text{-seitig}) = .01$), wodurch die bereinigten Werte der Animationsbedingung bei diesen Kriterien etwas höher liegen dürften. Die Variable „Räumliche Fähigkeiten“ hatte mit keiner der abhängigen Variablen einen signifikanten Zusammenhang, weshalb sie wahrscheinlich die Gruppenunterschiede in den Kriterien nicht maßgeblich beeinflusst hat. Durch den nachträglichen Einbezug der Kontrollvariablen, sollten somit die Benachteiligungen der Animationsgruppe in den Kriterien „Nachwissensskala“ und „Inferenzskala“ in den Haupteffekten etwas abgeschwächt betrachtet werden.

Mittelwertsvergleiche der Gruppen „Bild parallel“, „Bild seriell schnell“ und „Bild seriell langsam“

Zur Überprüfung der Hypothesen zu den Haupteffekten zwischen den Gruppen „Bild parallel“ und „Bild seriell schnell“ bzw. „Bild parallel“ und „Bild seriell langsam“ wurde ein analoges Verfahren wie bei den obigen Mittelwertsvergleichen gewählt.

Bei der folgenden Darstellung werden auch wieder zunächst die Hypothesen in kurzer Form genannt und die betreffenden Ergebnisse jeweils angeschlossen.

Hypothese 4a: Bild parallel ist anders als Bild seriell schnell; AV = Nachwissensskala

Hypothese 5a: Bild parallel ist anders als Bild seriell langsam; AV = Nachwissensskala

Tabelle 4.2.14

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Nachwissensskala“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild parallel	6	15.33	4.37
Bild seriell langsam	5	17.20	4.55
Bild seriell schnell	6	19.67	4.50

Die Mittelwertsunterschiede zwischen den drei Bedingungen waren nicht signifikant ($F(2, 13.58) = 1.41, p = .28, \eta^2 = .17$). Die Probanden der Bedingung „Bild parallel“ hatten im Schnitt die schlechtesten, die der Gruppe „Bild seriell schnell“ die besten Werte im Kriterium „Nachwissensskala“. Die Berechnung paarweiser Kontraste führte ebenfalls zu keinen signifikanten Ergebnissen („Bild parallel vs. Bild seriell schnell“: $w(9.99) = 1.69, p(2\text{-seitig}) = .12, d = .98$, „Bild parallel vs. Bild seriell langsam“: $w(8.49) = .69, p(2\text{-seitig}) = .51, d = .42$).

Hypothese 4b: Bild parallel ist anders als Bild seriell schnell; AV = Inferenzskala

Hypothese 5b: Bild parallel ist anders als Bild seriell langsam; AV = Inferenzskala

Tabelle 4.2.15

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Inferenzskala“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild parallel	6	16.83	4.49
Bild seriell langsam	5	22.60	7.09
Bild seriell schnell	6	25.67	4.76

Die Prüfgröße des Brown-Forsythe-Tests erreichte bei diesem Kriterium fast Signifikanz ($F(2, 10.19) = 3.80, p = .06, \eta^2 = .37$). Wie bei der abhängigen Variable „Nachwissensskala“, erzielte die Bedingung „Bild seriell schnell“ den höchsten Durchschnittswert, während die Probanden der Gruppe „Bild parallel“ die schlechtesten Inferierleistungen zeigten. Der Unterschied zwischen diesen beiden Gruppen wurde im paarweisen Kontrast signifikant ($w(9.97) = 3.31, p(2\text{-seitig}) = .008, d = 1.91$). Der Vergleich zwischen der Bedingung „Bild parallel“ und „Bild seriell langsam“ erbrachte hingegen kein statistisch bedeutsames Ergebnis ($w(6.54) = 1.57, p(2\text{-seitig}) = .16, d = .97$).

Hypothese 4c: Bild parallel ist anders als Bild seriell schnell; AV = Cognitive Load beim Lernen

Hypothese 5c: Bild parallel ist anders als Bild seriell langsam; AV = Cognitive Load beim Lernen

Tabelle 4.2.16

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load bei der Lernaufgabe“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild seriell schnell	6	9.50	1.87
Bild seriell langsam	5	9.80	2.39
Bild parallel	6	11.33	1.51

Zwischen den drei Gruppen bestanden bei dem Kriterium „Cognitive Load bei der Lernaufgabe“ keine signifikanten Unterschiede ($F(2, 10.93) = 1.47, p = .27, \eta^2 = .18$). Die Probanden der Gruppe „Bild seriell schnell“ empfanden beim Lernen den geringsten, die der Gruppe „Bild parallel“ den höchsten Cognitive Load. Der paarweise Vergleich zwischen den Gruppen „Bild parallel“ und „Bild seriell schnell“ ergab einen größeren, aber nicht signifikanten Mittelwertsunterschied ($w(9.56) = 1.87, p(2\text{-seitig}) = .09, d = 1.08$). Das Ergebnis des Kontrastes „Bild parallel vs. Bild seriell langsam“ war nicht signifikant ($w(6.52) = 1.25, p(2\text{-seitig}) = .26, d = .77$).

Hypothese 4d: Bild parallel ist anders als Bild seriell schnell; AV = Cognitive Load beim Nachttest

Hypothese 5d: Bild parallel ist anders als Bild seriell langsam; AV = Cognitive Load beim Nachttest

Tabelle 4.2.17

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load beim Nachttest“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild seriell schnell	6	11.83	1.94
Bild seriell langsam	5	13.20	1.92
Bild parallel	6	14.83	1.84

Bei dem Kriterium „Cognitive Load beim Nachtest“ waren die Unterschiede zwischen den Mittelwerten signifikant ($F(2, 13.59) = 3.74, p = .05, \eta^2 = .35$). Der schlechteste Wert wurde wiederum bei der Bedingung „Bild parallel“ verzeichnet. Die Differenz der Mittelwerte bei den Gruppen „Bild parallel“ und „Bild seriell schnell“ wies statistische Signifikanz auf ($w(9.97) = 2.75, p(2\text{-seitig}) = .02, d = 1.59$). Bei dem zweiten Kontrast „Bild parallel vs. Bild seriell langsam“ war dies hingegen nicht der Fall ($w(8.47) = 1.43, p(2\text{-seitig}) = .19, d = .87$).

Einfluss der Kovariaten

Mit Hilfe von Brown - Forsythe - Tests wurden signifikante Mittelwertsunterschiede bei den Variablen „Misserfolgsbefürchtung (FAM)“ ($F(2, 12.17) = 6.05, p = .02, \eta^2 = .45$) und „Merkfähigkeit“ ($F(2, 10.94) = 5.65, p = .02, \eta^2 = .43$) ermittelt. Bei beiden Variablen hatte die Bedingung „Bild parallel“ (Misserfolgsbefürchtung: $M = 17.33, SD = 5.13$; Merkfähigkeit: $M = 13.00, SD = 2.61$) schlechtere Werte als die Gruppen „Bild seriell schnell“ (Misserfolgsbefürchtung: $M = 9.00, SD = 4.60$; Merkfähigkeit: $M = 18.50, SD = 4.32$) und „Bild seriell langsam“ (Misserfolgsbefürchtung: $M = 12.40, SD = 2.41$; Merkfähigkeit: $M = 18.20, SD = 2.17$). „Misserfolgsbefürchtung“ korrelierte signifikant mit den abhängigen Variablen „Nachwissensskala“ ($r = -.52, p(1\text{-seitig}) = .02$) und „Cognitive Load beim Nachtest“ ($r = .67, p(1\text{-seitig}) = .002$). Bei der Variable „Merkfähigkeit“ lag ein signifikanter Zusammenhang zum Kriterium „Cognitive Load beim Nachtest“ vor ($r = -.57, p(1\text{-seitig}) = .01$). In der Interpretation der Haupteffekte ist somit zu berücksichtigen, dass die parallele Bildbedingung ungünstige Voraussetzungen in den Variablen „Misserfolgsbefürchtung“ und „Merkfähigkeit“ hatte, was zu ihren schlechteren Ergebnissen in den Kriterien „Nachwissensskala“ und „Cognitive Load beim Nachtest“ mit beigetragen haben könnte, wodurch die Unterschiede zwischen den Bedingungen zusätzlich vergrößert waren.

4.2.4 Diskussion

Die erste Forschungsfrage der Pilotstudie, ob der Versuchsablauf stimmig und das Experiment in der geplanten Form durchführbar war, konnte im Großen und Ganzen bejaht werden. Zu den einzelnen Testverfahren kamen kaum Fragen auf, was auf eine gute Verständlichkeit der Items und der Instruktionen hindeutete. Auch von technischer Seite her traten keine größeren Probleme in der Durchführung auf. Im Verlauf der Studie fiel allerdings auf, dass Personen, die kein Vorwissen zu Synapsen hatten, Schwierigkeiten bekundeten, sich im Lernmaterial orientieren zu können, da sie nicht über die Grundbedeutung einer Synapse Bescheid wussten. Daher wurde entschieden, in der Hauptstudie den Probanden vor der Betrachtung des Lernmaterials die Information zu geben, dass eine Synapse eine Verbindung zwischen zwei Nervenzellen ist, um eine grundlegende Orientierung zu gewährleisten. Zudem wurde in der Pilotstudie offensichtlich, dass einige Teilnehmer den Nachtest unvollständig ausfüllten. Um sicherzustellen, dass dies nur aus mangelndem Wissen und nicht aus fehlender Motivation heraus geschah, wurden die Versuchspersonen in der Hauptstudie stets explizit darum gebeten, sich bei der Bearbeitung des Nachtests zu bemühen. Da die durchschnittliche Versuchsdauer von $M = 95.86$ min angesichts der Zahl und des Anspruchs der Testverfahren akzeptabel war und da zudem die Dauer und der Ablauf des Versuchs von zumindest der Hälfte der Teilnehmer als „angemessen“ empfunden wurden, wurde beschlossen, das Experiment in der gleichen Abfolge wie in der Pilotstudie auch in der Hauptstudie durchzuführen.

Zu der zweiten Frage nach den Gütekriterien der einzelnen Verfahren war zu sagen, dass die Reliabilität des Vorwissenstests etwas schwächer, aber für eine Pilotstudie noch akzeptabel war ($\alpha = .53$). Das Schwierigkeitsniveau der Items war erneut relativ hoch, was ein mit verursachender Faktor für die teilweise schwächeren Trennschärfen der Aufgaben gewesen sein könnte. Ein weiterer dafür verantwortlicher Grund könnte die größere inhaltliche Heterogenität der Items des Vorwissenstests gewesen sein, die teilweise eher allgemein, teils aber auch sehr spezifisch nach den synaptischen Prozessen fragten. Da die Aufgaben wesentliche Aspekte des Lernmaterials und des späteren Nachtests inhaltlich abdeckten, hätte eine größere Itemelimination für die Hauptstudie eventuell einen Verlust der konzeptuellen Vollständigkeit der Erfassung des Vorwissens bedeutet. Zudem war die hohe Schwierigkeit der Aufgaben bei einer Stichprobe, die bewusst kein Vorwissen haben sollte, dem Test inhärent. Letztlich handelte es sich bei der Pilotstudie immer noch um eine

geringere Stichprobengröße ($n = 23$), bei der die Schätzungen der Kennwerte noch instabil sein konnten, wodurch alle Items des Vorwissenstests in die Hauptstudie übernommen wurden, um die Werte einer weiteren Überprüfung zu unterziehen.

Bei der Beschriftungsaufgabe zu begrifflichem Wissen war das Cronbachs Alpha ebenfalls akzeptabel ($\alpha = .66$). Sieben der zehn Aufgaben waren für die Probanden sehr leicht ($p > .80$), was an der für Studenten relativ einfachen Aufgabenstellung des Einprägens und Reproduzierens einzelner Begriffe liegen dürfte, zumal sich einige der Versuchspersonen, nach eigener Aussage, an ihr Schulwissen über Synapsen erinnerten. Bei der Beschriftungsaufgabe wurde keine weitere Itemelimination vorgenommen, da jeder Begriff in seiner Bedeutung den Score des Nachtests und den Cognitive Load mit beeinflusst. Da zudem die Reliabilität der Skala akzeptabel und die Trennschärfen der Items größtenteils zufriedenstellend bis gut waren, wurden alle Aufgaben ohne Änderung in die Hauptstudie übernommen.

Die Nachwissensskala wies in der Pilotstudie ein gutes Alpha von $\alpha = .84$ auf. Die Trennschärfen der Items zeigten zum Großteil zufriedenstellende bis gute Werte, und die Itemschwierigkeiten lagen, bis auf wenige Ausnahmen, im mittleren Bereich. Abweichungen von den empfohlenen Grenzwerten waren in ihrer Größe bei den entsprechenden Items tolerierbar, so dass alle Aufgaben, mit wenigen sprachlichen Änderungen, in den Nachtest der Hauptstudie eingingen.

Die Reliabilität der Inferenzskala hatte ebenfalls einen guten Wert von $\alpha = .78$. Bis auf eines der Items (Item 28) hatten die Aufgaben akzeptable bis gute Trennschärfen, und die Schwierigkeiten befanden sich überwiegend im empfohlenen Bereich. Da es sich bei der Pilotstudie noch um eine kleinere Stichprobe handelte, wurden alle Items in die Hauptstudie übernommen. Lediglich bei einem Item wurde eine Umformulierung vorgenommen, um seine Beantwortung sowie sein Rating eindeutiger zu machen.

Weiterhin sollten in der Pilotstudie die Hypothesen über die Haupteffekte zwischen den Treatmentstufen überprüft werden. Es ist zu betonen, dass es sich bei diesen Testungen um explorative Vorgänge handelte, bei denen aufgrund der geringen Stichprobengröße der Studie und der damit instabilen Schätzungen der einzelnen Kennwerte nicht zu weitreichende Interpretationen oder Schlussfolgerungen gezogen werden konnten, zumal der Einfluss der Kovariaten nicht ausreichend berücksichtigt werden konnte. In der Pilotstudie wurden sowohl die Hypothesen der ersten als auch die der zweiten Teilstudie der Hauptstudie separat voneinander getestet.

In den Hypothesen 1a – 1d der ersten Teilstudie wurde davon ausgegangen, dass Personen, die mit Animationen lernten, schlechtere Leistungen auf einer Wissens- und einer Inferenzskala zeigten sowie einen höheren Cognitive Load empfanden als Personen, die mit seriell schnellen Bildern lernten. Außer bei dem Kriterium „Inferenzskala“ konnten die Hypothesen nicht bestätigt werden. Die Unterschiede lagen zwar alle in der vorhergesagten Richtung, wurden aber nicht signifikant. Die Effektstärken waren geringer ausgeprägt. Eventuell könnte jedoch die statistische Bedeutsamkeit mit einer höheren Stichprobengröße erreicht werden. Bestätigt werden konnte die Annahme, dass Personen, die mit seriell schnellen Bildern lernten, signifikant bessere Werte auf einer Inferenzskala erzielten als Personen, die eine Animation sahen. Die Effektstärke wies diesbezüglich einen hohen Wert auf. Da gerade diese beiden Bedingungen direkt miteinander vergleichbar waren hinsichtlich ihrer Präsentationsgeschwindigkeit und der Zahl der Wiederholungen des synaptischen Prozesses, könnte dies als ein Hinweis darauf gewertet werden, dass sich die Theorie bestätigen könnte, dass eine Darstellung eines Vorganges mit Hilfe von Key-Frames die Konstruktion eines mentalen Modells und darauf aufbauende Inferierleistungen erleichtert. Allerdings muss einschränkend hinzugefügt werden, dass die Gruppe „Animation“ der Pilotstudie hinsichtlich der Variablen „Vorwissen“ und „Erfolgswahrscheinlichkeit“ gegenüber den Bildbedingungen benachteiligt war, was ihre Inferierfähigkeiten mit verschlechtert haben dürfte.

Die Hypothesen 2a – 2d verglichen die Animations- und die seriell langsame Bildbedingung miteinander und machten ebenfalls die Annahme, dass Personen, die mit einer Animation lernten, bei den Kriterien „Wissensskala“, „Inferenzskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachtest“ im Nachteil waren. Keine dieser Hypothesen konnte bestätigt werden, da die Ergebnisse der Mittelwertsvergleiche nicht signifikant wurden, zwei der Kontraste ergaben Resultate, die nicht in der erwarteten Richtung lagen. Bei den abhängigen Variablen „Nachwissensskala“ und „Cognitive Load beim Nachtest“ war die Animationsbedingung leicht im Vorteil, während bei den Kriterien „Inferenzskala“ und „Cognitive Load beim Lernen“ das Gegenteil der Fall war. Bei den Inferierfähigkeiten erreichte die Überlegenheit der Bildbedingung einen mittleren Effekt ($d = .47$), was allerdings auch wieder unter Vorbehalt der ungünstigen Voraussetzungen der Animationsgruppe zu betrachten ist. Ansonsten zeigten die Effektstärken kleine Werte. Insgesamt ist zu sagen, dass die Animationsbedingung nur in einem der acht Fälle die vorhergesagte signifikante Unterlegenheit gegenüber den Bildbedingungen zeigte. Ein

Grund hierfür könnte sein, dass die Animation eventuell motivational ansprechender gewirkt hat, was an manchen Rückmeldungen der Probanden innerhalb der Untersuchung deutlich wurde, die sich teilweise enttäuscht zeigten, wenn sie „nur“ eine der Bildbedingungen erhalten hatten. Oft wurde auch von den Teilnehmern angemerkt, dass sie sich während des Lernens an ihr Wissen über Synapsen zurückerinnert hätten, das sie auf dem Gymnasium erworben hatten. Dieses Vorwissen könnte eventuelle Nachteile einer Präsentationsform ausgeglichen und die Leistungen im Nachtest nivelliert haben. Zuletzt hat eine Animation den inhärenten Vorteil, vollständige Informationen zu liefern (Lewalter, 1997), zumal bei ihrem Design in der aktuellen Studie darauf geachtet wurde, keine konzeptuell irrelevanten Details perzeptuell zu betonen, was zu weiteren Nachteilen geführt hätte (Lowe, 1999a).

Es bleibt aber nochmals anzumerken, dass es sich hierbei um erste explorative Ergebnisse handelte, deren Stabilität in einer größer angelegten Hauptstudie geprüft werden muss.

In den Hypothesen 3a – 3d wurde die seriell schnelle mit der seriell langsamen Bildbedingung kontrastiert. Bei jedem der vier Kriterien wurde jeweils eine ungerichtete Hypothese formuliert, da es sowohl Gründe für die Überlegenheit der seriell schnellen als auch für die der seriell langsamen Bildbedingung gab. Die Ergebnisse der Mittelwertsvergleiche zeigten, dass die Differenzen nicht signifikant wurden, so dass keine der Hypothesen bestätigt werden konnte. Allerdings war die seriell schnelle Bildbedingung in allen vier abhängigen Variablen (Wissens- und Inferenzskala, Cognitive Load beim Lernen und beim Nachtest) der seriell langsamen Bildbedingung gegenüber überlegen. Bei der Wissens- und der Inferenzskala wurden ebenso wie bei dem Cognitive Load beim Nachtest mittlere Effektstärken festgestellt. Dies spricht dafür, dass sich zumindest im Ansatz die Vorzüge der seriell schnellen Bilder durchzusetzen scheinen, d.h., dass das Lernen durch die höhere Zahl an Wiederholungen des Prozesses und durch die Betonung der „interstates“ zwischen den Bildern erleichtert wird.

Bei den Hypothesen der zweiten Hauptteilstudie, die hier explorativ getestet werden sollten, ging es darum, die parallele mit den beiden seriellen Bildbedingungen jeweils zu vergleichen. Es wurden sowohl bei dem Kontrast zwischen der parallelen und der seriell schnellen als auch bei dem zwischen der parallelen und der seriell langsamen Bildbedingung ungerichtete Hypothesen hinsichtlich der vier abhängigen Variablen „Wissens- und Inferenzskala“, „Cognitive Load beim Lernen und beim Nachtest“ formuliert.

Bei dem Vergleich zwischen den parallelen und den seriell schnellen Bildern (Hypothesen 4a – 4d) konnten die ungerichteten Hypothesen bei den Kriterien

„Inferenzskala“ und „Cognitive Load beim Nachtest“ bestätigt werden, da ein signifikanter Unterschied zwischen den Bedingungen jeweils vorlag, wobei die Effektstärken ein sehr hohes Niveau hatten. In beiden Fällen war die seriell schnelle der parallelen Bildbedingung gegenüber überlegen. Hinsichtlich der beiden anderen abhängigen Variablen „Nachwissensskala“ und „Cognitive Load beim Lernen“ wiesen die Probanden der seriell schnellen Bildbedingung ebenfalls bessere Werte als die der parallelen Bildbedingung auf. Die Unterschiede erreichten annähernd Signifikanz, die sich durch die Verwendung einer größeren Stichprobe noch verdeutlichen könnte. Die Effektstärken hatten hohe Werte. Einschränkend muss hinzugefügt werden, dass die Probanden der parallelen Bildgruppe schlechtere Werte in den Variablen „Misserfolgsbefürchtung“ und „Merkfähigkeit“ hatten als die der seriell schnellen Bildgruppe, wodurch die Unterschiede in den Kriterien „Nachwissensskala“ und „Cognitive Load beim Nachtest“ vermutlich zusätzlich vergrößert waren.

Bei den Kontrasten zwischen den seriell langsamen und den parallelen Bildern stellte sich heraus, dass auch hier in allen vier Fällen die Personen der parallelen Bildbedingung die schwächeren Lernleistungen zeigten. Allerdings waren die Unterschiede zwischen den Gruppen nicht signifikant, so dass die Hypothesen 5a – 5d nicht bestätigt werden konnten. Die Maße der praktischen Signifikanz nahmen ein mittleres bis hohes Niveau an. Anzumerken bleibt, dass die parallele Bildbedingung wiederum über ungünstigere Lernvoraussetzungen hinsichtlich der Variablen „Misserfolgsbefürchtung“ und „Merkfähigkeit“ verfügte, was bei der Interpretation der Ergebnisse bei den Kriterien „Nachwissensskala“ und „Cognitive Load beim Nachtest“ zu berücksichtigen ist.

Zusammenfassend kann gesagt werden, dass von den vier ungerichteten Hypothesen, die signifikante Unterschiede zwischen parallelen und seriell schnellen Bildern annahmen, zwei bestätigt werden konnten, wobei sich zeigte, dass ein Lernprozess mit seriell schnellen Bildern zu besseren Ergebnissen führte. Bei den anderen beiden Kriterien stellte sich ebenfalls eine deskriptive Überlegenheit der seriell schnellen Bildbedingung heraus.

Von den ungerichteten Hypothesen 5a – 5d fand keine durch die Analysen Bestätigung, wobei aber auch hier stets mit den parallelen Bildern geringere Lernerfolge erzielt wurden.

Das bedeutet, dass in dieser explorativen Untersuchung in allen Fällen die seriellen der parallelen Bildbedingung gegenüber überlegen waren, wodurch die Vorteile sequentiell angeordneter Bilder ein größeres Gewicht zu haben scheinen. Veränderungen zwischen den Bildern werden bei einer solchen Darstellungsweise salienter, und der Verlauf des

Lernvorganges wird extern vorbestimmt, so dass einer eventuellen Überforderung des Lerners durch eine zu hohe Selbststeuerung vorgebeugt wird. Die Möglichkeit, aber auch die Anforderung zum selbstgesteuerten Lernen war bei den parallelen Bildern in der vorliegenden Studie durch das leicht interaktive Feature (Scrollen der Taskleiste, um alle Bilder zu sehen) noch verstärkt, was auch zu Lerndefiziten geführt haben könnte. In der Beobachtung der Probanden durch die Versuchsleiterin fiel auf, dass die Personen, die mit den parallelen Bildern lernten, oft nicht die gesamte Lernzeit zu nutzen schienen und aus eventuell nicht vorhandener Motivation oder aus Überschätzung des eigenen Verständnisses heraus die Bilder eher willkürlich und zufällig zu betrachten schienen, was durch ständiges Hoch- und Runterscrollen deutlich wurde.

Wie schon mehrmals betont wurde, handelte es sich bei den Ergebnissen der Mittelwertsvergleiche der Pilotstudie um erste explorative Hinweise, die in der Hauptstudie mit einer größeren Stichprobe und unter exakterem Einbezug der Kovariaten nochmals überprüft werden sollen.

4.3 Hauptstudie

4.3.1 Teilstudie 1

4.3.1.1 Fragestellung und Hypothesen

Haupteffekte

In der ersten Hauptteilstudie sollte die Animations- mit der seriell schnellen sowie mit der seriell langsamen Bildbedingung kontrastiert werden. Es wurde ebenfalls ein Vergleich zwischen seriell schnellen und seriell langsamen Bildern angestrebt. Die Hypothesen zwischen der Animations- und der seriell schnellen sowie der seriell langsamen Bildbedingung wurden jeweils zu Gunsten der Bildbedingungen gerichtet, die Annahmen zwischen den beiden seriellen Bildbedingungen wurden jeweils ungerichtet formuliert.

Animation vs. Seriell schnelle Key-Frames

Hypothese 1a: Personen, die mit seriell schnellen Key-Frames lernen, erzielen signifikant bessere Werte in den Wissensaufgaben eines Nachtests, als Personen, die mit Animationen lernen.

Hypothese 1b: Personen, die mit seriell schnellen Key-Frames lernen, erzielen signifikant bessere Werte in den Inferenzaufgaben eines Nachtests, als Personen, die mit Animationen lernen.

Hypothese 1c: Personen, die mit seriell schnellen Key-Frames lernen, empfinden signifikant weniger Cognitive Load beim Lernen, als Personen, die mit Animationen lernen.

Hypothese 1d: Personen, die mit seriell schnellen Key-Frames lernen, empfinden signifikant weniger Cognitive Load beim Lösen des Nachtests, als Personen, die mit Animationen lernen.

Animation vs. Seriell langsame Key-Frames

Hypothese 2a: Personen, die mit seriell langsamen Key-Frames lernen, erzielen signifikant bessere Werte in den Wissensaufgaben eines Nachttests, als Personen, die mit Animationen lernen.

Hypothese 2b: Personen, die mit seriell langsamen Key-Frames lernen, erzielen signifikant bessere Werte in den Inferenzaufgaben eines Nachttests, als Personen, die mit Animationen lernen.

Hypothese 2c: Personen, die mit seriell langsamen Key-Frames lernen, empfinden signifikant weniger Cognitive Load beim Lernen, als Personen, die mit Animationen lernen.

Hypothese 2d: Personen, die mit seriell langsamen Key-Frames lernen, empfinden signifikant weniger Cognitive Load beim Lösen des Nachttests, als Personen, die mit Animationen lernen.

Seriell schnelle Key-Frames vs. Seriell langsame Key-Frames

Hypothese 3a: Personen, die mit seriell schnellen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Wissensaufgaben eines Nachttests, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 3b: Personen, die mit seriell schnellen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Inferenzaufgaben eines Nachttests, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 3c: Personen, die mit seriell schnellen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lernen, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 3d: Personen, die mit seriell schnellen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lösen des Nachttests, als Personen, die mit seriell langsamen Key-Frames lernen.

Interaktionen

Interaktionen zwischen den Faktoren „Bedingung“ und „Vorwissen“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit dem Vorwissen von Personen

Bei diesen Hypothesen wurde keine genaue Richtung der Interaktionen angegeben, da es zum einen möglich war, dass das Vorwissen einer Person mehr das Lernen mit Animationen als das mit seriellen Bildern beeinflusst, es zum anderen aber auch plausibel war, dass das entgegengesetzte Interaktionsmuster sich durchsetzen könnte. Für seriell schnelle und seriell langsame Key-Frames wurden jeweils unabhängige Hypothesen aufgestellt, da durch die verschiedene Präsentationsdauer der Bilder in den beiden Bedingungen lernrelevante Merkmale hätten verändert sein können.

Hypothesen 4a - d: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *schnell*“ mit dem Vorwissen von Personen. Diese Annahme bezieht sich auf

a: die Leistung in Wissensaufgaben eines Nachttests

b: die Leistung in Inferenzaufgaben eines Nachttests

c: den Cognitive Load beim Lernen

d: den Cognitive Load beim Lösen des Nachttests

Hypothesen 5a - d: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *langsam*“ mit dem Vorwissen von Personen. Diese Annahme bezieht sich auf

a: die Leistung in Wissensaufgaben eines Nachttests

b: die Leistung in Inferenzaufgaben eines Nachttests

c: den Cognitive Load beim Lernen

d: den Cognitive Load beim Lösen des Nachtests

Interaktionen zwischen den Bedingungen „Key-Frames seriell schnell“ und „Key-Frames seriell langsam“ mit dem Vorwissen von Personen

Bezüglich der Interaktionen zwischen den Bedingungen „Bild seriell schnell“ und „Bild seriell langsam“ mit dem Vorwissen von Personen wurden ungerichtete Hypothesen aufgestellt, da aus theoretischer Sicht mehrere Interaktionsmuster plausibel waren.

Hypothesen 6a - d: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Bild seriell schnell“ und „Bild seriell langsam“ mit dem Vorwissen von Personen. Diese Annahme bezieht sich auf

a: die Leistung in Wissensaufgaben eines Nachtests

b: die Leistung in Inferenzaufgaben eines Nachtests

c: den Cognitive Load beim Lernen

d: den Cognitive Load beim Lösen des Nachtests

Interaktionen zwischen den Faktoren „Bedingung“ und „Räumliche Fähigkeiten“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit den räumlichen Fähigkeiten von Personen

In Analogie zu den obigen Interaktionen mit der Variable „Vorwissen“, wiesen auch die Hypothesen zu den Interaktionen zwischen der Animations- und der seriell schnellen bzw. der seriell langsamen Bildbedingung mit den räumlichen Fähigkeiten von Personen keine bestimmte Richtung auf. Der Lernvorteil durch Key-Frames könnte bei Personen mit geringeren räumlichen Fähigkeiten größer sein als bei Personen mit höherem räumlichem Vorstellungsvermögen, es könnte aber auch das gegenteilige Muster der Fall sein.

Hypothesen 7a - d: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *schnell*“ mit den räumlichen Fähigkeiten von Personen. Diese Annahme bezieht sich auf

a: die Leistung in Wissensaufgaben eines Nachttests

b: die Leistung in Inferenzaufgaben eines Nachttests

c: den Cognitive Load beim Lernen

d: den Cognitive Load beim Lösen des Nachttests

Hypothesen 8a - d: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *langsam*“ mit den räumlichen Fähigkeiten von Personen. Diese Annahme bezieht sich auf

a: die Leistung in Wissensaufgaben eines Nachttests

b: die Leistung in Inferenzaufgaben eines Nachttests

c: den Cognitive Load beim Lernen

d: den Cognitive Load beim Lösen des Nachttests

Interaktionen zwischen den Faktoren „Bedingung“ und „Merkfähigkeit“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit der Merkfähigkeit von Personen

Zuletzt wurden noch ungerichtete Hypothesen zu den Interaktionen zwischen der Animations- und der seriell schnellen bzw. der seriell langsamen Bildbedingung mit der Merkfähigkeit von Personen aufgestellt. Die Wechselwirkungen, die die seriell schnellen und die seriell langsamen Key-Frames miteinbeziehen, wurden wiederum in jeweils eigenständigen Hypothesen berücksichtigt, da durch die unterschiedlichen Präsentationsgeschwindigkeiten differentielle Effekte hätten entstehen können.

Hypothesen 9a - d: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *schnell*“ mit der Merkfähigkeit von Personen. Diese Annahme bezieht sich auf

a: die Leistung in Wissensaufgaben eines Nachttests

b: die Leistung in Inferenzaufgaben eines Nachttests

c: den Cognitive Load beim Lernen

d: den Cognitive Load beim Lösen des Nachttests

Hypothesen 10a - d: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell *langsam*“ mit der Merkfähigkeit von Personen. Diese Annahme bezieht sich auf

a: die Leistung in Wissensaufgaben eines Nachttests

b: die Leistung in Inferenzaufgaben eines Nachttests

c: den Cognitive Load beim Lernen

d: den Cognitive Load beim Lösen des Nachttests

4.3.1.2 Methoden

Die Durchführung der Teilstudien 1 und 2 der Hauptstudie wurde aus pragmatischen Gründen zusammengelegt, so dass das Gesamtdesign und die Messinstrumente anhand eines Methodenteils beschrieben werden.

4.3.1.2.1 Stichprobe

In der Stichprobe der gesamten Hauptstudie befanden sich 87 Personen, die größtenteils an der Universität Landau Lehramt, Pädagogik, Sozialwissenschaften oder Magisterkombinationen studierten. Die Anwerbung der Probanden erfolgte durch direkte Ansprache in einer Lehramtsvorlesung sowie durch mehrere Rundmails, die der gesamten Studentenschaft zugestellt wurden. Von der Teilnahme ausgeschlossen waren wiederum Psychologie-, Biologie- und Umweltwissenschaftsstudenten, ebenso wie Personen, die schon an der Vorpilot- oder an der Pilotstudie partizipiert hatten.

Nach der Durchführung der Studie wurden zwei Probanden aus den Analysen eliminiert, der eine aufgrund doppelter Versuchsteilnahme, der andere wegen einer auffällig inkonsistenten Bearbeitung des Nachtests. In der statistischen Ausreißeranalyse mittels Box Plots, die mit der gesamten Stichprobe durchgeführt wurde, zeigte sich, dass weitere sechs Personen durch Extremwerte in den Variablen in der Auswertung nicht berücksichtigt werden konnten, so dass die für die Mittelwertsvergleiche letztendlich maßgebliche Stichprobe zusammengenommen 79 Probanden umfasste.

Der Altersrange der Stichprobe lag zwischen 18 und 30 Jahren mit einem Mittelwert von 22.1 Jahren ($SD = 2.4$). Es nahmen 17 Männer und 62 Frauen an der Studie teil, was einem Geschlechterverhältnis von ca. einem Fünftel Männer (21,5%) zu vier Fünfteln (78,5%) Frauen entsprach. Somit war das weibliche Geschlecht in der Stichprobe wieder sehr stark repräsentiert.

4.3.1.2.2 Design

Da, wie oben bereits angedeutet, die beiden Teilstudien gemeinsam durchgeführt wurden, war das Design des Experiments, was die Durchführung betraf, einfaktoriell (Faktor = Präsentationsform der Lernmaterialien zu den synaptischen Prozessen) mit vier Treatmentstufen. Die Mittelwertsvergleiche in den beiden Teilstudien wurden jeweils mit einfaktoriellen Designs mit drei Treatmentstufen durchgeführt. Die Interaktionshypothesen, die als voneinander unabhängige Fragestellungen angesehen wurden, wurden in jeweils zweifaktoriellen Designs zusätzlich getestet.

Die Versuchspersonen wurden für die Durchführung den Gruppen randomisiert zugewiesen, so dass eine Zellenbesetzung von 21 Personen in der Animations-, 20 in der seriell schnellen Bild-, 20 in der seriell langsamen Bild-, sowie 18 in der parallelen Bildbedingung entstand. Abhängige Variablen waren wiederum die Wissens- und die Inferenzskala eines Nachttests sowie der Cognitive Load, der nach der Betrachtung des Lernmaterials und der Bearbeitung des Nachttests erhoben wurde.

Da ein Großteil der Kontrollvariablen, wie theoretisch erwartet, in der Pilotstudie mit den abhängigen Variablen größere bis signifikante Korrelationen gezeigt hatte und ihr Einfluss nochmals bei einer größeren Stichprobe ermittelt werden sollte, wurden als mögliche Kovariaten erneut das Vorwissen zur synaptischen Übertragung, die aktuelle Lernmotivation, das begriffliche Wissen zu Synapsen und die verbalen, räumlichen sowie Merkfähigkeiten der Probanden innerhalb der Untersuchung berücksichtigt.

4.3.1.2.3 Lernmaterialien

Das Lernmaterial wurde größtenteils in gleichbleibender Form aus der Pilotstudie übernommen. Das bedeutete auch, dass die Extraktion der Key-States und die graphische Gestaltung des Materials analog zu der vorherigen Untersuchung erfolgte, so dass an dieser Stelle aus Redundanzgründen auf eine nochmalige ausführliche Beschreibung verzichtet wird (für Details der Lernmaterialgestaltung, s. Abschnitt 4.2.2.3).

Die einzige Veränderung in den Materialien fand in der Bedingung statt, in der die Bilder mit langsamerer Geschwindigkeit aufeinander folgten („Bild seriell langsam“). Um den Kontrast zu der seriell schnellen Bildbedingung noch zu erhöhen, wurde der Gesamtvorgang nur noch zwei statt wie vorher drei Mal in der langsamen Bildbedingung wiederholt. Dies

hatte zur Folge, dass ein Lerndurchgang 120 s lang dauerte und ein einzelnes Bild eine Präsentationsdauer von 12,0 s hatte. Zwischen den beiden Durchgängen wurde 1 s Pause gelassen. Mit dieser Veränderung konnte mit größerer Deutlichkeit überprüft werden, ob eine ausführliche Betrachtung des Einzelbildes (wie bei „Bild seriell langsam“) oder die wiederholte Darstellung der größeren Zusammenhänge (wie bei „Bild seriell schnell“) wichtiger sind, um einen besseren Lernerfolg in Wissens- und Inferenzaufgaben zu erzielen (zur Ansicht der Lernmaterialien, s. Anhang C1 auf der beiliegenden CD).

4.3.1.2.4 Messinstrumente

In der Hauptstudie kamen die gleichen Messverfahren in derselben Abfolge wie in der Pilotstudie zum Einsatz, da die berücksichtigten Variablen ebenfalls identisch waren. Daher wird auf die nochmalige Nennung der standardisierten Instrumente verzichtet, und es werden nur noch die Verfahren angeführt, die für die Untersuchungen eigens konstruiert und für die Hauptstudie teilweise überarbeitet worden waren.

Soziodemographischer Fragebogen

Der soziodemographische Fragebogen der Hauptstudie entsprach größtenteils dem der Pilot- und der Vorpilotstudie (s. Abschnitte 4.1.2.4, 4.2.2.4). Allerdings wurden bei den Variablen „Schulabschluss“ und „Beruf“ im Vergleich zu den vorigen Untersuchungen Antwortalternativen eliminiert, da die Stichprobe streng homogen aus Studenten zusammengesetzt sein sollte, so dass Antwortmöglichkeiten wie „Mittlere Reife“ oder „Beamter“ überflüssig wurden (zur Ansicht des Fragebogens, s. Anhang C2 - 1).

Vorwissenstest

Der Vorwissenstest entsprach hinsichtlich seiner Gestaltungs- und Durchführungsmerkmale den Versionen aus den vorigen Untersuchungen (s. Abschnitte 4.1.2.4, 4.2.2.4 und Anhang A2 - 2).

Die Items wurden in gleicher Form und Reihenfolge aus der Pilotstudie übernommen (s. Anhang B2 - 2). Das Cronbachs Alpha der Skala hatte in der Pilotstudie $\alpha = .53$ betragen,

was einen mittelmäßigen Wert darstellte und für eine Weiterverwendung der Aufgaben vertretbar gewesen war. Ungefähr die Hälfte der Items hatte in der Pilotstudie sehr niedrige Schwierigkeitsindizes gehabt, die konzeptuell verständlich gewesen waren, da die Aufgaben an die Versuchspersonen, die bewusst über möglichst wenig Vorwissen verfügen sollten, hohe Anforderungen gestellt hatten. Daher war die Schwierigkeit ein weniger strenges Kriterium gewesen, um ungeeignete Items zu eliminieren. Die Trennschärfen der Aufgaben hatten in der Pilotstudie ein geringes bis mittelmäßiges Niveau gehabt, was an den eben erwähnten niedrigen Itemschwierigkeiten sowie an der inhaltlich heterogeneren Abfrage des Vorwissens durch den Test gelegen haben könnte. Da zudem der Stichprobenumfang der Pilotstudie eher gering und daher die Schätzungen der Kennwerte noch instabil gewesen waren, wurde die Entscheidung getroffen, alle Items in der Hauptstudie weiterzuverwenden. Eine zu starke Kürzung der Skala hätte zudem eine unvollständige Abbildung des Konstruktes und ebenfalls Minderungen in der Reliabilität zur Konsequenz gehabt.

Für das Rating der Antworten wurden zunächst die Ratingsysteme aus der Pilotstudie herangezogen (für eine genauere Beschreibung des Aufbaus der Ratingsysteme, s. Abschnitt 4.2.2.4). Die Summe der Punkte, die ein Proband bei den einzelnen Fragen erreicht hatte, stellte seinen Gesamtscore für das Vorwissen dar. Das Vorwissen wurde als eine potentielle Kovariate in den Analysen der beiden Hauptteilstudien berücksichtigt. Die Ratingsysteme wurden mit den Daten der ersten 60 Versuchspersonen der gesamten Hauptstudie inhaltlich weiterentwickelt. Danach wurden die Systeme in ihren Formulierungen vereinheitlicht und in ihren Strukturen möglichst standardisiert, um ihren Aufbau für einen zweiten Rater verständlicher zu machen. Da erste Berechnungen von Interraterreliabilitäten gute Kennwerte ergaben und auch inhaltlich kaum neue Lösungen in den Daten vorhanden waren, wurden die Ratingsysteme für die Auswertung der restlichen Versuchspersonen nicht mehr weiter abgeändert (zur Ansicht der nach den beiden Hauptteilstudien weiterentwickelten Ratingsysteme, s. Anhang C3, Datei: „Ratingsysteme_Vorwissenstest_Haupt.pdf“ auf der beigefügten CD.)

Es wurde für jedes Ratingsystem eine Intraclass-Korrelation (ICC) berechnet, da das Skalenniveau der einzelnen Stufen als metrisch angenommen wurde. Hierzu wurde die Hälfte der Bögen als Stichprobe herangezogen, die vor der Ausreißeranalyse vorhanden gewesen waren. Somit waren 43 von 86 Versuchspersonen für die Analyse maßgeblich. Als zweiten Rater stellte sich für die gesamte Stichprobe eine studentische Hilfskraft zur Verfügung.

Test zur Erfassung des begrifflichen Wissens zu Synapsen

Dieser Test war in seiner Form und Durchführung mit der Version, die in der Pilotstudie verwendet worden war, identisch (s. Abschnitt 4.2.2.4 und Anhang B2 - 5), und die Gesamtpunktzahl, die von den Probanden bei den Aufgaben jeweils erzielt worden war, ging, aggregiert gesehen, wieder als eine möglicherweise zu berücksichtigende Kovariate in die statistischen Analysen mit ein.

Es wurden alle Items aus der Pilotstudie übernommen, da das Cronbachs Alpha der Skala ($\alpha = .66$) sowie die Trennschärfen der Aufgaben größtenteils zufriedenstellend bis gut gewesen waren. Die Schwierigkeitsindizes der Items waren in der Pilotstudie fast durchgehend hoch gewesen ($p > .80$), was an der für Studenten relativ einfachen Aufgabenstellung des Reproduzierens gelegen haben könnte, zumal sie sich, nach eigener Aussage, beim Lernen des Standbildes an ihr Schulwissen erinnert hatten. Allerdings war jedes einzelne Item in seinem Einfluss auf die Scores der Wissens- und der Inferenzskala von Bedeutung gewesen, da konzeptuell gesehen, das Wissen um jeden Begriff die Beantwortung der Nachtestaufgaben eigenständig mit moderiert. Daher fanden bei diesem Test in der Hauptstudie keine weiteren Änderungen mehr statt.

Nachtest

Der Nachtest wurde in seiner Gestaltung und Durchführung analog zu den vorigen Studien konzeptualisiert (s. Abschnitte 4.1.2.4, 4.2.2.4 und Anhang A2 - 3). In Abänderung zur Pilotstudie wurde die konkrete Zahl der Aufgaben in die Instruktion eingefügt, zudem wurden die Teilnehmer mündlich darum gebeten, die Fragen „ernsthaft“ zu durchdenken und ggf. dann auch zu beantworten.

Die Items wurden hinsichtlich ihrer Reihenfolge in gleichbleibender Form aus der Pilotstudie übernommen.

Die Wissensskala hatte in der Pilotstudie ein Cronbachs Alpha von $\alpha = .84$ aufgewiesen, was einen hohen Wert dargestellt hatte. Die Trennschärfen der Items hatten, bis auf zwei Ausnahmen, im mittleren bis hohen Bereich (Range (r_{it}) = .28 - .85) gelegen. Auch die Schwierigkeitsindizes der Aufgaben hatten größtenteils mittlere Werte gezeigt.

Da es sich innerhalb der Pilotstudie um eine geringere Stichprobengröße gehandelt hatte, war die Itemanalyse eher konservativ durchgeführt worden, wodurch alle Aufgaben,

aufgrund der guten Reliabilität der Skala und der befriedigenden Trennschärfen, in die Hauptstudie mit aufgenommen wurden.

Die Items blieben in der Mehrheit in ihrer bestehenden Form erhalten. Lediglich bei zwei der Items wurden leichte Änderungen in ihrer Formulierung und Gestaltung im Vergleich zur Pilotstudie vorgenommen.

Die Inferenzskala hatte in der Pilotstudie ebenfalls einen zufriedenstellenden Alpha-Koeffizienten von $\alpha = .78$ gehabt. Die Aufgaben hatten vorwiegend mittlere bis gute Trennschärfen aufgewiesen. Bis auf wenige Ausnahmen, hatten die Schwierigkeitsindizes der Items im mittleren Bereich gestreut. Da die Kennwerte somit in ihrer Mehrheit zufriedenstellend gewesen waren und es sich um noch instabile Schätzungen aus der Pilotstudie gehandelt hatte, wurden wiederum sämtliche Items in den Nachtest der Hauptstudie integriert. Die Formulierungen der Aufgaben blieben, bis auf die von einem Item, in ihrer bestehenden Form erhalten (zur Ansicht der Nachtestitems, s. Anhang C2 - 2).

Die Ratingsysteme wurden, wie bei dem Vorwissenstest, in ihrem Anfangszustand aus den weiterentwickelten Versionen der Pilotstudie konstituiert. Die einzelnen Punktzahlen, die ein Teilnehmer bei den Aufgaben erzielt hatte, wurden nach Wissens- und Inferenzskala getrennt aufsummiert, und die Gesamtscores stellten zwei abhängige Variablen der beiden Teilstudien jeweils dar. Im Laufe der Hauptstudie wurden die Systeme mit Hilfe der vorliegenden Daten inhaltlich weiter ausgearbeitet, und die Formulierungen wurden vereinheitlicht. Da erste Berechnungen von Interraterreliabilitäten nach einer Stichprobe von ca. 60 Personen gute Ergebnisse zeigten, wurden die Ratingsysteme danach nicht mehr weiter durch empirische Daten angepasst (zur Ansicht der nach den beiden Hauptteilstudien weiterentwickelten Ratingsysteme, s. Anhang C3, Datei: „Ratingsysteme_Nachtest_Haupt.pdf“ auf der beigefügten CD.)

Im Rahmen der Itemanalyse in der Gesamtauswertung, wurde für jedes System, aufgrund einer angenommenen Metrik des Skalenniveaus, eine Intraclass-Korrelation mit Hilfe von 50% der Stichprobe berechnet, was einer Zahl von 43 Versuchspersonen entsprach. Das Gegenrating wurde bei allen Fällen wieder von einer studentischen Hilfskraft übernommen.

4.3.1.2.5 Durchführung

Die Durchführung der Hauptstudie fand im Wintersemester 2009 / 2010 sowie im Sommersemester 2010 statt. Durch einen Umzug der Abteilung erfolgten die Erhebungen schließlich in zwei unterschiedlichen Räumen, wobei die externen Bedingungen miteinander vergleichbar waren, da die gleichen Computer verwendet wurden und auch sonstige Einflussfaktoren, wie z.B. der Lärmpegel, ähnliche Ausmaße hatten. Die Versuchsleitung wurde von drei Hilfskräften übernommen, die zuvor eine Unterweisung erhalten und auch selbst an der Untersuchung teilgenommen hatten (außer Wertung), um mit den Testverfahren und mit eventuell auftretenden Problemen vertraut zu werden. Zudem lag ihnen ein detailliert beschriebenes, standardisiertes Ablaufschema des Experimentes während der Erhebungen vor, wodurch eine höhere Durchführungsobjektivität gewährleistet wurde. Die Testungen verliefen, je nach Zahl der Anmeldungen für einen Termin, in Einzel- oder Gruppensitzungen bis zu sechs Personen. Die Probanden wurden den Versuchsbedingungen randomisiert zugewiesen.

Der Ablauf der einzelnen Experimente erfolgte analog zur Durchführung in der Pilotstudie (s. auch Abschnitt 4.2.2.5). Eine Neuerung im Vergleich zu der vorigen Untersuchung fand insofern statt, als dass den Probanden direkt vor dem Lernen der Prozesse durch die Versuchsleiterin mündlich die Information gegeben wurde, dass „eine Synapse eine Verbindung zwischen zwei Nervenzellen ist“, da in der Pilotstudie aufgefallen gewesen war, dass manche Personen keinerlei Vorwissen zu Synapsen gehabt hatten, wodurch ihnen die Grundorientierung gefehlt hatte, um die Lerneinheit verstehen zu können.

Die äußeren Bedingungen wurden bei den Erhebungen der Hauptstudie möglichst gleich gehalten, und die Versuchsleiterinnen waren angehalten, größere Störungen und technische Probleme auf den entsprechenden Bögen der Versuchspersonen jeweils zu vermerken. Insgesamt gesehen, verliefen die Durchführungen weitgehend störungsfrei. Zudem beobachteten die Versuchsleiterinnen, ob eine mangelhafte Bearbeitung des Nachtests auf eine eventuell fehlende Motivation der Probanden zurückzuführen war, damit solche Fälle aus der späteren Auswertung eliminiert werden konnten, wie es auch bei einer Person schließlich von Nöten war. Im Unterschied zur Pilotstudie wurden die Probanden während der Durchführung der Hauptstudie mündlich darauf hingewiesen, den Nachtest „so gut wie möglich“ und „möglichst ernsthaft“ auszufüllen.

Die Computer wurden regelmäßig gewartet, und es wurde überprüft, ob die Lerneinheiten hinsichtlich ihrer Dauer und Bildqualität miteinander vergleichbar abliefen. Technische Probleme wurden registriert und behoben.

4.3.1.3 Ergebnisse

Sämtliche statistischen Analysen dieser Teilstudie wurden mit dem Programm „SPSS, Version 17.0“ durchgeführt.

Itemanalyse

Die Itemanalyse der Instrumente wurde für beide Teilstudien der Hauptstudie gemeinsam vorgenommen, um durch eine höhere Stichprobengröße stabilere Schätzungen der Kennwerte zu erhalten.

Zunächst wurden für den Vorwissenstest, die Wissens- und Inferenzskala des Nachttests sowie für den Test zu begrifflichem Wissen, die die selbst konstruierten Verfahren der Untersuchung darstellten, Reliabilitäten, Itemtrennschärfen und Itemschwierigkeiten berechnet. Für die drei ersten Instrumente waren Ratingsysteme erstellt worden, für die Intraclass-Korrelationen (ICCs) nach dem Kriterium der absoluten Übereinstimmung berechnet wurden, um ihre Reliabilität zu beurteilen. Hierzu wurde die Hälfte der Stichprobe ($n = 43$) herangezogen, wobei diese Zahl bei verschiedenen Items unterschiedlich sein konnte, je nachdem wie oft die entsprechende Aufgabe nicht beantwortet worden war.

In die weiteren Analysen gingen solche Items ein, deren Trennschärfen über $r_{it} = .30$ und deren Schwierigkeiten in einem mittleren Bereich ($p = .20 - p = .80$) lagen, wobei dieses Kriterium bei den Items des Vorwissenstests weniger streng angelegt wurde, da die Aufgaben für Novizen konzeptuell schwieriger waren und da sie in der Untersuchung größtenteils befriedigende Trennschärfen hatten. Ebenso gingen bei den Beschriftungsaufgaben zu den Begriffen aus theoretischen Gründen alle Items in die Analysen mit ein. Bei den Intraclass-Korrelationen lag der Mindestwert zur Beibehaltung bei $r_{icc} = .70$, wobei geringere Abweichungen bei ansonsten zufriedenstellenden Kennwerten der Items toleriert wurden.

Nach der Itemanalyse und Itemelimination stellten sich die Instrumente und ihre Gütekriterien wie folgt dar:

Tabelle 4.3.1.1

Gütekriterien des Vorwissenstests, der Wissens- und der Inferenzskala sowie der Begriffsaufgabe in der Hauptstudie

	<i>n</i> (Items)	α	<i>rit</i>	<i>p</i>	<i>rICC</i>
Vorwissenstest	9	.70	.25 – .55	.05 – .72	.63 – 1.00
Wissensskala	12	.92	.49 – .81	.47 – .74	.72 – .98
Inferenzskala	17	.89	.33 – .68	.32 – .76	.72 – 1.00
Begriffsaufgabe	10	.81	.19 – .62	.55 – .93	--

Die Tabelle zeigt, dass nach der Itemanalyse zwei Items des Vorwissenstests und ein Item der Inferenzskala eliminiert werden mussten. Ansonsten gingen alle Items in die weiteren Analysen mit ein. Die Reliabilitäten der Instrumente waren als zufriedenstellend bis sehr gut zu bezeichnen.

Bei den standardisierten Verfahren wurden die Reliabilitäten der Subskalen des FAM für die aktuelle Stichprobe berechnet, die nach der Itemelimination die folgenden Werte hatten:

Tabelle 4.3.1.2

Reliabilitäten der Subskalen des FAM in der Hauptstudie

	<i>n</i> (Items)	α
Erfolgswahrscheinlichkeit	4	.80
Misserfolgsbefürchtung	5	.84
Interesse	5	.80
Herausforderung	3	.57

Bei der Skala „Herausforderung“ war ein Item eliminiert worden, um ihr etwas schwächeres Alpha noch zu erhöhen und um somit eine bessere Reliabilität zu gewährleisten.

Ausreißeranalyse

Es wurde bei den Daten eine Ausreißeranalyse mit Hilfe von Box Plots durchgeführt. Als Auswertungsgrundlage wurde wiederum die Gesamtstichprobe mit allen vier Versuchsbedingungen herangezogen, um eine Vergleichbarkeit der Gruppen zwischen den beiden Teilstudien zu gewährleisten. Eine zusätzliche Ausreißeranalyse mit den beiden einzelnen Teilstichproben ergab zudem keine größeren Unterschiede zu den Ergebnissen der Gesamtanalyse, so dass ein solches Vorgehen gerechtfertigt war.

Die Ausreißeranalysen wurden für alle vier abhängigen Variablen und neun Kontrollvariablen durchgeführt. Bei den Kontrollvariablen wurde jeweils ein Gesamtplot, bei den abhängigen Variablen wurden zusätzlich noch bedingungsabhängige Plots erstellt, um die Lage der Daten zu beurteilen. Ausgeschlossen wurden solche Fälle, die jenseits der Whisker lagen und die von SPSS als Ausreißer- bzw. als Extremwerte angezeigt wurden. Generell wurde ein eher konservatives Auswahlverfahren gewählt, um eine möglichst hohe Fallzahl und eine somit höhere Power der nachfolgenden Verfahren zu gewährleisten. Es wurden sechs Fälle aufgrund der Ausreißeranalyse aus den Daten eliminiert, so dass 79 Personen in der Gesamtstichprobe verblieben.

Deskriptive Statistiken

Die Darstellung der deskriptiven Statistiken und alle weiteren statistischen Analysen beziehen sich auf die drei Versuchsbedingungen „Animation“, „Bild seriell schnell“ und „Bild seriell langsam“, die in der ersten Teilstudie miteinander verglichen werden sollten. Die Größe dieser Teilstichprobe betrug $n = 61$. Von diesen 61 Probanden waren 16 Männer und 45 Frauen. Das Durchschnittsalter belief sich auf 21.97 Jahre ($SD = 2.33$) mit einem Range von 18 bis 30 Jahren.

Nachfolgend seien die deskriptiven Statistiken der abhängigen und der Kontrollvariablen der ersten Hauptteilstudie aufgelistet:

Tabelle 4.3.1.3

Deskriptive Statistiken der abhängigen Variablen in der ersten Hauptteilstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Nachwissensskala	61	.00	24.00	15.67	7.02
Inferenzskala	61	.00	33.00	19.90	8.63
Cognitive Load (Lernaufgabe)	61	4.00	15.00	11.07	2.82
Cognitive Load (Nachttest)	61	9.00	17.00	13.75	2.14

Tabelle 4.3.1.4

Deskriptive Statistiken der Kontrollvariablen in der ersten Hauptteilstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Vorwissenstest	61	.00	12.00	4.54	3.25
Begriffswissen	61	1.00	10.00	7.56	2.60
IST verbal	61	14.00	32.00	23.59	4.06
IST räumlich	61	10.00	40.00	21.56	6.81
IST Merkfähigkeit	59	7.00	23.00	15.47	4.15
Erfolgswahrscheinlichkeit	61	9.00	26.00	19.52	3.98
Misserfolgsbefürchtung	61	5.00	31.00	14.72	6.71
Interesse	59	11.00	31.00	22.10	4.87
Herausforderung	60	11.00	20.00	16.07	2.39

Mittelwertsvergleiche

Testung auf Normalverteilung und Varianzhomogenität

Um die Hypothesen zu den Haupteffekten zwischen den drei Bedingungen zu testen, wurden Varianzanalysen (ANOVAs) und Kovarianzanalysen (ANCOVAs) angewandt. Diese Verfahren setzen Verteilungsmerkmale wie Normalverteilung der Daten und Varianzhomogenität zwischen den Bedingungen voraus, die in dieser Studie mittels Kolmogorov-Smirnov-Tests und Levene-Tests überprüft wurden.

Die Kolmogorov-Smirnov-Tests mit den abhängigen Variablen ergaben, dass bei der Inferenzskala, beim Cognitive Load während der Lerneinheit sowie beim Cognitive Load während des Nachtests Normalverteilungen in den Daten vorlagen. Lediglich bei der AV „Nachwissensskala“ wurde diese Voraussetzung nicht erfüllt (Kolmogorov-Smirnov-Z: 1.49, $p = .02$).

Durch eine Überprüfung anhand von Levene-Tests konnte bei allen vier Kriterien Varianzhomogenität zwischen den Bedingungen festgestellt werden.

Da Varianzhomogenität, Normalverteilung der Daten und gleiche Stichprobengrößen bei den abhängigen Variablen (bis auf eine Abweichung) gegeben waren, waren wesentliche Voraussetzungen für die Verfahren der Varianz- und der Kovarianzanalyse erfüllt, die somit für die Mittelwertsvergleiche angewandt werden konnten.

Bestimmung relevanter Kovariaten

Bei allen Kontrollvariablen wurde mit Hilfe von Varianzanalysen (ANOVAs) bzw. Brown-Forsyth-Tests (bei vorhandener Varianzheterogenität) geprüft, ob größere Mittelwertsunterschiede zwischen den Bedingungen vorlagen. Dies war nur bei den Variablen „Misserfolgsbefürchtung“ ($F(2, 58) = 6.34, p = .003, \eta^2 = .18$) und „Erfolgswahrscheinlichkeit“ ($F(2, 58) = 2.59, p = .08, \eta^2 = .08$) der Fall. Bei der Variable „Erfolgswahrscheinlichkeit“ zeigten sich signifikante Korrelationen zu den abhängigen Variablen „Inferenzskala“ ($r = .24, p$ (1-seitig) = .03) und „Cognitive Load beim Lernen“ ($r = -.26, p$ (1-seitig) = .03), bei der Variable „Misserfolgsbefürchtung“ lag eine signifikante Korrelation zu dem Kriterium „Cognitive Load beim Lernen“ vor ($r = .32, p$ (1-seitig) = .005).

Da die Variable „Erfolgswahrscheinlichkeit“ in der ANCOVA mit der AV „Cognitive Load beim Lernen“ keinen signifikanten Einfluss mehr hatte und das sparsamste Modell gewählt werden sollte, wurden folgende ANOVAs bzw. ANCOVAs durchgeführt:

1. AV: Nachwissensskala Verfahren: ANOVA
2. AV: Inferenzskala Verfahren: ANCOVA mit „Erfolgswahrscheinlichkeit“
3. AV: CL Lern Verfahren: ANCOVA mit „Misserfolgsbefürchtung“
4. AV: CL Nachtest Verfahren: ANOVA

„CL Lern“ steht für den Cognitive Load, den die Probanden bei der Betrachtung des Lernmaterials empfunden haben. Die AV „CL Nachtest“ bezeichnet hingegen den Cognitive Load, der nach der Beantwortung des Nachtests bei den Probanden gemessen wurde. In den folgenden Abschnitten werden die Abkürzungen „CL Lern“ und „CL Nachtest“ in den meisten Fällen gebraucht, um die entsprechenden Kriterien zu benennen.

Varianz- und Kovarianzanalysen

Die Ergebnisdarstellung der Haupt- und der späteren Interaktionseffekte wird in der Form aufgebaut, dass jeweils zunächst die Hypothesen des Theorieteils noch einmal in statistischer Kurzform genannt werden und die dazugehörigen Testungen und Ergebnisse direkt angeschlossen werden. Die Reihenfolge der Hypothesen ist hierbei verschieden zu der Auflistung im Theorieteil, was an der Systematik der statistischen Verfahren liegt.

Hypothese 1a: Bild seriell schnell > Animation; AV = Nachwissensskala

Hypothese 2a: Bild seriell langsam > Animation; AV = Nachwissensskala

Hypothese 3a: Bild seriell schnell ist anders als Bild seriell langsam;
AV = Nachwissensskala

Zur Überprüfung dieser drei Einzelhypothesen wurde zunächst eine globale Varianzanalyse gerechnet. Danach wurden anhand von A-priori-Kontrasten, die durch die theoretische Fundierung der Hypothesen gerechtfertigt waren, die genauen Unterschiede zwischen den einzelnen Bedingungen bestimmt. Als Verfahren wurden hierbei t- bzw. w-Tests mit einem Signifikanzniveau von jeweils $\alpha = .05$ herangezogen.

Bei den nachfolgenden Mittelwertsvergleichen, die hinsichtlich der drei anderen abhängigen Variablen angestellt wurden, wurde ein analoges Verfahren gewählt, wobei nach der Anwendung globaler Kovarianzanalysen auch zur Bestimmung der Kontraste Allgemeine lineare Modelle (ALMs), unter Berücksichtigung gleicher Kovariaten, mit jeweils zwei Gruppen zum Einsatz kamen.

Zunächst erfolgt die Ergebnisdarstellung der Testungen der oben genannten Hypothesen zum Kriterium „Nachwissensskala“:

Tabelle 4.3.1.5

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Nachwissensskala“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild seriell schnell	20	15.10	7.75
Bild seriell langsam	20	15.10	7.34
Animation	21	16.76	6.16

Eine einfaktorielle Varianzanalyse mit den drei Bedingungen ergab ein nicht signifikantes Ergebnis ($F(2, 58) = .38, p = .69, \eta^2 = .01$). Die besten Leistungen in der Nachwissensskala erzielten die Personen der Animationsgruppe, was einen Effekt in der nicht vorhergesagten Richtung darstellte. Bei zweiseitigen Testungen wären die Kontraste nicht signifikant geworden: „Animation vs. Bild seriell schnell“ ($t(39) = .76, p(2\text{-seitig}) = .45, d = .24$); „Animation vs. Bild seriell langsam“ ($t(39) = .79, p(2\text{-seitig}) = .44, d = .24$). Auch bei dem Kontrast „Bild seriell schnell vs. Bild seriell langsam“ ($t(38) = .00, p(2\text{-seitig}) = 1.00, d = .00$) ließ sich kein signifikanter Unterschied finden, da beide Bedingungen gleiche Werte zeigten.

Hypothese 1b: Bild seriell schnell > Animation; AV = Inferenzskala

Hypothese 2b: Bild seriell langsam > Animation; AV = Inferenzskala

Hypothese 3b: Bild seriell schnell ist anders als Bild seriell langsam; AV = Inferenzskala

Zunächst werden die adjustierten Mittelwerte der Gruppen dargestellt, wie sie sich nach einer ANCOVA mit der Kovariate „Erfolgswahrscheinlichkeit“ ergeben haben.

Tabelle 4.3.1.6

Adjustierte Mittelwerte und Standardfehler der Bedingungen in der AV „Inferenzskala“

	<i>n</i>	<i>M</i>	<i>SE</i>
Bild seriell schnell	20	18.52	1.91
Bild seriell langsam	20	20.33	1.94
Animation	21	20.81	1.91

In der seriell schnellen Bildbedingung wurden die wenigsten, in der Animationsbedingung die meisten Punkte erreicht, was wiederum ein nicht vorhergesagter Effekt war. Eine ANCOVA mit den drei Gruppen zeigte keine signifikanten Unterschiede ($F(2, 57) = .40$, $p = .67$, partielles $\eta^2 = .01$). Bei zweiseitigen Testungen wären die Kontraste zwischen der Animations- und den beiden seriellen Bildbedingungen nicht signifikant geworden („Animation vs. Bild seriell schnell“: $F(1, 38) = .76$, p (2-seitig) = .39, partielles $\eta^2 = .02$; „Animation vs. Bild seriell langsam“: $F(1, 38) = .001$, p (2-seitig) = .98, partielles $\eta^2 < .001$). Auch der Unterschied zwischen der seriell schnellen und der seriell langsamen Bildbedingung zeigte keine statistische Signifikanz. („Bild seriell schnell vs. Bild seriell langsam“: $F(1, 37) = .35$, p (2-seitig) = .56, partielles $\eta^2 = .01$).

Hypothese 1c: Bild seriell schnell < Animation; AV = CL Lern

Hypothese 2c: Bild seriell langsam < Animation; AV = CL Lern

Hypothese 3c: Bild seriell schnell ist anders als Bild seriell langsam; AV = CL Lern

Zur Überprüfung dieser Hypothesen wurde eine ANCOVA mit der Kontrollvariable „Misserfolgsbefürchtung“ eingesetzt, daher werden auch hier die adjustierten Mittelwerte berichtet.

Tabelle 4.3.1.7

Adjustierte Mittelwerte und Standardfehler der Bedingungen in der AV „CL Lern“

	<i>n</i>	<i>M</i>	<i>SE</i>
Animation	21	10.17	.60
Bild seriell schnell	20	10.97	.59
Bild seriell langsam	20	12.10	.63

Die Kovarianzanalyse mit den drei Bedingungen wies auf größere Unterschiede zwischen den Gruppen hin ($F(2, 57) = 2.27, p = .11, \text{partielles } \eta^2 = .07$), wobei gegenteilige Effekte wie vorhergesagt auftraten. Eine signifikante Differenz würde sich zeigen, wenn man die Animations- mit der seriell langsamen Bildbedingung in einer zweiseitigen Testung miteinander kontrastiert ($F(1, 38) = 4.94, p \text{ (2-seitig)} = .03, \text{partielles } \eta^2 = .12$). Deskriptiv gesehen, erlebte die Gruppe „Animation“ einen geringeren Cognitive Load beim Betrachten des Lernmaterials als die Gruppe „Bild seriell langsam“. In den beiden anderen Kontrasten konnten wiederum keine signifikanten Mittelwertsunterschiede gefunden werden („Animation vs. Bild seriell schnell“: $F(1, 38) = 1.01, p \text{ (2-seitig)} = .32, \text{partielles } \eta^2 = .03$; „Bild seriell schnell vs. Bild seriell langsam“: $F(1, 37) = 1.21, p \text{ (2-seitig)} = .28, \text{partielles } \eta^2 = .03$). Der Unterschied zwischen der Animations- und der seriell schnellen Bildbedingung lag ebenfalls nicht in der erwarteten Richtung.

Hypothese 1d: Bild seriell schnell < Animation; AV = CL Nachttest

Hypothese 2d: Bild seriell langsam < Animation; AV = CL Nachttest

Hypothese 3d: Bild seriell schnell ist anders als Bild seriell langsam; AV = CL Nachttest

Tabelle 4.3.1.8

Mittelwerte und Standardabweichungen der Bedingungen in der AV „CL Nachtest“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild seriell schnell	20	13.63	2.51
Bild seriell langsam	20	13.68	2.03
Animation	21	13.95	1.96

Hinsichtlich der AV „Cognitive Load beim Nachtest“ zeigte eine einfaktorielle Varianzanalyse keine signifikanten Unterschiede ($F(2, 58) = .14, p = .87, \eta^2 = .01$). Auch durch die Berechnung der drei Kontraste „Animation vs. Bild seriell schnell“ ($t(39) = .47, p(1\text{-seitig}) = .32, d = .15$); „Animation vs. Bild seriell langsam“ ($t(39) = .45, p(1\text{-seitig}) = .33, d = .14$) und „Bild seriell schnell vs. Bild seriell langsam“ ($t(38) = .07, p(2\text{-seitig}) = .95, d = .02$) ergaben sich keine größeren Differenzen. Die Unterschiede lagen alle in der vorhergesagten Richtung.

Interaktionsanalysen

Die Analyse der Interaktionen zwischen dem kategorialen Prädiktor „Bedingung“ und den metrisch skalierten Kontrollvariablen „Vorwissen“, „Räumliche Fähigkeiten“ und „Merkfähigkeit“ wurde mittels multipler Regressionen vorgenommen. Die drei Bedingungsvariablen „Animation vs. Bild seriell schnell“, „Animation vs. Bild seriell langsam“ und „Bild seriell schnell vs. Bild seriell langsam“ wurden jeweils gewichtet effektkodiert, um die unterschiedlichen Stichprobenumfänge bei den Gruppen zu berücksichtigen. Die metrischen Variablen wurden zentriert. In die einzelnen Modelle wurden zusätzlich die Kovariaten aus den vorhergehenden Analysen mit aufgenommen. Post-hoc wurden die Bedingungen jeweils dummykodiert, um den Einfluss der metrischen Variablen in den einzelnen Gruppen zu erfahren. Dadurch konnten die Richtungen der Interaktionen ermittelt werden. Um die Unterschiede zwischen den Gruppen auf den einzelnen Ausprägungen der metrischen Variablen festzustellen, wurden die Bedingungsvariablen ebenfalls dummykodiert (bei den Kontrasten „Animation vs. Bild seriell schnell“ und „Animation vs. Bild seriell langsam“ war jeweils die

Animationsbedingung Referenzgruppe, bei dem Kontrast „Bild seriell schnell vs. Bild seriell langsam“ war die seriell schnelle Bildbedingung Referenzgruppe). Die metrischen Variablen wurden einmal in Abweichungswerten von +1SD und einmal von -1SD der jeweiligen Verteilung kodiert und jeweils in ein Regressionsmodell aufgenommen. Das Vorgehen orientierte sich in weiten Teilen an der Struktur des Artikels von Richter (2007).

In der Ergebnisdarstellung werden größere oder signifikante Interaktionen anhand von Tabellen und Liniendiagrammen veranschaulicht.

Interaktionen zwischen den Faktoren „Bedingung“ und „Vorwissen“

Interaktionen zwischen den Gruppen „Animation“ und „Bild seriell schnell“ mit der Variable „Vorwissen“

Hypothese 4a: Interaktion „Animation vs. Bild seriell schnell“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = Nachwissensskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 57) = 6.50$, $p = .001$, $R^2 = .26$, $adj R^2 = .22$).

Zwischen der effektkodierten Bedingungsvariable „Animation vs. Bild seriell schnell“ und dem Vorwissen lag hinsichtlich des Kriteriums „Nachwissensskala“ keine signifikante Interaktion vor ($B = -.32$, $SE = .31$, $t = -1.02$, $p = .31$). Deskriptiv hatte das Vorwissen einen stärker positiven Einfluss in der Bild- als in der Animationsgruppe (Bild seriell schnell: $B = 1.27$, $p = .002$; Animation: $B = .60$, $p = .27$).

Hypothese 4b: Interaktion „Animation vs. Bild seriell schnell“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = Inferenzskala

Tabelle 4.3.1.9

Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild seriell schnell x Vorwissen“ bei der AV „Inferenzskala“

Variable	<i>B</i>	<i>SE B</i>	<i>t</i>	<i>p</i>
Bedingung: Animation vs. Bild seriell schnell	.83	1.08	.77	.45
Vorwissen	1.29	.29	4.47	< .001
Erfolgswahrscheinlichkeit	.36	.23	1.54	.13
Interaktion Bedingung x Vorwissen	-.72	.35	-2.09	.04

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 56) = 9.37$, $p < .001$, $R^2 = .40$, $adj R^2 = .36$).

Die Interaktion zwischen der Bedingungsvariable und dem Vorwissen wurde bei der AV „Inferenzskala“ signifikant ($B = -.72$, $p = .04$). Multiple Regressionen, die post-hoc durchgeführt wurden, zeigten, dass der Einfluss des Vorwissens in der Gruppe „Bild seriell schnell“ stärker positiv war ($B = 2.09$, $p < .001$) als in der Gruppe „Animation“ ($B = .64$, $p = .29$). Personen mit höherem Vorwissen profitierten etwas stärker von den seriell schnellen Bildern ($B = 3.01$, $p = .38$), während Personen mit niedrigerem Vorwissen signifikant bessere Werte mit den Animationen erzielten ($B = -6.40$, $p = .05$).

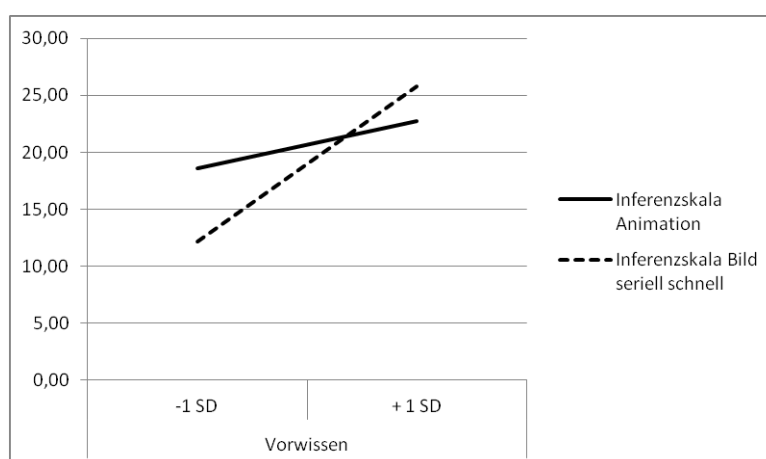


Abbildung 4.3.1.1: Regressionsgeraden des Vorwissens in den Gruppen „Animation“ und „Bild seriell schnell“ (AV: Inferenzskala)

Hypothese 4c: Interaktion „Animation vs. Bild seriell schnell“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = CL Lern

Tabelle 4.3.1.10

Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild seriell schnell x Vorwissen“ bei der AV „CL Lern“

Variable	<i>B</i>	<i>SE B</i>	<i>t</i>	<i>p</i>
Bedingung: Animation vs. Bild seriell schnell	-.28	.38	-.74	.46
Vorwissen	-.16	.10	-1.54	.13
Misserfolgsbefürchtung	.15	.05	3.13	.003
Interaktion Bedingung x Vorwissen	.34	.12	2.72	.009

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 56) = 5.32$, $p = .001$, $R^2 = .28$, $adj R^2 = .22$).

Zwischen der Bedingungsvariable und dem Vorwissen war eine sehr signifikante Interaktion festzustellen ($B = .34$, $p = .009$). In der Animationsgruppe war der Regressionskoeffizient des Vorwissens weniger stark negativ ($B = -.04$, $p = .84$) als in der seriell schnellen Bildgruppe ($B = -.63$, $p < .001$). Bei weniger Vorwissen war der Cognitive Load in der Animationsbedingung signifikant niedriger als in der seriell schnellen Bildbedingung ($B = 2.61$, $p = .02$), während bei höherem Vorwissen in geringerem Ausmaß das Gegenteil der Fall war ($B = -1.21$, $p = .29$).

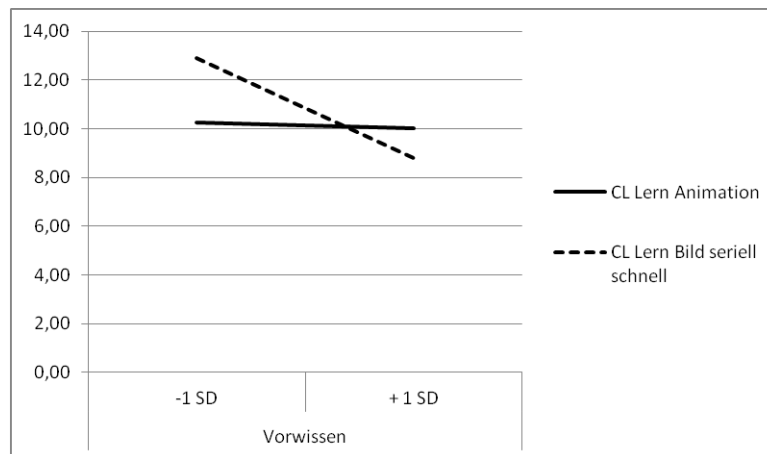


Abbildung 4.3.1.2: Regressionsgeraden des Vorwissens in den Gruppen „Animation“ und „Bild seriell schnell“ (AV: CL Lern)

Hypothese 4d: Interaktion „Animation vs. Bild seriell schnell“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = CL Nachttest

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 57) = 1.18, p = .33, R^2 = .06, adj R^2 = .009$).

Die Interaktion zwischen den Gruppen „Animation“ und „Bild seriell schnell“ mit der Variable „Vorwissen“ wurde in diesem Fall nicht signifikant ($B = .07, SE = .11, t = .64, p = .52$). Deskriptiv erlebten die Probanden mit höherem Vorwissen weniger Cognitive Load bei der seriell schnellen Bildbedingung ($B = -.73, p = .49$), während Probanden mit weniger Vorwissen sich bei beiden Bedingungen ungefähr gleich stark belastet fühlten ($B = .04, p = .96$).

Interaktionen zwischen den Gruppen „Animation“ und „Bild seriell langsam“ mit der Variable „Vorwissen“

Hypothese 5a: Interaktion „Animation vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = Nachwissensskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 57) = 6.57, p = .001, R^2 = .26, adj R^2 = .22$).

Es lag kein signifikanter Interaktionseffekt zwischen den Faktoren „Bedingung“ und „Vorwissen“ vor ($B = -.21, SE = .32, t = -.67, p = .51$). Deskriptiv war der Regressionskoeffizient des Vorwissens stärker positiv in der seriell langsamen Bild- als in

der Animationsbedingung (Bild seriell langsam: $B = 1.12$, $p = .01$; Animation: $B = .60$, $p = .27$).

Hypothese 5b: Interaktion „Animation vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = Inferenzskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 56) = 7.51$, $p < .001$, $R^2 = .35$, $adj R^2 = .30$).

Die Interaktion zwischen den beiden Bedingungen und dem Vorwissen der Probanden wurde nicht signifikant ($B = -.16$, $SE = .37$, $t = -.44$, $p = .66$). Auch hier zeigte sich kein größerer Unterschied im Einfluss des Vorwissens auf die abhängige Variable „Inferenzskala“ in den beiden Bedingungen (Animation: $B = .64$, $p = .29$; Bild seriell langsam: $B = 1.20$, $p = .01$).

Hypothese 5c: Interaktion „Animation vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = CL Lern

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 56) = 5.00$, $p = .002$, $R^2 = .26$, $adj R^2 = .21$).

Es trat keine signifikante Interaktion zwischen den Gruppen „Animation“ und „Bild seriell langsam“ mit dem Vorwissen auf ($B = -.10$, $SE = .13$, $t = -.76$, $p = .45$). Der Einfluss des Vorwissens war in beiden Bedingungen nicht sehr stark (Animation: $B = -.04$, $p = .84$; Bild seriell langsam: $B = .03$, $p = .88$).

Hypothese 5d: Interaktion „Animation vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = CL Nachttest

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 57) = 1.01$, $p = .40$, $R^2 = .05$, $adj R^2 < .001$).

Es konnte keine signifikante Interaktion zwischen der Bedingungsvariable „Animation vs. Bild seriell langsam“ und dem Vorwissen bezüglich des Kriteriums „CL Nachttest“ festgestellt werden ($B = -.03$, $SE = .11$, $t = -.28$, $p = .78$). Deskriptiv war sowohl bei geringerem als auch bei höherem Vorwissen der Cognitive Load in der seriell langsamen Bildbedingung leicht niedriger, wobei dieser Vorteil etwas größer bei dem geringeren

Vorwissen war (Höheres Vorwissen: $B = -.12$, $p = .91$; Niedrigeres Vorwissen: $B = -.31$, $p = .76$).

Interaktionen zwischen den Gruppen „Bild seriell schnell“ und „Bild seriell langsam“ mit der Variable „Vorwissen“

Hypothese 6a: Interaktion „Bild seriell schnell vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = Nachwissensskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 57) = 5.82$, $p = .002$, $R^2 = .23$, $adj R^2 = .19$).

Zwischen den beiden seriellen Bildbedingungen gab es keine signifikante Interaktion mit dem Vorwissen bei dem Kriterium „Nachwissensskala“ ($B = .11$, $SE = .29$, $t = .38$, $p = .71$). Der Einfluss des Vorwissens war leicht stärker in der seriell schnellen ($B = 1.27$, $p = .002$) als in der seriell langsamen Bildbedingung ($B = 1.12$, $p = .01$).

Hypothese 6b: Interaktion „Bild seriell schnell vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = Inferenzskala

Tabelle 4.3.1.11

Regressionsanalyse für die Bestimmung der Interaktion „Bild seriell schnell vs. Bild seriell langsam x Vorwissen“ bei der AV „Inferenzskala“

Variable	B	$SE B$	t	p
Bedingung: Bild seriell schnell vs. Bild seriell langsam	-.62	1.13	-.55	.58
Vorwissen	1.41	.29	4.94	< .001
Erfolgswahrscheinlichkeit	.30	.23	1.27	.21
Interaktion Bedingung x Vorwissen	.48	.32	1.50	.14

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 56) = 8.31$, $p < .001$, $R^2 = .37$, $adj R^2 = .33$).

Die Interaktion zwischen den beiden Gruppen „Bild seriell schnell“ und „Bild seriell langsam“ mit dem Vorwissen hatte hinsichtlich der abhängigen Variablen „Inferenzskala“ einen größeren, aber nicht signifikanten Wert ($B = .48, p = .14$). Der Regressionskoeffizient des Vorwissens lag in der Gruppe „Bild seriell schnell“ höher ($B = 2.09, p < .001$) als in der Gruppe „Bild seriell langsam“ ($B = 1.20, p = .01$), so dass Personen mit mehr Expertise deskriptiv bessere Werte auf der Inferenzkala mit seriell schnellen Bildern erzielten ($B = -1.86, p = .54$). Anfänger hingegen profitierten eher von seriell langsamen Bildern, wobei dieser Vorteil ebenfalls nicht signifikant wurde ($B = 3.94, p = .22$).

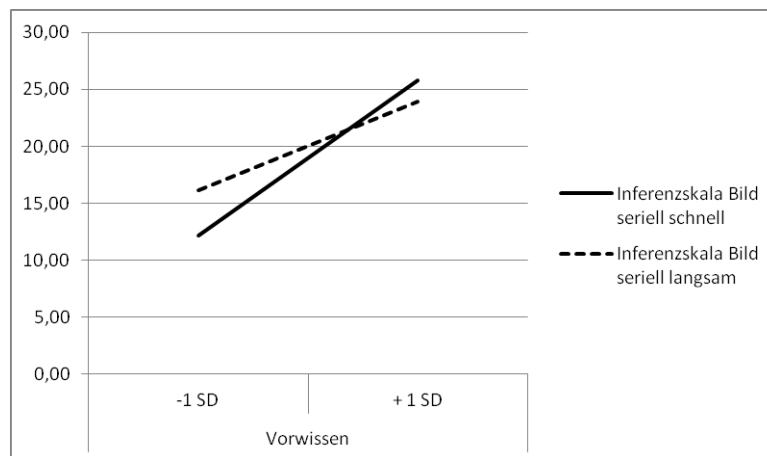


Abbildung 4.3.1.3: Regressionsgeraden des Vorwissens in den Gruppen „Bild seriell schnell“ und „Bild seriell langsam“ (AV: Inferenzskala)

Hypothese 6c: Interaktion „Bild seriell schnell vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = CL Lern

Tabelle 4.3.1.12

Regressionsanalyse für die Bestimmung der Interaktion „Bild seriell schnell vs. Bild seriell langsam x Vorwissen“ bei der AV „CL Lern“

Variable	<i>B</i>	<i>SE B</i>	<i>t</i>	<i>p</i>
Bedingung: Bild seriell schnell vs. Bild seriell langsam	-.51	.40	-1.29	.20
Vorwissen	-.22	.10	-2.28	.03
Misserfolgsbefürchtung	.15	.05	3.07	.003
Interaktion Bedingung x Vorwissen	-.35	.11	-3.23	.002

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 56) = 6.65$, $p < .001$, $R^2 = .32$, $adj R^2 = .27$).

Bei der abhängigen Variablen „Cognitive Load beim Lernen“ manifestierte sich eine sehr signifikante Interaktion zwischen der Bedingungsvariable und dem Vorwissen ($B = -.35$, $p = .002$). Der Regressionskoeffizient des Vorwissens war in der seriell schnellen Bildbedingung stark negativ, während er in der seriell langsamen Bildbedingung um Null lag (Bild seriell schnell: $B = -.63$, $p < .001$; Bild seriell langsam: $B = .03$, $p = .88$). Bei höherem Vorwissen empfanden die Probanden die seriell schnellen Key-Frames als signifikant weniger anstrengend als die seriell langsamen Key-Frames ($B = 3.40$, $p = .001$), während für Personen mit geringerem Vorwissen das Gegenteil galt, wobei dieser Unterschied allerdings nicht signifikant wurde ($B = -.85$, $p = .44$).

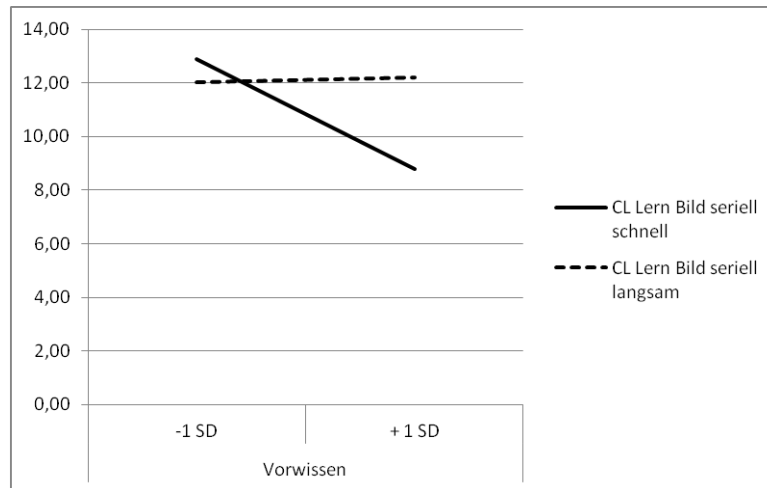


Abbildung 4.3.1.4: Regressionsgeraden des Vorwissens in den Gruppen „Bild seriell schnell“ und „Bild seriell langsam“ (AV: CL Lern)

Hypothese 6d: Interaktion „Bild seriell schnell vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = CL Nachttest

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 57) = 1.17, p = .33, R^2 = .06, adj R^2 = .008$).

Das Regressionsgewicht der Interaktion zwischen der Bedingungsvariable und dem Vorwissen erreichte bei der abhängigen Variablen „CL Nachttest“ keine statistische Signifikanz ($B = -.07, SE = .10, t = -.75, p = .46$). Deskriptiv beeinflusste das Vorwissen der Probanden den Cognitive Load stärker negativ in der seriell schnellen als in der seriell langsamen Bildbedingung (Bild seriell schnell: $B = -.22, p = .10$; Bild seriell langsam: $B = -.08, p = .60$)

Interaktionen zwischen den Faktoren „Bedingung“ und „Räumliche Fähigkeiten“

Interaktionen zwischen den Gruppen „Animation“ und „Bild seriell schnell“ mit der Variable „Räumliche Fähigkeiten“

Hypothese 7a: Interaktion „Animation vs. Bild seriell schnell“ x Räumliche Fähigkeiten ist signifikant, ungerichtete Fragestellung, AV = Nachwissensskala

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 57) = 1.15, p = .34, R^2 = .06, adj R^2 = .007$).

Bei den Gruppen „Animation“ und „Bild seriell schnell“ gab es keine signifikante Interaktion mit der Kontrollvariable „Räumliche Fähigkeiten“ bei dem Kriterium „Nachwissensskala“ ($B = -.19$, $SE = .18$, $t = -1.05$, $p = .30$). Deskriptiv profitierten Personen mit weniger räumlichen Fähigkeiten mehr vom Lernen mit Animationen ($B = -4.07$, $p = .22$), während für Probanden mit höheren räumlichen Fähigkeiten statische Bilder geeigneter waren, wobei auch diese Differenz nicht signifikant wurde ($B = 1.25$, $p = .72$).

Hypothese 7b: Interaktion „Animation vs. Bild seriell schnell“ x Räumliche Fähigkeiten ist signifikant, ungerichtete Fragestellung, AV = Inferenzskala

Tabelle 4.3.1.13

Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild seriell schnell x Räumliche Fähigkeiten“ bei der AV „Inferenzskala“

Variable	<i>B</i>	<i>SE B</i>	<i>t</i>	<i>p</i>
Bedingung: Animation vs. Bild seriell schnell	.89	1.28	.69	.49
Räumliche Fähigkeiten	.29	.16	1.79	.08
Erfolgswahrscheinlichkeit	.46	.29	1.61	.11
Interaktion Bedingung x Räuml. Fähigkeiten	-.30	.21	-1.44	.16

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 56) = 2.51$, $p = .05$, $R^2 = .15$, $adj R^2 = .09$).

Die Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell schnell“ mit den räumlichen Fähigkeiten nahm hinsichtlich der abhängigen Variablen „Inferenzskala“ einen größeren, aber nicht signifikanten Wert an ($B = -.30$, $p = .16$). Der Regressionskoeffizient der räumlichen Fähigkeiten lag bei der Gruppe „Bild seriell schnell“ höher ($B = .71$, $p = .03$) als bei der Gruppe „Animation“ ($B = .09$, $p = .77$). Bei höheren räumlichen Fähigkeiten wurden von den Probanden mehr Punkte mit seriell schnellen Bildern erreicht ($B = 2.48$, $p = .54$), während die Leistung bei niedrigeren räumlichen Fähigkeiten mit Animationen höher war ($B = -5.89$, $p = .15$). Die beiden Unterschiede waren jedoch nicht signifikant.

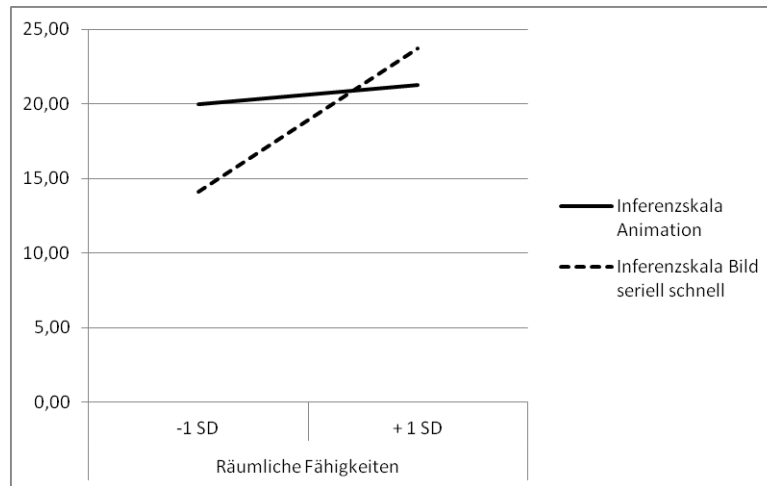


Abbildung 4.3.1.5: Regressionsgeraden der räumlichen Fähigkeiten in den Gruppen „Animation“ und „Bild seriell schnell“ (AV: Inferenzskala)

Hypothese 7c: Interaktion „Animation vs. Bild seriell schnell“ x Räumliche Fähigkeiten ist signifikant, ungerichtete Fragestellung, AV = CL Lern

Der multiple Korrelationskoeffizient des Gesamtmodells war annähernd signifikant ($F(4, 56) = 2.46, p = .06, R^2 = .15, adj R^2 = .09$).

Die Interaktion zwischen den Gruppen „Animation“ und „Bild seriell schnell“ mit den räumlichen Fähigkeiten war bei dem Kriterium „CL Lern“ nicht signifikant und lag nahe Null ($B = -.003, SE = .07, t = -.05, p = .96$). Der Einfluss der räumlichen Fähigkeiten war in beiden Bedingungen fast identisch und nur schwach ausgeprägt (Animation: $B = -.03, p = .72$; Bild seriell schnell: $B = -.04, p = .72$).

Hypothese 7d: Interaktion „Animation vs. Bild seriell schnell“ x Räumliche Fähigkeiten ist signifikant, ungerichtete Fragestellung, AV = CL Nachttest

Der multiple Korrelationskoeffizient des Gesamtmodells war annähernd signifikant ($F(3, 57) = 2.24, p = .09, R^2 = .11, adj R^2 = .06$).

Bei der abhängigen Variablen „CL Nachttest“ zeigte sich bei der Interaktion zwischen den Gruppen „Animation“ und „Bild seriell schnell“ und den räumlichen Fähigkeiten kein signifikanter Wert ($B = -.008, SE = .05, t = -.14, p = .89$). Der Regressionskoeffizient der räumlichen Fähigkeiten war in der Animationsbedingung leicht stärker negativ als in der

seriell schnellen Bildbedingung (Animation: $B = -.13$, $p = .09$; Bild seriell schnell: $B = -.12$, $p = .15$).

Interaktionen zwischen den Gruppen „Animation“ und „Bild seriell langsam“ mit der Variable „Räumliche Fähigkeiten“

Hypothese 8a: Interaktion „Animation vs. Bild seriell langsam“ x Räumliche Fähigkeiten ist signifikant, ungerichtete Fragestellung, AV = Nachwissensskala

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 57) = .82$, $p = .49$, $R^2 = .04$, $adj R^2 = -.009$).

Die Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell langsam“ mit den räumlichen Fähigkeiten stellte sich bei dem Kriterium „Nachwissensskala“ als nicht signifikant dar ($B = .03$, $SE = .16$, $t = .21$, $p = .83$). Der Einfluss der räumlichen Fähigkeiten war in beiden Bedingungen nicht stark und in der Animationsgruppe leicht höher als in der seriell langsamen Bildgruppe (Animation: $B = .09$, $p = .74$; Bild seriell langsam: $B = .07$, $p = .71$).

Hypothese 8b: Interaktion „Animation vs. Bild seriell langsam“ x Räumliche Fähigkeiten ist signifikant, ungerichtete Fragestellung, AV = Inferenzskala

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(4, 56) = 1.79$, $p = .14$, $R^2 = .11$, $adj R^2 = .05$).

Bei der abhängigen Variablen „Inferenzskala“ wies die Interaktion zwischen den Versuchsbedingungen und den räumlichen Fähigkeiten keinen signifikanten Wert auf und lag nahe Null ($B = -.004$, $SE = .19$, $t = -.02$, $p = .98$). Dementsprechend waren Unterschiede zwischen den Gruppen auf den einzelnen Ausprägungen der räumlichen Fähigkeiten nur sehr gering vorhanden (Niedrigere räumliche Fähigkeiten: $B = -.92$, $p = .82$; Höhere räumliche Fähigkeiten: $B = .45$, $p = .90$).

Hypothese 8c: Interaktion „Animation vs. Bild seriell langsam“ x Räumliche Fähigkeiten ist signifikant, ungerichtete Fragestellung, AV = CL Lern

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 56) = 3.86$, $p = .008$, $R^2 = .22$, $adj R^2 = .16$).

Auch hier zeigte die Interaktion zwischen den Gruppen „Animation“ und „Bild seriell langsam“ mit den räumlichen Fähigkeiten einen nicht signifikanten Wert ($B = .05$, $SE = .06$, $t = .92$, $p = .36$). Der Einfluss der räumlichen Fähigkeiten auf den Cognitive Load war in der Animationsbedingung weniger stark negativ ausgeprägt als in der seriell langsamen Bildbedingung (Animation: $B = -.03$, $p = .72$; Bild seriell langsam: $B = -.14$, $p = .07$).

Hypothese 8d: Interaktion „Animation vs. Bild seriell langsam“ x Räumliche Fähigkeiten ist signifikant, ungerichtete Fragestellung, AV = CL Nachtest

Der multiple Korrelationskoeffizient des Gesamtmodells war annähernd signifikant ($F(3, 57) = 2.31$, $p = .09$, $R^2 = .11$, $adj R^2 = .06$).

Der Wert der Interaktion zwischen der Bedingungsvariable „Animation vs. Bild seriell langsam“ und den räumlichen Fähigkeiten wies bei dem Kriterium „CL Nachtest“ keine statistische Signifikanz auf ($B = -.03$, $SE = .05$, $t = -.66$, $p = .51$). Deskriptiv war in der Animationsbedingung der Einfluss der räumlichen Fähigkeiten stärker negativ als in der seriell langsamen Bildbedingung (Animation: $B = -.13$, $p = .09$; Bild seriell langsam: $B = -.07$, $p = .23$)

Interaktionen zwischen den Faktoren „Bedingung“ und „Merkfähigkeit“

Interaktionen zwischen den Gruppen „Animation“ und „Bild seriell schnell“ mit der Variable „Merkfähigkeit“

Hypothese 9a: Interaktion „Animation vs. Bild seriell schnell“ x Merkfähigkeit ist signifikant, ungerichtete Fragestellung, AV = Nachwissensskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 55) = 5.77$, $p = .002$, $R^2 = .24$, $adj R^2 = .20$).

Zwischen den Faktoren „Bedingung“ und „Merkfähigkeit“ lag keine signifikante Interaktion beim Kriterium „Nachwissensskala“ vor ($B = .14$, $SE = .28$, $t = .49$, $p = .63$). Deskriptiv war bei höherer Merkfähigkeit die Animations- der seriell schnellen Bildbedingung gegenüber

überlegen ($B = -3.19$, $p = .32$), während bei niedrigerer Merkfähigkeit kein Unterschied zwischen den Gruppen zu sehen war ($B = -.04$, $p = .99$).

Hypothese 9b: Interaktion „Animation vs. Bild seriell schnell“ x Merkfähigkeit ist signifikant, ungerichtete Fragestellung, AV = Inferenzskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 54) = 3.39$, $p = .02$, $R^2 = .20$, $adj R^2 = .14$).

Es war kein signifikanter Interaktionseffekt zwischen den Gruppen „Animation“ und „Bild seriell schnell“ mit der Merkfähigkeit auf die AV „Inferenzskala“ festzustellen ($B = .36$, $SE = .35$, $t = 1.03$, $p = .31$). Der Einfluss der Gedächtnisfähigkeiten einer Person auf ihre Inferierfähigkeiten war beim Lernen mit Animationen deskriptiv stärker als beim Lernen mit seriell schnellen Key-Frames (Animation: $B = 1.28$, $p = .07$; Bild seriell schnell: $B = .44$, $p = .26$).

Hypothese 9c: Interaktion „Animation vs. Bild seriell schnell“ x Merkfähigkeit ist signifikant, ungerichtete Fragestellung, AV = CL Lern

Tabelle 4.3.1.14

Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild seriell schnell x Merkfähigkeit“ bei der AV „CL Lern“

Variable	<i>B</i>	<i>SE B</i>	<i>t</i>	<i>p</i>
Bedingung: Animation vs. Bild seriell schnell	-.31	.42	-.72	.47
Merkfähigkeit	-.19	.09	-2.11	.04
Misserfolgsbefürchtung	.15	.05	2.85	.006
Interaktion Bedingung x Merkfähigkeit	-.23	.12	-1.88	.07

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 54) = 3.32$, $p = .02$, $R^2 = .20$, $adj R^2 = .14$).

Die Interaktion zwischen den Gruppen „Animation“ und „Bild seriell schnell“ mit der Variable „Merkfähigkeit“ nahm einen fast signifikanten Wert an ($B = -.23, p = .07$). Der Einfluss des Gedächtnisses auf den Cognitive Load während der Lerneinheit war in der Animationsbedingung stärker negativ ($B = -.43, p = .06$) als in der seriell schnellen Bildbedingung ($B = .08, p = .54$). Personen mit höherer Merkfähigkeit verspürten signifikant mehr Cognitive Load in der seriell schnellen Bildbedingung ($B = 2.79, p = .04$), während für Personen mit niedrigerer Merkfähigkeit in nicht signifikantem Ausmaß das Lernen mit Animationen anstrengender war ($B = - 1.41, p = .32$).

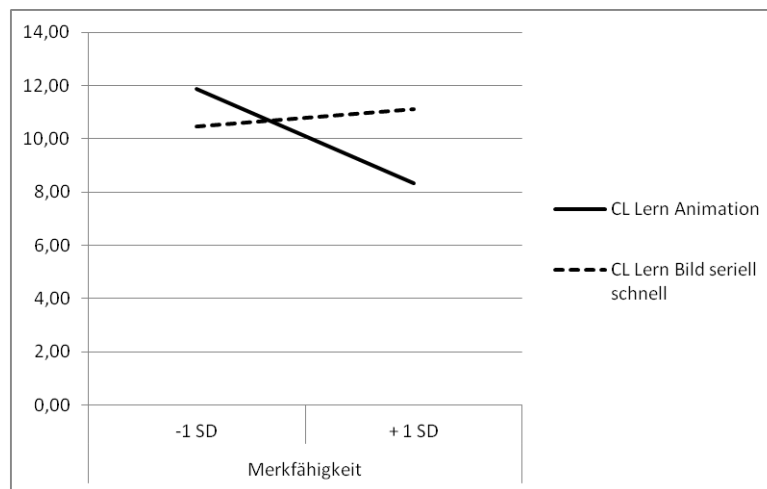


Abbildung 4.3.1.6: Regressionsgeraden der Merkfähigkeit in den Gruppen „Animation“ und „Bild seriell schnell“ (AV: CL Lern)

Hypothese 9d: Interaktion „Animation vs. Bild seriell schnell“ x Merkfähigkeit ist signifikant, ungerichtete Fragestellung, AV = CL Nachttest

Tabelle 4.3.1.15

Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild seriell schnell x Merkfähigkeit“ bei der AV „CL Nachtest“

Variable	<i>B</i>	<i>SE B</i>	<i>t</i>	<i>p</i>
Bedingung: Animation vs. Bild seriell schnell	.14	.33	.43	.67
Merkfähigkeit	-.17	.07	-2.36	.02
Interaktion Bedingung x Merkfähigkeit	-.16	.09	-1.74	.09

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 55) = 2.16, p = .10, R^2 = .11, adj R^2 = .06$).

Die Interaktion zwischen den Bedingungen „Animation“ und „Bild seriell schnell“ mit der Merkfähigkeit erreichte bei dem Kriterium „CL Nachtest“ einen größeren, aber nicht signifikanten Wert ($B = -.16, p = .09$). Wie bei „CL Lern“, beeinflusste die Merkfähigkeit einer Person den Cognitive Load beim Nachtest stärker negativ in der Animations- ($B = -.38, p = .04$) als in der seriell schnellen Bildbedingung ($B = -.02, p = .85$). Bei höherer Merkfähigkeit war der Cognitive Load größer in der seriell schnellen Bildgruppe ($B = 1.24, p = .25$), bei geringerer Merkfähigkeit war er in der Animationsgruppe stärker ausgeprägt ($B = -1.79, p = .12$). Die beiden Differenzen waren nicht signifikant.

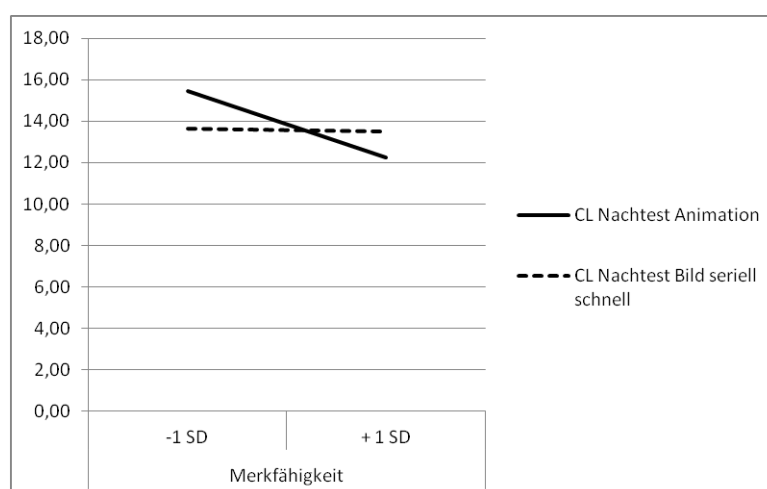


Abbildung 4.3.1.7: Regressionsgeraden der Merkfähigkeit in den Gruppen „Animation“ und „Bild seriell schnell“ (AV: CL Nachtest)

Interaktionen zwischen den Gruppen „Animation“ und „Bild seriell langsam“ mit der Variable „Merkfähigkeit“

Hypothese 10a: Interaktion „Animation vs. Bild seriell langsam“ x Merkfähigkeit ist signifikant, ungerichtete Fragestellung, AV = Nachwissensskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 55) = 5.56$, $p = .002$, $R^2 = .23$, $adj R^2 = .19$).

Zwischen den Bedingungen „Animation“ und „Bild seriell langsam“ und der Merkfähigkeit konnte bezüglich der abhängigen Variablen „Nachwissensskala“ keine signifikante Interaktion gefunden werden ($B = .14$, $SE = .28$, $t = .49$, $p = .63$). Deskriptiv war bei höherer Merkfähigkeit die Animationsgruppe im Vorteil ($B = -2.53$, $p = .46$), während bei geringerer Merkfähigkeit die seriell langsame Bildbedingung leicht bessere Ergebnisse erzielte ($B = .62$, $p = .85$).

Hypothese 10b: Interaktion „Animation vs. Bild seriell langsam“ x Merkfähigkeit ist signifikant, ungerichtete Fragestellung, AV = Inferenzskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 54) = 3.01$, $p = .03$, $R^2 = .18$, $adj R^2 = .12$).

Der Wert der Interaktion zwischen den Gruppen „Animation“ und „Bild seriell langsam“ mit der Merkfähigkeit war hinsichtlich des Kriteriums „Inferenzskala“ nicht signifikant ($B = .17$, $SE = .35$, $t = .47$, $p = .64$). Deskriptiv war der Regressionskoeffizient der Merkfähigkeit in der Animationsbedingung stärker positiv als in der seriell langsamen Bildbedingung (Animation: $B = 1.28$, $p = .07$; Bild seriell langsam: $B = .65$, $p = .10$).

Hypothese 10c: Interaktion „Animation vs. Bild seriell langsam“ x Merkfähigkeit ist signifikant, ungerichtete Fragestellung, AV = CL Lern

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 54) = 3.22$, $p = .02$, $R^2 = .19$, $adj R^2 = .13$).

Die Interaktion zwischen den Gruppen „Animation“ und „Bild seriell langsam“ mit der Merkfähigkeit wurde bei dem Kriterium „CL Lern“ nicht signifikant ($B = -.03$, $SE = .12$, $t = -.28$, $p = .78$). Deskriptiv war der Einfluss der Merkfähigkeit in der Animationsbedingung

stärker negativ als in der seriell langsamen Bildbedingung (Animation: $B = -.43$, $p = .06$; Bild seriell langsam: $B = -.16$, $p = .22$).

Hypothese 10d: Interaktion „Animation vs. Bild seriell langsam“ x Merkfähigkeit ist signifikant, ungerichtete Fragestellung, AV = CL Nachtest

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 55) = 1.25$, $p = .30$, $R^2 = .06$, $adj R^2 = .01$).

Die multiple Regression ergab keinen signifikanten Wert für die Interaktion zwischen den Gruppen „Animation“ und „Bild seriell langsam“ mit der Merkfähigkeit bezüglich des Kriteriums „CL Nachtest“ ($B = -.06$, $SE = .10$, $t = -.58$, $p = .57$). Deskriptiv war bei höherer Merkfähigkeit die Animationsbedingung im Vorteil ($B = .72$, $p = .53$), bei geringerer Merkfähigkeit war die mentale Belastung bei seriell langsamen Bildern niedriger ($B = -1.29$, $p = .23$).

4.3.1.4 Diskussion

Haupteffekte

Animation vs. Seriell schnelle und seriell langsame Key-Frames

In den beiden Hauptteilstudien sollten die theoretisch hergeleiteten Hypothesen mit im Vergleich zur Pilotstudie größeren Stichproben und somit vermutlich stabileren Schätzungen der Kennwerte überprüft werden. Ebenfalls konnte nun der Einfluss der Kontrollvariablen durch die Verwendung von Kovarianzanalysen besser berücksichtigt werden.

In dem ersten Hypothesenkomplex der ersten Hauptteilstudie zwischen Animation und seriell schnellen Key-Frames (Hypothesen 1a – 1d) war angenommen worden, dass mit seriell schnellen Bildern besser gelernt werden würde, was sich in besseren Leistungen auf einer Wissens- und einer Inferenzskala sowie in einem geringeren Cognitive Load, sowohl während des Lernens als auch nach während des Nachtests, äußern sollte. Keine dieser Hypothesen konnte in der ersten Hauptteilstudie bestätigt werden, da bei drei der Kriterien nicht vorhergesagte Ergebnisse errechnet wurden. Die Unterschiede zwischen den beiden Bedingungen wären auch bei zweiseitigen Testungen nicht signifikant geworden. Auch bei dem vierten Kriterium erreichte die Mittelwertsdifferenz keine statistische Signifikanz, lag aber in der vorhergesagten Richtung. Die Animations- und die seriell schnelle Bildbedingung erzielten in allen vier Kriterien ungefähr gleich gute Ergebnisse, wobei bei den abhängigen Variablen „Nachwissensskala“, „Inferenzskala“ und „Cognitive Load beim Lernen“ die Animationsgruppe leicht im Vorteil war. Die Effektstärken lagen auf einem geringen Niveau.

Analoge Hypothesen zu denen zwischen der Animation und den seriell schnellen Bildern wurden zwischen der Animation und den seriell langsamen Bildern aufgestellt (Hypothesen 2a – 2d). Auch hier wurde eine Überlegenheit der seriell langsamen Key-Frames in den Kriterien „Nachwissensskala“, „Inferenzskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachtest“ postuliert. Diese vier gerichteten Hypothesen fanden durch die Ergebnisse der ersten Hauptteilstudie keine Bestätigung. Bei den abhängigen Variablen „Nachwissensskala“ und „Inferenzskala“ lagen die Ergebnisse nicht in der vorhergesagten Richtung, wobei bei einer jeweils zweiseitigen Testung keine Signifikanz erreicht worden wäre. Bei dem Kriterium „Cognitive Load beim Nachtest“ wurde der Mittelwertsunterschied

nicht signifikant, zeigte aber das vorhergesagte Richtungsmuster. Die Effektstärken waren alle niedrig.

Ebenfalls nicht bestätigt werden konnte die Hypothese 2c hinsichtlich des Kriteriums „Cognitive Load beim Lernen“. Hier hätte zwar ein bei zweiseitiger Testung statistisch bedeutsamer Unterschied zwischen den Gruppen vorgelegen, allerdings wies dieser nicht die vorhergesagte Richtung auf. Die Effektstärke lag im mittleren bis hohen Bereich. Personen, die mit seriell langsamen Key-Frames lernten, verspürten mehr Cognitive Load als Personen, die eine Animation betrachteten.

Zusammenfassend kann gesagt werden, dass in keinem der Fälle die Animationsbedingung die vorhergesagte signifikante Unterlegenheit gegenüber den seriellen Bildbedingungen zeigte. Bei sieben der acht Vergleiche ergaben sich keine größeren Unterschiede zwischen den Gruppen. Deskriptiv gesehen zeichnete sich ein gemischtes Bild ab. In vier Fällen erreichten die Probanden der Animationsgruppe bessere Lernleistungen, in zwei Vergleichen waren die seriellen Bilder überlegen, und einmal waren die Mittelwerte der Bedingungen fast identisch. Das einzige signifikante Ergebnis der Analysen, allerdings nicht in der vorhergesagten Richtung, lag bei dem Kontrast zwischen der Animation und den seriell langsamen Bildern hinsichtlich der abhängigen Variablen „Cognitive Load beim Lernen“, wobei die Animationsbedingung hier im Vorteil war und diese Präsentationsform somit eine höhere „instruktionale Effizienz“ zeigte (Paas et al., 2007). Eine Erklärung dafür, warum die seriell langsamen Bilder als anstrengender und schwieriger empfunden wurden, könnte zum einen, motivational gesehen, darin liegen, dass gerade diese Bilder durch ihre hohe Präsentationsdauer als besonders eintönig und als wenig dynamisch empfunden wurden. Zudem war bei diesen Key-Frames die Inferenz der Zustände zwischen den Bildern besonders schwierig. Generell ist es so, dass man bei einer Animation „nur wahrnehmen“ muss, um einen Prozess nachzuvollziehen, da eine solche Präsentationsform vollständige dynamische Informationen liefert (Hegarty et al., 2003; Lewalter, 1997). Dies stellt sicherlich einen großen Vorteil einer Animation dar. Bei statischen Bildern hingegen muss man auch einen kognitiven Beitrag leisten, indem Zwischenzustände inferiert und indem der Prozess mental animiert werden muss (Hegarty, 1992). Bei seriell langsamen Bildern werden die einzelnen Key-Frames ziemlich lange präsentiert, wodurch der Aspekt der „intrastate“- und weniger der „interstate“-Information betont wird. Der Zusammenhang zwischen den Bildern verliert sich leichter, und es muss eine größere Gedächtnisleistung aufgebracht werden, um alle Key-States in Erinnerung zu behalten. Deshalb könnte der Aufbau des

gesamten mentalen Modells des synaptischen Vorganges mit seriell langsamen Bildern als schwierig und als anstrengend empfunden worden sein.

Der oben angesprochene Vorteil einer Animation, vollständige Informationen über einen Prozess zu liefern, wodurch der Lerner erst einmal „nur wahrnehmen“ muss, mag auch ein Grund für die Ergebnisse der anderen sieben Kontraste gewesen sein, bei denen die Animation nicht die vorhergesagte bedeutsame Unterlegenheit, sondern eher eine Ebenbürtigkeit mit den seriellen Bildern zeigte. Die nicht signifikanten Unterschiede zwischen den Bedingungen weisen darauf hin, dass zum einen bei den seriellen Bildern inhärente Nachteile, bei der Animation hingegen ihre positiven Aspekte ein größeres Gewicht beim Lernen zu haben schienen. Beim Lernen mit Bildern muss der Betrachter beispielsweise stets die Zustände zwischen den Key-Frames selbst im Arbeitsgedächtnis inferieren (Hegarty et al., 2003). Diese Konstruktions- und mentalen Animationsprozesse belasten das Arbeitsgedächtnis und stellen eine Schwierigkeit bei einem Lernvorgang mit statischen Bildern dar. Umgekehrt könnte ein weiterer Grund für die überraschend guten Ergebnisse der Animationsbedingung die motivational höhere Attraktivität einer solchen Präsentationsform gewesen sein. Wie schon bei der Pilotstudie erwähnt, zeigten sich ein paar der Probanden enttäuscht, wenn sie mit statischen Bildern anstatt mit der Animation lernen sollten. Eine Animation wirkt durch ihre Neuartigkeit und ihren modernen Charakter interessanter (Perez & White, 1985), und die permanente Dynamik auf dem Screen reizt eher die Aufmerksamkeit des Betrachters. Ein weiterer, öfter in der Literatur (Lowe, 1999a) genannter, Nachteil einer Animation ist der, dass ein Lerner häufig seine Aufmerksamkeit auf konzeptuell irrelevante, dafür aber perzeptuell betonte Details der Darstellung lenkt. Dies war beim Design der Animation der vorliegenden Studie wahrscheinlich nicht der Fall, da bei der Materialentwicklung darauf geachtet wurde, kein Teil der Synapse besonders hervorstechend zu zeichnen. Dadurch wurde eventuell der konzeptuelle Strukturierungsvorgang beim Lernen mit der Animation erleichtert. Das vielleicht größte Defizit einer Animation ist ihre Flüchtigkeit (Lowe, 1999a). Dieser Faktor war in der aktuellen Studie abgemildert, zum einen dadurch, dass der synaptische Vorgang sieben Mal wiederholt wurde, zum anderen durch die Präsentation des beschrifteten Standbildes am Anfang der Lerneinheit, bei dem sich die strukturellen Eigenheiten der Synapse eingepägt werden konnten. Die genannten Besonderheiten der vorliegenden Untersuchung könnten die inhärenten Defizite einer Animation abgefedert und somit die Lernleistungen der entsprechenden Probanden begünstigt haben.

Zuletzt brachten viele Probanden, wie auch schon in der Pilotstudie beschrieben, zum Ausdruck, dass sie sich während des Lernens an ihr Wissen über Synapsen aus dem Gymnasium erinnert hätten, obwohl sie im Vorwissenstest die Aufgaben zunächst weniger gut beantworten konnten. Dieses Wissen aus der Schulzeit könnte ebenfalls die Unterschiede zwischen den Gruppen nivelliert haben, da eventuell vorhandene Schwächen einer Darstellungsform beim Lernen dadurch ausgeglichen werden konnten.

Zusammenfassend kann gesagt werden, dass sich die Annahme, dass mit Hilfe von Key-States besser gelernt wird als mit Animationen, wenn ein dynamisches mentales Modell aufgebaut und wenn im Anschluss an den Lernvorgang Inferenzaufgaben gelöst werden sollen, in der aktuellen Studie erst einmal nicht bestätigt hat. Dies könnte unter anderem auch an der Thematik des Lernmaterials gelegen haben, das eventuell zu einfach für die Zielgruppe war und das daher Unterschiede zwischen den Bedingungen nicht adäquat widerspiegeln konnte. Ein absolutes Verwerfen der Annahmen der „Key-State-Theorie“ wäre an der Stelle wahrscheinlich verfrüht, da es zwar sein mag, dass die Darstellung eines Prozesses mit Hilfe von Key-Frames nicht der menschlichen kognitiven Architektur kongruent ist und den Aufbau dynamischer mentaler Modelle nicht begünstigt. Es könnte aber auch sein, dass dies durchaus der Fall sein mag, dass aber der Vorteil der Key-Frames, den sie eventuell durch eine bessere kognitive Passung haben, durch inhärente Nachteile des Lernens mit statischen Bildern bzw. durch die Vorzüge eines Lernvorganges mit Animationen ausgeglichen wird.

Seriell schnelle Key-Frames vs. Seriell langsame Key-Frames

Zusätzlich wurden in der ersten Hauptteilstudie die seriell schnellen mit den seriell langsamen Bildern bezüglich der Kriterien „Nachwissensskala“, „Inferenzskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachtest“ kontrastiert. Es wurden jeweils ungerichtete Hypothesen aufgestellt (Hypothesen 3a – 3d), da sowohl die Möglichkeit bestand, dass die seriell schnellen Key-Frames überlegen waren, es aber auch Gründe gab, die dafür sprachen, dass durch seriell langsame Bilder bessere Lernergebnisse erzielt werden konnten.

Nach den statistischen Analysen stellte sich heraus, dass keine der vier ungerichteten Hypothesen bestätigt werden konnte. Bei den Kriterien „Nachwissensskala“ und „Cognitive Load beim Nachtest“ waren die Lernergebnisse der beiden Bedingungen sogar fast bzw.

vollkommen identisch. Auch bei den abhängigen Variablen „Inferenzskala“ und „Cognitive Load beim Lernen“ waren die Unterschiede zwischen den Gruppen nicht signifikant. Dementsprechend wiesen die Effektstärken auch nur geringe Werte auf. Deskriptiv betrachtet, zeigten die Probanden der seriell langsamen Bildbedingung bessere Inferierfähigkeiten als die der seriell schnellen Bildbedingung. Allerdings wurde in der seriell langsamen Bildbedingung auch mehr Cognitive Load beim Lernen erlebt. Generell schienen sich somit die Vor- bzw. Nachteile von seriell schnellen und seriell langsamen Key-Frames auszugleichen, so dass die Variation der Präsentationsgeschwindigkeit der Bilder zumindest in dieser Untersuchung zu keinen größeren Unterschieden geführt hat. Lernerleichternd bei den seriell schnellen Bildern dürfte die größere Zahl an Wiederholungen des synaptischen Prozesses gewesen sein, ebenso wie die stärkere Betonung des Gesamtkontextes. Die Inferenz der Zustände zwischen den Key-Frames dürfte ebenfalls einfacher gewesen sein als bei den seriell langsamen Bildern. Die seriell langsamen Bilder haben den Vorzug, dass der Lerner ausreichend Zeit hat, sich die Details der Teilprozesse einzuprägen und sie zu verstehen. Der Fokus liegt auf der „intrastate“-Information, so dass das mentale Modell schrittweise aufgebaut werden kann (Mayer & Chandler, 2001). Insgesamt gesehen, scheinen sich allerdings die Vorzüge beider Präsentationsarten aufzuwiegen. Dass die Probanden der seriell schnellen Bildbedingung weniger Cognitive Load beim Lernen empfanden als die der seriell langsamen Bildbedingung, könnte, wie auch schon beim Kontrast zur Animationsbedingung beschrieben, daran gelegen haben, dass die langsamen Bilder eventuell als zu monoton empfunden wurden und dass die Inferenz der Zustände zwischen den Bildern den Probanden bei den seriell schnellen Key-Frames leichter gefallen war.

Interaktionen

Interaktionen zwischen den Faktoren „Bedingung“ und „Vorwissen“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit dem Vorwissen von Personen

Bei den Interaktionshypothesen wurden zunächst die Gruppen „Animation“ und „Bild seriell schnell“ miteinander verglichen, und es wurde davon ausgegangen, dass signifikante Interaktionen zwischen diesen Bedingungen und der Variable „Vorwissen“ im Hinblick auf die vier Kriterien „Nachwissensskala“, „Inferenzskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachtest“ bestehen (Hypothesen 4a – 4d). Welches Richtungsmuster diese Interaktionen jeweils annehmen, wurde in den Hypothesen offen gelassen, da das Vorwissen sowohl beim Lernen mit statischen Bildern, als auch beim Lernen mit Animationen einen moderierenden Einfluss nehmen kann (Hegarty & Kriz, 2008).

Die Hypothesen 4a und 4d konnten durch die Analysen nicht bestätigt werden, da bei den abhängigen Variablen „Nachwissensskala“ und „Cognitive Load beim Nachtest“ keine signifikanten Interaktionen vorlagen. Bei den Kriterien „Inferenzskala“ und „Cognitive Load beim Lernen“ bestätigten sich die ungerichteten Hypothesen jedoch, bei dem Cognitive Load beim Lernen konnte eine sehr signifikante Interaktion gefunden werden. Durch Post-hoc-Analysen stellte sich bei der Inferenzskala das Muster heraus, dass der Einfluss des Vorwissens der Probanden auf dieses Kriterium bei seriell schnellen Bildern stärker positiv war als bei der Animation. Personen mit höherem Vorwissen profitierten etwas stärker von den seriell schnellen Bildern, während Personen mit niedrigerem Vorwissen signifikant bessere Werte mit den Animationen erzielten. Bei dem Cognitive Load während des Lernens wurde dieses Ergebnis verfestigt, d.h., auch hier hatte das Vorwissen einen stärker negativen Effekt auf die von den Probanden empfundene mentale Belastung in der seriell schnellen Bild- als in der Animationsbedingung. Der Cognitive Load war bei geringerem Vorwissen in der Animationsgruppe signifikant niedriger als in der seriell schnellen Bildgruppe, während bei höherem Vorwissen in geringerem Ausmaß das Gegenteil der Fall war.

Das bedeutet, dass durch die Interaktionsanalysen sog. „Expertise Reversal Effects“ (Kalyuga, Ayres, Chandler & Sweller, 2003) bei den Kriterien „Inferenzskala“ und „Cognitive Load beim Lernen“ gefunden werden konnten, die beide in die gleiche Richtung wiesen. Personen mit höherem Vorwissen lernten deskriptiv besser mit seriell schnellen Key-Frames als mit einer Animation, während bei Probanden mit geringerem Vorwissen das

Gegenteil der Fall war, wobei die Unterschiede hier sogar signifikant wurden. Das Interessante bei dem Kontrast zwischen den seriell schnellen Bildern und der Animation war die gute Vergleichbarkeit zwischen den Bedingungen, da sowohl die Zahl der Wiederholungen als auch die Präsentationsgeschwindigkeit des synaptischen Prozesses innerhalb des Lernmaterials bei den beiden Gruppen gleich waren. Der maßgebliche Unterschied zwischen den Bedingungen war, dass durch die Animation vollständige Informationen geliefert wurden (Lewalter, 1997), während bei den Key-Frames zwischen den Bildern naturgemäß Informationslücken waren. Die Probanden der Bildbedingung mussten die Zustände zwischen den Bildern inferieren und den Prozess mental animieren, um ein vollständiges dynamisches mentales Modell zu erhalten (Hegarty et al., 2003; Hegarty, 1992). Diese Notwendigkeit beim Lernen mit statischen Bildern scheint für Personen mit höherem Vorwissen eine kognitiv angemessene Anforderung zu sein, da sie, zumindest deskriptiv, besser mit den Key-Frames lernten, während es für Personen mit niedrigerem Vorwissen schwierig zu sein scheint, die Inferenzen sowie die mentale Animation zu leisten (Narayanan & Hegarty, 1998), was zu ihren signifikant schlechteren Lernergebnissen mit statischen Bildern geführt haben könnte. Die Animation könnte für sie hingegen eine „Supplantationsfunktion“ (Salomon, 1979) gehabt haben. In der Literatur (Schnotz & Lowe, 2008) wird angemerkt, dass beim Lernen mit Animationen das Vorwissen einer Person deshalb oft eine Rolle spielt, damit sie nicht konzeptuell irrelevanten, aber perzeptuell salienten Details ihre Aufmerksamkeit schenkt. Eventuell war der Einfluss des Vorwissens in der Animationsbedingung in der vorliegenden Studie deshalb nicht so stark wie in der seriell schnellen Bildbedingung, da beim Design der Animation darauf geachtet wurde, die Synapse und die an ihr ablaufenden Vorgänge klar und übersichtlich zu zeichnen und nichts Unwichtiges zu betonen. Zudem war die Flüchtigkeit der dynamischen Darstellung durch die mehrfache Wiederholung des synaptischen Prozesses abgemildert, so dass auch Personen mit geringem Vorwissen mehrere Gelegenheiten hatten, das Material angemessen zu strukturieren, wodurch sie eventuell signifikant besser mit den Animationen als mit den statischen Bildern lernten.

Relativierend ist anzumerken, dass in der Stichprobe generell kein überdurchschnittlich hohes Vorwissen vorhanden war. Der Mittelwert lag bei $M = 4.54$ von 18.00 Punkten ($SD = 3.25$), so dass die Ergebnisse nicht ohne Weiteres auf Personen mit sehr hohem Vorwissen oder auf Experten ausgeweitet werden können.

Ebenfalls ungerichtete Hypothesen bestanden bei dem Kontrast zwischen der Animation und den seriell langsamen Bildern hinsichtlich ihrer Interaktionen mit der Variable „Vorwissen“ (Hypothesen 5a – 5d). Auch hier wurden die Richtungen der jeweiligen Interaktionsmuster offen gelassen, da, laut Literatur (Hegarty & Kriz, 2008), das Vorwissen einer Person sowohl das Lernen mit statischen Bildern als auch das Lernen mit Animationen beeinflussen kann. Bei allen vier Kriterien („Nachwissensskala“, „Inferenzskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachtest“) konnten die ungerichteten Hypothesen nicht bestätigt werden, da keine signifikanten Interaktionen vorlagen. Deskriptiv gesehen, zeigten die Interaktionsmuster ein gemischtes Bild, so dass auch ansatzweise keine eindeutige Aussage gemacht werden konnte. Dieses Ergebnis war insofern überraschend, da auch bei seriell langsamen Bildern Inferenzen auf die Zwischenzustände gezogen werden müssen und diese Inferenzen eventuell sogar schwieriger sind als bei seriell schnellen Bildern. Daher hätte nach den obigen Ergebnissen erwartet werden können, dass Personen mit höherem Vorwissen bei seriell langsamen Bildern im Vorteil sind im Vergleich zu Personen, die über ein geringeres Vorwissen verfügen. Allerdings könnten seriell langsame Bilder gerade für Personen mit niedrigerem Vorwissen den Vorzug haben, dass sie die einzelnen Bilder lang betrachten und daher die einzelnen Teilprozesse im Detail verstehen können. Dadurch kann das mentale Modell schrittweise aufgebaut werden, ohne eine Überforderung zu verursachen (Mayer & Chandler, 2001). Dieser lernerleichternde Aspekt könnte die Notwendigkeit zur Inferenz von Zwischenzuständen ausgeglichen haben, wodurch sich nicht das gleiche Ergebnismuster wie bei den seriell schnellen Bildern eingestellt hat.

Interaktionen zwischen den Bedingungen „Key-Frames seriell schnell“ und „Key-Frames seriell langsam“ mit dem Vorwissen von Personen

Die Hypothesen 6a – 6d fokussierten sich auf die Interaktionen zwischen der seriell schnellen und der seriell langsamen Bildbedingung mit dem Vorwissen der Probanden hinsichtlich der vier in der Untersuchung verwendeten Kriterien. Es wurde davon ausgegangen, dass signifikante Interaktionen bestehen, wobei deren Richtungen zunächst unklar waren, da sowohl bei seriell schnellen, als auch bei seriell langsamen Bildern lernerleichternde, aber auch -erschwerende Faktoren vorhanden sind, auf die sich das Vorwissen einer Person differentiell auswirken kann.

Bei den Kriterien „Nachwissensskala“ und „Cognitive Load beim Nachttest“ fanden sich keine statistisch bedeutsamen Interaktionen. Hinsichtlich der Inferenzskala nahm der Regressionskoeffizient der Interaktion einen größeren, aber nicht signifikanten Wert an. Die entsprechenden Hypothesen (6a, 6b, 6d) konnten somit nicht bestätigt werden. Deskriptiv betrachtet, hatte das Vorwissen einen stärkeren Einfluss auf die Scores der Inferenzskala des Nachttests in der seriell schnellen Bildbedingung im Vergleich zur seriell langsamen Bildbedingung. Personen mit mehr Expertise erzielten deskriptiv bessere Werte auf der Inferenzskala mit seriell schnellen Key-Frames. Anfänger hingegen profitierten eher von seriell langsamen Bildern, wobei dieser Vorteil ebenfalls nicht signifikant wurde. Das gleiche Muster wurde bei dem Kriterium „Cognitive Load beim Lernen“ festgestellt. Die Interaktion war hier sehr signifikant, so dass die Hypothese 6c bestätigt werden konnte. Das Vorwissen der Probanden beeinflusste die mentale Belastung beim Lernen stärker negativ in der seriell schnellen als in der seriell langsamen Bildbedingung. Bei höherem Vorwissen empfanden die Probanden die seriell schnellen Key-Frames als signifikant weniger anstrengend als die seriell langsamen Key-Frames, während für Personen mit geringerem Vorwissen das Gegenteil galt, wobei dieser Unterschied allerdings nicht signifikant wurde. Das bedeutet, dass Personen mit höherem Vorwissen signifikant weniger Cognitive Load beim Lernen empfanden, wenn sie seriell schnelle Bilder betrachteten als wenn sie mit seriell langsamen Bildern konfrontiert waren. Bei Personen mit weniger Vorwissen war, in nur deskriptivem Ausmaß, das Gegenteil der Fall, so dass man hier auch wiederum von einem „Expertise Reversal Effect“ (Kalyuga et al., 2003) sprechen konnte. Der gleiche Sachverhalt galt für die Inferierfähigkeiten der Probanden, obwohl die Interaktion hier nicht signifikant war: Personen mit höherem Vorwissen lernten besser mit seriell schnellen als mit seriell langsamen Key-Frames, während für Personen mit niedrigerem Vorwissen seriell langsame Bilder geeigneter waren, wobei diese Unterschiede nicht signifikant wurden. Eine Erklärung für diese Ergebnisse könnte darin liegen, dass seriell langsame Bilder den Vorteil haben, dass ein einzelnes Key-Frame relativ lange gezeigt wird. Der Lerner hat mehr Zeit, sich im Bild zu orientieren, die Teile der Synapse zu benennen und die Vorgänge zu verbalisieren. Dadurch wird ein gutes Verständnis der Teilprozesse gewährleistet, das eine Person mit weniger Vorwissen erst einmal erlangen muss, um das mentale Modell des Gesamtvorganges letztendlich aufbauen zu können (Mayer & Chandler, 2001), weshalb seriell langsame Bilder, zumindest deskriptiv, für sie von Vorteil waren. Die „intrastate“-Information wird bei dieser Darstellungsform betont. Bei seriell schnellen Bildern hingegen werden die einzelnen

Key-Frames nur kurz gezeigt, und es werden eher der Gesamtkontext des Prozesses und die „interstate“-Information hervorgehoben. Für Personen mit höherem Vorwissen scheint die geringere Präsentationsdauer bei seriell schnellen Bildern angemessen zu sein, um die Prozesse lernen bzw. um sich an sie erinnern zu können. Eine höhere Präsentationsdauer, wie sie bei seriell langsamen Bildern der Fall ist, könnte sie zum einen motivational unterfordern, zum anderen ihnen redundante Informationen über die Einzelbilder geben, die sie nicht benötigen. Daher wird ihr Arbeitsgedächtnis bei einer längeren Präsentationsdauer der Bilder unnötig belastet, weshalb sie vielleicht die seriell langsamen Key-Frames als signifikant anstrengender empfanden. Bei fortgeschrittenen Lernern scheint es wichtiger zu sein, den Gesamtkontext und die Beziehungen zwischen den Teilprozessen zu betonen, da dieser höhere Abstraktionsgrad eher ihrem Kenntnisstand entspricht und ihn erweitert. Wie schon weiter oben erwähnt, sind die Ergebnisse allerdings auf den Vorwissensbereich erst einmal einzuschränken, der in der Stichprobe vorhanden war und der auf einem unterdurchschnittlichen Niveau lag.

Interaktionen zwischen den Faktoren „Bedingung“ und „Räumliche Fähigkeiten“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit den räumlichen Fähigkeiten von Personen

In den Hypothesen 7a – 7d wurden die Bedingungen „Animation“ und „Bild seriell schnell“ miteinander verglichen, auf die Frage hin, inwiefern eine Interaktion zwischen ihnen und der Variable „Räumliche Fähigkeiten“ hinsichtlich der Kriterien „Nachwissensskala“, „Inferenzskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachtest“ jeweils vorliegt. Die Hypothesen waren ungerichtet formuliert, da räumliches Vorstellungsvermögen sowohl die Wahrnehmung und Verarbeitung einer Animation beeinflusst, als auch die mentale Animation beim Lernen mit statischen Bildern unterstützt (Hegarty & Kriz, 2008; Sims & Hegarty, 1997).

Von den vier Hypothesen konnte keine bestätigt werden. Lediglich bei dem Kriterium „Inferenzskala“ lag eine etwas größere deskriptive Interaktion vor, in der Richtung, dass die räumlichen Fähigkeiten einer Person bedeutender für ihre Inferierfähigkeiten in der seriell schnellen Bild- als in der Animationsbedingung waren. Das bedeutet, dass Personen mit höheren räumlichen Fähigkeiten deskriptiv besser mit seriell schnellen Bildern lernten als

mit einer Animation, während bei Personen mit geringeren räumlichen Fähigkeiten das Gegenteil der Fall war, wobei auch dieser Unterschied nicht signifikant wurde. Ein ähnliches Richtungsmuster konnte deskriptiv bei dem Kriterium „Nachwissensskala“ gefunden werden, das in den beiden „Cognitive Load“-Variablen allerdings nicht erhalten blieb.

Die Befunde in den beiden Leistungsmaßen sind aufgrund der nicht vorhandenen Signifikanzen nicht zu weit zu interpretieren. Allerdings können sie als Hinweis darauf gewertet werden, dass für Personen mit geringeren räumlichen Fähigkeiten der Aufbau eines mentalen Modells offenbar leichter ist, wenn sie mit Animationen lernen, da diese Darstellungsform den Prozess vollständig präsentiert und den Personen daher eine vollkommen eigenständige mentale Animation und mentale Rotation der Vorgänge abgenommen wird, für die räumliches Vorstellungsvermögen notwendig ist („Supplantationsfunktion“ der Animation, Salomon, 1979). Umgekehrt scheint die Anforderung zur mentalen Animation beim Lernen mit statischen Bildern für Personen mit höherem räumlichem Vorstellungsvermögen angemessen zu sein. Das bloße Betrachten einer Animation könnte eventuell eine Unterforderung für diese Personen darstellen.

Die Hypothesen 8a – 8d, in denen von signifikanten Interaktionen zwischen der Animations- und der seriell langsamen Bildbedingung mit den räumlichen Fähigkeiten ausgegangen wurde, waren ebenfalls ungerichtet aufgestellt und ließen das Interaktionsmuster somit offen. Die Analysen zeigten keine signifikanten Interaktionen zwischen den beiden Gruppen und dem räumlichen Vorstellungsvermögen der Probanden, so dass keine der Hypothesen bestätigt werden konnte. Deskriptiv gesehen, gab es kein einheitliches Interaktionsmuster.

Insgesamt nahm von den acht überprüften Interaktionen zwischen der Animations- und den seriellen Bildbedingungen mit der Variable „Räumliche Fähigkeiten“ keine einen statistisch bedeutsamen Wert an, lediglich in einem Fall konnte eine etwas größere deskriptive Interaktion festgestellt werden.

In der Herleitung der Hypothesen war davon ausgegangen worden, dass die räumlichen Fähigkeiten einer Person sowohl in den seriellen Bild- als auch in der Animationsbedingung einen stärkeren Einfluss auf die Lernleistung haben können als in der jeweils anderen Bedingung, so dass signifikante Interaktionen entstehen.

Für einen stärkeren Einfluss des räumlichen Vorstellungsvermögens auf das Lernen mit Animationen würde sprechen, dass räumliche Fähigkeiten benötigt werden, um die räumliche Information innerhalb einer Animation verarbeiten zu können. Da eine Animation

flüchtig ist, muss die räumliche Verarbeitung mit einer gewissen Geschwindigkeit ablaufen, und das räumliche Arbeitsgedächtnis muss genügend Kapazität haben, um die Eindrücke sammeln und miteinander verbinden zu können (Salthouse, 1996; Shah & Miyake, 1996). Diese Anforderungen an den Lerner waren in der aktuellen Studie abgemildert: Die Geschwindigkeit der Animation wurde bei der Materialentwicklung empirisch auf ihre Angemessenheit überprüft. Der Übertragungsprozess an der Synapse wurde sieben Mal wiederholt, so dass der Lerner nicht so viele Informationen im Gedächtnis speichern musste und auch im Falle einer geringeren räumlichen Verarbeitungsgeschwindigkeit mehrere Möglichkeiten hatte, den Prozess in seiner Gesamtheit zu verstehen. Außerdem war die Darstellung zweidimensional, weshalb eventuell der Einfluss der räumlichen Fähigkeiten beim Lernen mit der Animation, insgesamt gesehen, und auch im Vergleich zum Lernen mit statischen Bildern, nicht so stark war. Die Regressionskoeffizienten lagen bei drei der Kriterien um Null.

Bei den seriellen Bildbedingungen war in den Hypothesen ein größerer Effekt des räumlichen Vorstellungsvermögens darauf begründet, dass hier der Lerner das mentale Modell eigenständiger aufbauen und dafür den Prozess mental animieren und rotieren muss. Für die mentale Animation sind räumliche Fähigkeiten unabdingbar (Hegarty, 1992; Sims & Hegarty, 1997). Eventuell waren allerdings in der vorliegenden Studie die internalen Konstruktionsprozesse etwas erleichtert, da die Darstellung der Synapse zweidimensional war und auch ein zweidimensionales mentales Modell ausreichte, um die Aufgaben des Nachtests lösen zu können. Daher waren die mentalen Animations- und Rotationsprozesse räumlich nicht so schwierig, so dass die internalen Konstruktionsvorgänge des dynamischen mentalen Modells auch für Personen mit geringerem räumlichem Vorstellungsvermögen machbar waren und der Einfluss der figuralen Fähigkeiten in den statischen Bildbedingungen nicht viel stärker war als in der Animationsbedingung.

Interaktionen zwischen den Faktoren „Bedingung“ und „Merkfähigkeit“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit der Merkfähigkeit von Personen

Bei den Hypothesen 9a – 9d ging es darum, die Animations- mit der seriell schnellen Bildbedingung zu kontrastieren, um zu überprüfen, ob signifikante Interaktionen mit der

Variable „Merkfähigkeit“ bezüglich der abhängigen Variablen „Nachwissensskala“, „Inferenzskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachtest“ vorlagen. Die Hypothesen waren wiederum ungerichtet formuliert, da es zum einen Gründe gab, die dafür sprachen, dass für das Lernen mit Animationen mehr Gedächtniskapazität benötigt wird, es zum anderen aber auch möglich war, dass das Lernen mit statischen Bildern mehr Merkfähigkeit beansprucht.

Bei den Kriterien „Nachwissensskala“ und „Inferenzskala“ ergaben sich keine signifikanten Interaktionen, so dass die entsprechenden Hypothesen nicht bestätigt werden konnten. Statistisch gesehen, mussten auch die Hypothesen 9c und 9d zu den mentalen Belastungen beim Lernen und während des Nachtests verworfen werden. Allerdings erreichten die Regressionsgewichte der Interaktionen hier annähernd Signifikanz (ihre p -Werte lagen bei $p = .07$ bzw. $p = .09$), so dass das Ergebnismuster zumindest im Ansatz interpretiert werden kann. In beiden Fällen beeinflusste die Merkfähigkeit der Probanden den Cognitive Load stärker negativ in der Animations- als in der seriell schnellen Bildbedingung. Das heißt, dass Personen mit geringeren Gedächtnisfähigkeiten deskriptiv mehr Cognitive Load beim Betrachten des Lernmaterials und bei der Bearbeitung des Nachtests in der Animations- als in der seriell schnellen Bildbedingung empfanden. Personen mit höherer Merkfähigkeit hingegen verspürten eine stärkere mentale Belastung bei seriell schnellen Key-Frames als bei Animationen, wobei dieser Unterschied bei der Variable „Cognitive Load beim Lernen“ auch signifikant war.

Es scheint somit so zu sein, dass beim Lernen mit einer Animation vergleichsweise bessere Gedächtnisfähigkeiten benötigt werden als beim Lernen mit seriell schnellen Key-Frames. Dies könnte daran liegen, dass eine Animation vollständige Informationen über einen Prozess liefert, was aber auch nachteilig sein kann, da mehr Eindrücke auf den Lerner einströmen, die er filtern und im Arbeitsgedächtnis speichern muss (Lewalter, 1997; Sweller et al., 1998). Teilweise sind diese Informationen für das Verständnis des Vorganges nicht unbedingt notwendig und belasten das Arbeitsgedächtnis unnötig. Zum anderen muss der Lerner aufgrund der Flüchtigkeit der Animation, die in der aktuellen Studie durch die höhere Zahl an Wiederholungen des Prozesses abgemildert war, bei dieser Präsentationsform Vergleiche zwischen einzelnen Zuständen oder genauere Inspektionen internal durchführen, was bei statischen Bildern eher external unterstützt wird (Lowe, 1999a). Bei Personen mit höherer Merkfähigkeit scheint die Belastung des Arbeitsgedächtnisses durch eine Animation

angemessen zu sein, da sie diese als signifikant weniger anstrengend empfanden als die seriell schnellen Bilder.

Zuletzt waren auch die Hypothesen 10a – 10d zu den Interaktionen zwischen der Animations- und der seriell langsamen Bildbedingung mit der Variable „Merkfähigkeit“ ungerichtet formuliert.

Von diesen Hypothesen konnte keine durch die Analysen bestätigt werden, alle Regressionsgewichte der Interaktionen waren nicht signifikant. Allerdings wiesen die Interaktionsmuster bei drei der Kriterien deskriptiv die Richtung auf, dass Personen mit geringerer Merkfähigkeit mehr von den Key-Frames, Personen mit höheren Gedächtnisfähigkeiten jedoch mehr von den Animationen profitierten, was konsistent zu den obigen Befunden mit seriell schnellen Key-Frames war. Auf alle abhängigen Variablen übte die Merkfähigkeit der Probanden einen deskriptiv stärkeren Einfluss in der Animations- als in der seriell langsamen Bildbedingung aus, bei den Leistungsmaßen geschah dies in positiver, bei den Belastungsmaßen in negativer Richtung.

Dies weist zusätzlich im Ansatz darauf hin, dass durch die hohe Informationsdichte der Animation und durch die Notwendigkeit der internalen Vergleiche der einzelnen Zustände, bei dynamischen Darstellungen vergleichsweise mehr Merkfähigkeit gebraucht wird als bei statischen Darstellungen.

4.3.2 Teilstudie 2

4.3.2.1 Fragestellung und Hypothesen

Haupteffekte

In der zweiten Hauptteilstudie sollten Vergleiche zwischen parallel angeordneten und seriell schnellen sowie seriell langsamen Key-Frames gezogen werden. Die Hypothesen zwischen parallelen und seriell schnellen Bildern wurden ungerichtet formuliert, da theoretische Vorüberlegungen nahegelegt hatten, dass es bei beiden Präsentationsarten Vor- und Nachteile beim Lernprozess geben kann. In analoger Weise wurden auch die Hypothesen zwischen parallelen und seriell langsamen Key-Frames ungerichtet aufgestellt. Die Kontraste zwischen parallelen und seriell schnellen bzw. seriell langsamen Key-Frames waren unabhängig voneinander, da durch die unterschiedlichen Darstellungsgeschwindigkeiten und die damit einhergehenden teilweise verschiedenen Lernmerkmale der seriellen Bilder differentielle Effekte hätten entstehen können.

Die Hypothesen lauteten wie folgt:

Parallele Key-Frames vs. Seriell schnelle Key-Frames

Hypothese 1a: Personen, die mit parallelen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Wissensaufgaben eines Nachtests, als Personen, die mit seriell schnellen Key-Frames lernen.

Hypothese 1b: Personen, die mit parallelen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Inferenzaufgaben eines Nachtests, als Personen, die mit seriell schnellen Key-Frames lernen.

Hypothese 1c: Personen, die mit parallelen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lernen, als Personen, die mit seriell schnellen Key-Frames lernen.

Hypothese 1d: Personen, die mit parallelen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lösen des Nachttests, als Personen, die mit seriell schnellen Key-Frames lernen.

Parallele Key-Frames vs. Seriell langsame Key-Frames

Hypothese 2a: Personen, die mit parallelen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Wissensaufgaben eines Nachttests, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 2b: Personen, die mit parallelen Key-Frames lernen, erzielen signifikant unterschiedliche Werte in den Inferenzaufgaben eines Nachttests, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 2c: Personen, die mit parallelen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lernen, als Personen, die mit seriell langsamen Key-Frames lernen.

Hypothese 2d: Personen, die mit parallelen Key-Frames lernen, empfinden einen signifikant unterschiedlichen Cognitive Load beim Lösen des Nachttests, als Personen, die mit seriell langsamen Key-Frames lernen.

Interaktionen

Interaktionen zwischen den Faktoren „Bedingung“ und „Vorwissen“

Interaktionen zwischen den Bedingungen „Key-Frames parallel“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit dem Vorwissen von Personen

Bei diesen Wechselwirkungen wurden ebenfalls ungerichtete Hypothesen aufgestellt, da möglich war, dass Personen mit einem geringeren Vorwissen mehr von parallelen als von seriellen Key-Frames profitieren im Vergleich zu Personen, die über ein höheres Vorwissen verfügen. Umgekehrt war aber auch ein gegenteiliges Interaktionsmuster plausibel. Für die seriell schnellen und die seriell langsamen Key-Frames wurden aus denselben Gründen wie bei den Haupteffekten jeweils eigene Hypothesen aufgestellt.

Hypothesen 3a - d: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Bild parallel“ und „Bild seriell *schnell*“ mit dem Vorwissen von Personen. Diese Annahme bezieht sich auf

a: die Leistung in Wissensaufgaben eines Nachttests

b: die Leistung in Inferenzaufgaben eines Nachttests

c: den Cognitive Load beim Lernen

d: den Cognitive Load beim Lösen des Nachttests

Hypothesen 4a - d: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Bild parallel“ und „Bild seriell *langsam*“ mit dem Vorwissen von Personen. Diese Annahme bezieht sich auf

a: die Leistung in Wissensaufgaben eines Nachttests

b: die Leistung in Inferenzaufgaben eines Nachttests

c: den Cognitive Load beim Lernen

d: den Cognitive Load beim Lösen des Nachttests

4.3.2.2 Methoden

Die methodischen Aspekte der zweiten Hauptteilstudie wurden im Methodenteil der ersten Hauptteilstudie mit beschrieben, da die beiden Studien aus pragmatischen Gründen gemeinsam durchgeführt wurden. Daher sei für die Darstellung des methodischen Vorgehens bei der zweiten Teilstudie auf das Kapitel 4.3.1.2 der vorliegenden Dissertation verwiesen.

4.3.2.3 Ergebnisse

Auch bei dieser Teilstudie wurden die gesamten statistischen Analysen mit dem Programm „SPSS, Version 17.0“ vorgenommen.

Item- und Ausreißeranalyse

Wie oben schon erwähnt, wurden die Item- und Ausreißeranalysen der beiden Hauptteilstudien gemeinsam durchgeführt, so dass für die betreffenden Ergebnisse auf den Abschnitt 4.3.1.3 der ersten Hauptteilstudie verwiesen sei.

Deskriptive Statistiken

In der Stichprobe der zweiten Hauptteilstudie befanden sich 58 Personen. Davon waren 14 Männer und 44 Frauen. Die Probanden waren im Durchschnitt 22.12 Jahre alt ($SD = 2.25$), und der Altersrange betrug 18 bis 28 Jahre.

Im Folgenden werden die deskriptiven Statistiken der abhängigen und der Kontrollvariablen dargestellt.

Tabelle 4.3.2.1

Deskriptive Statistiken der abhängigen Variablen in der zweiten Hauptteilstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Nachwissensskala	58	.00	24.00	14.66	7.23
Inferenzskala	58	.00	32.00	18.78	8.62
Cognitive Load (Lernaufgabe)	58	4.00	16.00	11.22	2.82
Cognitive Load (Nachttest)	58	9.00	17.00	13.78	2.30

Tabelle 4.3.2.2

Deskriptive Statistiken der Kontrollvariablen in der zweiten Hauptteilstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Vorwissenstest	58	.00	12.00	4.19	3.17
Begriffswissen	58	1.00	10.00	7.43	2.64
IST verbal	58	14.00	32.00	23.34	4.12
IST räumlich	58	8.00	40.00	21.55	6.68
IST Merkfähigkeit	57	7.00	23.00	15.28	4.77
Erfolgswahrscheinlichkeit	58	9.00	26.00	19.90	4.09
Misserfolgsbefürchtung	57	5.00	28.00	13.09	6.27
Interesse	56	7.00	31.00	21.70	5.51
Herausforderung	57	11.00	20.00	15.88	2.36

*Mittelwertsvergleiche**Testung auf Normalverteilung und Varianzhomogenität*

Die Testung auf Normalverteilung der Daten in den abhängigen Variablen fand mit Hilfe von Kolmogorov-Smirnov-Tests statt. Normalverteilungen lagen bei allen Kriterien, außer bei der Nachwissensskala, vor (Kolmogorov-Smirnov-Z: 1.35, $p = .05$).

Die Varianzhomogenitäten in den Bedingungen wurden mittels Levene-Tests überprüft. Die Varianzen der drei Gruppen waren in allen abhängigen Variablen homogen.

Wichtige Voraussetzungen der ANOVA sowie der ANCOVA waren somit erfüllt, so dass diese Verfahren zur Anwendung kommen konnten.

Bestimmung relevanter Kovariaten

Varianzanalysen bzw. Brown-Forsyth-Tests der Kontrollvariablen zeigten, dass ein größerer Mittelwertsunterschied zwischen den Gruppen nur bei der FAM-Subskala „Misserfolgsbefürchtung“ vorhanden war ($F(2, 54) = 2.82, p = .07, \eta^2 = .09$). Diese Variable korrelierte lediglich mit dem Kriterium „CL Lern“ signifikant ($r = .27, p$ (1-seitig) = .02), so dass für die Mittelwertsvergleiche folgende Verfahren zum Einsatz kamen:

- | | |
|-------------------------|--|
| 1. AV: Nachwissensskala | Verfahren: ANOVA |
| 2. AV: Inferenzskala | Verfahren: ANOVA |
| 3. AV: CL Lern | Verfahren: ANCOVA mit „Misserfolgsbefürchtung“ |
| 4. AV: CL Nachtest | Verfahren: ANOVA |

Varianz- und Kovarianzanalysen

Hypothese 1a: Bild parallel ist anders als Bild seriell schnell; AV = Nachwissensskala

Hypothese 2a: Bild parallel ist anders als Bild seriell langsam; AV = Nachwissensskala

Die Testung dieser Hypothesen erfolgte mit einer ANOVA und anschließenden t-Tests, mit denen die A-priori-Kontraste überprüft werden sollten. Bei den nachfolgenden Mittelwertsvergleichen, die sich auf die anderen Kriterien bezogen, wurde analog vorgegangen, wobei eine global angewandte Kovarianzanalyse paarweise ALMs unter Verwendung gleicher Kovariaten zur Folge hatte. Das Signifikanzniveau betrug bei den Kontrasten jeweils $\alpha = .05$.

Zur klareren Strukturierung werden bei jedem Vergleich zunächst die Hypothesen des Theorieteils in statistischer Kurzform nochmals dargestellt, und die entsprechenden Ergebnisse werden jeweils hinzugefügt.

Tabelle 4.3.2.3

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Nachwissensskala“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild parallel	18	13.67	6.83
Bild seriell schnell	20	15.10	7.75
Bild seriell langsam	20	15.10	7.34

Die Bedingung „Bild parallel“ erzielte den geringsten Wert in der AV „Nachwissensskala“, wobei die Unterschiede zwischen den Gruppen insgesamt keine Signifikanz erreichten ($F(2,55) = .24$, $p = .79$, $\eta^2 = .01$). Ebenfalls keine bedeutsamen Unterschiede wiesen die Kontraste zwischen den Gruppen „Bild parallel“ und „Bild seriell schnell“ ($t(36) = .60$, p (2-seitig) = .55, $d = .20$) sowie „Bild parallel“ und „Bild seriell langsam“ ($t(36) = .62$, p (2-seitig) = .54, $d = .20$) auf.

Hypothese 1b: Bild parallel ist anders als Bild seriell schnell; AV = Inferenzskala

Hypothese 2b: Bild parallel ist anders als Bild seriell langsam; AV = Inferenzskala

Tabelle 4.3.2.4

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Inferenzskala“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild parallel	18	16.39	6.61
Bild seriell schnell	20	18.70	10.60
Bild seriell langsam	20	21.00	7.80

In der Tabelle ist zu erkennen, dass die besten Inferierleistungen von der Bedingung „Bild seriell langsam“, die schlechtesten von der Gruppe „Bild parallel“ erbracht wurden. Eine einfaktorielles Varianzanalyse zeigte keine signifikanten Unterschiede ($F(2,55) = 1.38$, $p = .26$, $\eta^2 = .05$). Jedoch erreichte die Größe der Mittelwertsdifferenz im Kontrast zwischen

der parallelen und der seriell langsamen Bildbedingung annähernd Signifikanz ($t(36) = 1.96$, $p(2\text{-seitig}) = .06$, $d = .64$), wobei sich die parallele Bildbedingung als unterlegen erwies. In dem anderen Kontrast zwischen der parallelen und der seriell schnellen Bildbedingung ließ sich hingegen kein größerer Unterschied finden ($w(32,2) = .82$, $p(2\text{-seitig}) = .42$, $d = .26$).

Hypothese 1c: Bild parallel ist anders als Bild seriell schnell; AV = CL Lern

Hypothese 2c: Bild parallel ist anders als Bild seriell langsam; AV = CL Lern

Die Testung dieser Hypothesen erfolgte mit einer ANCOVA mit der Kontrollvariable „Misserfolgsbefürchtung“, daher werden hier adjustierte Mittelwerte berichtet.

Tabelle 4.3.2.5

Adjustierte Mittelwerte und Standardfehler der Bedingungen in der AV „CL Lern“

	<i>n</i>	<i>M</i>	<i>SE</i>
Bild seriell schnell	20	10.77	.63
Bild parallel	17	11.38	.67
Bild seriell langsam	20	11.71	.63

Eine Kovarianzanalyse unter Einbezug der drei Bedingungen ergab keine signifikanten Unterschiede ($F(2, 53) = .55$, $p = .58$, partielles $\eta^2 = .02$). Bei den Probanden der seriell schnellen Bildbedingung wurde der geringste, bei denen der seriell langsamen Bildbedingung der höchste Cognitive Load gemessen. Sowohl der Mittelwertsunterschied im Kontrast zwischen der parallelen und der seriell schnellen ($F(1, 34) = .45$, $p(2\text{-seitig}) = .51$, partielles $\eta^2 = .01$), als auch der zwischen paralleler und seriell langsamer Bildbedingung ($F(1, 34) = .09$, $p(2\text{-seitig}) = .77$, partielles $\eta^2 = .003$), erreichte keine Signifikanz.

Hypothese 1d: Bild parallel ist anders als Bild seriell schnell; AV = CL Nachtest

Hypothese 2d: Bild parallel ist anders als Bild seriell langsam; AV = CL Nachtest

Tabelle 4.3.2.6

Mittelwerte und Standardabweichungen der Bedingungen in der AV „CL Nachtest“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild seriell schnell	20	13.63	2.51
Bild seriell langsam	20	13.68	2.03
Bild parallel	18	14.06	2.44

Zwischen den Mittelwerten der Bedingungen in der abhängigen Variable „Cognitive Load beim Nachtest“ ließen sich nur geringe Differenzen finden ($F(2,55) = .19, p = .83, \eta^2 = .01$). Die parallele Bildbedingung hatte einen leicht höheren Wert als die seriell schnelle ($t(36) = .54, p(2\text{-seitig}) = .60, d = .17$) und als die seriell langsame Bildbedingung ($t(36) = .53, p(2\text{-seitig}) = .60, d = .17$).

Interaktionsanalysen

Die Berechnungen der Interaktionen wurden in analoger Weise zu dem Vorgehen in der ersten Hauptteilstudie vorgenommen. Für eine genauere Erläuterung der Analysen sei daher auf das Kapitel 4.3.1.3 verwiesen.

Ebenso wie in der ersten Teilstudie werden auch hier größere oder signifikante Interaktionen mit Tabellen und Liniendiagrammen veranschaulicht.

Interaktionen zwischen den Faktoren „Bedingung“ und „Vorwissen“

Interaktionen zwischen den Gruppen „Bild parallel“ und „Bild seriell schnell“ mit der Variable „Vorwissen“

Hypothese 3a: Interaktion „Bild parallel vs. Bild seriell schnell“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = Nachwissensskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 54) = 5.81$, $p = .002$, $R^2 = .24$, $adj R^2 = .20$).

Zwischen den Gruppen „Bild parallel“ und „Bild seriell schnell“ und der Variable „Vorwissen“ lag bezüglich des Kriteriums „Nachwissensskala“ keine signifikante Interaktion vor ($B = -.38$, $SE = .44$, $t = -.86$, $p = .39$). Der Einfluss des Vorwissens war stärker in der seriell schnellen als in der parallelen Bildbedingung (Bild seriell schnell: $B = 1.27$, $p = .003$; Bild parallel: $B = .40$, $p = .66$).

Hypothese 3b: Interaktion „Bild parallel vs. Bild seriell schnell“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = Inferenzskala

Tabelle 4.3.2.7

Regressionsanalyse für die Bestimmung der Interaktion „Bild parallel vs. Bild seriell schnell x Vorwissen“ bei der AV „Inferenzskala“

Variable	<i>B</i>	<i>SE B</i>	<i>t</i>	<i>p</i>
Bedingung: Bild parallel vs. Bild seriell schnell	-.74	1.23	-.60	.55
Vorwissen	1.37	.32	4.24	< .001
Interaktion Bedingung x Vorwissen	-.86	.47	-1.83	.07

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 54) = 12.17$, $p < .001$, $R^2 = .40$, $adj R^2 = .37$).

Die Interaktion zwischen den Bedingungen „Bild parallel“ und „Bild seriell schnell“ mit dem Vorwissen nahm bei der abhängigen Variablen „Inferenzskala“ einen annähernd

signifikanten Wert an ($B = -.86, p = .07$). Das Vorwissen hatte einen höheren Einfluss auf die Inferierfähigkeiten einer Person in der seriell schnellen ($B = 2.18, p < .001$) als in der parallelen Bildbedingung ($B = .35, p = .71$). Bei höherem Vorwissen war die seriell schnelle Bildgruppe im Vorteil ($B = 7.38, p = .13$), während bei geringerem Vorwissen parallele Key-Frames geeigneter waren ($B = -4.21, p = .21$). Diese Unterschiede erreichten allerdings keine Signifikanz.

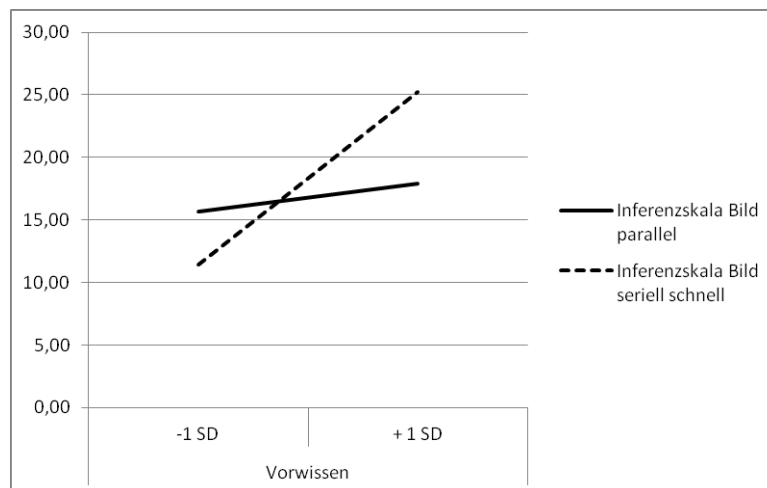


Abbildung 4.3.2.1: Regressionsgeraden des Vorwissens in den Gruppen „Bild parallel“ und „Bild seriell schnell“ (AV: Inferenzskala)

Hypothese 3c: Interaktion „Bild parallel vs. Bild seriell schnell“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = CL Lern

Tabelle 4.3.2.8

Regressionsanalyse für die Bestimmung der Interaktion „Bild parallel vs. Bild seriell schnell x Vorwissen“ bei der AV „CL Lern“

Variable	B	SE B	t	p
Bedingung: Bild parallel vs. Bild seriell schnell	.44	.47	.95	.35
Vorwissen	-.16	.12	-1.29	.20
Misserfolgsbefürchtung	.15	.06	2.70	.009
Interaktion Bedingung x Vorwissen	.40	.17	2.33	.02

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 52) = 4.47$, $p = .004$, $R^2 = .26$, $adj R^2 = .20$).

Bei dem Kriterium „CL Lern“ zeigte sich das Muster, dass das Vorwissen in der seriell schnellen Bildbedingung einen stärker negativen Regressionskoeffizienten hatte ($B = -.62$, $p < .001$) als in der parallelen Bildbedingung ($B = -.33$, $p = .33$). Der Interaktionsterm war signifikant ($B = .40$, $p = .02$). Personen mit höherem Vorwissen verspürten weniger Cognitive Load bei seriell schnellen Key-Frames ($B = -1.08$, $p = .54$), von Personen mit niedrigerem Vorwissen hingegen wurden parallele Key-Frames als weniger anstrengend beurteilt ($B = .77$, $p = .52$). Diese Unterschiede waren jedoch nicht signifikant.

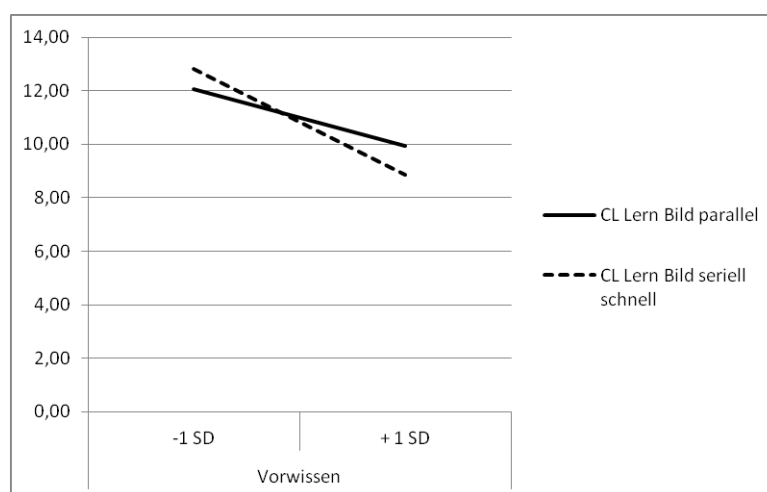


Abbildung 4.3.2.2: Regressionsgeraden des Vorwissens in den Gruppen „Bild parallel“ und „Bild seriell schnell“ (AV: CL Lern)

Hypothese 3d: Interaktion „Bild parallel vs. Bild seriell schnell“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = CL Nachttest

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 54) = 1.05, p = .38, R^2 = .06, adj R^2 = .002$).

Der Regressionskoeffizient der Interaktion zwischen den Bedingungen „Bild parallel“ und „Bild seriell schnell“ mit dem Vorwissen erreichte bei der abhängigen Variablen „CL Nachttest“ kein signifikantes Niveau ($B = .07, SE = .16, t = .42, p = .68$). Deskriptiv war der Einfluss des Vorwissens auf den Cognitive Load stärker negativ in der seriell schnellen als in der parallelen Bildbedingung (Bild seriell schnell: $B = -.22, p = .13$; Bild parallel: $B = -.20, p = .55$).

Interaktionen zwischen den Gruppen „Bild parallel“ und „Bild seriell langsam“ mit der Variable „Vorwissen“

Hypothese 4a: Interaktion „Bild parallel vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = Nachwissensskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 54) = 5.58, p = .002, R^2 = .24, adj R^2 = .19$).

Die Interaktion zwischen den Gruppen „Bild parallel“ und „Bild seriell langsam“ mit der Variable „Vorwissen“ wurde bezüglich des Nachwissens nicht signifikant ($B = -.15, SE = .45, t = -.33, p = .74$). Deskriptiv war der Regressionskoeffizient des Vorwissens stärker positiv in der seriell langsamen als in der parallelen Bildbedingung (Bild seriell langsam: $B = 1.12, p = .01$; Bild parallel: $B = .40, p = .66$).

Hypothese 4b: Interaktion „Bild parallel vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = Inferenzskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 54) = 10.65, p < .001, R^2 = .37, adj R^2 = .34$).

Der Wert der Interaktion zwischen den Gruppen „Bild parallel“ und „Bild seriell langsam“ mit dem Vorwissen wurde nicht signifikant ($B = .04, SE = .49, t = .08, p = .94$). Deskriptiv war bei höherem Vorwissen die seriell langsame Bildbedingung stärker im Vorteil als bei

niedrigerem Vorwissen (Höheres Vorwissen: $B = 5.98$, $p = .21$; Niedrigeres Vorwissen: $B = .52$, $p = .88$).

Hypothese 4c: Interaktion „Bild parallel vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = CL Lern

Tabelle 4.3.2.9

Regressionsanalyse für die Bestimmung der Interaktion „Bild parallel vs. Bild seriell langsam x Vorwissen“ bei der AV „CL Lern“

Variable	<i>B</i>	<i>SE B</i>	<i>t</i>	<i>p</i>
Bedingung: Bild parallel vs. Bild seriell langsam	-.71	.46	-1.54	.13
Vorwissen	-.43	.12	-3.62	.001
Misserfolgsbefürchtung	.14	.05	2.65	.01
Interaktion Bedingung x Vorwissen	-.42	.18	-2.38	.02

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 52) = 4.89$, $p = .002$, $R^2 = .27$, $adj R^2 = .22$).

Der Interaktionswert zwischen den Gruppen „Bild parallel“ und „Bild seriell langsam“ mit der Variable „Vorwissen“ erreichte bezüglich des Kriteriums „Cognitive Load beim Lernen“ statistische Signifikanz ($B = -.42$, $p = .02$). Das Vorwissen beeinflusste den Cognitive Load in der Bedingung „Bild parallel“ stärker negativ ($B = -.33$, $p = .33$) als in der seriell langsamen Bildbedingung ($B = .05$, $p = .75$). Personen mit höherem Vorwissen empfanden weniger Cognitive load bei den parallelen Key-Frames, während bei Personen mit niedrigerem Vorwissen das Gegenteil der Fall war, wobei die Differenzen keine statistische Signifikanz erreichten (Höheres Vorwissen: $B = 1.89$, $p = .28$; Niedrigeres Vorwissen: $B = -.55$, $p = .67$).

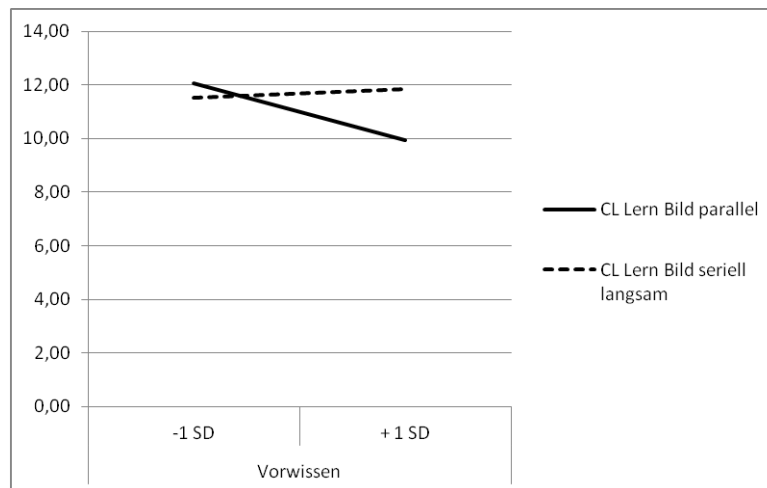


Abbildung 4.3.2.3: Regressionsgeraden des Vorwissens in den Gruppen „Bild parallel“ und „Bild seriell langsam“ (AV: CL Lern)

Hypothese 4d: Interaktion „Bild parallel vs. Bild seriell langsam“ x Vorwissen ist signifikant, ungerichtete Fragestellung, AV = CL Nachttest

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 54) = 1.11, p = .36, R^2 = .06, adj R^2 = .005$).

Zwischen den Bedingungen „Bild parallel“ und „Bild seriell langsam“ lag mit dem Vorwissen bei der abhängigen Variablen „CL Nachttest“ keine statistisch signifikante Interaktion vor ($B = -.10, SE = .16, t = -.65, p = .52$). Der Einfluss des Vorwissens war stärker negativ in der parallelen als in der seriell langsamen Bildbedingung (Bild parallel: $B = -.20, p = .55$; Bild seriell langsam: $B = -.08, p = .63$).

Post-hoc-Ergebnisse der Interaktionsanalysen

Explorativ wurden weitere Interaktionsanalysen zwischen den Bedingungs- und den Kontrollvariablen durchgeführt. Es zeigte sich, dass bei den Gruppen „Bild parallel“ und „Bild seriell langsam“ größere bis signifikante Interaktionen mit der motivationalen Variable „Herausforderung“ bei allen vier Kriterien vorlagen (Nachwissensskala: $B = 1.28, p = .03$; Inferenzskala: $B = 1.27, p = .06$; CL Lern: $B = -.39, p = .08$; CL Nachttest: $B = -.44, p = .009$). In der Bedingung „Bild parallel“ waren die Regressionskoeffizienten der Variable „Herausforderung“ bei den Kriterien „Nachwissensskala“ und „Inferenzskala“ stärker positiv (Nachwissensskala: $B = 1.17, p = .17$; Inferenzskala: $B = 1.28, p = .19$), bei den abhängigen Variablen „CL Lern“ und „CL Nachttest“ waren sie jedoch negativ (CL Lern: $B = -.53,$

$p = .12$; CL Nachtest: $B = -.82$, $p = .001$). In der seriell langsamen Bildbedingung war der Einfluss der Variable „Herausforderung“ auf die Werte in der Nachwissens- und in der Inferenzskala hingegen negativ (Nachwissensskala: $B = -1.18$, $p = .10$; Inferenzskala: $B = -1.10$, $p = .19$). Bei allen vier Kriterien profitierten Personen mit höheren Werten in der Variable „Herausforderung“ mehr von den parallelen Key-Frames, während Personen mit geringeren Herausforderungsscores besser mit seriell langsamen Key-Frames umgehen konnten. In einigen Fällen wurden diese Unterschiede annähernd bzw. vollständig signifikant (z.B. „CL Nachtest“: Höhere Herausforderung: $B = 2.19$, $p = .06$; Niedrigere Herausforderung: $B = -1.87$, $p = .06$). Dies ließ darauf schließen, dass sich eine intrinsisch motivierte Einstellung förderlich in der selbst zu steuernden, leicht interaktiven, parallelen Bildbedingung ausgewirkt hat. Für Personen, die die Aufgabe nicht als Herausforderung empfanden, waren seriell langsame Key-Frames geeigneter, die extern gesteuert waren und keine zusätzliche Eigentätigkeit des Betrachters erforderten.

4.3.2.4 Diskussion

Haupteffekte

Parallele Key-Frames vs. Seriell schnelle und seriell langsame Key-Frames

In der zweiten Hauptteilstudie ging es darum, die parallele mit der seriell schnellen und der seriell langsamen Bildbedingung miteinander zu vergleichen.

Bei dem Kontrast zwischen der parallelen und der seriell schnellen Bildbedingung wurden die Hypothesen zu den vier Kriterien „Nachwissensskala“, „Inferenzskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachtest“ jeweils ungerichtet formuliert (Hypothesen 1a – 1d). Dies lag daran, dass es zum einen plausibel war, von einer Überlegenheit der parallelen Bilder auszugehen, da die Probanden hier besser selbstgesteuert den Lernvorgang bestimmen konnten. Zum anderen bieten jedoch seriell schnelle Bilder den Vorteil, dass Unterschiede zwischen den Key-Frames perzeptuell salienter werden.

Von den vier ungerichteten Hypothesen konnte keine bestätigt werden, da keine signifikanten Unterschiede zwischen den beiden Gruppen in allen vier abhängigen Variablen vorlagen. Deskriptiv gesehen, erwies sich die parallele Bildbedingung in allen Fällen der seriell schnellen Bildbedingung gegenüber als unterlegen, wobei die Effektstärken ein geringes Ausmaß annahmen.

Um die parallelen mit den seriell langsamen Key-Frames zu vergleichen, wurden wiederum ungerichtete Hypothesen bei allen vier Kriterien gewählt (Hypothesen 2a – 2d). Die Gründe hierfür sind analog zu der Argumentation bezüglich des Kontrastes zwischen den parallelen und den seriell schnellen Bildern zu sehen. Nach den Analysen zeigte sich, dass keine signifikanten Differenzen zwischen den Gruppenmittelwerten hinsichtlich der abhängigen Variablen „Nachwissensskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachtest“ gefunden werden konnten, wodurch die entsprechenden Hypothesen verworfen werden mussten. Die Effektstärken lagen auf niedrigem Niveau. Mit Ausnahme des Kriteriums „Cognitive Load beim Lernen“, erreichten die Probanden der parallelen Bildgruppe schlechtere Lernergebnisse als die der seriell langsamen Bildgruppe. Statistisch gesehen, musste auch die Hypothese 2b bezüglich der Inferierfähigkeiten der Teilnehmer verworfen werden. Allerdings lag der p-Wert hier sehr nahe an der Grenze zur statistischen Signifikanz ($p = .06$), und die Effektstärke wies einen mittleren Wert auf, so dass dieses

Ergebnismuster, wenn auch mit Einschränkungen, näher interpretiert werden kann. Die Probanden der seriell langsamen Bildbedingung zeigten hier bessere Leistungen als die der parallelen Bildbedingung. Es scheint somit so zu sein, dass sich bei diesem Kontrast die Nachteile der parallelen Bilder stärker ausgewirkt haben.

Bei parallelen Key-Frames ist es generell so, dass der Proband seinen Lernprozess selbst steuern kann bzw. muss. Er kann selbst aussuchen, welche Bilder er näher inspizieren und miteinander vergleichen möchte. Diese höhere Eigenständigkeit im Lernen wurde in der aktuellen Studie durch das leichte Interaktivitätsfeature noch verstärkt, bei dem die Lerner an der rechten Seite des Screens an der Taskleiste mit der Maus hoch- und runterscrollen konnten, um alle zehn Frames zu sehen. Eine parallele Anordnung der Bilder hat aber auch Nachteile für den Lerner: Er wählt seine Lernsequenz selbst aus, was aber auch das Risiko in sich birgt, dass er sich für eine eher ungeeignete Reihenfolge des Lernens entscheidet und weniger relevanten Details mehr Beachtung schenkt (Schnotz, 2006). Er muss sich selbst navigieren und seinen Lernprozess stärker selbst überwachen, was ein höheres Maß an Metakognition erfordert, die die Kapazität des Arbeitsgedächtnisses mit beansprucht. Auch muss der Lerner eine gewisse intrinsische Motivation haben, um von der Selbststeuerung bei parallelen Bildern profitieren zu können, da er bei dem Lernprozess stärker aktiv eingebunden ist. Schon in der Pilotstudie war aufgefallen, dass die Personen der parallelen Bildbedingung die Interaktivität nicht voll in Anspruch zu nehmen schienen. Gegen Ende der Lernzeit wirkte das Scrollen der Probanden eher unsystematisch und unkonzentriert. Eventuell überschätzten die Personen auch ihr Verständnis während des Lernvorganges, da sie sich, nach eigener Aussage, an ihr „Wissen aus der Schule“ erinnerten. Bei den seriellen Bildern hingegen wird die Aufmerksamkeit des Lerners durch die Dynamik der Prozesse auf dem Screen eher die volle Lernzeit lang auf das Geschehen auf dem Bildschirm gelenkt.

Zudem ergaben statistische Analysen größere oder signifikante Interaktionen zwischen den Bedingungen „Bild parallel“ und „Bild seriell langsam“ mit der motivationalen Variable „Herausforderung“, die angibt, wie stark sich eine Person bei der Aufgabe anstrengen möchte und wie sehr sie die Aufgabe als „Herausforderung“ für sich bewertet. Es zeigte sich, dass eine solche, eher intrinsische Motivation, sich bei der Aufgabe beweisen zu wollen, einen stärker positiven Effekt auf die Kriterien „Nachwissensskala“ und „Inferenzskala“ sowie einen negativen Einfluss auf die mentale Belastung in der parallelen Bildbedingung hatte. In der seriell langsamen Bildbedingung war der Einfluss der Variable „Herausforderung“ auf die Werte in der Nachwissens- und der Inferenzskala hingegen

negativ. Bei allen vier Kriterien erreichten Personen mit höheren Werten in der Variable „Herausforderung“ bessere Ergebnisse mit den parallelen Key-Frames, während Personen mit geringeren Herausforderungsscores mehr von seriell langsamen Key-Frames profitierten. In einigen Fällen wurden diese Unterschiede annähernd bzw. vollständig signifikant.

Diese Resultate stehen im Einklang mit der Beobachtung aus den Experimenten, dass eine mangelnde intrinsische Motivation zu schlechteren Lernergebnissen in der parallelen Bildbedingung mit geführt haben könnte. Bei seriell langsamen Bildern hingegen wirkte sich eine solche Motivation eher negativ aus, was eventuell durch den monotonen Charakter der Darstellung erklärt werden könnte.

Im Großen und Ganzen schienen die Versuchspersonen die Vorteile der parallelen Key-Frames nicht vollständig genutzt zu haben.

Die Vorzüge der seriell langsamen Bilder liegen hingegen darin, dass den Lernern eine bestimmte Lernreihenfolge der Key-Frames vorgegeben wird, die eventuell qualitativ besser ist als die selbst gewählten Lernsequenzen der Probanden in der parallelen Bildbedingung. Selbstgesteuertes Lernen ist bei seriellen Bildern weniger erforderlich, wodurch auch die höheren Anforderungen an die Metakognition und die Motivation des Lernalters entfallen. Zudem werden bei seriell langsamen Key-Frames die Unterschiede zwischen den Bildern perzeptuell salienter, da die Frames direkt hintereinander gezeigt werden, wodurch dem Lerner eher auffällt, welche Details sich von Bild zu Bild verändern.

Es zeigte sich bei den Ergebnissen, dass bei sieben von acht Vergleichen zwischen parallelen und seriellen Key-Frames, die Probanden der parallelen Bildbedingung unterlegen waren. Hiermit bestätigt sich das Bild, dass bei den parallelen Bildern innerhalb des Lernprozesses eher die inhärenten, oben genannten, Nachteile zum Tragen kamen. Allerdings lag die Größe der Unterschiede zwischen den Bedingungen in den meisten Fällen jenseits der Signifikanzgrenze, so dass diese Argumentation nur beschränkt vertreten werden kann.

Insgesamt gesehen, fielen die Unterschiede zwischen der parallelen und den seriellen Bildbedingungen eher gering aus. Es scheint so zu sein, dass sich die Vor- und Nachteile der drei Darstellungsformen im Großen und Ganzen aufwiegen, wobei, wie schon erwähnt, die Defizite der parallelen Bilder ein leicht höheres Gewicht zu haben scheinen. Eventuell hat hier die erforderliche Selbststeuerung, die durch die Interaktivität noch verstärkt war, eher zu Lernproblemen geführt, als dass sie das Verständnis der Prozesse verbessert hätte, da die Probanden zu einer optimalen Nutzung gewisse motivationale und kognitive

Voraussetzungen mitbringen mussten. Auch in anderen Studien konnte eine solche hemmende Wirkung von Selbststeuerung bzw. Interaktivität beobachtet werden (Schnotz, Böckheler & Grzondziel, 1999; Bétrancourt & Realini, 2005). Letztendlich könnte auch wieder die Thematik des Lernmaterials einen Einfluss gehabt haben, da sie für Studenten eher leicht ist, die die „Signalübertragung an Synapsen“ oft schon in der gymnasialen Oberstufe gelernt haben. Dieses beim Lernen getriggerte Vorwissen könnte zu einer Nivellierung der Lernleistungen trotz unterschiedlicher Darstellungsformen geführt haben.

Interaktionen

Interaktionen zwischen den Faktoren „Bedingung“ und „Vorwissen“

Interaktionen zwischen den Bedingungen „Key-Frames parallel“ und „Key-Frames seriell schnell“ sowie „Key-Frames seriell langsam“ mit dem Vorwissen von Personen

In den Hypothesen 3a – 3d wurde von signifikanten Interaktionen zwischen der parallelen und der seriell schnellen Bildbedingung mit der Variable „Vorwissen“ bezüglich der vier Kriterien „Nachwissensskala“, „Inferenzskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachtest“ ausgegangen. Es handelte sich hierbei um ungerichtete Hypothesen, da zum einen denkbar war, dass die höhere Eigenaktivität des Lernalters bei parallelen Bildern für Personen mit niedrigerem Vorwissen förderlich ist, da sie den Lernvorgang ihren Bedürfnissen anpassen können. Zum anderen könnte gerade diese höhere Selbststeuerung eine kognitive Überforderung für Novizen darstellen, so dass sie mehr von seriell schnellen Bildern profitieren (Schnotz & Lowe, 2008).

Bei den Kriterien „Nachwissensskala“ und „Cognitive Load beim Nachtest“ konnten die Hypothesen nicht bestätigt werden, da die Regressionsgewichte der Interaktionen nicht signifikant wurden. Bei exaktem Anlegen des Signifikanzkriteriums musste auch die Hypothese 3b bezüglich der Inferenzfähigkeiten der Probanden verworfen werden. Die Interaktion erreichte einen größeren, aber nicht signifikanten Wert ($p = .07$). Das Interaktionsmuster verlief in der Form, dass Personen mit höheren Vorwissenswertes deskriptiv bessere Ergebnisse mit seriell schnellen als mit parallelen Bildern erzielten, während sich bei Personen mit geringerem Vorwissen das Gegenteil herausstellte, wobei auch dieser Unterschied nicht signifikant wurde. Diese Resultate verfestigten sich bei der Betrachtung des Kriteriums „Cognitive Load beim Lernen“. Hier wurde die Interaktion

signifikant, so dass die entsprechende Hypothese bestätigt werden konnte. Ebenso wie bei der Variable „Inferenzskala“, profitierten Personen mit höherem Vorwissen mehr von seriell schnellen als von parallelen Bildern, bei Personen mit geringerem Vorwissen stellte sich ein gegenteiliges Bild dar. Die Differenzen waren allerdings auch hier nicht signifikant.

Bevor diese Ergebnisse näher diskutiert werden, sollen zunächst die Interaktionshypothesen zwischen den Gruppen „Bild parallel“ und „Bild seriell langsam“ mit der Variable „Vorwissen“ nochmals genannt werden.

In diesen Hypothesen (4a – 4d) war die Annahme, dass signifikante Interaktionen zwischen der parallelen und der seriell langsamen Bildbedingung mit dem Vorwissen der Probanden hinsichtlich der vier abhängigen Variablen bestehen. Die Hypothesen waren aus größtenteils denselben Gründen wie bei dem Kontrast zwischen parallelen und seriell schnellen Bildern ungerichtet formuliert.

Aufgrund der statistischen Analysen, konnten die Hypothesen, außer bei dem Kriterium „Cognitive Load beim Lernen“, nicht bestätigt werden. Bei dem Cognitive Load, der nach dem Lernen gemessen wurde, zeigte sich eine signifikante Interaktion zwischen den beiden Bedingungen mit der Variable „Vorwissen“, in der Richtung, dass Personen mit höherem Vorwissen das Lernen mit parallelen Bildern als weniger anstrengend und schwierig empfanden als das Lernen mit seriell langsamen Bildern, während sich der Sachverhalt für Personen mit geringerem Vorwissen als umgekehrt darstellte, wobei die Unterschiede jedoch keine Signifikanz erreichten.

Schaut man sich die Ergebnisse des Kontrastes zwischen der parallelen und der seriell schnellen Bildbedingung an, scheint es so zu sein, dass für Personen mit geringerem Vorwissen parallel angeordnete Bilder geeigneter sind, da sie durch die höhere Selbststeuerung den Lernprozess mehr nach ihren Bedürfnissen gestalten können und dass eine fremdgesteuerte serielle Präsentation der Key-Frames für sie eher nachteilig ist, wobei einschränkend hinzugefügt werden muss, dass dieser Unterschied nicht signifikant war.

Eine solche Interpretation würde jedoch eher im Widerspruch zu den Resultaten der Haupteffekte und der Interaktionsanalysen zwischen der parallelen und der seriell langsamen Bildgruppe stehen, bei denen Novizen deskriptiv besser mit seriell langsamen Bildern zurechtkamen. Eventuell könnten die signifikanten Interaktionen bezüglich des Cognitive Loads während des Lernens zwischen der parallelen und der seriell schnellen bzw. zwischen der parallelen und der seriell langsamen Bildgruppe, die unterschiedliche Richtungsmuster

aufwiesen, durch die verschiedenen Präsentationsgeschwindigkeiten der seriellen Bilder entstanden sein.

Für Personen mit geringerem Vorwissen könnte die Präsentationsdauer der einzelnen Key-Frames bei der seriell schnellen Variante zu kurz sein, um die einzelnen Teilprozesse verstehen und sich einprägen zu können. Da sie aber nichts bzw. fast nichts über die synaptischen Vorgänge wissen, brauchen sie erst einmal eine solche Grundlage, um schrittweise das gesamte mentale Modell des Prozesses aufbauen zu können (Mayer & Chandler, 2001). Bei parallelen Key-Frames können sie die Bilder in Ruhe inspizieren und anschauen und sich vor allem den Frames widmen, die sie nicht ganz verstanden haben. Auf diese Art und Weise scheint ein Verstehen einzelner Teilprozesse und deren Zusammenfügen eher möglich zu sein. Die Nachteile, die sich durch die höhere Selbststeuerung bei den parallelen Bildern ergeben, werden ausgeglichen, weshalb diese, zumindest deskriptiv, für Personen mit geringerem Vorwissen geeigneter waren als seriell schnelle Key-Frames.

Bei der Interaktion zwischen simultanen und seriell langsamen Bildern mit dem Vorwissen der Probanden bezüglich des Kriteriums „Cognitive Load beim Lernen“ war der Sachverhalt genau umgekehrt, so dass Personen mit niedrigerem Vorwissen die sequentiellen Frames als weniger schwierig und anstrengend empfanden als die parallelen Frames, wobei diese Differenz auch nicht signifikant war. Dieses Ergebnis könnte, aufgrund der fehlenden statistischen Bedeutsamkeit nur ansatzweise, so interpretiert werden, dass bei einer langsameren Abfolge der seriellen Bilder, ein Anfänger genügend Zeit hat, sich in den einzelnen Frames zu orientieren und die Teilabläufe zu verstehen, was bei seriell schnellen Bildern nicht der Fall ist. Auf diese Weise kann das mentale Modell schrittweise aufgebaut werden. Im Vergleich zu der parallel angeordneten Variante entfallen auch die höheren Anforderungen an die Selbststeuerung und die Metakognition des Lerners, der eine extern konstruierte Sequenz der Abläufe vorgesetzt bekommt, die eventuell einer selbstgewählten gegenüber überlegen ist. Daher war vielleicht bei diesem Kontrast eine sequentielle Abfolge für Personen mit niedrigerem Vorwissen etwas geeigneter als eine simultane Präsentation, während die Ergebnisse bei der schnelleren Darbietung der seriellen Key-Frames genau umgekehrt waren. Diese Schlussfolgerungen sind auch hier auf den Vorwissensrange einzuschränken, der innerhalb der Stichprobe bestand und der, insgesamt gesehen, auf einem niedrigeren Niveau lag ($M = 4.19$ von 18.00 Punkten, $SD = 3.17$).

5 Empirischer Teil von Studie 2: Vergleich von Animationen und statischen Bildern hinsichtlich ihrer Wirksamkeit auf die Lösung von bildlichen Kategorisierungsaufgaben

5.1 Pilotstudie

5.1.1 Fragestellung und Hypothesen

Auch bei der zweiten Studie wurde zunächst eine Pilotstudie durchgeführt, um Hinweise darauf zu erhalten, ob der geplante Versuchsablauf praktisch durchführbar war und ob die Lern- und Testmaterialien verständlich für die Teilnehmer waren. Daraus ergab sich als erste Fragestellung für die Pilotstudie:

Forschungsfrage 1: Ist der Versuchsablauf stimmig und das Experiment in der geplanten Form durchführbar?

Zudem wurden für das Experiment ein paar der verwendeten Testverfahren eigens konstruiert, so dass für sie noch keine Gütekriterien vorlagen. Um die Reliabilitäten, Trennschärfen und Schwierigkeiten für den Vorerfahrungs- und Vorwissenstest sowie für den Nachtest mit der in ihm enthaltenen Animations- und statischen Bildskala (s.u.) zu erhalten, wurden für diese Messinstrumente jeweils Itemanalysen gerechnet. Diese sollten auf die folgende Frage eine Antwort liefern:

Forschungsfrage 2: Sind die Gütekriterien des Wahrnehmungsvorerfahrungstests, des Vorwissenstests und des Nachtests, sowie die der in ihm enthaltenen Animations- und Bildskala, die die selbst konstruierten Verfahren des Experimentes darstellen, zufriedenstellend?

Weiterhin sollten in der Pilotstudie zum einen die Hypothesen zu den Haupteffekten zwischen der Animations- und der statischen Bildbedingung sowie zum anderen die Hypothese zur Interaktion zwischen dem Format der Lern- und dem der Testbedingungen jeweils explorativ untersucht werden. Aufgrund der in der Pilotstudie verwendeten geringeren Stichprobengröße, sollen die Ergebnisse im Diskussionsteil noch nicht zu weitreichend interpretiert werden.

Aufgrund der theoretischen Überlegungen, die bei der Herleitung der Grundannahmen gemacht wurden (s. Abschnitt 3.2), wurden die Haupteffektshypothesen bezüglich der Lösung von bildlichen Kategorisierungsaufgaben und bezüglich des Cognitive Loads zugunsten der Animationsbedingung jeweils gerichtet formuliert:

Hypothese 1a: Personen, die mit Animationen lernen, erzielen signifikant bessere Werte in animierten und statischen bildlichen Kategorisierungsaufgaben eines Nachtests, als Personen, die mit parallelen Key-Frames lernen.

Hypothese 1b: Personen, die mit Animationen lernen, erzielen signifikant bessere Werte in animierten bildlichen Kategorisierungsaufgaben eines Nachtests, als Personen, die mit parallelen Key-Frames lernen.

Hypothese 1c: Personen, die mit Animationen lernen, erzielen signifikant bessere Werte in statischen bildlichen Kategorisierungsaufgaben eines Nachtests, als Personen, die mit parallelen Key-Frames lernen.

Hypothese 1d: Personen, die mit Animationen lernen, empfinden signifikant weniger Cognitive Load beim Lernen, als Personen, die mit parallelen Key-Frames lernen.

Hypothese 1e: Personen, die mit Animationen lernen, empfinden signifikant weniger Cognitive Load beim Lösen des Nachtests, als Personen, die mit parallelen Key-Frames lernen.

Die Interaktionshypothese zwischen dem Präsentationsformat der Lernbedingung und dem der Nachtestskalen wurde ebenfalls gerichtet formuliert:

Hypothese 2: Es besteht eine signifikante Interaktion zwischen dem Präsentationsformat des Lernmaterials und dem der Nachtestaufgaben, in der Richtung, dass der Lernvorteil durch Animationen bei animierten bildlichen Kategorisierungsaufgaben größer ist als bei statischen bildlichen Kategorisierungsaufgaben.

5.1.2 Methoden

5.1.2.1 Stichprobe

Die Stichprobe der Pilotstudie umfasste 24 Probanden, die Studenten der Universität Landau waren. Von der Fachrichtung her, handelte es sich hierbei um Studenten der Psychologie, der Lehramtsstudiengänge, der Erziehungswissenschaften sowie der Sozial- und Umweltwissenschaften. Eine Person der Stichprobe füllte die Tests eher genervt und unmotiviert aus, was sich auch in auffällig schlechten Werten des Intelligenztests äußerte. Der Proband wurde aus den weiteren Analysen ausgeschlossen, wodurch die statistischen Rechnungen mit 23 Personen durchgeführt wurden. Die Anwerbung der Probanden erfolgte über eine Rundmail an alle Studenten der Universität, die sich über die Lernplattform „OLAT“ zu dem Versuch anmelden konnten. Es gab bei der Ansprache keine vorherigen Restriktionen bezüglich bestimmter Fachrichtungen, es wurden lediglich solche Personen ausgeschlossen, die zu viele Kenntnisse zu Pferden oder dem Thema „Reiten“ hatten, z. B. Leute, die intensiv Reitsport betrieben oder betrieben hatten. Zudem wurde solchen Personen abgesagt, die schon an den Studien über Synapsen teilgenommen hatten, da bei ihnen hinsichtlich des Intelligenztests IST-2000-R ein Transfereffekt hätte bestehen können, da dieser auch in der aktuellen Studie zum Einsatz kam. Das Durchschnittsalter innerhalb der Stichprobe betrug 22,26 Jahre ($SD = 1.94$) bei einer Altersspanne von 20 bis 26 Jahren. Wie an der Universität zu erwarten war, waren Frauen in der Stichprobe sehr stark repräsentiert: Von 23 Teilnehmern waren 18 Personen weiblich und nur 5 männlich. Dies entsprach einem Geschlechterverhältnis von 78,3% Frauen zu 21,7% Männern.

5.1.2.2 Design

Da neben den einfachen Haupteffekten des Faktors „Bedingung“ (Animation vs. Parallele Key-Frames) auf die Leistung in Kategorisierungsaufgaben, auch die Interaktion zwischen dem Lernformat (animiert vs. statisch) und dem Präsentationsformat der Items des Nachtests (animiert vs. statisch) getestet werden sollte, handelte es sich bei der Studie um ein 2 x 2-Design mit Messwiederholung. Die Messwiederholung bestand darin, dass die Teilnehmer im Nachtest sowohl Aufgaben einer animierten, als auch einer statischen Subskala lösen mussten. Desweiteren wurde einfaktoriell der Einfluss der beiden Bedingungen „Animation“ und „Key-Frames parallel“ auf den Cognitive Load der Probanden überprüft, den sie

während der Lerneinheit und während des Nachtests empfanden. Das Thema des Lernmaterials der Studie waren „Die vier Gangarten des Pferdes“. Die Gruppe „Animation“ bekam Lernmaterial gezeigt, bei dem die Pferdegangarten durch Animationen präsentiert wurden. Jede der vier Gangarten wurde jeweils durch eine Animation dargestellt, die aufeinander folgten. Die Gruppe „Key-Frames parallel“ lernte die Pferdebewegungen mit Hilfe von jeweils acht statisch dargebotenen Key-Frames pro Gangart. Die Bilder waren parallel auf dem Bildschirm angeordnet, so dass die Lerner alle Frames simultan und nach eigenen Bedürfnissen betrachten konnten. Ein Interaktivitätsfeature musste nicht eingebaut werden. Eine parallele Anordnung der Key-Frames wurde einer seriellen gegenüber bevorzugt, da somit ein größerer Kontrast zu der Darstellung mit Animationen erreicht wurde. Die Probanden wurden den beiden Bedingungen randomisiert zugewiesen, so dass sich in der Animationsgruppe 12 und in der statischen Bildgruppe 11 Teilnehmer befanden, die in jeweils zwei Messsituationen untersucht wurden.

Innerhalb der beiden Bedingungen „Animation“ und „Bild parallel“ wurde zudem die Abfolge variiert, mit der die Pferdegangarten gezeigt wurden, so dass vier Sequenzen pro Untersuchungsbedingung entstanden. Diese Variation stellte jedoch keinen zusätzlichen Faktor dar. Auch zu diesen Subgruppen wurden die Probanden randomisiert zugewiesen.

Die abhängigen Variablen der Studie waren der Gesamtscore in einem Nachtest zu bildlichen Kategorisierungsaufgaben, die Scores der in ihm enthaltenen Animations- und statischen Bildskala sowie der Cognitive Load der Probanden, der nach der Lerneinheit und nach dem Nachtest gemessen wurde.

Neben den abhängigen Variablen wurden auch Kontrollvariablen im Design berücksichtigt, von denen ein größerer Einfluss auf die Kriterien anzunehmen war. Hierzu zählte das Vorwissen der Teilnehmer über Pferdegangarten, aber auch ihre bisherigen Wahrnehmungserfahrungen mit Pferdebewegungen. Diese wurden beispielsweise durch die Fragen operationalisiert, ob die Teilnehmer viel mit Pferden zu tun haben oder sich häufiger Pferdefilme ansehen. Weitere mögliche Kovariaten waren die aktuelle Lernmotivation der Probanden, sowie ihre verbalen, räumlichen und Merkfähigkeiten.

5.1.2.3 Lernmaterialien

Extraktion der Key-States der einzelnen Pferdegangarten

Innerhalb der Materialentwicklung wurden für die parallele Bildbedingung die Key-States der einzelnen Gangarten: Schritt, Trab, Galopp und Tölt ausgewählt, um die Bewegungen zu repräsentieren. Dies geschah wiederum in Analogie zu den theoretischen Vorüberlegungen, dass es sich bei den Key-States zum einen um mittlere Bilder der Unterereignisse, aber auch um die Anfangs- und Endpunkte dieser Subereignisse, die meist auch kausale Schnittstellen darstellen, handelt.

Biologische Bewegungen wie Pferdegangarten lassen sich generell schwieriger in Teilprozesse gliedern als lineare Vorgänge, da die einzelnen Bewegungskomponenten viel mehr ineinander übergehen (Schnotz & Lowe, 2008). Dennoch war bei genauerer Betrachtung eine Unterteilung der Gangarten in mehrere Phasen möglich, wie sie auch von Reitportlern gemacht wird (Brandl, 1982; <http://www.geysir.com/de/islandpferde/gangarten>). In der aktuellen Studie wurde die Segmentierung der Gangbewegungen anhand des Kriteriums der „maximalen physikalischen Veränderungen“ vorgenommen (Zacks & Tversky, 2001). Ereignisgrenzen befinden sich also dort, wo augenscheinlich größere Bewegungsänderungen innerhalb des Gangzyklus stattfinden. Beispielsweise wäre beim Schritt der Anfang des ersten Unterereignisses dort, wo das linke Vorderbein sich anfängt zu heben und der Abschluss dieses Ereignisses an der Stelle, wo die Hebebewegung aufhört und in eine Senkbewegung übergeht.

Wenn man diese Art der Gliederung konsequent anwandte, ergaben sich für die einzelnen Gangarten die folgenden Unterereignisse bzw. Phasen:

Schritt:

Phase 1: Linkes Vorderbein hebt sich

Phase 2: Linkes Vorderbein senkt sich

Phase 3: Rechtes Hinterbein hebt sich

Phase 4: Rechtes Hinterbein senkt sich

Phase 5: Rechtes Vorderbein hebt sich

Phase 6: Rechtes Vorderbein senkt sich

Phase 7: Linkes Hinterbein hebt sich

Phase 8: Linkes Hinterbein senkt sich

Überlappungen bestehen zwischen den Phasen 2 und 3, 4 und 5, 6 und 7 sowie 8 und 1.

Trab:

Phase 1: Linkes Vorderbein und rechtes Hinterbein heben ab

Phase 2: Linkes Vorderbein und rechtes Hinterbein bewegen sich nach vorne / 1. Hälfte

Phase 3: Linkes Vorderbein und rechtes Hinterbein bewegen sich nach vorne / 2. Hälfte

Phase 4: Linkes Vorderbein und rechtes Hinterbein senken sich

Phase 5: Rechtes Vorderbein und linkes Hinterbein heben ab

Phase 6: Rechtes Vorderbein und linkes Hinterbein bewegen sich nach vorne / 1. Hälfte

Phase 7: Rechtes Vorderbein und linkes Hinterbein bewegen sich nach vorne / 2. Hälfte

Phase 8: Rechtes Vorderbein und linkes Hinterbein senken sich

Der Wendepunkt für die Unterteilung der Phasen 2 und 3 bzw. 6 und 7 in jeweils zwei Hälften war, wenn das Hinterbein auf gleicher Höhe mit dem anderen Hinterbein stand. Überlappungen sind zwischen den Phasen 4 und 5 sowie zwischen den Phasen 8 und 1 vorhanden. Diese stellen in der Reitlehre auch Sprungphasen dar.

Galopp:

Phase 1: Rechtes Hinterbein senkt sich

Phase 2: Linkes Hinterbein und rechtes Vorderbein senken sich

Phase 3: Linkes Vorderbein senkt sich

Phase 4: Rechtes Hinterbein hebt sich

Phase 5: Linkes Hinterbein und rechtes Vorderbein heben sich

Phase 6: Linkes Vorderbein hebt sich

Phase 7: Sprungphase

Eine Überlappung besteht zwischen den Phasen 3 und 4.

Tölt:

Phase 1: Linkes Vorderbein hebt sich

Phase 2: Linkes Vorderbein senkt sich

Phase 3: Rechtes Hinterbein hebt sich

Phase 4: Rechtes Hinterbein senkt sich

Phase 5: Rechtes Vorderbein hebt sich

Phase 6: Rechtes Vorderbein senkt sich

Phase 7: Linkes Hinterbein hebt sich

Phase 8: Linkes Hinterbein senkt sich

Die Phasen 2 und 3, 4 und 5, 6 und 7 sowie 8 und 1 gehen ineinander über.

Ein Vergleich dieser Einteilung anhand des Kriteriums großer perzeptueller Veränderungen mit der Phasenfolge, wie sie in der Reitlehre vorgenommen wird, zeigt, dass die Gliederungen einander ähnlich, aber nicht identisch sind.

Für die Extraktion der Key-Frames galt generell, dass für die Repräsentation der einzelnen Phasen mittlere Bilder genommen wurden, da diese prototypisch und sehr informationsreich sind. Die Anfangs- und Endpunkte der Phasen schienen von geringerer Wichtigkeit zu sein, da diese bei den Pferdegangarten keine kausalen Schnittstellen im eigentlichen Sinne darstellen. Die Abfolge zwischen den Phasen ist eher temporär zu sehen. Daher wurden die Schnittstellen zwischen den Phasen nur an den Stellen repräsentiert, die laut Lehrbuch charakteristische Merkmale der jeweiligen Gangart sind. Diese Übergänge wurden in den meisten Fällen in die mittleren Bilder integriert, um die Zahl der Key-Frames gering zu halten und um beschwerliche Such- und Vergleichsprozesse zu vermeiden. Die Zahl der Key-Frames wurde für jede Gangart konstant (8 Key-Frames) gehalten, um gleiche Lernbedingungen für jede Bewegungsform zu schaffen.

Die folgende Aufzählung stellt kurz die Merkmale dar, nach denen die Key-Frames konkret für die Gangarten jeweils extrahiert wurden:

Schritt:

- Es wurde für jede der acht Phasen ein mittleres Bild verwendet.
- In den mittleren Bildern waren die laut Lehrbuch wichtigen Übergänge zwischen den Phasen 2 und 3, 4 und 5, 6 und 7 sowie 8 und 1 integriert.

Trab:

- Für jede der acht Phasen wurde ein mittleres Bild gewählt.
- Die Überlappungen zwischen den Phasen 4 und 5 sowie 8 und 1 wurden in den entsprechenden mittleren Bildern angedeutet.

Galopp:

- Es wurde für jede der sieben Phasen je ein mittleres Bild verwendet.
- Die Sprungphase wurde durch zwei Frames repräsentiert, da sie konzeptuell wichtig ist und somit die Gesamtzahl von acht Key-Frames für die Gangart erreicht werden konnte.
- Der inhaltlich wichtigere Übergang zwischen den Phasen 3 und 4 wurde in dem zugehörigen mittleren Bild dargestellt.

Tölt:

- Jede der acht Phasen wurde durch ein mittleres Bild repräsentiert. Die Phasen gehen so stark ineinander über, dass hierfür nur vier Key-Frames benötigt wurden.
- Zusätzlich wurden die Anfangsbilder der Hehebewegungen hinzugenommen, da diese charakteristische Merkmale dieser Gangart sind (sog. „Einbeinstützen“, <http://www.geysir.com/de/islandpferde/gangarten>).

Für eine Ansicht der einzelnen Frames, sei auf die Lerndateien mit statischen Bildern im Anhang D1 verwiesen, z.B. auf die Datei „Bild_zuerst_Galopp.swf“.

Beschreibung der Materialien

Das Lernmaterial wurde von der Doktorandin konzeptuell erarbeitet und von einem Graphikdesigner produziert (für eine Ansicht der Lernmaterialien, s. Anhang D1 auf der beiliegenden CD). Hierzu wurde die Software „Macromedia Flash 8“ verwendet, so dass die Dateien im „.swf“-Format als „Shockwave Flash-Objekte“ vorlagen und bei den Untersuchungen mit dem Browser „Mozilla Firefox“ abgespielt werden konnten. Als Zeichenvorlage für die Animations- und die Bilddateien dienten von der Doktorandin und dem Graphikdesigner selbstaufgezeichnete Videos von Pferden aus regionalen Reitställen. Sowohl in der Animations- als auch in der Bildbedingung startete die Lerneinheit mit einer Instruktionssseite, die den Probanden die anstehende Aufgabe erklärte. Rechts unten auf der Seite war ein Pfeil eingebaut, mit dem die Versuchspersonen die Lernphase durch einen Klick mit der Maus starten konnten, wenn sie die Anleitung verstanden hatten. Danach

wurden, je nach Bedingung, die vier Animationen bzw. die vier Bildgruppen zu den Pferdegangarten nacheinander jeweils 20 s lang gezeigt. Die Lernzeit für eine Gangart von 20 s war auf den Ergebnissen informeller empirischer Testungen begründet, bei denen Personen mit Animationen ca. 10 bis 15 s lang lernten, bis sie das Gefühl hatten, sich eine Gangart eingeprägt zu haben. Danach lösten sie Beispielitems zur Kategorisierung der Gangarten mittelmäßig bis gut. Personen, die mit statischen Bildern lernten, benötigten hingegen ca. 30 bis 40 s, bis sie die Bewegungen nachvollziehen konnten und lösten anschließende Aufgaben dann überdurchschnittlich gut. Da die Lernzeiten für beide Bedingungen gleich sein sollten, wurde als Kompromisslösung eine Dauer von 20 s gewählt. Ein solches empirisches Vorgehen zur Ermittlung der Lernzeit wurde als zuverlässiger erachtet als ein einfaches Ableiten aus der Literatur.

Sowohl in der Animations- als auch in der Bildbedingung waren die Pferde in der Lernphase schwarz gefärbt, um einen größtmöglichen Kontrast zum weißen Hintergrund und ein somit besseres Lernen zu gewährleisten. Ebenfalls wurde in beiden Gruppen die Abfolge variiert, mit denen die vier Pferdegangarten jeweils gezeigt wurden, um Sequenzeffekte beim Lernen zu vermeiden. Die erste Reihenfolge lautete: „Schritt – Trab – Galopp – Tölt“, die dann mittels eines „Lateinischen Quadrates“ umgestellt wurde, so dass jede der Gangarten einmal am Anfang und auch einmal am Ende der Lerneinheit stand. Auf diese Weise wurden jeweils vier verschiedene Animations- und Bilddateien erzeugt.

Vor jeder Animation bzw. vor jeder Bildgruppe der beiden Lerneinheiten wurde eine Schrift mit dem jeweiligen Gangartnamen in der Mitte des Bildschirms ca. 4 s lang eingeblendet, damit die Probanden im Nachtest die Bewegungen benennen konnten. Nach jeder Gangart konnten die Teilnehmer selbst entscheiden, ob sie eine Pause brauchten, bevor sie die nächste Gangart sahen. Daher wurde nach jeder Animation bzw. nach jeder Bildgruppe jeweils der Satz: „Klicken Sie bitte mit der Maus auf den Pfeil, wenn Sie bereit sind, die nächste Pferdegangart zu lernen“ eingeblendet. Auf derselben Seite befand sich rechts unten ein Pfeil, auf den die Teilnehmer mit der Maus klicken konnten, damit die nächste Gangart erschien. Nach der letzten Animation oder Bildgruppe wurde der Satz „Bitte wenden Sie sich an den Versuchsleiter“ gezeigt, was eine kurze Pause für die meisten Probanden bedeutete. Sie wurden dann von der Versuchsleiterin instruiert, die Items zum Cognitive Load im Testheft zu beantworten. Auf der letzten Seite der Lerneinheit war auch rechts unten wieder ein Pfeil zu sehen, durch den die Teilnehmer zur Instruktion des Nachtests gelangen

konnten. Die Aktivität des Pfeils war allerdings 10 s lang gesperrt, damit die Versuchspersonen nicht aus Versehen weiterklicken konnten.

Bei den Animationen der Lerneinheit wurde bei ihrer Zeichnung darauf geachtet, dass bei allen Gangarten die Merkmale der Pferde, wie z.B. ihr Schweif oder ihre Mähne, sehr ähnlich waren, damit die Probanden die Gangarten nur an den Beinen und nicht an einem anderen Detail voneinander unterscheiden konnten. Aus demselben Grund waren auch die Geschwindigkeiten der Gangarten annähernd, soweit technisch möglich, miteinander vergleichbar, da sie ein wichtiger diagnostischer Hinweis für die Unterscheidung der Gangarten sind. Spätere Diskriminationen wären bei Hinzunahme dieser Information zu einfach gewesen. Die Geschwindigkeiten befanden sich ungefähr auf mittlerem Niveau. Die folgenden Abbildungen 5.1.1 bis 5.1.4 zeigen Beispielbilder aus den vier Animationen der Gangarten:

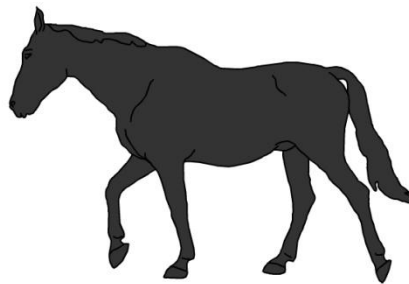


Abbildung 5.1.1: Beispielbild für die Animation der Gangart „Schritt“

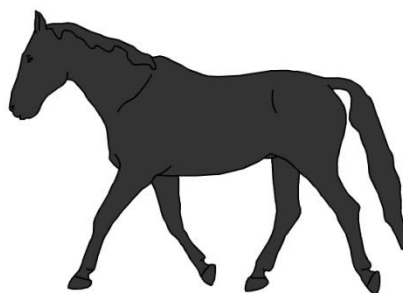


Abbildung 5.1.2: Beispielbild für die Animation der Gangart „Trab“

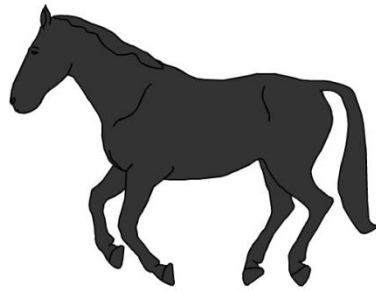


Abbildung 5.1.3: Beispielbild für die Animation der Gangart „Galopp“

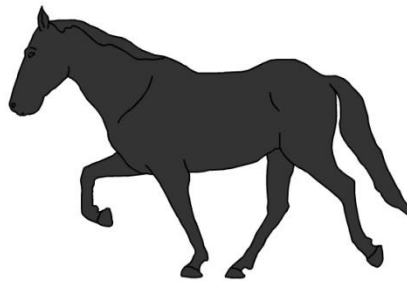


Abbildung 5.1.4: Beispielbild für die Animation der Gangart „Tölt“

Bei der statischen Bildbedingung wurden die Key-Frames aus der jeweiligen Animation der entsprechenden Gangart extrahiert, so dass es sich immer um das gleiche Pferd mit denselben Merkmalen handelte. Für die Darstellung jeder Gangart wurden acht charakteristische Key-Frames ausgewählt und in zwei Spalten mit je vier Bildern auf dem Bildschirm angeordnet. Die Frames waren durchnummeriert und hatten ihren Anfang im linken oberen Bild. Die Darstellung der Frames war somit parallel bzw. simultan, da alle Bilder gleichzeitig zu sehen waren, so dass der Lerner nach eigenen Bedürfnissen entscheiden konnte, welchen Bewegungszustand er näher betrachten mochte. Ein Interaktivitätsfeature war bei dieser Anordnung nicht erforderlich. Durch die Präsentation der Bilder in zwei Spalten, musste das Auge des Betrachters nur an zwei Stellen (von Frame 4 zu 5 bzw. von Frame 8 zu 1) größere Blickbewegungen vollziehen, ansonsten lagen die Frames nah beieinander. Abbildung 5.1.5 verdeutlicht die Darstellung der statischen Bilder nochmals graphisch am Beispiel „Schritt“.

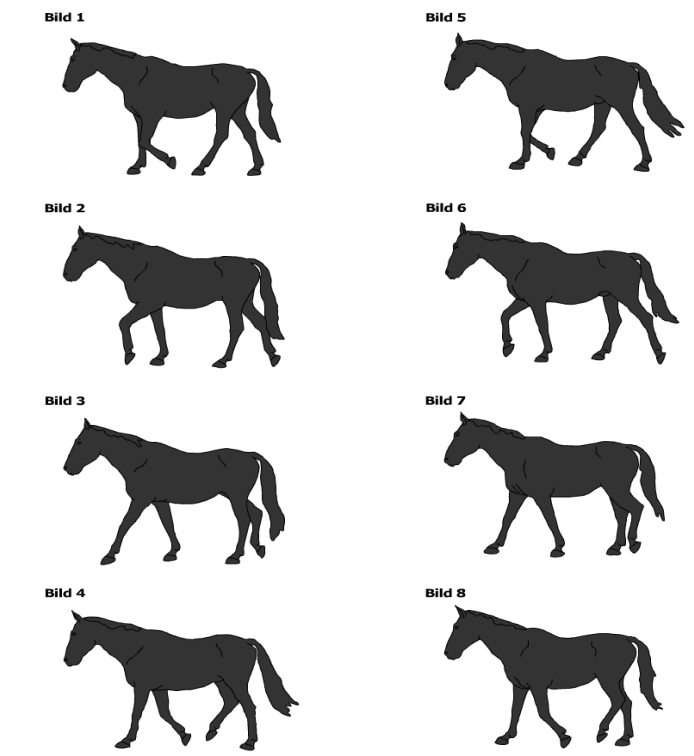


Abbildung 5.1.5: Anordnung der Key-Frames im Lernmaterial am Beispiel ‚Schritt‘

5.1.2.4 Messinstrumente

Die abhängigen Variablen der Pilotstudie waren, wie oben bereits genannt, der Gesamtscore eines Nachtests mit bildlichen Kategorisierungsaufgaben, die Scores der in ihm enthaltenen Animations- und statischen Bildskala sowie der Cognitive Load, der nach der Lerneinheit und nach dem Nachtest gemessen wurde.

Die Kontrollvariablen waren die Wahrnehmungsvorerfahrungen mit Pferdebewegungen, das Vorwissen zu Pferdegangarten, die aktuelle Lernmotivation sowie die verbalen, räumlichen und Merkfähigkeiten der Probanden.

Die zur Erfassung dieser Variablen benötigten Testverfahren werden im Folgenden in der Abfolge vorgestellt, in der sie von den Teilnehmern in der Studie bearbeitet wurden. Im Allgemeinen wurde eine zu den Synapsenstudien möglichst analoge Reihenfolge der Verfahren eingehalten, um eine bessere Vergleichbarkeit zwischen den Studien zu schaffen. Die standardisierten Verfahren (z.B. der FAM, Rheinberg et al., 2001) waren ebenfalls die

gleichen wie in den Studien zu den Inferenzaufgaben und sollen deshalb hier nur noch kurz beschrieben werden.

Soziodemographischer Fragebogen

Da sich der soziodemographische Fragebogen, der für die Studien zu den Inferenzaufgaben konstruiert worden war, in den dortigen Experimenten bewährt hatte, wurde er in der gleichen Form, in der er in der Hauptstudie letztlich vorgelegen hatte, auch in der Pilotstudie zu den Kategorisierungsaufgaben verwendet (s. Anhang C2 - 1). Die Probanden wurden nach ihrem Alter, Geschlecht, Schulabschluss und Beruf befragt. Studenten, auf die sich die Erhebungen konzentrierten, sollten weiterhin ihr Studienfach und ihre Semesterzahl angeben. Um Anonymität und eine spätere Zuordnung der Testmaterialien zu gewährleisten, konnten die Probanden einen achtstelligen Code auf dem Bogen kreieren.

Test zur Erfassung der Wahrnehmungsvorerfahrungen mit Pferdebewegungen

Der Test bestand aus elf Items, die nach den bisherigen Wahrnehmungserfahrungen der Probanden mit Pferden in Realität oder in verschiedenen Medien fragten (zur Ansicht des Tests, s. Anhang D2 - 1). Die Begründung zur Erfassung dieser Variable lag darin, dass angenommen wurde, dass Personen, die in Natur häufiger Pferde sehen oder sie sich in Medien, z.B. in Filmen, betrachten, eher dynamische Schemata zu den Pferdegangarten schon aufgebaut haben bzw. dass diese Schemata schon elaborierter und leichter abrufbar sind. Dies könnte die Qualität des Lernprozesses und die Ergebnisse des Nachtests mit beeinflussen.

Ein Beispielitem dazu, wie in der aktuellen Studie nach den Wahrnehmungserfahrungen mit Pferden in Natur gefragt wurde, war: „Haben oder hatten Sie früher häufiger mit Pferden zu tun?“, während ein Item zur Erfassung von Wahrnehmungserfahrungen in Medien beispielsweise: „Sehen oder sahen Sie sich früher im Fernsehen öfters Filme und Serien mit Pferden an?“ lautete. Diese Items wurden den Teilnehmern in gemischter Reihenfolge vorgegeben. Die Fragen waren offen und persönlich formuliert und konnten mit „Ja“ oder „Nein“ beantwortet werden. Die Probanden mussten hierzu das entsprechende Kästchen jeweils ankreuzen (Multiple Choice-Format). Es handelte sich somit um ein dichotomes Antwortformat. Die Fragen bezogen sich sowohl auf gegenwärtige als auch auf vergangene

Erfahrungen mit Pferden, da angenommen wurde, dass Personen, die zumindest in früheren Jahren mit Pferden häufiger in Berührung gekommen waren, eher dynamische Schemata aufgebaut haben als solche, die nie mit Pferden näher zu tun hatten.

Für die statistischen Analysen wurde jede „Ja“-Antwort mit „1“ und jede „Nein“-Antwort mit „0“ gewertet. Die einzelnen Werte wurden aufsummiert, so dass ein Gesamtscore gebildet wurde, der als eine mögliche Kovariate in den weiteren Rechnungen berücksichtigt wurde.

Vorwissenstest

Der Vorwissenstest sollte erfassen, wie viel die Probanden vorab konzeptuell über Pferdegangarten wussten (zur Ansicht des Tests, s. Anhang D2 - 2). Auch hier wurde davon ausgegangen, dass solche Kenntnisse die Ergebnisse des späteren Nachtests mit bestimmen können und dass ihre Ausprägungen bei den Probanden daher kontrolliert werden sollten. Es mochte gut möglich sein, dass das Vorwissen mit den Wahrnehmungsvorerfahrungen korreliert war, da Personen, die viel über Pferde wissen auch oft mehr mit Pferden zu tun haben oder hatten.

Die Vorwissensskala bestand aus zwei Items, die ein halboffenes Antwortformat aufwiesen. Die Probanden sollten ihre Antworten schriftlich in ganzen Sätzen oder in Stichworten festhalten. Es handelte sich um einen Power-Test, was bedeutete, dass die Versuchspersonen keine Zeitbegrenzung hatten, um die Fragen zu beantworten. Sie wurden zudem angewiesen, bei den Items, die sie nicht lösen konnten, ein „?“ oder ein „/“ zu machen.

Die erste Frage des Vorwissenstests wies ein konzeptuell eher niedrigeres Schwierigkeitsniveau auf und konnte teilweise aus dem Alltagswissen heraus beantwortet werden. Die Versuchspersonen sollten drei Pferdegangarten aufzählen (*Item*: „Nennen Sie drei Pferdegangarten.“). Die zweite Frage war etwas schwieriger gestellt, hier sollten die Probanden eine der genannten Gangarten beschreiben (*Item*: „Beschreiben Sie eine der von Ihnen eben genannten Pferdegangarten.“). Es wurden keine weiteren Items hinzugefügt, da konzeptuell genauere Fragen wahrscheinlich zu schwierig für Laien gewesen wären und Ratenendenzen provoziert hätten. Zudem konnten keine Itemstämme verwendet werden, in denen schon Gangartnamen enthalten waren, da dies eine Hilfe zur Beantwortung der ersten Frage gewesen wäre.

Die Antworten auf die Fragen wurden mittels eigens konstruierter Ratingsysteme in „richtig = 2“, „zum Teil richtig = 1“ und „falsch / nicht beantwortet = 0“ unterteilt. Die Kodierung war somit dreistufig und wurde als metrisch angenommen. Die Kategorie „richtig“ enthielt jeweils eine Musterantwort, die mit Hilfe von Reitlehrbüchern und dem Internet (z.B. Brandl, 1982; <http://www.geysir.com/de/islandpferde/gangarten>) erstellt worden war. Unter die Kategorie „zum Teil richtig“ wurden richtige, aber unvollständige Antworten gezählt. Alle anderen Lösungen galten als „falsch“. Jede Kategorie umfasste zudem Ankerbeispiele und Kodierregeln, damit die Lösungen trennschärfer eingeordnet werden konnten. Die Ratingsysteme sollten in der Pilotstudie erprobt und durch die dort erhaltenen Daten empirisch weiterentwickelt werden (zur Ansicht der nach der Pilotstudie weiterentwickelten Ratingsysteme, s. Anhang D3 auf der beigefügten CD).

Auf eine Berechnung der Intraclass-Korrelationen der Ratingsysteme wurde in der Pilotstudie noch verzichtet, da die Stichprobengröße hierfür zu gering war. Die erzielten Punkte auf den beiden Items wurden für jede Person aufsummiert und stellten den Gesamtscore des Vorwissens dar.

Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen (FAM)
(Rheinberg et al., 2001)

In Analogie zur ersten Studie der vorliegenden Dissertation, die Animationen und Bilder bezüglich der Lösung von Inferenzaufgaben miteinander verglich, wurde auch bei diesem Experiment zu den Kategorisierungsaufgaben davon ausgegangen, dass die aktuelle Lernmotivation die Lernleistungen der Probanden beeinflusst, da diese generell nicht nur von kognitiven Fähigkeiten, sondern auch von Motivationsfaktoren determiniert werden (Atkinson, 1974). Ebenso wie in der vorherigen Studie zu den Inferenzaufgaben, wurde auch hier die Motivation der Probanden mit Hilfe des standardisierten Verfahrens „FAM“ (Rheinberg et al., 2001) bestimmt (s. auch Anhang B2 - 3). Für eine Beschreibung seiner Gütekriterien sei daher auf den Abschnitt 4.2.2.4 bzw. auf die entsprechende Literatur verwiesen (s. Rheinberg et al., 2001). In der Pilotstudie wurde der FAM den Probanden nach dem Lesen einer Instruktion vorgelegt, in der sie erfuhren, um welche Art von Lernaufgabe es sich handelte. Die Summenscores der vier Subskalen des Tests: „Misserfolgsbefürchtung“, „Erfolgswahrscheinlichkeit“, „Interesse“ und „Herausforderung“, wurden in den statistischen Analysen als weitere potentielle Kovariaten angesehen.

Erfassung des Cognitive Loads

Ebenfalls in Analogie zu der Studie zu Inferenzaufgaben, wurde auch in der Studie zu Kategorisierungsaufgaben der Cognitive Load der Probanden erhoben. Auch hier fand die Erfassung des Cognitive Loads zu zwei Messzeitpunkten statt. Im ersten Fall sollten die Probanden beurteilen, wie schwierig sie die Lernaufgabe fanden, beim zweiten Mal ging es hingegen um die Einschätzung der mentalen Belastung, die der Nachtest bei ihnen verursacht hat. Wie auch bei den Studien zu Inferenzaufgaben, wurden zwei Items verwendet, um den Cognitive Load zu bestimmen (zur Ansicht der Items, s. Anhang D2 - 3). Das erste Item fragte nach der wahrgenommenen Aufgabenschwierigkeit und war an die Vorlage von Bratfisch, Borg und Dornic (1972) angelehnt. Das zweite Item stellte eine Adaptation der „Cognitive load rating scale“ von Paas (1992) dar. Hier sollten die Teilnehmer angeben, wie stark sie sich beim Bearbeiten der Lernaufgabe bzw. des Nachtests angestrengt haben. Beide Items hatten ein neunstufiges Antwortformat. Aus Redundanzgründen wird an dieser Stelle auf eine genauere Beschreibung der Items und ihrer Gütekriterien verzichtet, und es sei diesbezüglich auf Abschnitt 4.2.2.4 der ersten Studie dieser Dissertation und auf die entsprechende Literatur verwiesen (Paas, van Merriënboer & Adam, 1994).

Für die statistische Auswertung wurden die beiden von den Probanden angekreuzten Werte auf den Items jeweils addiert, so dass zwei Summenscores entstanden, die die mentale Belastung angaben, die zum einen durch die Lerneinheit und zum anderen durch den Nachtest verursacht wurde. Diese Scores stellten neben den Leistungsindizes zwei weitere abhängige Variablen der Studie dar, die Hinweise zur „instruktionalen Effizienz“ (Paas et al., 2007) der beiden Präsentationsarten (Animation vs. Bild parallel) gaben und somit eine Ergänzung zu den objektiven Leistungsmaßen waren.

Nachtest

Sowohl der Animations- wie auch der Bildgruppe wurde der gleiche Nachtest vorgelegt (s. Anhang D2 – 4 und Anhang D1 auf der beiliegenden CD, der Nachtest folgt in den Lerndateien direkt auf das Lernmaterial). Der Test war in seiner Darbietungsform zweigeteilt: Zum einen sahen die Probanden am Computer Animationen oder statische Bilder der Gangarten, die sie gerade gelernt haben, und sollten anschließend jeweils

beurteilen, um welche Gangart es sich dabei handelte. Ihre Antworten sollten sie in einem Testheft geben. Dadurch wurde der Ablauf des Versuchs unabhängiger von den technischen Gegebenheiten. Der Nachtest war somit sowohl computerbasiert als auch als Paper-Pencil-Verfahren konzipiert.

Am Anfang des Tests stand zunächst wieder eine Instruktion, zu der die Probanden von der letzten Seite der Lerneinheit aus durch einen Klick mit der Maus hin gelangen konnten. In der Instruktion wurde der Nachtest den Probanden kurz erklärt. Nachdem die Versuchsleiterin das Prinzip des Antwortbogens in der Durchführung nochmals kurz mündlich erläutert hatte, konnten die Versuchspersonen den Test mit einem Klick mit der Maus auf einen Pfeil starten.

Der Test bestand aus 32 Items, davon waren 16 Items Animationen und 16 Items statische Bildgruppen, die jeweils eine der vier Gangarten zeigten. Vor jeder Aufgabe wurde ihre jeweilige Nummer ca. 3 s lang eingeblendet, damit die Probanden bei der späteren Beantwortung eine bessere Orientierung im Testheft hatten. Nach jeder Aufgabe wurden die Probanden aufgefordert, im Testheft anzukreuzen, welche Gangart sie gerade gesehen haben. Durch einen Klick mit der Maus auf einen Pfeil auf derselben Seite konnten sie zu der nächsten Aufgabe gelangen. Dieser Pfeil war stets zunächst 8 s lang gesperrt, damit die Teilnehmer nicht aus Versehen gleich weiterklicken konnten, ohne das Item zuvor beantwortet zu haben. Für die Lösung der Aufgaben hatten die Probanden so viel Zeit, wie sie benötigten (Power-Test). Nach der letzten Aufgabe wurden die Teilnehmer aufgefordert, sich an die Versuchsleiterin zu wenden, da dann die Bearbeitung der Items zum Cognitive Load erfolgte.

Die Animations- sowie die Bilditems hatten jeweils vier Schwierigkeitsstufen: Bei den Animationsaufgaben wurde die Schwierigkeit über die Präsentationszeit variiert. Auf der einfachsten Stufe sahen die Probanden die Aufgaben 10 s lang, dann nur noch 75% der Zeit (7,5 s), dann 50% (5 s) und letztlich noch 25% der anfänglichen Darbietungszeit (2,5 s). Die Bilditems hingegen wurden immer 10 s lang gezeigt, ihre Schwierigkeit wurde durch die Zahl der präsentierten Key-Frames verändert. Im einfachsten Fall sahen die Probanden vollständig die acht Key-Frames der jeweiligen Gangart, mit denen sie auch in der Lerneinheit konfrontiert waren. Die Anordnung der Bilder war die gleiche wie die in der Lernphase. Auf den drei schwierigeren Stufen wurden jeweils zwei der Key-Frames weggenommen, so dass nur noch sechs, vier oder zwei der Frames zu sehen waren, wobei die Positionen der Bilder konstant blieben. Bei der Elimination der Key-Frames wurde

darauf geachtet, dass die charakteristischen Zustände einer Gangart auch auf den schwierigeren Niveaus bestehen blieben (beispielsweise die Sprungphase beim „Galopp“), so dass die Aufgaben lösbar blieben. Abbildung 5.1.6 zeigt ein Beispiel für die Bilditem der dritten Schwierigkeitsstufe mit vier Key-Frames des Nachttests.

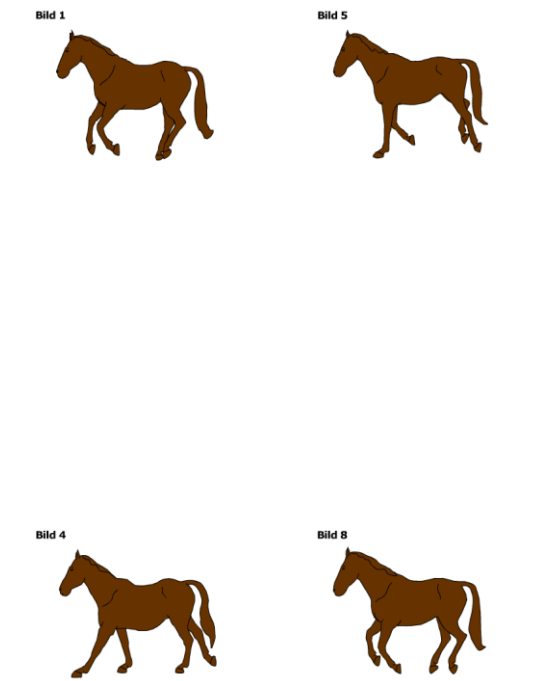


Abbildung 5.1.6: Beispiel für ein Bilditem mit vier Key-Frames anhand der Gangart „Galopp“

Die Herleitung der Präsentationszeit von 10 s auf der einfachsten Stufe bei den Animationsitems geschah wiederum aus empirischen Daten heraus, die aus informellen Testungen stammten, was als ein zuverlässigeres Vorgehen angesehen wurde als wenn die Testzeiten aus der Literatur heraus bestimmt worden wären. In den informellen Testungen benötigten die Teilnehmer 8 bis 9 s, um Aufgaben mit Animationen mittelmäßig bis gut beantworten zu können. Sie brauchten hingegen 12 bis 13 s, um Gangarten, die mit statischen Bildern dargestellt wurden, mittelmäßig bis gut voneinander unterscheiden zu können. Eine Testzeit von 10 s war somit ein Kompromiss zwischen diesen beiden Bearbeitungszeiten.

Die Pferde der Testphase waren braun oder weiß angefärbt. Dadurch sollte vermieden werden, dass die Probanden die Gangarten durch reine Gedächtniseffekte aus der Lernphase heraus erkennen konnten. Dort hatten die Pferde eine schwarze Färbung. Ansonsten waren

die Merkmale der Pferde, wie beispielsweise der Schweif oder die Mähne, bei dem Nachtest und bei der Lerneinheit gleich gezeichnet.

Da auf jeder der vier Schwierigkeitsstufen jede Gangart einmal animiert und einmal statisch abgefragt werden sollte, ergaben sich somit acht Items für jedes Schwierigkeitsniveau. Ebenfalls sollten die Pferde bei jeder Gangart eines Schwierigkeitszyklus sowohl braun als auch weiß gezeigt werden, so dass eine Kombinatorik zwischen der Farbe der Pferde und ihrer Präsentationsform entstand. Der erste Zyklus des Nachtests bestand aus den konzeptuell acht einfachsten Items, der zweite Zyklus aus den Aufgaben des zweiteinfachsten Niveaus, usw., wobei die Schwierigkeitsstufen etwas durchmischt wurden, damit die Probanden keine Systematik in der Abfolge der Items erkennen konnten. Insgesamt gesehen, wurde der Test in seinem Verlauf immer schwieriger. Bei der Testkonstruktion wurde bei der Aufgabensequenz ebenfalls darauf geachtet, dass nicht die gleiche Gangart zwei Mal hintereinander gezeigt wurde. Aufgrund dieser Überlegungen ergab sich das folgende Testschema:

Zyklus 1 (100 + 75 %)	Zyklus 2 (100 + 75 %)	Zyklus 3 (50 + 25 %)	Zyklus 4 (50 + 25 %)
a s s a a s s a	a s a s a s a s	a s s a a s s a	a s a s a s a s
b b w w b w w b	b w w b w b w b	b b w w w b b w	b w w b b w b w

Anmerkungen. a = Animationsitem, s = Bilditem, b = braun gefärbtes Pferd, w = weiß gefärbtes Pferd

Abbildung 5.1.7: Die Itemabfolge des Nachtests in der Pilotstudie

Im Antwortbogen des Nachtests wurden für jede der 32 Aufgaben die vier Gangarten als Antwortalternativen im Multiple Choice-Format vorgegeben, so dass die Probanden die richtige Lösung jeweils ankreuzen konnten. Um Sequenzeffekte bei der Beantwortung zu vermeiden, wurde die Reihenfolge der Gangarten bei den Antwortalternativen variiert. In der späteren Durchführung wurden die Probanden zudem gebeten, nicht mehr zurückzublättern, wenn sie eine Seite bearbeitet hatten.

Für die Auswertung wurde jede richtige Antwort des Nachtests mit einem Punkt gewertet, für eine falsche Lösung wurde hingegen kein Punkt vergeben. Jede nicht beantwortete Aufgabe galt als „missing value.“ Somit hatten die Lösungen der einzelnen Items nominales Skalenniveau („richtig“ / „falsch“). Die Summe der bei den einzelnen Items erzielten Punkte stellte den Gesamtscore des Nachtests dar, der als metrische Variable angenommen wurde.

Die Scores der Subskalen „Animation“ und „Bild“ ergaben sich jeweils dadurch, indem man lediglich die einzelnen Punkte der Animationsaufgaben sowie die der statischen Bildaufgaben getrennt voneinander aufsummierte. Die Animationsskala bestand aus den Items 1, 4, 5, 8, 9, 11, 13, 15, 17, 20, 21, 24, 25, 27, 29 und 31, die Bildskala umfasste hingegen die Items 2, 3, 6, 7, 10, 12, 14, 16, 18, 19, 22, 23, 26, 28, 30 und 32. Die Scores der Animations- sowie der Bildskala waren zwei weitere abhängige Variablen des Experimentes.

IST-2000-R (Liepmann et al., 2007)

In Analogie zu der ersten Studie der vorliegenden Dissertation wurden auch hier die räumlichen, verbalen und die Merkfähigkeiten der Probanden erhoben, da davon ausgegangen wurde, dass diese Intelligenzkomponenten Lernprozesse generell und speziell auch solche mit Animationen und statischen Bildern beeinflussen (Hegarty & Kriz, 2008). Ebenso wie in den Experimenten zu den Inferenzaufgaben, wurde auch in dieser Studie der „IST-2000-R“ (Liepmann et al., 2007) eingesetzt. Die verwendeten Subskalen waren abermals die „Figurenauswahl“ (FA) und die „Würfelaufgaben“ (WÜ) zur Bestimmung der räumlichen Fähigkeiten, die „Analogien“ (AN) und die „Gemeinsamkeiten“ (GE) zur Messung der sprachlichen Intelligenz sowie die Aufgabengruppen „Merkfähigkeit verbal“ und „Merkfähigkeit figural“ zur Erhebung der Gedächtnisfähigkeiten der Probanden. Für eine genauere Beschreibung dieses standardisierten Verfahrens und für eine Ansicht der Aufgabengruppen sei auf den Abschnitt 4.2.2.4 der vorliegenden Dissertation und auf das Testmanual sowie auf das Testheft (Grundmodul, Form A) von Liepmann et al. (2007) verwiesen.

Für die Auswertung wurden die Punktzahlen der beiden Subskalen, die sich auf den gleichen Intelligenzbereich bezogen, jeweils addiert, so dass drei Gesamtscores für figurale und verbale Intelligenz sowie für Gedächtnisleistungen generiert wurden. Diese Scores wurden als mögliche Kovariaten in den weiteren Analysen berücksichtigt.

Interview

Am Ende der Experimente wurde mit den Probanden jeweils ein kurzes, halbstandardisiertes Interview geführt, um ihre subjektiven Meinungen zur Verbesserung des Versuchs und der Testmaterialien zu verwenden. Für das Interview wurde vorher ein Leitfaden mit vier Fragen erstellt. Dies diente dazu, das Gespräch besser zu strukturieren, wobei aber nach Bedarf auch spontan durch die Versuchsleiterin weiter nachgefragt werden konnte. Die erste Frage des Interviews zielte darauf ab, zu erfahren, an was die Probanden die Pferdegangarten im Nachtest jeweils erkannt hatten (für den genauen Wortlaut der Fragen, s. Anhang D2 - 5). Dies war deshalb notwendig zu fragen, um ausschließen zu können, dass die Probanden die Gangarten an anderen Merkmalen als an den Beinen (z.B. an der Geschwindigkeit der Pferde) erkannten, was bei der Interpretation der Ergebnisse sonst hätte berücksichtigt werden müssen. Bei der zweiten Frage konnten die Teilnehmer den Versuch insgesamt beurteilen. Bei der dritten Frage konnten sie angeben, ob und wenn ja, an welcher Stelle des Versuchs sie besondere Schwierigkeiten hatten. Auf die vierte Frage hin konnten sie schließlich noch Verbesserungsvorschläge hinsichtlich der Versuchsgestaltung machen.

Die Antworten der Teilnehmer wurden stichwortartig mit protokolliert. Bei Gruppentestungen wurden die Probanden gebeten, möglichst eigenständig und unabhängig voneinander zu antworten.

Da es sich lediglich um ein Interview zur Verbesserung der Materialien handelte, wurde auf eine Auswertung durch aufwändigere Verfahren, wie beispielsweise durch Kategoriensysteme, verzichtet. Häufig genannte Aussagen wurden zusammengefasst und bei der Weiterentwicklung der einzelnen Verfahren für die Hauptstudie berücksichtigt.

5.1.2.5 Durchführung

Die Untersuchungen der Pilotstudie wurden im Sommersemester 2011 innerhalb eines Zeitraums von zwei Wochen durchgeführt. Die Versuchsleitung wurde bei allen Experimenten von der Doktorandin selbst übernommen, um eventuell auftretende Schwierigkeiten bei der Durchführung erkennen zu können. Der Versuchsablauf war zuvor in einem Schema schriftlich festgehalten worden, was zu einer höheren Durchführungsobjektivität und Standardisierung der Experimente führte. Auch die räumlichen Bedingungen waren für alle Versuchspersonen gleich, als Erhebungsort diente das Labor der Abteilung „Allgemeine und Pädagogische Psychologie“ der Universität

Landau. Die Versuchspersonen wurden durch eine Rundmail angeworben, die an alle Studenten der Universität ging. Die Anmeldung zur Teilnahme erfolgte über die Lernplattform „OLAT“. Es wurden keine Personen bestimmter Fachrichtungen vorab ausgeschlossen, lediglich Personen mit zu viel Vorwissen über das Thema „Pferde“ oder „Reitsport“ wurden gebeten, sich nicht zur Teilnahme zu melden. Die Experimente wurden in Gruppentestungen mit bis zu vier Personen durchgeführt. Im Vorfeld der Untersuchungen wurde die Technik überprüft und sichergestellt, dass alle Lerndateien an den verwendeten Computern vergleichbar abgespielt wurden.

Zu Beginn der Experimente wurden die Versuchspersonen begrüßt, und es wurde ihnen kurz der Ablauf und die Dauer des Versuchs erklärt. Danach füllten sie den soziodemographischen Fragebogen aus, wobei ihnen die Anonymität ihrer Daten garantiert wurde. Die Probanden wurden generell gebeten, nach dem Ausfüllen eines Tests so lange zu warten, bis die letzte Person mit der Bearbeitung fertig war und dann erst im Testheft weiterzublättern. So wurde ein simultanes Ausfüllen der Fragebögen sichergestellt. Nach der Angabe ihrer persönlichen Daten bearbeiteten die Teilnehmer den Wahrnehmungsvorerfahrungstest zu Pferdebewegungen sowie den anschließenden Vorwissenstest. Bei letzterem wurden die Probanden darauf hingewiesen, ein „?“ oder ein „/“ bei Fragen zu machen, die sie nicht beantworten konnten, und es wurde betont, dass es nicht schlimm sei, wenn sie zu der Thematik noch nicht viel wüssten. Für diese ersten beiden Tests gab es keine Zeitbegrenzungen. Daraufhin lasen die Probanden die Instruktion im Testheft, in der ihnen der Versuch am Computer erklärt wurde, und darauf aufbauend konnten sie den FAM ausfüllen. Die Teilnehmer wurden sodann einem Computer zugewiesen, der ihrer jeweiligen Versuchsbedingung entsprach. Die Zuteilung zu den Versuchsbedingungen erfolgte randomisiert. Nachdem die Probanden am Computer nochmals eine kurze Instruktion sahen, konnten sie den Versuch selbständig starten und die Pferdebewegungen mittels Animationen oder statischen Bildern lernen. Diese Lernphase dauerte ca. 2 min. Im Anschluss sollten sie im Testheft auf der Skala zur Bestimmung des Cognitive Loads beurteilen, wie anstrengend und wie schwierig das Lernen für sie war. Die folgende Testphase startete auch wieder mit einer Instruktion, die von der Versuchsleiterin ergänzend mündlich erklärt wurde. Die Versuchspersonen konnten dann den Test wiederum selbständig starten und die Aufgaben nacheinander bearbeiten. Dabei sollten sie im Testheft nicht mehr zurückblättern, wenn sie die Items einer Seite gelöst hatten. Auch hier wurde im Nachhinein der Cognitive Load gemessen, den die Testphase verursacht hat. Zuletzt wurden

den Teilnehmern sechs Subskalen des IST-2000-R vorgelegt, deren Bearbeitung ca. 40 min in Anspruch nahm. Den Versuchspersonen wurde gesagt, dass dieser Test „räumliche und sprachliche Fähigkeiten“ messe, das Wort „Intelligenztest“ wurde bewusst vermieden, um nicht zu viel Druck in der Untersuchungssituation zu erzeugen. Zudem wurde der Aufbau des Tests und die jeweiligen Instruktionen nochmals erklärt. Bei dem IST-2000-R handelte es sich um einen Speed-Test, das heißt, die Probanden hatten nur begrenzt Zeit, die Aufgaben zu lösen. In einem kurzen, halbstandardisierten Interview konnten die Teilnehmer ihre Meinung zu dem Versuch äußern und Verbesserungsvorschläge machen. Die Probanden wurden in den Gruppentestungen darum gebeten, die Aussagen der jeweils anderen Teilnehmer möglichst zu ignorieren und ihre eigene, selbständige Meinung kundzutun. Danach wurde ihnen ihre Belohnung in Form von 10 Euro oder 1,5 Versuchspersonenstunden ausgehändigt. Die Dauer der Experimente wurde mit einer Stoppuhr jeweils mit erfasst, um eine Zeitschätzung für die Hauptstudie zu erhalten. Die durchschnittliche Bearbeitungszeit lag bei 1 h 09 min ($SD = 2.38$ min, $Range = 1$ h 05 min – 1 h 12 min). Die Zeiten waren sehr homogen, da die Tests im Großen und Ganzen zeitlich vorstrukturiert waren. Störungen, die sich während der Durchführung ergaben, wurden mit notiert, und ihre Tragweite für eventuelle Verzerrungen der Ergebnisse wurde mit berücksichtigt. Generell kamen eher kleinere Störungen vor, wie dass beispielsweise die Taskleiste beim Lernen am Computer herunterklappte oder die Jalousien des Labors automatisch herunterfahren, was aber objektiv und subjektiv, die Teilnehmer wurden anschließend dazu befragt, keine Beeinträchtigung darstellte. Lediglich eine Versuchsperson füllte die Tests offensichtlich unmotiviert und genervt aus, was sich auch in auffällig schlechten Werten des IST-2000-R äußerte, wodurch die Person aus den weiteren Analysen ausgeschlossen wurde. Im Allgemeinen war der Ablauf der Experimente eher problemlos, zu den einzelnen Verfahren gab es kaum Nachfragen, was auf eine gute Verständlichkeit der Aufgaben und Instruktionen hindeutete.

5.1.3 Ergebnisse

Die statistischen Analysen der Pilotstudie wurden mit der Computersoftware „SPSS, Version 19.0“ durchgeführt.

Beobachtung der Durchführbarkeit

Forschungsfrage 1: Ist der Versuchsablauf stimmig und das Experiment in der geplanten Form durchführbar?

Im Großen und Ganzen verliefen die Experimente der Pilotstudie relativ problemlos. Dies wurde auch an den Äußerungen der Probanden im Interview deutlich, in dem mehr als die Hälfte von ihnen positive Kommentare über den Versuch abgaben (der Versuch war „gut“, „interessant“, „angemessen“, etc.) und auch fanden, dass man „nichts besser machen könnte.“ Es kamen kaum Fragen zu den einzelnen Verfahren, was auf eine gute Verständlichkeit der Aufgaben und der Instruktionen hindeutete. Während der Durchführung tauchten lediglich kleinere Probleme auf, wie dass beispielsweise die Taskleiste im Lernprogramm öfters herunterklappte oder dass die Pausen innerhalb des Nachtests zu lange geschaltet waren, was auch von den Probanden im Interview geäußert wurde. Der Nachtest zu den Pferdegangarten wurde häufig als „nicht so schwer“ empfunden, manche Versuchspersonen erwähnten einen Übungseffekt innerhalb des Tests, während der IST-2000-R häufiger als „schwer“ bewertet wurde. Die durchschnittliche Versuchszeit belief sich auf $M = 1 \text{ h } 09 \text{ min}$ ($SD = 2.38 \text{ min}$, $Min = 1 \text{ h } 05 \text{ min}$, $Max = 1 \text{ h } 12 \text{ min}$).

Itemanalyse

Forschungsfrage 2: Sind die Gütekriterien des Wahrnehmungsvorerfahrungstests, des Vorwissenstests und des Nachtests, sowie die der in ihm enthaltenen Animations- und Bildskala, die die selbst konstruierten Verfahren des Experimentes darstellen, zufriedenstellend?

Für die oben genannten Verfahren wurde jeweils der Cronbachs Alpha-Koeffizient sowie die Trennschärfen und Schwierigkeiten ihrer jeweiligen Items berechnet. Die Bezugsgruppe der Analysen war stets die Gesamtstichprobe ($n = 23$) der Pilotstudie. Auf die Berechnung der

Intraclasskorrelationen (ICCs) der Ratingsysteme des Vorwissenstests wurde noch verzichtet, da der Stichprobenumfang hierfür nicht ausreichend war.

Bei der Itemelimination für die weiteren statistischen Analysen lag der Grenzwert der Trennschärfe bei $r_{it} = .30$, bei den Itemschwierigkeiten wurde ein mittlerer Wertebereich zwischen $p = .20$ und $p = .80$ präferiert. Generell wurden diese Richtlinien nicht übermäßig streng angelegt, da es sich bei dem vorliegenden Experiment um eine Pilotstudie mit geringerer Stichprobengröße und noch instabilen Kennwertschätzungen handelte.

Zunächst wird die Itemanalyse zu dem Wahrnehmungsvorerfahrungstest dargestellt.

a) Wahrnehmungsvorerfahrungstest zu Pferdebewegungen

Tabelle 5.1.1

Trennschärfen und Schwierigkeiten der Items des Wahrnehmungsvorerfahrungstests in der Pilotstudie

Item	r_{it}	p
Item 1	.83	.22
Item 2	.60	.09
Item 3	-.10	.04
Item 4	.22	.04
Item 5	.24	.26
Item 6	.36	.74
Item 7	.51	.09
Item 8	.35	.17
Item 9	.00	.00
Item 10	.60	.13
Item 11	.83	.22

Anmerkungen. Die Itemnummern entsprechen denen des Wahrnehmungsvorerfahrungstests in der Pilotstudie (s. Anhang D2 - 1 für den genauen Wortlaut der Items)

Das Cronbachs Alpha des Vorerfahrungstests lag bei $\alpha = .77$ und war somit zufriedenstellend. Für die weiteren Analysen der Pilotstudie wurden die Items 3 und 9 ausgeschlossen, da sie sehr niedrige Trennschärfen und niedrige Schwierigkeitswerte aufwiesen. Durch ihre Elimination konnte das Cronbachs Alpha auf $\alpha = .80$ gesteigert werden. Die Items 2, 4, 7, 8 und 10 wurden hingegen beibehalten, da sie zwar auch niedrige Schwierigkeitswerte hatten, aber ihre Trennschärfen mittlere bis gute Werte zeigten. Zudem waren bei der Anwerbung zuvor Personen mit zu viel Reiterfahrung ausgeschlossen worden, so dass auch konzeptuell gesehen, keine hohen Zustimmungsraten bei den Items zu erwarten gewesen waren.

b) Vorwissenstest

Tabelle 5.1.2

Trennschärfen und Schwierigkeiten der Items des Vorwissenstests in der Pilotstudie

Item	<i>rit</i>	<i>p</i>
Item 1	.23	.74
Item 2	.23	.48

Anmerkungen. Die Itemnummern entsprechen denen des Vorwissenstests in der Pilotstudie (s. Anhang D2 - 2 für den genauen Wortlaut der Items)

Das Cronbachs Alpha des Vorwissenstests betrug $\alpha = .38$. Dieser schwächere Wert mochte darin begründet sein, dass die Skala nur aus zwei Items bestand und dass die Items in ihrer Aufgabenstellung heterogener waren (Gangarten aufzählen vs. Bewegungsabläufe beschreiben). Die Trennschärfen waren für eine Pilotstudie noch akzeptabel, und die Schwierigkeiten befanden sich im mittleren Bereich. Für die weiteren Analysen wurde daher keines der Items eliminiert.

c) Nachtest (gesamt)

Tabelle 5.1.3

Trennschärpen und Schwierigkeiten der Items des Nachtests (gesamt) in der Pilotstudie

Item	<i>rit</i>	<i>p</i>
1: A, br, Schr, 10s	.24	.91
2: B, br, Tr, 8	.33	.45
3: B, w, Ga, 8	.49	.91
4: A, w, Tö, 10s	.28	.77
5: A, br, Ga, 10s	.29	.68
6: B, w, Tö, 6	.13	.64
7: B, w, Schr, 8	.03	.91
8: A, br, Tr, 7.5s	.67	.86
9: A, br, Tö, 7.5s	.37	.86
10: B, w, Tr, 6	.44	.68
11: A, w, Schr, 7.5s	.00	1.00
12: B, br, Ga, 6	.61	.82
13: A, w, Tr, 10s	.61	.86
14: B, br, Tö, 8	.70	.86
15: A, w, Ga, 7.5s	.53	.73
16: B, br, Schr, 6	-.04	.73
17: A, br, Ga, 5s	.59	.82
18: B, br, Tö, 4	.34	.73
19: B, w, Schr, 4	.04	.73

Tabelle wird auf der nächsten Seite fortgesetzt

Fortsetzung Tabelle 5.1.3

Item	<i>rit</i>	<i>p</i>
20: A, w, Tr, 5s	.50	.73
21: A, w, Schr, 2.5s	.15	.91
22: B, br, Tr, 4	.30	.64
23: B, br, Ga, 2	.65	.91
24: A, w, Tö, 5s	.52	.77
25: A, br, Tr, 2.5s	.72	.73
26: B, w, Tö, 2	.56	.82
27: A, w, Ga, 2.5s	.44	.82
28: B, br, Schr, 2	.19	.86
29: A, br, Tö, 2.5s	.27	.86
30: B, w, Tr, 2	.63	.55
31: A, br, Schr, 5s	.00	.91
32: B, w, Ga, 4	.31	.95

Anmerkungen. A = Animationsitem, B = Bilditem, w = weiß gefärbtes Pferd, br = braun gefärbtes Pferd, Schr = „Schritt“, Tö = „Tölt“, Tr = „Trab“, Ga = „Galopp“. Die Zahlen an vierter Stelle bedeuten bei den Animationsitems die jeweilige Präsentationszeit, bei den Bilditems hingegen die Zahl der Key-Frames.

Die Reliabilitätsanalyse ergab einen Wert für Cronbachs Alpha von $\alpha = .86$, der als gut bezeichnet werden konnte. Wegen ihrer geringen Trennschärfen und ihrer zudem hohen Schwierigkeiten wurden die Items 7, 11 und 31 nicht weiter in den statistischen Rechnungen berücksichtigt. Bei den Items 16 und 19 waren die Schwierigkeiten akzeptabel, aber auch hier lagen zu geringe Trennschärfen vor, weshalb sie ebenfalls in den Analysen nicht weiterverwendet wurden. Nach dieser Itemelimination stieg das Cronbachs Alpha des Tests auf einen Wert von $\alpha = .88$.

15 weitere Items zeigten auch zu hohe Schwierigkeitswerte, was bedeutete, dass der Test für die Probanden insgesamt zu einfach war („Ceiling-Effekt“). Allerdings waren die

Trennschärpen dieser Aufgaben akzeptabel bis gut, so dass sie im Itempool für die weiteren Rechnungen belassen wurden.

d) Animationsskala

Tabelle 5.1.4

Trennschärpen und Schwierigkeiten der Items der Animationsskala in der Pilotstudie

Item	<i>rit</i>	<i>p</i>
1: A, br, Schr, 10s	.27	.91
4: A, w, Tö, 10s	.25	.77
5: A, br, Ga, 10s	.17	.68
8: A, br, Tr, 7.5s	.61	.86
9: A, br, Tö, 7.5s	.25	.86
11: A, w, Schr, 7.5s	.00	1.00
13: A, w, Tr, 10s	.55	.86
15: A, w, Ga, 7.5s	.43	.73
17: A, br, Ga, 5s	.62	.82
20: A, w, Tr, 5s	.51	.73
21: A, w, Schr, 2.5s	.15	.91
24: A, w, Tö, 5s	.51	.77
25: A, br, Tr, 2.5s	.73	.73
27: A, w, Ga, 2.5s	.38	.82

Tabelle wird auf der nächsten Seite fortgesetzt

Fortsetzung Tabelle 5.1.4

Item	<i>rit</i>	<i>p</i>
29: A, br, Tö, 2.5s	.35	.86
31: A, br, Schr, 5s	.09	.91

Anmerkungen. A = Animationsitem, w = weiß gefärbtes Pferd, br = braun gefärbtes Pferd, Schr = „Schritt“, Tö = „Tölt“, Tr = „Trab“, Ga = „Galopp“. Die Zahlen an vierter Stelle bedeuten die jeweiligen Präsentationszeiten der Items.

Für das Cronbachs Alpha der Animationsskala wurde ein Wert von $\alpha = .78$ errechnet. Die Aufgaben 11 und 31 wurden für die anschließenden Analysen eliminiert, da ihre Trennschärfen zu niedrig und ihre Schwierigkeitswerte zu hoch waren, wodurch die Reliabilität der Skala auf $\alpha = .79$ erhöht werden konnte. Bei acht weiteren Items konnten ebenfalls zu hohe Schwierigkeitswerte festgestellt werden. Aufgrund ihrer akzeptablen bis guten Trennschärfen wurden diese Aufgaben jedoch für die Rechnungen weiter berücksichtigt. Insgesamt betrachtet, war die Subskala somit sehr leicht für die Teilnehmer zu lösen gewesen.

e) Bildskala

Tabelle 5.1.5

Trennschärfen und Schwierigkeiten der Items der Bildskala in der Pilotstudie

Item	<i>rit</i>	<i>p</i>
2: B, br, Tr, 8	.19	.48
3: B, w, Ga, 8	.31	.91
6: B, w, Tö, 6	.20	.61
7: B, w, Schr, 8	-.10	.91

Tabelle wird auf der nächsten Seite fortgesetzt

Fortsetzung Tabelle 5.1.5

Item	<i>rit</i>	<i>p</i>
10: B, w, Tr, 6	.40	.65
12: B, br, Ga, 6	.44	.83
14: B, br, Tö, 8	.58	.87
16: B, br, Schr, 6	-.04	.74
18: B, br, Tö, 4	.31	.70
19: B, w, Schr, 4	.03	.74
22: B, br, Tr, 4	.38	.61
23: B, br, Ga, 2	.64	.87
26: B, w, Tö, 2	.56	.78
28: B, br, Schr, 2	.07	.87
30: B, w, Tr, 2	.56	.52
32: B, w, Ga, 4	.33	.96

Anmerkungen. B = Bilditem, w = weiß gefärbtes Pferd, br = braun gefärbtes Pferd, Schr = „Schritt“, Tö = „Tölt“, Tr = „Trab“, Ga = „Galopp“. Die Zahlen an vierter Stelle bedeuten die Zahl der Key-Frames bei den jeweiligen Items.

Das Cronbachs Alpha der Bildskala lag bei einem Wert von $\alpha = .69$. Item 7 und Item 28 wurden wegen ihrer geringen Trennschärfen und ihrer hohen Schwierigkeitswerte für die weiteren Analysen eliminiert. Weiterhin wurden die Items 16 und 19 aufgrund ihrer niedrigen Trennschärfen ausgeschlossen, die Schwierigkeitswerte waren hier akzeptabel. Das Cronbachs Alpha der Skala wurde durch diese Eliminationen auf $\alpha = .78$ gesteigert. Fünf weitere Items zeigten ebenfalls zu hohe Schwierigkeitswerte, wiesen aber gleichzeitig mittlere bis gute Trennschärfen auf, wodurch sie in die weiteren Rechnungen mit eingingen. Die Skala lag somit wiederum auf einem eher einfachen Schwierigkeitsniveau.

Zusätzlich zu den selbst konstruierten Verfahren wurden auch für die Subskalen des FAM Itemanalysen durchgeführt. Nach der Itemelimination ergaben sich die folgenden Reliabilitäten:

Tabelle 5.1.6

Reliabilitäten der Subskalen des FAM in der Pilotstudie

	<i>n</i> (Items)	α
Erfolgswahrscheinlichkeit	3	.68
Misserfolgsbefürchtung	4	.81
Interesse	5	.88
Herausforderung	4	.45

Bei den Skalen „Erfolgswahrscheinlichkeit“ und „Misserfolgsbefürchtung“ war jeweils ein Item für die weiteren Analysen ausgeschlossen worden, da dadurch die Reliabilitäten der Skalen erhöht werden konnten. Das Cronbachs Alpha der Skala „Herausforderung“ zeigte einen niedrigeren Wert, der jedoch durch den Ausschluss von Items nicht weiter gesteigert werden konnte.

Deskriptive Statistiken

Bei der Berechnung der deskriptiven Statistiken für die abhängigen und die Kontrollvariablen der Studie wurde auf vorherige Ausreißeranalysen mittels Box Plots verzichtet, da der Stichprobenumfang für genauere Verteilungsschätzungen noch zu gering war und auch die Power der nachfolgenden Verfahren durch den Ausschluss zu vieler Personen gelitten hätte.

Der Stichprobenumfang in der Pilotstudie betrug $n = 23$. Das Durchschnittsalter lag bei 22.26 Jahren ($SD = 1.94$, $Min = 20.00$, $Max = 26.00$). In der Stichprobe befanden sich 5 Männer und 18 Frauen.

Zunächst werden die deskriptiven Statistiken der abhängigen Variablen, dann diejenigen der Kontrollvariablen dargestellt.

Tabelle 5.1.7

Deskriptive Statistiken der abhängigen Variablen in der Pilotstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Nachtest (gesamt)	22	6.00	27.00	21.14	5.39
Animationsskala	22	4.00	14.00	11.32	2.89
Bildskala	23	1.00	12.00	8.78	2.78
Cognitive Load (Lernaufgabe)	23	3.00	18.00	11.39	3.74
Cognitive Load (Nachtest)	23	6.00	18.00	13.17	2.73

Tabelle 5.1.8

Deskriptive Statistiken der Kontrollvariablen in der Pilotstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Wahrnehmungs- vorerfahrungstest	23	.00	7.00	1.96	2.08
Vorwissenstest	23	1.00	4.00	2.43	.84
IST verbal	23	16.00	31.00	24.52	3.90
IST räumlich	23	16.00	33.00	23.52	4.92
IST Merkfähigkeit	21	11.00	22.00	18.33	3.32
Erfolgswahrscheinlichkeit	23	14.00	21.00	18.26	2.18
Misserfolgsbefürchtung	23	4.00	22.00	9.17	4.63

Tabelle wird auf der nächsten Seite fortgesetzt

Fortsetzung Tabelle 5.1.8

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Interesse	23	9.00	35.00	22.43	6.27
Herausforderung	23	10.00	25.00	18.04	3.28

Mittelwertsvergleiche

Testung auf Normalverteilung und Varianzhomogenität

Um die Hypothesen zu den Haupteffekten zwischen den beiden Bedingungen zu testen, wurde entschieden, t-Tests bzw., bei Hinzunahme relevanter Kovariaten, „Allgemeine lineare Modelle“ (ALMs) zur Anwendung kommen zu lassen. Diese Verfahren machen die Annahmen der Normalverteilung der Daten und der Varianzhomogenität zwischen den Bedingungen, die mittels Kolmogorov-Smirnov-Tests und Levene-Tests überprüft wurden, was aufgrund der Zellengrößen von 11 bzw. 12 Probanden in der Durchführung vertretbar war.

Die Kolmogorov-Smirnov-Tests mit den abhängigen Variablen zeigten, dass bei allen Kriterien Normalverteilungen vorlagen. Auch die Varianzhomogenität zwischen den Bedingungen war bei allen abhängigen Variablen vorhanden.

Da Varianzhomogenität, Normalverteilung der Daten und gleiche Stichprobengrößen bei den abhängigen Variablen gegeben waren, konnten t-Tests und ALMs zur Bestimmung der Mittelwertsvergleiche angewandt werden.

Bestimmung relevanter Kovariaten

Um relevante Kovariaten für die weiteren Rechnungen herauszufinden, wurden für die Kontrollvariablen t- und w-Tests zwischen den Bedingungen einerseits und ihre bivariaten Korrelationen zu den abhängigen Variablen andererseits gerechnet. Bei keiner der Kontrollvariablen konnten sowohl größere bis signifikante Mittelwertsunterschiede als auch signifikante Korrelationen zu einem der Kriterien festgestellt werden. Daher wurde keine der

Kontrollvariablen in die Mittelwertsvergleiche aufgenommen, die ausschließlich mit t-Tests durchgeführt wurden.

t-Tests

Die Überprüfung der Haupteffekte des Faktors „Bedingung“ (Animation vs. Bild parallel) auf die fünf Kriterien der Pilotstudie wurde jeweils mit einem t-Test mit einem Signifikanzniveau von $\alpha = .05$ durchgeführt.

Die Darstellung der Resultate erfolgt in der Form, dass zunächst nochmals die Hypothesen aus dem Theorieteil in statistischer Kurzform genannt und dazu komplementär die jeweiligen Ergebnisse hinzugefügt werden.

Das erste Kriterium ist das objektive Leistungsmaß „Nachtest gesamt“:

Hypothese 1a: Animation > Bild parallel; AV = Nachtest gesamt (Animations- + Bildskala)

Tabelle 5.1.9

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Nachtest gesamt“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild parallel	11	20.55	6.25
Animation	11	21.73	4.61

Der Mittelwertsunterschied zwischen den beiden Gruppen war nicht signifikant ($t(20) = .51$, $p(1\text{-seitig}) = .31$, $d = .21$). Die Animationsgruppe erzielte leicht bessere Ergebnisse als die Gruppe, die mit den statischen Bildern lernte.

Hypothese 1b: Animation > Bild parallel; AV = Animationsskala

Tabelle 5.1.10

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Animationsskala“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild parallel	11	10.73	3.29
Animation	11	11.91	2.43

Der t-Test erbrachte bei dem Kriterium „Animationsskala“ keinen signifikanten Mittelwertsunterschied zwischen den Bedingungen ($t(20) = .96$, $p(1\text{-seitig}) = .18$, $d = .41$). Die Animationsbedingung erreichte wiederum bessere Resultate als die parallele Bildbedingung.

Hypothese 1c: Animation > Bild parallel; AV = Bildskala

Tabelle 5.1.11

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Bildskala“

	<i>n</i>	<i>M</i>	<i>SD</i>
Animation	12	8.67	2.61
Bild parallel	11	8.91	3.08

Die Mittelwerte der beiden Bedingungen lagen nahe beieinander, wobei der Unterschied in der nicht vorhergesagten Richtung war. ($t(21) = -.20$, $p(2\text{-seitig}) = .84$, $d = .08$). Die Gruppe, die mit parallelen Bildern lernte, war bei dem Kriterium ganz leicht im Vorteil.

Hypothese 1d: Animation < Bild parallel; AV = Cognitive Load beim Lernen

Tabelle 5.1.12

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load bei der Lernaufgabe“

	<i>n</i>	<i>M</i>	<i>SD</i>
Animation	12	9.42	3.29
Bild parallel	11	13.55	3.01

Bei dem Kriterium „Cognitive Load beim Lernen“ wurde der Unterschied zwischen den beiden Bedingungen sehr signifikant ($t(21) = -3.13$, $p(1\text{-seitig}) = .003$, $d = 1.31$). Die Probanden, die die Gangarten mit Animationen lernten, empfanden das Bearbeiten der Lerneinheit als weniger schwierig und anstrengend als die Teilnehmer, die hierfür die parallelen Bilder nutzten.

Hypothese 1e: Animation < Bild parallel; AV = Cognitive Load beim Nachtest

Tabelle 5.1.13

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load beim Nachtest“

	<i>n</i>	<i>M</i>	<i>SD</i>
Animation	12	12.58	3.09
Bild parallel	11	13.82	2.23

Die Mittelwertsdifferenz zwischen der Animations- und der parallelen Bildgruppe wurde bei dem Kriterium „Cognitive Load beim Nachtest“ nicht signifikant ($t(21) = -1.09$, $p(1\text{-seitig}) = .15$, $d = .46$). Die Personen in der Animationsbedingung verspürten beim Bearbeiten des Nachtests weniger Cognitive Load als die Probanden der statischen Bildbedingung.

Analyse der Interaktion zwischen dem Format der Lernbedingung und dem Format der Nachtestskalen

Hypothese 2: Interaktion Lernformat x Nachtestformat ist signifikant, gerichtete Fragestellung

Um den Interaktionseffekt zwischen dem Format der Lernbedingung und dem Format der Nachtestskalen zu testen, wurde eine zweifaktorielle Varianzanalyse mit Messwiederholung durchgeführt. Der Zwischensubjektfaktor war hierbei die Versuchsbedingung (Animation vs. Bild parallel), den Innersubjektfaktor stellten die Subskalen des Nachtests dar (Animationsskala vs. Bildskala).

Tabelle 5.1.14

Zellenmittelwerte und Standardfehler in der zweifaktoriellen ANOVA mit Messwiederholung

Bedingung	Subskala	<i>n</i>	<i>M</i>	<i>SE</i>
Animation	Animationsskala	11	11.91	.87
	Bildskala	11	9.00	.84
Bild parallel	Animationsskala	11	10.73	.87
	Bildskala	11	8.91	.84

Der Interaktionseffekt wurde annähernd signifikant (Wilks lambda = .89, $F(1, 20) = 2.40$, $p(1\text{-seitig}) = .07$, partielles $\eta^2 = .11$). Das Interaktionsmuster entsprach der vorhergesagten Richtung: Der Lernvorteil durch Animationen im Vergleich zu statischen Bildern war auf der Animationsskala stärker ausgeprägt als auf der statischen Bildskala des Nachtests.

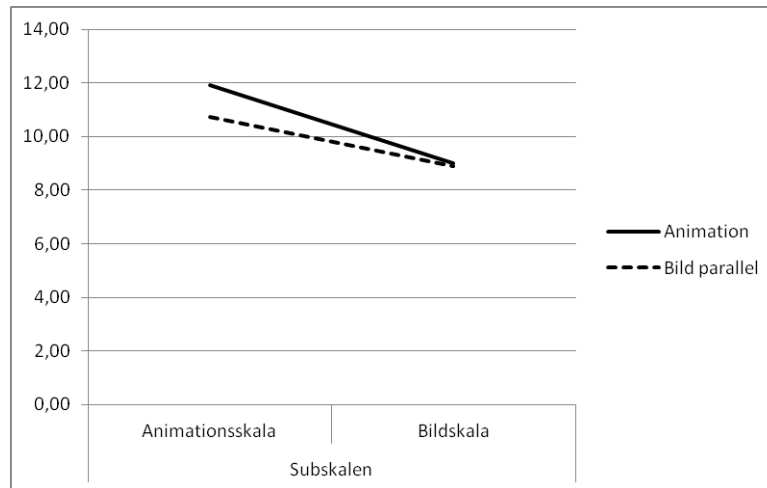


Abbildung 5.1.8: Graphische Veranschaulichung der Interaktion zwischen dem Format der Lernbedingung und dem Format der Nachtestskalen in der Pilotstudie

5.1.4 Diskussion

Im Rahmen der Pilotstudie sollte zunächst die Frage untersucht werden, ob der Versuchsablauf stimmig und das Experiment in der geplanten Form durchführbar war. Im Großen und Ganzen konnte diese Frage bejaht werden, da die Experimente vorwiegend problemlos verliefen. Zu den einzelnen Tests sowie zu den Instruktionen kamen kaum Fragen auf, was auf deren gute Verständlichkeit hindeutete. Über die Hälfte der Versuchspersonen äußerte sich im abschließenden Interview positiv über das Experiment insgesamt, wodurch an seiner Grundkonzeption sowie an seiner Testabfolge nichts Wesentliches geändert wurde. Der Versuch konnte zudem relativ zeitökonomisch durchgeführt werden, was sich in einer durchschnittlichen Erhebungsdauer von 1 h 09 min, die zudem eine geringe Streuung aufwies ($SD = 2.38$ min), manifestierte.

Wie sich auch in den Schwierigkeitsindizes zeigte, war lediglich der Nachtest zu einfach für die Probanden, wodurch die Lernzeiten und die Itemreihenfolge im Test für die Hauptstudie geändert werden mussten. Zudem war bei der Durchführung aufgefallen, dass die Pausen, die im Nachtest zwischen den Aufgaben waren, zu lange geschaltet waren. Da dies trotz vorheriger Hinweise in der Instruktion eher zu Konzentrationsmängeln bei den Teilnehmern führte, wurden die Pfeile, durch die die Probanden mit einem Klick zur jeweils nächsten Aufgabe gelangen konnten, in der Hauptstudie statt 8 s nur noch 4 s lang gesperrt. Die Taskleiste, die bei manchen Teilnehmern beim Lernen erschien, wurde hingegen nicht entfernt, da dies technisch aufwändig gewesen wäre und die Taskleiste für die Probanden subjektiv und auch objektiv keine größere Störung dargestellt hatte.

Im Interview war häufiger angeklungen, dass der IST-2000-R „schwer“ war und dass die Probanden Probleme mit den einzelnen Aufgabengruppen hatten. Da die Testergebnisse jedoch im Normbereich lagen und da der IST-2000-R ein standardisiertes Verfahren ist, wurde er in derselben Form auch in die Hauptstudie übernommen. Wahrscheinlich wurde er von den Teilnehmern im Kontrast zu den eher einfachen Aufgaben zu den Pferdegangarten als schwierig empfunden. In der Hauptstudie wurde er weiterhin am Ende des Experimentes eingesetzt, damit die konzeptuell wichtigeren Nachtestaufgaben nicht durch Ermüdungseffekte verzerrt wurden.

Aus diesen Ausführungen wird deutlich, dass wie bereits erwähnt, der Versuch in seiner Grundkonzeption stimmig war und kein erheblicher Änderungsbedarf bestand.

Die zweite Forschungsfrage der Pilotstudie fokussierte sich auf die Gütekriterien der eigens für die Experimente konstruierten Verfahren und darauf, ob diese ausreichend waren, damit die Tests weiter verwendet werden konnten.

Das Cronbachs Alpha des Wahrnehmungsvorerfahrungstests lag bei $\alpha = .77$ und war somit zufriedenstellend. Die Mehrzahl der Items hatte niedrige Schwierigkeitswerte, was gleichbedeutend war mit niedrigen Zustimmungsraten zu den Items. Dies war insofern zu erwarten gewesen, da in der Anwerbemail Personen mit zu viel Reiterfahrung gebeten wurden, sich nicht zur Teilnahme zu melden, so dass in der vorliegenden Stichprobe Probanden mit weniger Erfahrungen mit Pferden waren. Die Trennschärfen der Items waren, bis auf zwei Ausnahmen, zufriedenstellend. Diese wurden trotzdem in die Hauptstudie mit aufgenommen, da die Aufgaben ökonomisch durchgeführt werden konnten und da dadurch ihre Gütekriterien nochmals an einer größeren Stichprobe getestet werden konnten. Lediglich Item 6 („Gibt oder gab es in Ihrer Familie oder Ihrem Freundeskreis aktive Reiter?“) wurde aus inhaltlichen Gründen eliminiert, da nach nochmaliger Durchsicht entschieden wurde, dass das Vorhandensein von Reitern in der Familie oder im Freundeskreis zu wenige Hinweise auf die Wahrnehmungserfahrungen der Probanden selbst mit Pferden liefert.

Beim Vorwissenstest wies die Reliabilität einen schwächeren Wert auf ($\alpha = .38$), was wahrscheinlich durch die geringe Skalenlänge und die inhaltlich größere Heterogenität der Items begründet war. Bei der ersten Aufgabe mussten die Teilnehmer lediglich Gangarten aufzählen, während sie bei dem zweiten Item Bewegungsabläufe beschreiben sollten. Die Trennschärfen der beiden Items waren für eine Pilotstudie akzeptabel, und die Schwierigkeiten lagen im mittleren Bereich. Zudem fragte gerade das zweite Item explizit nach bei den Probanden schon vorhandenem Bewegungswissen zu Pferdegangarten, das wahrscheinlich die Lösung des Nachtests mit beeinflusst. Das Vorwissen korrelierte zu $r = .32$ mit den Wahrnehmungsvorerfahrungen der Probanden, was einen fast signifikanten Zusammenhang darstellte, der aufgrund theoretischer Vorüberlegungen zu erwarten gewesen war. Allerdings war die Korrelation nicht als derart hoch einzuschätzen, dass die beiden Skalen zusammengefasst werden sollten, da sie, konzeptuell gesehen, nach unterschiedlichen Inhalten fragten. Daher wurde die Vorwissensskala in der vorliegenden Form eigenständig in die Hauptstudie übernommen.

Der Nachtest insgesamt zeigte eine hohe Reliabilität mit einem Cronbachs Alpha von $\alpha = .86$. Die Trennschärfen der meisten Items konnten als akzeptabel bis gut bezeichnet

werden. Allerdings waren 18 der 32 Items des Nachttests für die Probanden zu einfach, was an ihren höheren Schwierigkeitswerten ($p > .80$) zu sehen war. Ähnliche Ergebnismuster zeichneten sich auch in der Animations- und der Bildskala des Nachttests ab. Die Animationsskala, in der alle Animationsitems zusammengefasst waren, wies eine zufriedenstellende Reliabilität ($\alpha = .78$) mit überwiegend akzeptablen bis guten Trennschärfen der Items auf, wobei 10 der 16 Items zu leicht waren. Die Reliabilität der Bildskala, die alle Bilditems beinhaltete, war im Vergleich zu der der beiden anderen Variablen leicht schwächer ($\alpha = .69$), wobei zumindest hier ungefähr die Hälfte der Aufgaben Schwierigkeiten im mittleren Bereich hatte. Die Trennschärfen waren größtenteils zufriedenstellend.

Auffällig an der Itemanalyse des Nachttests war somit sein relativ niedriger Schwierigkeitsgrad für die Probanden, so dass man von einem Deckeneffekt innerhalb des Designs sprechen konnte. Um einen solchen Effekt in der Hauptstudie zu vermeiden, wurden dort die Lernzeiten von 20 s auf 10 s für eine Gangart reduziert. Informelle Testungen hatten gezeigt, dass eine solche Lernzeit zumindest für das Lernen mit Animationen ausreichend war, um Nachttestitems lösen zu können. Zudem wurde die Aufgabenreihenfolge innerhalb des Tests umgedreht, so dass die konzeptuell schwierigeren Items (kürzere Animationsitems, weniger Bilder bei den Bilditems) am Anfang des Tests bearbeitet werden mussten. Der Grund für diese Änderung lag darin, dass manche der Versuchspersonen im Interview erwähnten, dass der Nachttest selbst einen Lern - bzw. einen Übungseffekt für sie gehabt hätte, so dass sie die inhaltlich schwierigeren Items am Ende des Tests nicht als komplexer als die leichteren Aufgaben am Anfang gefunden hätten. Dies spiegelte sich auch in der Itemanalyse in konstant ungefähr gleichbleibenden Schwierigkeitswerten der Items über den Test hinweg wider. Daher wurden die konzeptuell schwierigeren Aufgaben in der Hauptstudie an den Anfang des Tests gesetzt, um das Verfahren insgesamt schwieriger zu machen und um klarere Effekte, zumindest in der ersten Hälfte des Tests, zu erhalten. Die Gestaltung der einzelnen Items wurde hingegen beibehalten, da die Annahme war, dass durch die oben genannten Änderungen auch die Gütekriterien der Aufgaben im Vergleich zur Pilotstudie differieren würden, so dass eine Elimination oder Umstrukturierung einzelner Items als nicht sinnvoll erschien.

In der Pilotstudie sollten zudem explorativ die Haupteffektshypothesen zwischen der Animations- und der parallelen Bildbedingung bezüglich der abhängigen Variablen der Untersuchung getestet werden, sowie der womöglich vorhandene Interaktionseffekt

zwischen dem Format des Lernmaterials und dem Format der Nachtestskalen bestimmt werden. Hierbei ist vorab anzumerken, dass es sich bei den Resultaten um erste Ergebnisse handelte, die noch nicht zu weitreichend interpretiert werden dürfen. Durch die geringere Stichprobengröße in der Pilotstudie waren die Schätzungen der Kennwerte noch instabil. Zudem war die Berechnung einer Ausreißeranalyse nicht möglich. Deshalb war die Durchführung einer Hauptstudie mit einer größeren Stichprobe unabdingbar, um stabilere Erkenntnisse zu erhalten. Nichtsdestotrotz sollen die Ergebnisse der Pilotstudie hier näher betrachtet und zumindest im Ansatz diskutiert werden.

In den Hypothesen zu den Haupteffekten zwischen den Bedingungen „Animation“ und „Bild parallel“ (Hypothesen 1a – 1e) wurde die Annahme gemacht, dass Personen die mit Animationen lernten, bessere Ergebnisse bei bildlichen Kategorisierungsaufgaben erzielten als Personen, die mit parallel angeordneten, statischen Key-Frames lernten. Diese Überlegenheit sollte sich in besseren Ergebnissen in einem Nachtest und der in ihm enthaltenen Animations- und statischen Bildskala, sowie in einem geringeren Cognitive Load während der Lerneinheit und während des Nachtests manifestieren.

Darüber hinaus wurde in einer gerichteten Interaktionshypothese (Hypothese 2) davon ausgegangen, dass der Lernvorteil durch Animationen, aufgrund des gleichen Darstellungsformates, auf der Animationsskala stärker ausgeprägt war als auf der statischen Bildskala des Nachtests.

Nach den Analysen zeigte sich, dass vier der fünf Haupteffektshypothesen nicht bestätigt werden konnten, da die Unterschiede zwischen den beiden Bedingungen nicht signifikant wurden. Lediglich beim Kriterium „Cognitive Load beim Lernen“ konnte die entsprechende Hypothese bestätigt werden. Personen, die sich die Pferdegangarten mit Hilfe von Animationen aneigneten, empfanden die Lerneinheit als viel weniger schwierig und anstrengend, als die Personen, die mit statischen Key-Frames lernten. Der Unterschied war sehr signifikant und wies einen sehr großen Effekt auf. Deskriptiv gesehen, erreichte bei den abhängigen Variablen „Nachtest gesamt“, „Animationsskala“ und „Cognitive Load beim Nachtest“ die Animationsbedingung bessere Lernergebnisse als die statische Bildbedingung, wie in den Hypothesen vorhergesagt worden war. In der abhängigen Variable „Bildskala“ lagen die beiden Bedingungen fast gleichauf. Bezüglich der Kriterien „Animationsskala“ und „Cognitive Load beim Nachtest“ wurden fast mittlere Effektstärken erreicht.

Die Testung der Interaktionshypothese zwischen dem Präsentationsformat des Lernmaterials und dem Format der Nachtestskalen verfehlte knapp die statistische Signifikanz

(p (1-seitig) = .07), wodurch die Hypothese ebenfalls verworfen werden musste. Allerdings verlief das deskriptive Muster der Interaktion in der vorhergesagten Richtung und wies dabei einen mittleren bis großen Effekt auf. Der positive Effekt des Lernens mit Animationen auf die Leistung in Kategorisierungsaufgaben war auf der Animationsskala stärker ausgeprägt als auf der statischen Bildskala.

Das bedeutet, dass sich die vorhergesagten Richtungen der Unterschiede in den Haupteffektshypothesen im Trend der Ergebnisse zumindest im Ansatz bestätigten. Eventuell könnten die Differenzen durch die Verwendung einer größeren Stichprobe in der Hauptstudie Signifikanz erreichen. Zudem lag im Nachtest der Pilotstudie ein Deckeneffekt vor, das heißt, die Schwierigkeiten der Items waren so niedrig, dass die Aufgaben von beiden Gruppen sehr gut gelöst werden konnten. Das Schwierigkeitsniveau war somit so gering, dass zwischen den Bedingungen nicht gut differenziert werden konnte, wodurch die Ergebnisse gegebenenfalls nicht signifikant wurden. Das Design der Lern- und Testmaterialien wurde für die Hauptstudie überarbeitet und in seiner Gestaltung schwieriger angesetzt (verkürzte Lernzeit, umgekehrte Aufgabenreihenfolge im Nachtest), so dass dann eventuell ein mittlerer Schwierigkeitsgrad erreicht werden kann, bei dem die Gruppenunterschiede deutlicher werden.

Jedoch wurde in der Pilotstudie zumindest bei der abhängigen Variablen „Cognitive Load beim Lernen“ ein sehr signifikantes Ergebnis in der vorhergesagten Richtung gefunden. Der Lernprozess mit Animationen wurde von den Probanden als einfacher und weniger anstrengend erlebt als der mit statischen Bildern, was zudem auch im Interview von manchen Probanden geäußert wurde. Dieses Ergebnis und auch die in den meisten Fällen deskriptive Überlegenheit der Animationsbedingung könnten erste Hinweise auf eine höhere „instruktionale Effizienz“ (Paas et al., 2007) dieses Präsentationsformates geben. Die Unterschiede in den Leistungsmaßen zwischen der Animations- und der statischen Bildgruppe wurden zwar nicht signifikant, aber das Lernen mit Animationen schien angenehmer zu sein und den Teilnehmern leichter zu fallen. Dies könnte auch im Ansatz darauf hindeuten, dass die theoretischen Annahmen, dass aus einer Animation die dynamischen perzeptuellen Schemata von Bewegungen leichter, dem alltäglichen Wahrnehmungsprozess analog, heraus konstruiert und angewandt werden können, Bestätigung finden könnten. Allerdings hing der Lernvorteil, der durch Animationen erzielt werden konnte, auch davon ab, ob nach dem Lernprozess animierte oder statische Kategorisierungsaufgaben gelöst werden mussten. Die Interaktion zwischen den

Präsentationsarten von Lerneinheit und Nachtest war fast signifikant, in der Richtung, dass die Animationsbedingung der statischen Bildbedingung gegenüber auf der Animationsskala deutlich überlegen war, während auf der statischen Bildskala zwischen den Gruppen kaum ein Unterschied zu sehen war. Vorsichtig interpretiert, würde dies eventuell bedeuten, dass die Nachteile des Lernens mit statischen Bildern, bei dem im Vergleich zum Lernen mit Animationen die Bewegungszwischenzustände kognitiv inferiert werden müssen, zumindest auf der statischen Bildskala durch die Verwendung des gleichen Präsentationsformates nivelliert werden, da dann Gedächtnistransfereffekte und die Notwendigkeit der gleichen mentalen Operationen wie im Lernprozess das Lösen der Kategorisierungsaufgaben unterstützen.

Insgesamt ist zu sagen, dass auch die Lernergebnisse der statischen Bildgruppe auf einem guten Niveau waren, was eventuell daran liegen könnte, dass der Nachtest selbst einen Übungs- und zusätzlichen Lerneffekt für die Teilnehmer hatte, was von manchen Probanden im abschließenden Interview angemerkt wurde. Dadurch könnten sich die Unterschiede zwischen den Gruppen nivelliert haben und der schwierigere Lernvorgang mit statischen Bildern kompensiert worden sein.

Wie oben schon erwähnt, handelte es sich bei den vorliegenden Ergebnissen um erste explorative Resultate einer Pilotstudie, die aufgrund der geringeren Stichprobengröße noch nicht zu weit interpretiert werden dürfen und denen Ergebnisse einer Hauptstudie angeschlossen werden sollten.

5.2 Hauptstudie

5.2.1 Fragestellung und Hypothesen

Haupteffekte

In der Hauptstudie sollte mit Hilfe einer größeren Stichprobe der Vergleich zwischen Animationen und parallel angeordneten Key-Frames hinsichtlich ihrer Wirksamkeit auf die Lösung von bildlichen Kategorisierungsaufgaben sowie auf die mentale Belastung von Probanden angestellt werden. Die Haupteffektshypothesen auf die fünf Kriterien wurden jeweils zugunsten der Animationsbedingung gerichtet formuliert.

Animation vs. Parallele Key-Frames

Hypothese 1a: Personen, die mit Animationen lernen, erzielen signifikant bessere Werte in animierten und statischen bildlichen Kategorisierungsaufgaben eines Nachttests, als Personen, die mit parallelen Key-Frames lernen.

Hypothese 1b: Personen, die mit Animationen lernen, erzielen signifikant bessere Werte in animierten bildlichen Kategorisierungsaufgaben eines Nachttests, als Personen, die mit parallelen Key-Frames lernen.

Hypothese 1c: Personen, die mit Animationen lernen, erzielen signifikant bessere Werte in statischen bildlichen Kategorisierungsaufgaben eines Nachttests, als Personen, die mit parallelen Key-Frames lernen.

Hypothese 1d: Personen, die mit Animationen lernen, empfinden signifikant weniger Cognitive Load beim Lernen, als Personen, die mit parallelen Key-Frames lernen.

Hypothese 1e: Personen, die mit Animationen lernen, empfinden signifikant weniger Cognitive Load beim Lösen des Nachttests, als Personen, die mit parallelen Key-Frames lernen.

Interaktion zwischen dem Präsentationsformat der Lernbedingung und dem Präsentationsformat der Nachtestaufgaben

Zudem wurde untersucht, ob der eventuell vorhandene Haupteffekt der Animationen bei Nachtestaufgaben verschiedenen Präsentationsformates auch unterschiedlich groß war. Die Interaktionshypothese wurde ebenfalls gerichtet formuliert:

Hypothese 2: Es besteht eine signifikante Interaktion zwischen dem Präsentationsformat des Lernmaterials und dem der Nachtestaufgaben, in der Richtung, dass der Lernvorteil durch Animationen bei animierten bildlichen Kategorisierungsaufgaben größer ist als bei statischen bildlichen Kategorisierungsaufgaben.

Weitere Interaktionen

Interaktionen zwischen den Faktoren „Bedingung“ und „Wahrnehmungsvorerfahrung“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames parallel“ mit der Wahrnehmungsvorerfahrung von Personen

Es wurde bei diesen Wechselwirkungen von gerichteten Hypothesen ausgegangen, da sich die günstigen Eigenschaften einer Animation im Vergleich zu parallelen Key-Frames vor allem für Personen mit geringerer Vorerfahrung mit Pferdegangarten positiv auswirken sollten.

Hypothesen 3a - e: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Animation“ und „Bild parallel“ mit der Wahrnehmungsvorerfahrung von Personen, in der Richtung, dass der Lernvorteil durch Animationen bei Personen mit geringerer Wahrnehmungsvorerfahrung größer ist als bei Personen mit höherer Wahrnehmungsvorerfahrung. Diese Annahme bezieht sich auf

a: die Leistung in animierten und statischen bildlichen Kategorisierungsaufgaben eines Nachtests

b: die Leistung in animierten bildlichen Kategorisierungsaufgaben eines Nachttests

c: die Leistung in statischen bildlichen Kategorisierungsaufgaben eines Nachttests

d: den Cognitive Load beim Lernen

e: den Cognitive Load beim Lösen des Nachttests

Interaktionen zwischen den Faktoren „Bedingung“ und „Räumliche Fähigkeiten“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames parallel“ mit den räumlichen Fähigkeiten von Personen

Es sollte weiterhin untersucht werden, ob sich zwischen den beiden Darstellungsformen „Animation“ und „Bild parallel“ unterschiedliche Effekte einstellen, je nachdem ob die Personen über ein höheres bzw. niedrigeres räumliches Vorstellungsvermögen verfügen. Die diesbezüglichen Hypothesen wurden gerichtet aufgestellt:

Hypothesen 4a - e: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Animation“ und „Bild parallel“ mit den räumlichen Fähigkeiten von Personen, in der Richtung, dass der Lernvorteil durch Animationen bei Personen mit geringeren räumlichen Fähigkeiten größer ist als bei Personen mit höheren räumlichen Fähigkeiten. Diese Annahme bezieht sich auf

a: die Leistung in animierten und statischen bildlichen Kategorisierungsaufgaben eines Nachttests

b: die Leistung in animierten bildlichen Kategorisierungsaufgaben eines Nachttests

c: die Leistung in statischen bildlichen Kategorisierungsaufgaben eines Nachttests

d: den Cognitive Load beim Lernen

e: den Cognitive Load beim Lösen des Nachttests

Interaktionen zwischen den Faktoren „Bedingung“ und „Merkfähigkeit“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames parallel“ mit der Merkfähigkeit von Personen

Zuletzt stellte sich die Frage, ob differentielle Treatmentwirkungen in Abhängigkeit von der Merkfähigkeit von Personen bestanden. Auch hier wiesen die Hypothesen ein gerichtetes Muster auf.

Hypothesen 5a - e: Es besteht eine signifikante Interaktion zwischen den Bedingungen „Animation“ und „Bild parallel“ mit der Merkfähigkeit von Personen, in der Richtung, dass der Lernvorteil durch Animationen bei Personen mit geringerer Merkfähigkeit größer ist als bei Personen mit höherer Merkfähigkeit. Diese Annahme bezieht sich auf

a: die Leistung in animierten und statischen bildlichen Kategorisierungsaufgaben eines Nachttests

b: die Leistung in animierten bildlichen Kategorisierungsaufgaben eines Nachttests

c: die Leistung in statischen bildlichen Kategorisierungsaufgaben eines Nachttests

d: den Cognitive Load beim Lernen

e: den Cognitive Load beim Lösen des Nachttests

5.2.2 Methoden

5.2.2.1 Stichprobe

In der Hauptstudie wurden 59 Studenten der Fachrichtungen Psychologie, Lehramt, Erziehungswissenschaften, u. a., der Universität Landau untersucht. Die Teilnehmer wurden durch Rundmails angeworben, die an die gesamte Studierendenschaft der Universität versandt wurden. Die Studenten konnten sich dann selbständig über die Lernplattform „OLAT“ zu dem Versuch anmelden. In den Mails wurden Personen mit zu viel Erfahrung im Reit- oder Pferdebereich darum gebeten, sich nicht für die Studie zu melden. Zudem wurden solche Personen ausgeschlossen, die schon an der Pilotstudie oder der Pilot- und Hauptstudie des Experimentes zu Inferenzaufgaben teilgenommen hatten. Letztere hätten beim IST-2000-R, der in allen Studien verwendet wurde, eventuell einen Übungsvorteil bei den Aufgaben gehabt.

Nach der Durchführung der Studie wurde ein Proband aus den weiteren Analysen ausgeschlossen, da er während der kurzen Lernzeit teilweise nicht auf den Bildschirm geschaut hatte und abgelenkt war. Durch die statistische Ausreißeranalyse mittels Box Plots mussten sechs weitere Versuchspersonen eliminiert werden, da sie, gemessen an der übrigen Stichprobe, auffällig hohe bzw. niedrige Werte in den Variablen hatten. Somit gingen 52 Personen in die Mittelwertvergleiche und die darauf folgenden Analysen mit ein.

Das Durchschnittsalter der Stichprobe lag bei $M = 23.08$ Jahren ($SD = 4.03$, $Min = 20.00$, $Max = 45.00$). Unter den Teilnehmern waren 41 Frauen und 11 Männer, was einer Geschlechterverteilung von 78,8% Frauen zu 21,2% Männern entsprach.

5.2.2.2 Design

Bei der Versuchsanordnung handelte es sich um ein 2 x 2-Design mit Messwiederholung hinsichtlich der Lösung der Kategorisierungsaufgaben (Faktor = Präsentationsform der Lernmaterialien zu den Pferdegangarten; Messwiederholung: Animations- und statische Bildsubskala eines Nachttests). Der Einfluss der beiden Bedingungen „Animation“ und „Bild parallel“ auf den Cognitive Load der Teilnehmer wurde mit einem einfaktoriellen Design untersucht. Die „Aptitude-Treatment-Interaktionen“ zwischen den beiden Lernbedingungen und der Wahrnehmungsvorerfahrung, den räumlichen Fähigkeiten sowie der Merkfähigkeit

der Teilnehmer wurden zusätzlich in jeweils zweifaktoriellen Designs untersucht, da sie als eigenständige Fragestellungen angesehen wurden.

Die Probanden wurden den beiden Lernbedingungen in der Durchführung randomisiert zugewiesen. In der Animationsbedingung befanden sich 27 Teilnehmer, in der parallelen Bildbedingung waren 25 Versuchspersonen.

Innerhalb der beiden Bedingungen wurden wie in der Pilotstudie die Sequenzen der zu lernenden Gangarten variiert, wobei diese Variation keinen zusätzlichen Faktor in der Studie darstellte. Auch zu diesen Teilgruppen erfolgte eine randomisierte Zuteilung.

Die abhängigen Variablen des Designs waren, wie in der Pilotstudie, der Score eines Nachtests zu bildlichen Kategorisierungsaufgaben, sowie die Scores der in ihm enthaltenen Animations- und Bildskala. Zudem wurde der Cognitive Load der Probanden nach der Lerneinheit und nach dem Nachtest erhoben.

Die Kontrollvariablen hatten geringere Korrelationen zu den Kriterien in der Pilotstudie gezeigt. Da dies aber auch an der geringen Stichprobengröße und an dem teilweise vorhandenen Deckeneffekt in den Daten gelegen haben könnte, wurden die Variablen in die Hauptstudie übernommen, in der eine größere Stichprobe und eine schwierigere Versuchsanordnung verwendet wurde. Die Kontrollvariablen waren somit: die bisherigen Wahrnehmungserfahrungen der Teilnehmer mit Pferden, das Vorwissen zu Pferdegangarten, die aktuelle Lernmotivation und die räumlichen, sprachlichen sowie die Gedächtnisfähigkeiten der Teilnehmer.

5.2.2.3 Lernmaterialien

Die Auswahl der Key-States und die Gestaltung des Lernmaterials waren mit denen in der Pilotstudie identisch, so dass für eine genaue Beschreibung auf den Abschnitt 5.1.2.3 der vorliegenden Dissertation verwiesen sei.

Die einzige Veränderung von Pilot- zu Hauptstudie fand hinsichtlich der Lernzeiten statt. Da der Nachtest der Pilotstudie für die Probanden sehr einfach gewesen war, wurden die Lernzeiten von 20 s auf jeweils 10 s für eine Gangart reduziert. Entsprechend wurde auch die Instruktion der Lernphase überarbeitet. In informellen Testungen hatte sich gezeigt, dass diese Präsentationsdauer zumindest für das Lernen mit Animationen ausreichend war, um bildliche Kategorisierungsaufgaben lösen zu können (für eine Ansicht des Lernmaterials, s. Anhang E1 auf der beiliegenden CD).

5.2.2.4 Messinstrumente

Da der soziodemographische Fragebogen, der FAM, die Items zu Cognitive Load und der IST-2000-R in identischer Form wie in der Pilotstudie eingesetzt wurden, sei aus Redundanzgründen an dieser Stelle auf eine ausführliche Beschreibung dieser Verfahren verzichtet und auf den Abschnitt 5.1.2.4 verwiesen.

Erwähnt und erläutert hingegen werden der Test zur Erfassung der Wahrnehmungsvorerfahrungen mit Pferdebewegungen, der Vorwissenstest und der Nachtest, bei denen für die Hauptstudie kleinere bis mittlere Änderungen vorgenommen wurden.

Test zur Erfassung der Wahrnehmungsvorerfahrungen mit Pferdebewegungen

Der Wahrnehmungsvorerfahrungstest wurde in seinen Gestaltungsmerkmalen in der gleichen Form aus der Pilotstudie übernommen (s. Abschnitt 5.1.2.4 und die Anhänge D2 - 1 sowie E2 - 1). Die Skala hatte eine zufriedenstellende Reliabilität von $\alpha = .77$ in der Pilotstudie gehabt. Die Trennschärfen waren, bis auf zwei Ausnahmen, zufriedenstellend bis gut gewesen. Einige der Schwierigkeitswerte waren sehr niedrig gewesen, was teilweise zu erwarten gewesen war, da die Probanden der Stichprobe bewusst wenige Erfahrungen mit Pferden hatten haben sollen, was sich in geringen Zustimmungsraten zu den Items des Tests geäußert hatte. Aus diesen Gründen wurden die Fragen in der gleichen Form wie in der Pilotstudie in die Hauptstudie übernommen, um sie nochmal an einer größeren Stichprobe mit stabileren Kennwertschätzungen durchzuführen. Die einzige Änderung innerhalb des Tests bestand darin, dass Item 6 („Gibt oder gab es in Ihrer Familie oder Ihrem Freundeskreis aktive Reiter?“) eliminiert wurde, so dass die Skala in der Hauptstudie nur noch aus zehn Items bestand. Nach nochmaliger Durchsicht der Fragen bestand die Vermutung, dass das Vorhandensein von Reitern in der Familie oder im Freundeskreis zu wenig über die eigenen Wahrnehmungserfahrungen der Probanden mit Pferden aussagt.

Wurde eine Frage mit „Ja“ beantwortet, so wurde dies in der Auswertung mit 1 Punkt gewertet, eine „Nein“-Antwort führte hingegen zu keinem Punkt. Der Gesamtscore des Tests, der die Summe aller Antworten darstellte, ging als eine womöglich miteinzubeziehende Kovariate in die statistischen Analysen ein.

Vorwissenstest

Die Items des Vorwissenstests wurden in identischer Form aus der Pilot- in die Hauptstudie übernommen und auch in vergleichbarer Weise durchgeführt (s. auch Anhang D2 - 2). Die Reliabilität der Skala war in der Pilotstudie schwächer gewesen ($\alpha = .38$), was eventuell durch die geringere Skalenzahl und durch die inhaltlich stärkere Heterogenität der Items erklärt werden könnte. Die Trennschärfen waren akzeptabel gewesen, und die Schwierigkeiten hatten im mittleren Bereich gelegen, so dass die Skala in eigenständiger Form in die Hauptstudie übernommen wurde, um ihre Güte an einer größeren Stichprobe zu testen.

Allerdings wurden die Ratingsysteme, mit denen die Antworten der Probanden in „richtig“ (2 Punkte), „zum Teil richtig“ (1 Punkt) oder „falsch / nicht beantwortet“ (0 Punkte) eingestuft wurden, kleineren Überarbeitungen aufgrund der empirischen Daten der Pilotstudie unterzogen. Dort hatte sich herausgestellt, dass die Probanden auf die zweite Frage, bei der sie eine Pferdegangart beschreiben mussten, eher undifferenzierte Antworten gegeben hatten, die häufig das Geschwindigkeitsmerkmal der jeweiligen Gangart (z. B. „Galopp ist die schnellste Gangart“) enthalten hatten. Daher wurde in den Ratingsystemen für jede Gangart die jeweilige Geschwindigkeitsinformation hinzugefügt, und die Ankerbeispiele wurden entsprechend überarbeitet. Ansonsten waren die Ratingsysteme zu Beginn der Hauptstudie in ihrem Inhalt und den Gestaltungsmerkmalen gleich mit denen der Pilotstudie (s. Abschnitt 5.1.2.4). Im Verlauf der Hauptstudie wurden, wie in der Pilotstudie, die Ratingsysteme weiter überarbeitet (zur Ansicht der nach der Hauptstudie weiterentwickelten Ratingsysteme, s. Anhang E3 auf der beigefügten CD).

Für die Ratingsysteme der beiden Items wurde jeweils eine Intraclass-Korrelation (ICC) bestimmt, da das Skalenniveau der einzelnen Stufen als metrisch angenommen wurde. Die Grundlage für die Berechnung der Werte war die Hälfte der Stichprobe, die vor der Ausreißeranalyse vorhanden gewesen war. Die Bögen wurden zufällig ausgewählt. Somit gingen in die Bestimmung der ICCs die Daten von 30 Versuchspersonen mit ein. Als zweiter Rater fungierte eine studentische Hilfskraft.

Für die weiteren Analysen wurden die Punkte, die auf den beiden Items jeweils erreicht wurden, addiert und stellten somit den Gesamtscore des Vorwissens einer Person dar.

Nachtest

Wie die Ergebnisse der Itemanalyse gezeigt hatten, waren die Aufgaben des Nachtests für die Probanden in der Pilotstudie sehr einfach gewesen. Über die Hälfte der Items hatte zu hohe Schwierigkeitsindizes gehabt ($p > .80$), wobei die Reliabilität der Skala gut ($\alpha = .86$) und die Trennschärfen größtenteils akzeptabel bis gut gewesen waren. Bei den Subskalen, die ausschließlich Animations- bzw. statische Bilditems enthielten, hatte sich ein ungefähr analoges Ergebnismuster abgezeichnet.

Daher wurde neben der Reduktion der Lernzeit auch die Itemreihenfolge des Nachtests umgekehrt (Aufgabe 32 wurde zu Aufgabe 1, Aufgabe 31 wurde zu Aufgabe 2, usw.), so dass in der Hauptstudie die konzeptuell schwierigeren Aufgaben am Anfang des Tests waren. Durch eine einfache Umkehrung der Itemreihenfolge wurden zwar die Positionen der vier Schwierigkeitszyklen verändert, aber die Sequenz der Aufgaben innerhalb der Zyklen sowie die Kombinatorik zwischen der Präsentationsart und der Farbe der Pferde, auf die bei der Testkonstruktion geachtet worden war, blieben unangetastet. Es war in der Pilotstudie aufgefallen, dass die inhaltlich schwierigeren Items keine niedrigeren Schwierigkeitswerte in der Itemanalyse gezeigt hatten und auch von den Probanden offensichtlich nicht als komplexer empfunden worden waren, da der Nachtest selbst anscheinend einen Übungseffekt für die Teilnehmer gehabt hatte, wie von manchen im Interview bekundet worden war. Die größere Schwierigkeit der Aufgaben war also durch ihre Position am Ende des Tests ausgeglichen worden, da die Teilnehmer innerhalb des Tests das Lösen der Aufgaben immer besser erlernt hatten. Deshalb wurden diese Aufgaben in der Hauptstudie an den Anfang des Nachtests plazierte, damit sie noch nicht von dem Übungseffekt betroffen und damit schwieriger für die Teilnehmer waren. Mit dieser Maßnahme sollte das Verfahren insgesamt diffiziler gemacht und der in der Pilotstudie aufgetretene „Ceiling-Effekt“ vermieden werden. Daher wurde auch nach der Pilotstudie keine Elimination einzelner Items vorgenommen, da angenommen wurde, dass sich durch diese Überarbeitungen die Gütekriterien der Aufgaben ändern würden, so dass eine vorzeitige Auswahl von Items als nicht sinnvoll erschien. Die Items an sich wurden also in ihren Merkmalen in gleichbleibender Form aus der Pilotstudie übernommen (s. Abschnitt 5.1.2.4).

Analog zu den beschriebenen Änderungen wurde auch die Instruktion des Nachtests überarbeitet.

Weitere, geringfügigere, Verbesserungen des Designs bestanden darin, dass die Pausen zwischen den Aufgaben des Nachtests, die dadurch entstanden, dass die Probanden nicht

gleich weiterklicken konnten, um die jeweils nächste Aufgabe zu sehen, verringert wurden. Die Pfeile zum Weiterklicken waren in der Pilotstudie immer 8 s lang gesperrt gewesen, um ein versehentliches zu schnelles Weiterkommen zur nächsten Aufgabe, ohne die vorherige gelöst zu haben, zu verhindern. Da die Teilnehmer in der Pilotstudie recht wenig Zeit zum Lösen der Aufgaben benötigt hatten, wurde diese Sperrzeit auf 4 s verringert.

Zuletzt wurde die Instruktion des Nachttests dahingehend abgeändert, dass die Versuchspersonen gebeten wurden, mit dem Starten der Testphase zu warten, bis der Versuchsleiter den Antwortbogen nochmals mündlich erklärt hatte. Ansonsten wurde an der Gestaltung und dem Ablauf des Tests sowie an dem Antwortbogen im Vergleich zur Pilotstudie nichts geändert (für eine Beschreibung der Testmerkmale, s. ebenfalls Abschnitt 5.1.2.4).

Für die Auswertung des Nachttests wurde jede richtige Antwort mit 1 Punkt gewertet, jede falsche Antwort hingegen mit 0 Punkten. Ein nicht beantwortetes Item stellte ein „missing value“ dar. Die bei den einzelnen Aufgaben erreichten Punkte wurden aufsummiert und waren als Gesamtscore eine abhängige Variable der Untersuchung. Um die Gesamtscores der beiden weiteren Kriterien „Animationsskala“ und „Bildskala“ zu erhalten, wurden die bei den Animationsitems und die bei den statischen Bilditems erzielten Punkte getrennt voneinander aufsummiert. Die Animationsskala bestand aus den Items 2, 4, 6, 8, 9, 12, 13, 16, 18, 20, 22, 24, 25, 28, 29 und 32, die Bildskala umfasste die Aufgaben 1, 3, 5, 7, 10, 11, 14, 15, 17, 19, 21, 23, 26, 27, 30 und 31 (für eine Ansicht des Nachttests, s. die Anhänge D2 – 4, E2 – 2 und E1 auf der beiliegenden CD, der Nachttest folgt in den Lerndateien direkt auf das Lernmaterial).

5.2.2.5 Durchführung

Die Experimente der Hauptstudie fanden im Wintersemester 2011 / 2012 in einem Zeitraum von ungefähr vier Wochen statt. Die Versuchsleitung wurde zwischen der Doktorandin selbst und einer vorab geschulten studentischen Hilfskraft aufgeteilt. Zudem stand ein detailliert beschriebenes Durchführungsschema zur Verfügung, das den Ablauf standardisieren und eine höhere Durchführungsobjektivität gewährleisten sollte. In räumlicher Hinsicht herrschten bei allen Untersuchungen vergleichbare Bedingungen, alle Experimente wurden im Labor der Abteilung „Allgemeine und Pädagogische Psychologie“ der Universität Landau abgehalten. Die Anwerbung der Probanden erfolgte über Rundmails,

die an alle Studenten der Universität versandt wurden. In den Mails wurden Personen mit zu viel Reit- oder Pferdeerfahrung darum gebeten, sich nicht zur Teilnahme zu melden, bestimmte Fachrichtungen wurden hingegen nicht vorab ausgeschlossen. Die Probanden konnten sich über die Lernplattform „OLAT“ selbständig für die Termine eintragen. Die Experimente fanden, je nach Zahl der Anmeldungen, in Einzel- oder Gruppentestungen bis zu vier Personen statt. Die Computer wurden vor dem Beginn der Studie technisch gewartet, und es wurde sichergestellt, dass alle Bedingungen vergleichbar abgespielt wurden.

Der Ablauf der Experimente an sich entsprach im Wesentlichen dem der Pilotstudie, weshalb für eine genauere Beschreibung auf Abschnitt 5.1.2.5 im Methodenteil der Pilotstudie verwiesen sei. Unterschiede bestanden lediglich in der kürzeren Lernzeit, die jetzt insgesamt nur noch ca. 1 min betrug, und darin, dass am Ende der Durchführungen kein Interview mit den Teilnehmern mehr geführt wurde. Die Experimente dauerten in den meisten Fällen ca. 1 h 15 min. Störungen innerhalb des Ablaufs wurden notiert und in ihrer Tragweite für die Auswertung der Ergebnisse mit berücksichtigt. Es kamen fast ausschließlich kleinere Störungen vor (z.B. Herunterfahren der Jalousien des Labors, Zuspätkommen von Probanden), die aber für das Lernen und für die Bearbeitungen der Tests keine größeren Beeinträchtigungen darstellten. Lediglich eine Versuchsperson schaute während der kurzen Lernphase zeitweise nicht auf den Bildschirm, was zu ihrem Ausschluss aus den weiteren Analysen führte. Generell verlief die Durchführung der Studie größtenteils problemlos, zu den einzelnen Verfahren wurden kaum Fragen gestellt, was auf deren gute Verständlichkeit hindeutete.

5.2.3 Ergebnisse

Die statistischen Berechnungen der Hauptstudie wurden mit dem Programm „SPSS 19“ durchgeführt.

Itemanalyse

Itemanalysen wurden in der Hauptstudie für den Vorerfahrungs- und den Vorwissenstest sowie für den Nachtest insgesamt und für die in ihm enthaltene Animations- und statische Bildskala gerechnet. Grundlage für die Analysen war stets die Gesamtstichprobe ($n = 58$). Es wurde für jede Skala ihr Cronbachs Alpha sowie die Trennschärfen und Schwierigkeiten ihrer Items bestimmt. Bei den beiden Items des Vorwissenstests, die ein halboffenes Antwortformat hatten, wurde für jedes der Ratingsysteme eine Intraclass-Korrelation (ICC) nach dem Kriterium der absoluten Übereinstimmung berechnet. In diese Analysen gingen die Daten von 30 Teilnehmern mit ein, wobei sich, je nach Zahl der „missing values“, die Stichprobe bei den Fragen etwas verringerte.

Für die weiteren Rechnungen wurden prinzipiell solche Items bevorzugt, deren Trennschärfen über $r_{it} = .30$ und deren Schwierigkeiten im mittleren Bereich ($p = .20 - p = .80$) lagen. Allerdings stellte sich heraus, dass im Vorerfahrungstest einige Items mit niedrigen Schwierigkeitswerten waren, was zu erwarten gewesen war, da die angeworbenen Teilnehmer nicht zu viel Vorerfahrung haben sollten. Im Nachtest, der Animations- und der Bildskala zeigte sich hingegen ein „Ceiling-Effekt“, das heißt, dass ungefähr die Hälfte der Aufgaben zu leicht für die Probanden war. Daher wurden die Schwierigkeiten der Aufgaben nicht als ein übermäßig strenges Kriterium zur Itemelimination angelegt, da die Skalenlängen sonst, konzeptuell gesehen, sehr kurz geworden wären. Ein maßgeblicheres Kriterium zum Ausschluss von Items waren deshalb die jeweiligen Trennschärfen der Aufgaben, die aufgrund des Deckeneffektes ebenfalls toleranter interpretiert wurden. Bei der Intraclass-Korrelation lag der Mindestwert eines reliablen Ratingsystems bei $r_{icc} = .70$.

Nach der Itemanalyse und Itemelimination zeigten die Skalen folgende Gütekriterien:

Tabelle 5.2.1

Gütekriterien des Vorerfahrungs- und des Vorwissenstests, des Nachtests sowie der Animations- und der Bildskala in der Hauptstudie

	<i>n</i> (Items)	α	<i>rit</i>	<i>p</i>	<i>rICC</i>
Vorerfahrungstest	7	.77	.32 - .67	.10 - .38	--
Vorwissenstest	2	.56	.39 - .39	.37 - .72	.39 – 1.00
Nachtest (gesamt)	30	.90	.21 - .73	.51 - .95	--
Animationsskala	16	.86	.26 - .75	.67 - .95	--
Bildskala	14	.78	.21 - .64	.51 - .88	--

An der Tabelle ist zu erkennen, dass für die weiteren Analysen drei Items des Vorerfahrungstests, sowie jeweils zwei Aufgaben des Nachtests und der Bildskala ausgeschlossen werden mussten. Ein Item des Vorwissenstests hatte eine schwächere Intraclass-Korrelation, wobei die Aufgabe aufgrund ihrer ansonsten zufriedenstellenden Gütekriterien beibehalten wurde. Die Reliabilitäten der Skalen waren, bis auf die des Vorwissenstests, als zufriedenstellend bis gut zu bezeichnen.

Hinsichtlich der standardisierten Verfahren ergaben sich für die Subskalen des FAM die folgenden Reliabilitäten nach der Itemelimination:

Tabelle 5.2.2

Reliabilitäten der Subskalen des FAM in der Hauptstudie

	<i>n</i> (Items)	α
Erfolgswahrscheinlichkeit	4	.76
Misserfolgsbefürchtung	5	.82
Interesse	5	.67
Herausforderung	3	.54

Bei der Subskala „Herausforderung“ war ein Item für die weiteren Rechnungen eliminiert worden, um ihre etwas schwächere Reliabilität zu steigern.

Ausreißeranalyse

Für die fünf Kriterien und die neun Kontrollvariablen der Studie wurden jeweils Ausreißeranalysen mittels Box Plots durchgeführt. Für jede der Kontrollvariablen wurde ein Plot für die gesamte Stichprobe erstellt, um die Lage der Daten zu beurteilen. Bei den Kriterien wurden zusätzlich bedingungsabhängige Plots herangezogen, um mögliche Ausreißer zu eliminieren. Es wurden solche Fälle von den weiteren Analysen ausgeschlossen, die jenseits der Whisker der Box lagen und die von SPSS als Ausreißerwerte angezeigt wurden, wobei die Auswahl eher konservativ getätigt wurde, um eine möglichst hohe Fallzahl in der Stichprobe und eine höhere Power der nachfolgenden Verfahren zu gewährleisten. Somit verringerte sich die Zahl der Fälle nach der Ausreißeranalyse um sechs Personen, so dass 52 Teilnehmer in die weiteren Analysen eingingen.

Deskriptive Statistiken

Das Durchschnittsalter der Stichprobe ($n = 52$) lag bei $M = 23.08$ Jahren ($SD = 4.03$) bei einem Range von 20 bis 45 Jahren. Unter den Teilnehmern waren 41 weiblich und 11 männlich, was einem Geschlechterverhältnis von 78,8% Frauen zu 21,2% Männern entsprach.

Zunächst werden die deskriptiven Statistiken der abhängigen Variablen, dann die der Kontrollvariablen aufgelistet:

Tabelle 5.2.3

Deskriptive Statistiken der abhängigen Variablen in der Hauptstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Nachtest (gesamt)	51	8.00	30.00	23.14	6.17
Animationsskala	52	3.00	16.00	13.13	3.53
Bildskala	51	2.00	14.00	10.06	3.13
Cognitive Load (Lernaufgabe)	52	4.00	16.00	11.25	3.33
Cognitive Load (Nachtest)	52	6.00	16.00	12.52	2.50

Tabelle 5.2.4

Deskriptive Statistiken der Kontrollvariablen in der Hauptstudie

	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>
Wahrnehmungs- vorerfahrungstest	52	.00	5.00	1.19	1.56
Vorwissenstest	52	.00	4.00	2.15	1.04
IST verbal	52	16.00	33.00	25.21	4.06
IST räumlich	52	5.00	35.00	22.48	6.26
IST Merkfähigkeit	51	10.00	23.00	17.29	3.23
Erfolgswahrscheinlichkeit	50	10.00	28.00	21.90	3.76
Misserfolgsbefürchtung	50	5.00	23.00	11.50	4.81
Interesse	50	14.00	30.00	22.26	3.82
Herausforderung	51	9.00	20.00	15.10	2.35

Mittelwertsvergleiche

Testung auf Normalverteilung und Varianzhomogenität

Da für die Mittelwertsvergleiche zwischen den Bedingungen „Animation“ und „Bild parallel“ hinsichtlich ihrer Haupteffekte auf die fünf Kriterien der Studie t-Tests bzw., bei Hinzunahme relevanter Kovariaten, „Allgemeine lineare Modelle (ALMs)“ gerechnet werden sollten, wurden zunächst die Normalverteilungen der Kriterien und die Varianzhomogenitäten zwischen den Bedingungen als dafür notwendige Voraussetzungen berechnet. Hierzu wurden Kolmogorov-Smirnov-Tests und Levene-Tests herangezogen.

Die Ergebnisse der Kolmogorov-Smirnov-Tests zeigten, dass bei den Kriterien, bis auf die abhängige Variable „Animationsskala“, Normalverteilungen vorlagen. Bei der Animationsskala wich die Verteilung der Daten deutlich von der Normalität ab (Kolmogorov-Smirnov-Z: 1.81, $p = .003$).

Die Berechnung von Levene-Tests ergab, dass die Varianzen zwischen den Bedingungen bei vier der fünf Kriterien homogen waren. Lediglich bei der abhängigen Variablen „Cognitive Load beim Nachtest“ waren die Varianzen knapp heterogen (Levene-Statistik: 4.23, $p = .05$).

Da bei keinem der Kriterien sowohl Varianzheterogenität als auch Nichtnormalverteilung der Daten vorlag, konnten t-Tests und ALMs angewandt werden.

Bestimmung relevanter Kovariaten

Bei den Kontrollvariablen wurden zunächst t-Tests bzw., bei gegebener Varianzheterogenität, w-Tests gerechnet, um größere oder signifikante Mittelwertsunterschiede zwischen den Gruppen herauszufiltern. Die einzige, fast signifikante, Mittelwertsdifferenz lag bei der Kontrollvariable „Interesse“ vor ($t(48) = -1.97$, p (2-seitig) = .06, $d = .55$). Diese korrelierte signifikant mit den abhängigen Variablen „Nachtest gesamt“ ($r = .27$, p (1-seitig) = .03) und „Bildskala“ ($r = .27$, p (1-seitig) = .03). Daher wurden bei den Kriterien zur Bestimmung ihrer Mittelwertsunterschiede folgende Verfahren gerechnet:

- | | |
|------------------------|--------------------------------|
| 1. AV: Nachtest gesamt | Verfahren: ALM mit „Interesse“ |
| 2. AV: Animationsskala | Verfahren: t-Test |
| 3. AV: Bildskala | Verfahren: ALM mit „Interesse“ |
| 4. AV: CL Lern | Verfahren: t-Test |
| 5. AV: CL Nachtest | Verfahren: t-Test |

„CL Lern“ steht für den Cognitive Load, den die Probanden bei der Betrachtung des Lernmaterials empfunden haben. Die AV „CL Nachtest“ bezeichnet hingegen den Cognitive Load, der nach der Beantwortung des Nachtests bei den Probanden gemessen wurde. In den folgenden Abschnitten werden die Abkürzungen „CL Lern“ und „CL Nachtest“ in den meisten Fällen gebraucht, um die entsprechenden Kriterien zu benennen.

t-Tests und ALMs

Die t-Tests und ALMs wurden jeweils mit einem Signifikanzniveau von $\alpha = .05$ durchgeführt.

Zunächst werden nochmals die Hypothesen des Theorieteils in statistischer Kurzform genannt und die dazugehörigen Resultate jeweils direkt angeschlossen.

Hypothese 1a: Animation > Bild parallel; AV = Nachtest gesamt (Animations- + Bildskala)

Tabelle 5.2.5

Adjustierte Mittelwerte und Standardfehler der Bedingungen in der AV „Nachtest gesamt“

	<i>n</i>	<i>M</i>	<i>SE</i>
Bild parallel	24	22.41	1.18
Animation	25	24.25	1.16

Der Unterschied zwischen den beiden Bedingungen wurde nicht signifikant ($F(1, 46) = 1.20$, p (1-seitig) = .14, partielles $\eta^2 = .03$). Wie vorhergesagt, erreichte die Animationsbedingung einen höheren Wert im Nachtest als die statische Bildbedingung.

Hypothese 1b: Animation > Bild parallel; AV = Animationsskala

Tabelle 5.2.6

Mittelwerte und Standardabweichungen der Bedingungen in der AV „Animationsskala“

	<i>n</i>	<i>M</i>	<i>SD</i>
Bild parallel	25	12.48	3.87
Animation	27	13.74	3.12

Die Mittelwertsdifferenz war zwar größer, aber nicht signifikant ($t(50) = 1.30$, p (1-seitig) = .10, $d = .36$). Die Gruppe „Animation“ erzielte bessere Ergebnisse als die Gruppe „Bild parallel“.

Hypothese 1c: Animation > Bild parallel; AV = Bildskala

Tabelle 5.2.7

Adjustierte Mittelwerte und Standardfehler der Bedingungen in der AV „Bildskala“

	<i>n</i>	<i>M</i>	<i>SE</i>
Bild parallel	24	9.76	.63
Animation	25	10.47	.62

Zwischen den Bedingungen lag keine signifikante Mittelwertsdifferenz vor ($F(1, 46) = .64$, p (1-seitig) = .22, partielles $\eta^2 = .01$). Die Personen, die mit Animationen lernten, waren gegenüber denen, die mit parallelen Key-Frames lernten, leicht im Vorteil.

Hypothese 1d: Animation < Bild parallel; AV = CL Lern

Tabelle 5.2.8

Mittelwerte und Standardabweichungen der Bedingungen in der AV „CL Lern“

	<i>n</i>	<i>M</i>	<i>SD</i>
Animation	27	9.85	2.97
Bild parallel	25	12.76	3.07

Der Unterschied zwischen den beiden Gruppen wurde in der vorhergesagten Richtung hochsignifikant ($t(50) = -3.47$, $p(1\text{-seitig}) < .001$, $d = .96$). Die Gruppe „Animation“ verspürte beim Lernen viel weniger mentale Belastung als die Gruppe „Bild parallel“.

Hypothese 1e: Animation < Bild parallel; AV = CL Nachtest

Tabelle 5.2.9

Mittelwerte und Standardabweichungen der Bedingungen in der AV „CL Nachtest“

	<i>n</i>	<i>M</i>	<i>SD</i>
Animation	27	12.07	2.95
Bild parallel	25	13.00	1.85

Auch hier war der von den Probanden empfundene Cognitive Load beim Lösen des Nachtests in der Animationsbedingung geringer als in der statischen Bildbedingung. Allerdings war die Mittelwertsdifferenz nicht signifikant ($t(50) = -1.34$, $p(1\text{-seitig}) = .10$, $d = .38$).

Analyse der Interaktion zwischen dem Format der Lernbedingung und dem Format der Nachtestskalen

Hypothese 2: Interaktion Lernformat x Nachtestformat ist signifikant, gerichtete Fragestellung

Um die Stärke und die Richtung der Interaktion zwischen dem Darstellungsformat der Lernmaterialien und dem Format der Nachtestskalen zu bestimmen, wurde eine zweifaktorielle Kovarianzanalyse mit Messwiederholung gerechnet. Der Zwischenssubjektfaktor war die Versuchsbedingung (Animation vs. Bild parallel), die beiden Subskalen des Nachtests (Animationsskala vs. Bildskala) stellten den Innersubjektfaktor dar. In die Analyse wurde die Variable „Interesse“ als Kovariate mit aufgenommen, da sie einfaktoriellement die Ergebnisse der abhängigen Variablen „Bildskala“ mit beeinflusste.

Tabelle 5.2.10

Zellenmittelwerte und Standardfehler in der zweifaktoriellen ANCOVA mit Messwiederholung

Bedingung	Subskala	<i>n</i>	<i>M</i>	<i>SE</i>
Animation	Animationsskala	25	13.78	.65
	Bildskala	25	10.47	.62
Bild parallel	Animationsskala	24	12.65	.67
	Bildskala	24	9.76	.63

Die Interaktion zwischen dem Format der Lernbedingung und dem der Nachtestskalen wurde nicht signifikant (Wilks lambda = .99, $F(1,46) = .30$, $p(1\text{-seitig}) = .29$, partielles $\eta^2 = .01$). Allerdings lag das deskriptive Muster der Interaktion in der vorhergesagten Richtung: Der Lernvorteil durch Animationen war auf der Animationsskala des Nachtests stärker ausgeprägt als auf der statischen Bildskala.

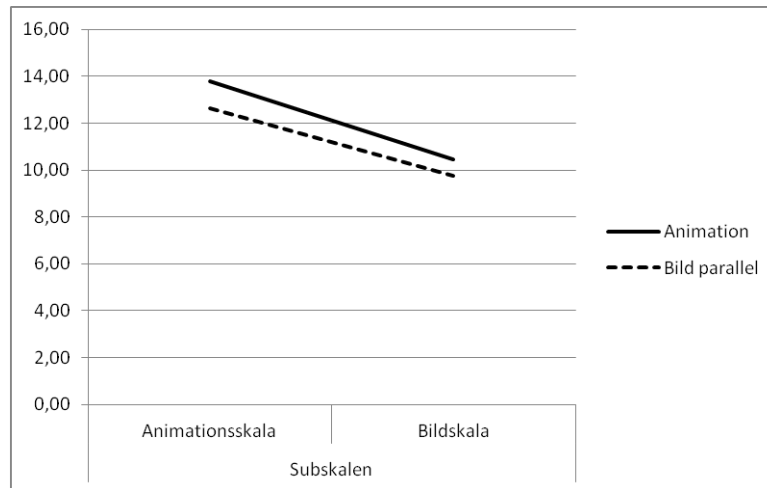


Abbildung 5.2.1: Graphische Veranschaulichung der Interaktion zwischen dem Format der Lernbedingung und dem Format der Nachtestskalen in der Hauptstudie

Weitere Interaktionsanalysen

Die Interaktionsanalysen wurden wiederum mit Hilfe multipler Regressionen durchgeführt, wobei das Vorgehen analog zu der Beschreibung von Richter (2007) war. Die Bedingungsvariable „Animation vs. Bild parallel“ wurde gewichtet effektkodiert, um die leicht unterschiedlichen Stichprobenumfänge in den beiden Gruppen zu berücksichtigen. Die metrischen Variablen wurden jeweils zentriert. Durch Eingabe dieser Variablen sowie der zugehörigen Kovariaten aus den vorangegangenen Analysen in ein Regressionsmodell wurden die jeweiligen Interaktionen bestimmt. Post-hoc wurde die Bedingungsvariable auf zwei Arten dummykodiert, wobei jede Bedingung in jeweils einem Fall Referenzgruppe war. Anschließende Regressionsmodelle ergaben den Einfluss der metrischen Variablen in den beiden Gruppen. Um die Unterschiede zwischen der Animations- und der statischen Bildgruppe bei höheren oder niedrigeren Ausprägungen der metrischen Variablen zu errechnen, wurde die Bedingungsvariable ebenfalls dummykodiert (die Animationsbedingung war Referenzgruppe), die metrische Variable wurde in Abweichungswerten von +1SD und -1SD der Verteilung kodiert und in jeweils ein Regressionsmodell eingegeben. Bei den Interaktionen mit der Variable „Wahrnehmungsvorerfahrung“ wurden Abweichungswerte von -0.76 SD verwendet, um die Gruppenunterschiede bei einer geringeren Ausprägung dieses Merkmals zu bestimmen, da bei -1 SD Aussagen über einen negativen Testwert gemacht worden wären, was empirisch nicht sinnvoll ist. -0.76 SD entsprachen einem Vorerfahrungsscore von Null.

Größere und signifikante Interaktionen werden in der Ergebnisdarstellung durch Tabellen und Liniendiagramme veranschaulicht.

Interaktionen zwischen den Faktoren „Bedingung“ und „Wahrnehmungsvorerfahrung“

Interaktionen zwischen den Gruppen „Animation“ und „Bild parallel“ mit der Variable „Wahrnehmungsvorerfahrung“

Hypothese 3a: Interaktion „Animation vs. Bild parallel“ x Wahrnehmungsvorerfahrung ist signifikant, gerichtete Fragestellung, AV = Nachtest gesamt (Animations- + Bildskala)

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(4, 44) = 2.14, p = .09, R^2 = .16, adj R^2 = .09$).

Das Interaktionsmuster lag nicht in der vorhergesagten Richtung und war bei zweiseitiger Testung auch nicht signifikant ($B = .04, SE = .49, t = .08, p$ (2-seitig) = .94). Deskriptiv war der Lernvorteil durch Animationen größer bei höherer als bei niedrigerer Wahrnehmungsvorerfahrung (Höhere Vorerfahrung: $B = -1.51, p$ (2-seitig) = .51; Niedrigere Vorerfahrung: $B = -1.28, p$ (2-seitig) = .55).

Hypothese 3b: Interaktion „Animation vs. Bild parallel“ x Wahrnehmungsvorerfahrung ist signifikant, gerichtete Fragestellung, AV = Animationsskala

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 48) = 2.84, p = .05, R^2 = .15, adj R^2 = .10$).

Die Interaktion zwischen den Gruppen „Animation“ und „Bild parallel“ mit der Wahrnehmungsvorerfahrung der Probanden war bei dem Kriterium „Animationsskala“ nicht signifikant, lag aber in der vorhergesagten Richtung ($B = -.34, SE = .29, t = -1.15, p$ (1-seitig) = .13). Deskriptiv zeigten die Probanden mit höherer Vorerfahrung ungefähr gleich gute Ergebnisse in den beiden Bedingungen ($B = .04, p$ (2-seitig) = .98), während bei geringerer Vorerfahrung die parallele Bildgruppe stärker im Nachteil war ($B = -1.88, p$ (2-seitig) = .12).

Hypothese 3c: Interaktion „Animation vs. Bild parallel“ x Wahrnehmungsvorerfahrung ist signifikant, gerichtete Fragestellung, AV = Bildskala

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(4,44) = 1.87, p = .13, R^2 = .15, adj R^2 = .07$).

Die Interaktion zwischen den beiden Gruppen und der Wahrnehmungsvorerfahrung der Probanden lag nicht in der vorhergesagten Richtung und war auch bei zweiseitiger Testung nicht signifikant ($B = .26, SE = .26, t = .10, p$ (2-seitig) = .33). Bei höheren Vorerfahrungswerten war die Animationsbedingung deskriptiv im Vorteil, während bei niedrigerer Vorerfahrung in geringem Ausmaß das Gegenteil der Fall war (Höhere Vorerfahrung: $B = -1.35, p$ (2-seitig) = .27; Niedrigere Vorerfahrung: $B = .15, p$ (2-seitig) = .90).

Hypothese 3d: Interaktion „Animation vs. Bild parallel“ x Wahrnehmungsvorerfahrung ist signifikant, gerichtete Fragestellung, AV = CL Lern

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 48) = 4.01, p = .01, R^2 = .20, adj R^2 = .15$).

Die Interaktion zwischen den beiden Gruppen und der Wahrnehmungsvorerfahrung der Probanden lag nicht in der vorhergesagten Richtung und erreichte bei zweiseitiger Testung keine Signifikanz ($B = -.05, SE = .27, t = -.18, p$ (2-seitig) = .86). Deskriptiv war der Lernvorteil durch Animationen kleiner bei weniger Vorerfahrung als bei höherer Vorerfahrung (Niedrigere Vorerfahrung: $B = 2.74, p$ (2-seitig) = .01; Höhere Vorerfahrung: $B = 3.02, p$ (2-seitig) = .02).

Hypothese 3e: Interaktion „Animation vs. Bild parallel“ x Wahrnehmungsvorerfahrung ist signifikant, gerichtete Fragestellung, AV = CL Nachtest

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 48) = 3.15, p = .03, R^2 = .17, adj R^2 = .11$).

Die Interaktion zwischen den beiden Bedingungen und der Wahrnehmungsvorerfahrung lag nicht in der vorhergesagten Richtung und nahm bei zweiseitiger Testung keinen signifikanten Wert an ($B = -.24, SE = .20, t = -1.15, p$ (2-seitig) = .26). Sowohl Personen mit höherer als auch solche mit niedrigerer Vorerfahrung verspürten deskriptiv mehr Cognitive Load in der parallelen Bildbedingung, wobei diese Differenz bei höherer Vorerfahrung

stärker ausgeprägt war (Höhere Vorerfahrung: $B = 1.54$, p (2-seitig) = .11; Niedrigere Vorerfahrung: $B = .19$, p (2-seitig) = .82).

Interaktionen zwischen den Faktoren „Bedingung“ und „Räumliche Fähigkeiten“

Interaktionen zwischen den Gruppen „Animation“ und „Bild parallel“ mit der Variable „Räumliche Fähigkeiten“

Hypothese 4a: Interaktion „Animation vs. Bild parallel“ x Räumliche Fähigkeiten ist signifikant, gerichtete Fragestellung, AV = Nachtest gesamt (Animations- + Bildskala)

Tabelle 5.2.11

Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild parallel x Räumliche Fähigkeiten“ bei der AV „Nachtest gesamt“

Variable	B	$SE B$	t	p
Bedingung: Animation vs. Bild parallel	1.06	.80	1.33	.19
Räumliche Fähigkeiten	.20	.13	1.52	.14
Interesse	.46	.22	2.09	.04
Interaktion Bedingung x Räuml. Fähigkeiten	-.22	.13	-1.73	.05

Der multiple Korrelationskoeffizient des Gesamtmodells war fast signifikant ($F(4, 44) = 2.49$, $p = .06$, $R^2 = .18$, $adj R^2 = .11$).

Die Interaktion zwischen den beiden Gruppen und den räumlichen Fähigkeiten lag bei der AV „Nachtest gesamt“ in der vorhergesagten Richtung und war signifikant ($B = -.22$, p (1-seitig) = .05). Die räumlichen Fähigkeiten hatten einen stärker positiven Einfluss in der parallelen Bildbedingung als in der Animationsbedingung (Bild parallel: $B = .45$, p (2-seitig) = .03; Animation: $B = -.02$, p (2-seitig) = .91). Personen mit geringeren räumlichen Fähigkeiten hatten signifikant niedrigere Werte in der parallelen Bildbedingung ($B = -5.11$, p (2-seitig) = .04), während Personen mit höheren räumlichen Fähigkeiten bei den Animationen leicht im Nachteil waren ($B = .71$, p (2-seitig) = .77).

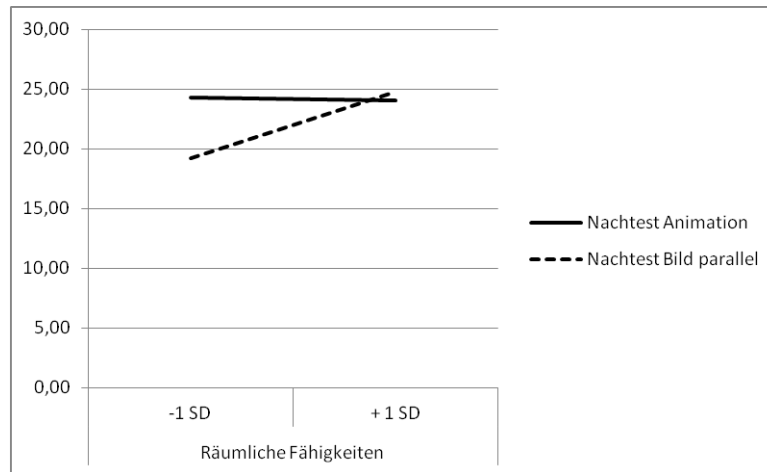


Abbildung 5.2.2: Regressionsgeraden der räumlichen Fähigkeiten in den Gruppen „Animation“ und „Bild parallel“ (AV: Nachtest gesamt)

Hypothese 4b: Interaktion „Animation vs. Bild parallel“ x Räumliche Fähigkeiten ist signifikant, gerichtete Fragestellung, AV = Animationsskala

Tabelle 5.2.12

Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild parallel x Räumliche Fähigkeiten“ bei der AV „Animationsskala“

Variable	<i>B</i>	<i>SE B</i>	<i>t</i>	<i>p</i>
Bedingung: Animation vs. Bild parallel	.69	.44	1.57	.12
Räumliche Fähigkeiten	.19	.07	2.56	.01
Interaktion Bedingung x Räuml. Fähigkeiten	-.10	.07	-1.45	.08

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 48) = 3.30$, $p = .03$, $R^2 = .17$, $adj R^2 = .12$).

Die Interaktion lag bei dem Kriterium „Animationsskala“ in der vorhergesagten Richtung und nahm einen annähernd signifikanten Wert an ($B = -.10$, p (1-seitig) = .08). Der Regressionskoeffizient der Variable „Räumliche Fähigkeiten“ war in der Gruppe „Bild parallel“ stärker positiv als in der Gruppe „Animation“ (Bild parallel: $B = .30$, p (2-seitig) = .01; Animation: $B = .09$, p (2-seitig) = .39). Personen mit niedrigeren

räumlichen Fähigkeiten hatten signifikant weniger Punkte auf der Animationsskala in der parallelen Bildbedingung ($B = -2.80$, p (2-seitig) = .04), während bei höheren räumlichen Fähigkeiten die parallele Bildbedingung nur ganz leicht im Nachteil war ($B = -.09$, p (2-seitig) = .95).

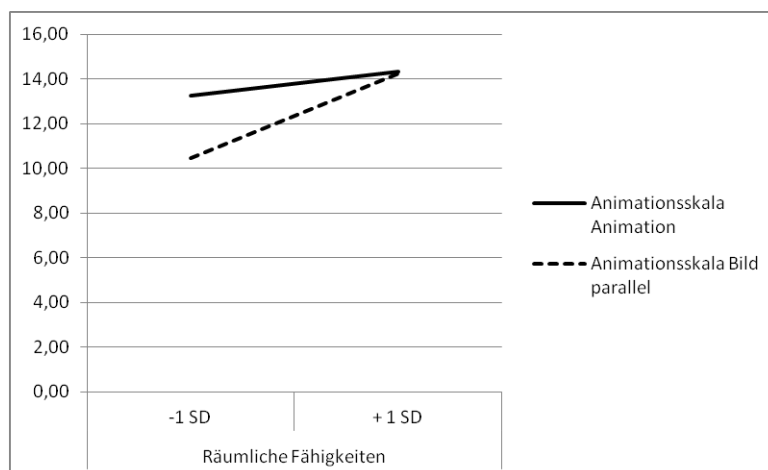


Abbildung 5.2.3: Regressionsgeraden der räumlichen Fähigkeiten in den Gruppen „Animation“ und „Bild parallel“ (AV: Animationsskala)

Hypothese 4c: Interaktion „Animation vs. Bild parallel“ x Räumliche Fähigkeiten ist signifikant, gerichtete Fragestellung, AV = Bildskala

Tabelle 5.2.13

Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild parallel x Räumliche Fähigkeiten“ bei der AV „Bildskala“

Variable	<i>B</i>	<i>SE B</i>	<i>t</i>	<i>p</i>
Bedingung: Animation vs. Bild parallel	.41	.42	.98	.33
Räumliche Fähigkeiten	.08	.07	1.13	.27
Interesse	.24	.11	2.07	.04
Interaktion Bedingung x Räuml. Fähigkeiten	-.16	.07	-2.28	.02

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(4, 44) = 2.68$, $p = .04$, $R^2 = .20$, $adj R^2 = .12$).

Die Interaktion zwischen den Gruppen „Animation“ und „Bild parallel“ mit den räumlichen Fähigkeiten der Probanden lag bei dem Kriterium „Bildskala“ in der vorhergesagten Richtung und war signifikant ($B = -.16$, p (1-seitig) = .02). Die räumlichen Fähigkeiten beeinflussten die Leistung stärker positiv in der parallelen Bild- als in der Animationsbedingung (Bild parallel: $B = .25$, p (2-seitig) = .03; Animation: $B = -.08$, p (2-seitig) = .42). Bei höheren räumlichen Fähigkeiten war die parallele Bildbedingung deskriptiv im Vorteil ($B = 1.16$, p (2-seitig) = .35), während für Personen mit niedrigeren räumlichen Fähigkeiten Animationen in signifikantem Ausmaß geeigneter waren ($B = -2.86$, p (2-seitig) = .03).

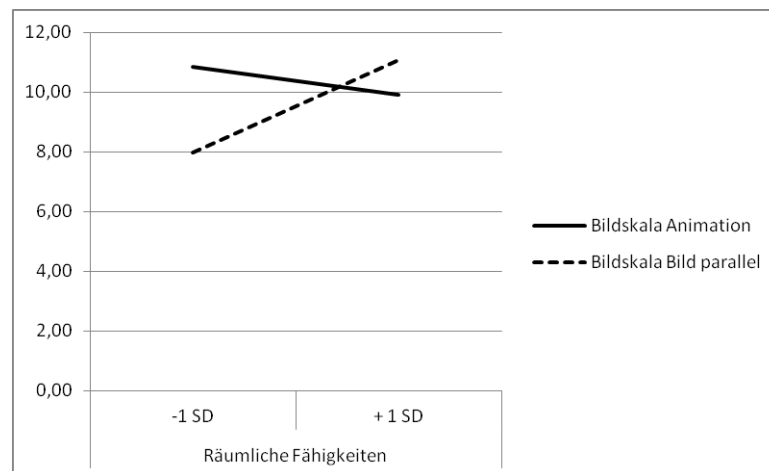


Abbildung 5.2.4: Regressionsgeraden der räumlichen Fähigkeiten in den Gruppen „Animation“ und „Bild parallel“ (AV: Bildskala)

Hypothese 4d: Interaktion „Animation vs. Bild parallel“ x Räumliche Fähigkeiten ist signifikant, gerichtete Fragestellung, AV = CL Lern

Der multiple Korrelationskoeffizient des Gesamtmodells war sehr signifikant ($F(3, 48) = 5.46$, $p = .003$, $R^2 = .26$, $adj R^2 = .21$).

Die Interaktion zwischen den beiden Bedingungen und der Variable „Räumliche Fähigkeiten“ lag bei dem Kriterium „CL Lern“ in der vorhergesagten Richtung, nahm aber keinen signifikanten Wert an ($B = .06$, $SE = .06$, $t = .90$, p (1-seitig) = .19). Bei höheren räumlichen Fähigkeiten verspürten die Probanden bei parallelen Key-Frames fast signifikant

mehr Cognitive Load als bei Animationen ($B = 2.27$, p (2-seitig) = .06). Bei geringeren räumlichen Fähigkeiten war der Nachteil durch die Key-Frames noch stärker ausgeprägt, so dass der Unterschied sehr signifikant wurde ($B = 3.78$, p (2-seitig) = .002).

Hypothese 4e: Interaktion „Animation vs. Bild parallel“ x Räumliche Fähigkeiten ist signifikant, gerichtete Fragestellung, AV = CL Nachtest

Tabelle 5.2.14

Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild parallel x Räumliche Fähigkeiten“ bei der AV „CL Nachtest“

Variable	<i>B</i>	<i>SE B</i>	<i>t</i>	<i>p</i>
Bedingung: Animation vs. Bild parallel	-.49	.32	-1.51	.14
Räumliche Fähigkeiten	-.09	.05	-1.64	.11
Interaktion Bedingung x Räuml. Fähigkeiten	.08	.05	1.60	.06

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 48) = 2.24$, $p = .10$, $R^2 = .12$, $adj R^2 = .07$).

Die Interaktion zwischen den beiden Gruppen und der Variable „Räumliche Fähigkeiten“ lag in der vorhergesagten Richtung und wurde fast signifikant ($B = .08$, p (1-seitig) = .06). Der Einfluss der räumlichen Fähigkeiten der Probanden auf den Cognitive Load war stärker negativ in der parallelen Bild- als in der Animationsgruppe (Bild parallel: $B = -.18$, p (2-seitig) = .03; Animation: $B = -.01$, p (2-seitig) = .94). Bei niedrigeren räumlichen Fähigkeiten war die parallele Bildbedingung signifikant im Nachteil ($B = 2.11$, p (2-seitig) = .03), während bei höheren räumlichen Fähigkeiten die Probanden in beiden Gruppen ungefähr gleich viel Cognitive Load verspürten ($B = -.08$, p (2-seitig) = .94).

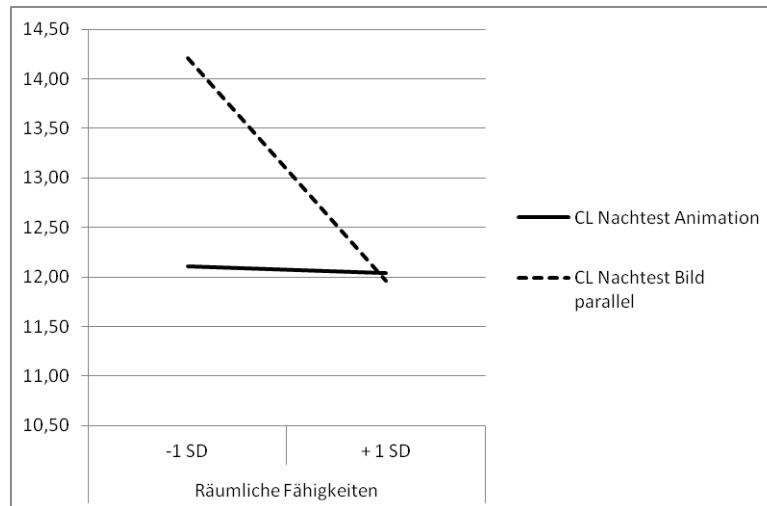


Abbildung 5.2.5: Regressionsgeraden der räumlichen Fähigkeiten in den Gruppen „Animation“ und „Bild parallel“ (AV: CL Nachtest)

Interaktionen zwischen den Faktoren „Bedingung“ und „Merkfähigkeit“

Interaktionen zwischen den Gruppen „Animation“ und „Bild parallel“ mit der Variable „Merkfähigkeit“

Hypothese 5a: Interaktion „Animation vs. Bild parallel“ x Merkfähigkeit ist signifikant, gerichtete Fragestellung, AV = Nachtest gesamt (Animations- + Bildskala)

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(4, 43) = 1.69, p = .17, R^2 = .14, adj R^2 = .06$).

Die Interaktion zwischen den Bedingungen „Animation“ und „Bild parallel“ mit der Merkfähigkeit lag bei der AV „Nachtest gesamt“ in der vorhergesagten Richtung, erreichte aber keinen signifikanten Wert ($B = -.17, SE = .25, t = -.68, p$ (1-seitig) = .25). Der Lernvorteil durch Animationen war bei geringerer Merkfähigkeit stärker ausgeprägt als bei höherer Merkfähigkeit (Niedrigere Merkfähigkeit: $B = -2.80, p$ (2-seitig) = .24; Höhere Merkfähigkeit: $B = -.47, p$ (2-seitig) = .85).

Hypothese 5b: Interaktion „Animation vs. Bild parallel“ x Merkfähigkeit ist signifikant, gerichtete Fragestellung, AV = Animationskala

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 47) = .85, p = .47, R^2 = .05, adj R^2 = -.01$).

Die Interaktion zwischen den Gruppen „Animation“ und „Bild parallel“ mit der Merkfähigkeit der Probanden lag bei dem Kriterium „Animationsskala“ in der vorhergesagten Richtung, zeigte aber keine statistische Signifikanz ($B = -.06$, $SE = .16$, $t = -.39$, p (1-seitig) = .35). Der Lernnachteil durch parallele Key-Frames war bei geringerer Merkfähigkeit deskriptiv größer als bei höherer Merkfähigkeit (Niedrigere Merkfähigkeit: $B = -1.66$, p (2-seitig) = .26; Höhere Merkfähigkeit: $B = -.85$, p (2-seitig) = .56).

Hypothese 5c: Interaktion „Animation vs. Bild parallel“ x Merkfähigkeit ist signifikant, gerichtete Fragestellung, AV = Bildskala

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(4, 43) = 1.48$, $p = .22$, $R^2 = .12$, $adj R^2 = .04$).

Zwischen den Gruppen „Animation“ und „Bild parallel“ und der Merkfähigkeit der Teilnehmer lag eine Interaktion in der vorhergesagten Richtung vor, die nicht signifikant wurde ($B = -.06$, $SE = .14$, $t = -.47$, p (1-seitig) = .32). Deskriptiv war bei niedrigerer Merkfähigkeit der Vorteil durch das Lernen mit Animationen höher als bei stärker ausgeprägten Gedächtnisfähigkeiten (Niedrigere Merkfähigkeit: $B = -1.01$, p (2-seitig) = .42; Höhere Merkfähigkeit: $B = -.17$, p (2-seitig) = .90).

Hypothese 5d: Interaktion „Animation vs. Bild parallel“ x Merkfähigkeit ist signifikant, gerichtete Fragestellung, AV = CL Lern

Der multiple Korrelationskoeffizient des Gesamtmodells war signifikant ($F(3, 47) = 4.15$, $p = .01$, $R^2 = .21$, $adj R^2 = .16$).

Die Interaktion zwischen der Bedingungsvariable und der Merkfähigkeit der Probanden entsprach beim Kriterium „CL Lern“ der vorhergesagten Richtung, wies aber keine Signifikanz auf ($B = .04$, $SE = .13$, $t = .32$, p (1-seitig) = .38). Sowohl die Personen mit höherer als auch die mit niedrigerer Merkfähigkeit verspürten signifikant mehr Cognitive Load beim Lernen mit parallelen Bildern, wobei dieser Effekt stärker bei geringerer Merkfähigkeit war (Höhere Merkfähigkeit: $B = 2.70$, p (2-seitig) = .04; Niedrigere Merkfähigkeit: $B = 3.28$, p (2-seitig) = .01).

Hypothese 5e: Interaktion „Animation vs. Bild parallel“ x Merkfähigkeit ist signifikant, gerichtete Fragestellung, AV = CL Nachtest

Der multiple Korrelationskoeffizient des Gesamtmodells war nicht signifikant ($F(3, 47) = .67, p = .58, R^2 = .04, adj R^2 = -.02$).

Die Interaktion zwischen den beiden Gruppen mit der Variable „Merkfähigkeit“ lag bei der AV „CL Nachtest“ in der vorhergesagten Richtung, wies aber keinen signifikanten Wert auf ($B = .06, SE = .11, t = .54, p$ (1-seitig) = .30). Deskriptiv war der Lernvorteil durch Animationen größer bei geringerer als bei höherer Merkfähigkeit (Niedrigere Merkfähigkeit: $B = 1.28, p$ (2-seitig) = .22; Höhere Merkfähigkeit: $B = .49, p$ (2-seitig) = .64).

5.2.4 Diskussion

Haupteffekte unter Einbezug der Interaktion zwischen dem Präsentationsformat der Lernbedingung und dem Präsentationsformat der Nachtestskalen

Animation vs. Parallele Key-Frames

In den Hypothesen 1a – e war davon ausgegangen worden, dass Personen, die die Pferdegangarten mit Hilfe von Animationen erlernten, bessere Leistungen in einem Nachtest mit bildlichen Kategorisierungsaufgaben und der in ihm enthaltenen Animations- und Bildskala zeigen sollten als Personen, die hierfür statische Key-Frames betrachteten. Zudem sollten sie weniger Cognitive Load beim Lernen und beim Lösen des Nachtests verspüren. Ein Grund für diese Unterschiede sollte die Annahme sein, dass mit Hilfe von Animationen perzeptuelle dynamische Schemata leichter aufgebaut werden. Die Hypothese 2 thematisierte hingegen die Interaktion zwischen dem Präsentationsformat des Lernmaterials und dem der Nachtestskalen. Hier wurde angenommen, dass aufgrund von Transfereffekten, der Lernvorteil durch Animationen auf der Animationsskala des Nachtests stärker ausgeprägt wäre als auf der statischen Bildskala des Nachtests.

Die Ergebnisse der Hauptstudie zeigten, dass vier der fünf Haupteffektshypothesen nicht bestätigt werden konnten. Deskriptiv gesehen, lagen die Unterschiede in der vorhergesagten Richtung, d.h., die Personen, die mit Animationen lernten, erzielten bessere Lernergebnisse als die Personen, die mit parallelen Key-Frames lernten, allerdings wurden die Differenzen nicht signifikant. Die Größe der Effektstärken für die Kriterien „Nachtest gesamt“, „Animationsskala“, „Bildskala“ und „Cognitive Load beim Nachtest“ lag jeweils im kleinen bis mittleren Bereich. Lediglich bei der abhängigen Variablen „Cognitive Load beim Lernen“ wurde der Unterschied zwischen den Bedingungen in der vorhergesagten Richtung hochsignifikant, und die Effektstärke wies einen großen Wert auf. Die Probanden der Animationsgruppe empfanden den Lernvorgang als weniger schwierig und anstrengend als die Probanden der statischen Bildgruppe. Die Interaktion zwischen dem Darstellungsformat des Lernmaterials und dem der Nachtestskalen wurde nicht signifikant und zeigte auch nur einen kleinen Effekt. Das deskriptive Muster der Interaktion lag jedoch in der vorhergesagten Richtung: Der Lernvorteil durch Animationen fiel auf der animierten Skala des Nachtests leicht stärker aus als auf der statischen Bildskala.

Vor der Interpretation dieser Effekte ist zu sagen, dass auch in der Hauptstudie, obwohl die Lernzeiten im Vergleich zur Pilotstudie halbiert wurden, der Nachtest ziemlich leicht für die Probanden war. Ungefähr die Hälfte der Aufgaben hatte zu hohe Schwierigkeitswerte, wodurch der Test eventuell wiederum nicht optimal zwischen den beiden Gruppen differenzieren konnte, so dass vielleicht auch dadurch die Mehrzahl der Unterschiede nicht signifikant wurde. Zum einen war die hohe Leichtigkeit der Nachtestaufgaben für die Probanden ein etwas überraschendes Ergebnis, da informelle empirische Voruntersuchungen ergeben hatten, dass 10 s Lernzeit zwar für Personen, die mit Animationen die Gangarten lernten, ausreichend war, um Kategorisierungsaufgaben zu lösen, nicht jedoch für Personen, die mit statischen Bildern lernten. Diese hatten erst nach ca. 30 bis 40 s Betrachtungszeit das Gefühl gehabt, die Gangarten voneinander unterscheiden zu können, da der Lernvorgang mit statischen Bildern wahrscheinlich kognitiv anspruchsvoller war. Auf der anderen Seite stand der Befund der niedrigen Schwierigkeit des Nachtests mit theoretischen Überlegungen aus der Literatur (z.B. Troje, 2002) in Einklang. Der Mensch ist evolutionär dazu ausgestattet, biologische Bewegungen in seinem Umfeld zu erkennen und zu analysieren, worauf auch die wissenschaftliche Vermutung hinweist, dass es ein eigenes Hirnareal gibt, den Sulcus temporalis superior, um speziell biologische Bewegungen zu verarbeiten. Dieses Areal zeigt eine stärkere Aktivität bei biologischen als bei durcheinandergewürfelten Bewegungen (Grossman & Blake, 2001). Daher kann beispielsweise menschliche Bewegung, mit der man oft konfrontiert ist, in 0,1 s erkannt werden (Wade & Swanston, 2001). Da die Hirnareale, die an der Wahrnehmung tatsächlicher Bewegungen beteiligt sind, auch bei der Betrachtung von stationären Bildern, die Bewegung implizieren, aktiv sind, könnte ein Lernvorgang mit Hilfe von Key-Frames auch in wenigen Sekunden funktionieren, allerdings mit mehr Anstrengung. Eine weitere Vermutung zur Begründung der konzeptuell geringen Itemschwierigkeiten könnte darin liegen, dass vielleicht auch nur wenige Features eines dynamischen Schemas gebraucht werden, um Kategorisierungsaufgaben lösen zu können. In einem solchen Fall bedarf es hierfür wahrscheinlich nur einer relativ kurzen Lernzeit, die niedriger liegt als 10 s. Nichtsdestotrotz war sich in der aktuellen Studie auf die empirischen Voruntersuchungen verlassen worden, die letztendlich eine Lernzeit von 10 s ergeben hatten. Es war zudem davon ausgegangen worden, dass das Erlernen von vier Pferdegangarten eine schwierigere Form des Lernens biologischer Bewegungen sein würde, zumal viele der Probanden keine oder nur kaum Erfahrungen mit Pferden hatten. Der Gedanke war zudem, dass die Aufgabenstellung auch für Personen, die mit statischen Bildern konfrontiert waren,

lösbar sein sollte, da dieser Lernprozess in einiger Hinsicht kognitiv anspruchsvoller ist als der mit Animationen. Zusammenfassend kann gesagt werden, dass in weiteren Studien mit biologischen Bewegungen eher geringere Lern- und Testzeiten verwendet werden sollten, um deutlichere Effekte zu erhalten.

Betrachtet man, ungeachtet des Faktors der Testschwierigkeit, die vorhandenen Haupteffekte zwischen den Bedingungen, zeigte sich ein hochsignifikanter Unterschied zwischen der Gruppe „Animation“ und der Gruppe „Bild parallel“ hinsichtlich der mentalen Belastung, die beim Lernen empfunden wurde, in der Richtung, dass der Cognitive Load beim Lernen mit Animationen deutlich geringer ausgeprägt war. Dies weist auf eine größere „instruktionale Effizienz“ (Paas et al., 2007) der Animationen bei diesem Lernvorgang hin, was bedeutet, dass mit Animationen zwar keine signifikant besseren Leistungsergebnisse erzielt werden konnten als mit statischen Key-Frames, dass aber die Probanden jedoch den Lernprozess als viel weniger anstrengend und schwierig empfanden. Das Lernen mit Animationen war angenehmer und ging mit geringeren „kognitiven Kosten“ vonstatten. Dieses Ergebnis und die sonstige deskriptive Überlegenheit der Animationen gegenüber den parallelen Key-Frames könnten zumindest ein Hinweis darauf sein, dass Bewegungsmuster aus Animationen leichter extrahiert und dynamische perzeptuelle Schemata dadurch einfacher aufgebaut werden können. Der Erkennungsprozess bei Animationen läuft analog zur alltäglichen Wahrnehmung von biologischen Bewegungen ab, für die der Mensch eine evolutionäre Prädisposition und eventuell ein eigenes Hirnareal hat (Wade & Swanston, 2001; Grossman & Blake, 2001). Beim Lernen mit statischen Bildern hingegen müssen die Zustände zwischen den Frames inferiert und der Vorgang mental animiert werden, um ein dynamisches perzeptuelles Schema zu konstruieren (Hegarty et al., 2003; Hegarty, 1992). Der Lernvorgang läuft bei Animationen wahrscheinlich eher implizit und auf perzeptueller Ebene ab, wohingegen er bei statischen Bildern mehr unter Beteiligung der Kognition und explizit vonstatten geht. Bei Animationen werden Prozessinformationen, bei statischen Bildern eher strukturelle Informationen betont. Dies könnten Gründe dafür sein, weshalb das Lernen mit Animationen von den Probanden der Studie als einfacher und weniger anstrengend erlebt wurde. Zudem hatten die Animationen der Pferdegangarten in vermindertem Maß die klassischen Nachteile dynamischer Darstellungen, wie sie in der Literatur genannt werden (z.B. Lowe, 1999a; Paas et al., 2007). Der Vorgang war repetitiv, die Gangzyklen wurden oft gezeigt, daher waren die Animationen weniger flüchtig und haben das Arbeitsgedächtnis weniger mit immer neuen Eindrücken überlastet, so dass das

Lernen mit dieser dynamischen Darstellungsform von den Teilnehmern als angenehmer erlebt wurde als das mit statischen Bildern.

Es bleibt dennoch deutlich anzumerken, dass bei vier der fünf Kriterien die Unterschiede zwischen den Bedingungen nicht signifikant wurden, so dass die Probanden der statischen Bildgruppe ebenfalls Ergebnisse auf einem guten Niveau erzielten. Dies könnte daran liegen, dass der perzeptuelle Apparat des Menschen im Alltag häufig biologische Bewegungen verarbeiten muss und daher die Inferenzen von Zuständen zwischen einzelnen Frames bei solchen Bewegungen leichter fallen als bei linearen Abläufen. Der Grad der allgemeinen Vorerfahrung mit biologischen Bewegungen ist eventuell höher. Wenn vielleicht zudem die Kategorisierungsaufgaben anhand weniger Bewegungsfeatures gelöst werden können, wofür die geringe Schwierigkeit der Kategorisierungsaufgaben ein Hinweis sein könnte, mag es sein, dass keine kompletten und elaborierten dynamischen Schemata für die Bearbeitung des Nachtests gebraucht wurden, weshalb ein theoretisch schwierigerer Lernvorgang anhand von statischen Bildern in dem Fall auch ausreichend war, um ein gutes Ergebnis zu erzielen. Letztendlich könnte auch der schon in der Pilotstudie angesprochene Übungseffekt innerhalb des Nachtests auch in dieser Studie aufgetreten sein, wodurch eventuell der qualitativ schwierigere Lernvorgang mit statischen Bildern kompensiert worden sein könnte.

Bei der Interaktionshypothese zwischen dem Lernformat und dem Nachtestformat sollte untersucht werden, ob der Lernvorteil durch Animationen auf der Animationsskala des Nachtests stärker ausgeprägt war als auf der statischen Bildskala. Die Interaktion wurde nicht signifikant, und der Lernvorteil durch Animationen war bei den animierten Items nur wenig größer als bei den statischen Bilditems. Dieses Ergebnis steht im Widerspruch zu dem der Pilotstudie, bei der die Interaktion fast signifikant wurde, wobei dieses Resultat aufgrund der geringeren Stichprobengröße vorsichtig zu interpretieren ist. Offenbar fanden in der Hauptstudie weniger Transfereffekte statt, die durch die Verwendung des gleichen Darstellungsformates in der Lernphase und im Nachtest verursacht werden können. Diese Effekte können dadurch entstehen, dass bei der Verwendung gleicher Präsentationsformen die enkodierten Stimuli den Teststimuli ähneln oder dass beide identisch sind. Dadurch können Gedächtnisprozesse und die Anwendung gleicher mentaler Operationen dabei helfen, solche Aufgaben, die das gleiche Format haben wie das Lernmaterial, besser zu lösen. Vielleicht waren in der Hauptstudie im Vergleich zur Pilotstudie 10 s Lernzeit zu wenig, um sich neben dem fokussierten Erlernen der Gangarten, äußere Merkmale der Darstellung

einprägen und mentale Operationen, wie das Inferieren von Zwischenzuständen bei den Key-Frames, einüben zu können, was bei der Lösung von Items mit gleichem Präsentationsformat geholfen hätte. Gerade für das Lernen und Wiedererkennen von statischem Material ist das menschliche Gedächtnis weniger effizient als für dynamisches (Goldstein et al., 1982), weshalb eventuell eine gewisse Zeit benötigt wird, um sich statische Darstellungsmerkmale einprägen zu können. Dies könnte ein Faktor sein, weshalb in der Hauptstudie auf der statischen Bildskala die statische Bildgruppe der Animationsgruppe gegenüber unterlegen war und sich ihr nicht wie in der Pilotstudie aufgrund von eventuellen Transfereffekten annähern konnte. Andererseits steht das in der aktuellen Studie erhaltene Ergebnis in Einklang mit den Resultaten von Goldstein et al. (1982), die fanden, dass Personen, die mit einem dynamischen Format gelernt hatten, auch darin enthaltene statische Bilder noch relativ gut wiedererkennen konnten und sich nicht signifikant von Personen unterschieden, die mit statischen Bildern gelernt hatten. Hier konnte somit auch kein großer Vorteil durch die Verwendung von einem gleichen Format im Lern- und Testmaterial gefunden werden.

Weitere Interaktionen

Interaktionen zwischen den Faktoren „Bedingung“ und „Wahrnehmungsvorerfahrung“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames parallel“ mit der Wahrnehmungsvorerfahrung von Personen

In den Hypothesen 3a – e war davon ausgegangen worden, dass zwischen den beiden Bedingungen „Animation“ und „Bild parallel“ und der Wahrnehmungsvorerfahrung der Probanden signifikante Interaktionen hinsichtlich der Kriterien „Nachttest gesamt“, „Animationsskala“, „Bildskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachttest“ bestehen. Die Hypothesen waren gerichtet formuliert, wobei die Annahme war, dass der Lernvorteil durch Animationen bei geringerer Vorerfahrung größer sein würde als bei höherer Vorerfahrung, da das Lernen mit Animationen eher implizit und auf perzeptueller Ebene abläuft und daher für Anfänger geeigneter ist.

Von den fünf Hypothesen konnte keine bestätigt werden, da alle Interaktionen keine statistische Signifikanz erreichten. Deskriptiv verliefen zudem vier der fünf Interaktionsmuster nicht in der vorhergesagten Richtung, der Lernvorteil durch Animationen

war dort somit größer bei höherer als bei niedrigerer Vorerfahrung, der Einfluss dieses Merkmals war in der Animationsbedingung stärker als in der parallelen Bildbedingung. Bei höherer Vorerfahrung war die Animationsbedingung der Bildgruppe gegenüber deskriptiv bis signifikant in vier der Kriterien überlegen, in einem Fall lagen die beiden Gruppen ungefähr gleichauf. Bei dem Kriterium „Bildskala“ war die Animationsbedingung bei geringerer Vorerfahrung sogar ganz leicht im Nachteil, während sie ansonsten ebenfalls deskriptiv bis signifikant der Bildbedingung gegenüber im Vorteil war, wobei diese Unterschiede bei drei der Kriterien aber nicht so stark ausgeprägt waren wie bei den Probanden mit einem größeren Erfahrungshintergrund.

Da die Ergebnisse keine Signifikanz aufwiesen, kann man nur im Ansatz interpretieren, warum Personen mit geringerer Vorerfahrung im Vergleich zu Animationen besser mit den parallelen Key-Frames zurechtkamen als Personen mit höherer Vorerfahrung. Für unerfahrene Betrachter haben statische Bilder den Vorteil, dass sie sie in Ruhe inspizieren können, da sie vollkommen stationär sind (Kosslyn, 1994). Die Animationen haben eine höhere Geschwindigkeit, was es eventuell etwas schwieriger macht, das dynamische Schema aus ihnen zu extrahieren. Allerdings waren die Animationen in ihrer Flüchtigkeit stark herabgemindert, da es sich bei den Gangarten um repetitive Vorgänge handelte, so dass der statische Vorzug der Bilder im Vergleich zu dieser dynamischen Darstellung nicht so ausgeprägt gewesen sein dürfte. Zudem waren in der statischen Bildbedingung die Bewegungen der einzelnen Gangarten zuvor in sinnvolle Einheiten zergliedert und je durch ein Key-Frame präsentiert worden, so dass dem Lerner eigene Segmentierungsvorgänge erspart blieben, was für Anfänger gut geeignet war.

Jedoch bleibt zu betonen, dass generell keine signifikanten Interaktionen zwischen den Bedingungen und der Wahrnehmungsvorerfahrung der Probanden vorlagen, was als das hervorstechendere Ergebnis gewertet werden kann. Das bedeutet, dass die beiden Treatmentstufen keine stark differentiellen Wirkungen in Abhängigkeit von der Wahrnehmungsvorerfahrung der Probanden zeigten und dass der Lernvorteil durch Animationen sich nicht wie vorhergesagt bei geringerer Vorerfahrung stärker vergrößert hat, er war in den meisten Fällen sogar leicht niedriger. Dass somit unerfahrene Lerner vergleichsweise gut auch mit parallelen Key-Frames zurechtkamen, könnte daran gelegen haben, dass Menschen allgemein oft mit biologischen Bewegungen konfrontiert sind und ihnen die Inferenzen ihrer Zwischenzustände, wie sie bei einem Lernvorgang mit statischen Key-Frames zu leisten sind, eher leicht fallen. Zudem könnte Anfängern der Übungseffekt

des Nachttests geholfen haben, auch nach einem schwierigeren Lernvorgang mit statischen Bildern, noch relativ gute Ergebnisse zu erzielen. Wenn vielleicht nur wenige Features der Bewegungsmuster zur Lösung der Kategorisierungsaufgaben benötigt wurden, hatten Personen mit geringerer Vorerfahrung eine höhere Chance, auch mit Hilfe von statischen Bildern den Test gut zu bewältigen, weshalb sie beim Lernen mit parallelen Key-Frames nicht so stark benachteiligt waren wie vorhergesagt.

Ein weiterer Grund, weshalb sich die Hypothesen nicht bestätigt haben, könnte darin liegen, dass die Items des Tests zur Erfassung der Wahrnehmungsvorerfahrung breiter konstruiert waren. In den Fragen wurden sowohl gegenwärtige als auch vergangene Erfahrungen thematisiert („Haben oder hatten Sie früher häufiger mit Pferden zu tun?“), und das Antwortformat war dichotom („ja / nein“), so dass keine genauen Häufigkeiten der einzelnen Aktivitäten gemessen wurden und die Items von den einzelnen Probanden unterschiedlich interpretiert werden konnten. Auf diese Weise hatten eventuell Personen mit empirisch unterschiedlichen Erfahrungshintergründen den gleichen Punktwert in dem Test und wurden zu wenig trennscharf eingruppiert. Diese möglicherweise mangelnde Differenzierung zwischen den Probanden könnte ein Grund dafür sein, dass der Lernvorteil durch Animationen auf den beiden Vorerfahrungsstufen nicht stark unterschiedlich wurde. Das vorhergesagte Interaktionsmuster könnte vielleicht dann gefunden werden, wenn man die Fragen differenzierter formuliert, indem gegenwärtige und vergangene Erfahrungen getrennt behandelt sowie indem ein mehrstufiges Antwortformat verwendet wird.

Die Probanden der Stichprobe hatten einen eher niedrigen Erfahrungshintergrund ($M = 1.19$, $SD = 1.56$ von 7.00 möglichen Punkten), was bei der Interpretation der Ergebnisse ebenfalls zu berücksichtigen ist.

Interaktionen zwischen den Faktoren „Bedingung“ und „Räumliche Fähigkeiten“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames parallel“ mit den räumlichen Fähigkeiten von Personen

Die Hypothesen 4a – e machten die Annahme, dass zwischen den Bedingungen „Animation“ und „Bild parallel“ und den räumlichen Fähigkeiten der Probanden signifikante Interaktionen hinsichtlich der fünf abhängigen Variablen der Studie bestehen. Es lagen hierbei gerichtete Hypothesen vor, da davon ausgegangen wurde, dass der Lernvorteil durch

Animationen bei geringeren räumlichen Fähigkeiten größer ausfallen würde als bei höheren räumlichen Fähigkeiten.

Durch die statistischen Analysen konnten die Hypothesen bezüglich der Kriterien „Nachtst gesamt“ und „Bildskala“ bestätigt werden, da die Ergebnisse jeweils signifikant wurden. Die Interaktionen, die hinsichtlich der abhängigen Variablen „Animationsskala“ und „Cognitive Load beim Nachtst“ errechnet wurden, erreichten annähernd Signifikanz ($p = .08$ bzw. $p = .06$), was dennoch zur Verwerfung der jeweiligen Hypothese führte. Ebenfalls nicht bestätigt werden konnte die Annahme zum Kriterium „Cognitive Load beim Lernen“. Dennoch verliefen die Interaktionsmuster alle in der vorhergesagten Richtung: Der Lernvorteil durch Animationen war bei geringeren räumlichen Fähigkeiten stärker ausgeprägt als bei höheren räumlichen Fähigkeiten. Generell war auffällig, dass das räumliche Vorstellungsvermögen beim Lernen mit Animationen auf alle Kriterien kaum einen Effekt hatte, die entsprechenden Regressionskoeffizienten lagen um Null, während bei den parallelen Key-Frames die Leistungsindizes positiv durch die figuralen Fähigkeiten beeinflusst wurden. Probanden, die mit geringeren räumlichen Fähigkeiten ausgestattet waren, erzielten in allen fünf Kriterien signifikant schlechtere Ergebnisse in der parallelen Bild- als in der Animationsbedingung. Verfügten die Teilnehmer jedoch über höhere räumliche Fähigkeiten, waren sie bei den abhängigen Variablen „Nachtst gesamt“ und „Bildskala“ durch das Lernen mit parallelen Bildern deskriptiv im Vorteil, während hinsichtlich der „Animationsskala“ und dem „Cognitive Load beim Nachtst“ ungefähr ein Gleichstand zwischen den beiden Gruppen herrschte.

Dass der Lernvorteil durch Animationen bei geringeren räumlichen Fähigkeiten stärker ausgeprägt war als bei höheren räumlichen Fähigkeiten, so dass Personen mit niedrigerem figuralem Vorstellungsvermögen mit parallelen Key-Frames signifikant schlechter lernten, könnte, wie auch schon in der Herleitung der Hypothesen ausgeführt, daran liegen, dass der Lernvorgang mit statischen Bildern bei biologischen Bewegungen kognitiv anspruchsvoller ist und eher explizit abläuft. Die Zustände zwischen den Frames müssen inferiert, der Vorgang mental animiert, und das dynamische Schema muss internal konstruiert werden, wofür räumliche Fähigkeiten benötigt werden (Hegarty et al., 2003; Hegarty, 1992; Sims & Hegarty, 1997). Wie in der Diskussion zu den Interaktionen mit der „Wahrnehmungsvorerfahrung“ ausgeführt wurde, waren die Inferenzen zwar allgemein vermutlich nicht übermäßig schwierig, schienen aber zumindest in räumlicher Hinsicht

anspruchsvoller gewesen zu sein, da sich die Bewegungen und die Stellungen der vier Beine des Pferdes im Raum zueinander bei gleich mehreren Gangarten vorgestellt werden mussten. Beim Lernen mit Animationen hingegen wurde im vorliegenden Fall vermutlich ein geringeres räumliches Vorstellungsvermögen gebraucht. Die Animationen stellten biologische Bewegungen und somit repetitive Vorgänge dar, bei denen sich die einzelnen Bewegungszyklen ständig wiederholen und an denen weniger Komponenten beteiligt sind als an vielen linearen Prozessen. Die Animation verliert dadurch an inhärenter Flüchtigkeit, weshalb die räumliche Verarbeitungsgeschwindigkeit und die räumliche Speicherkapazität, die sonst zur Wahrnehmung und zum Verständnis einer Animation benötigt werden, nicht mehr so wichtig sind (Salthouse, 1996; Shah & Miyake, 1996). Die Gangzyklen wiederholen sich öfters, weshalb der Lerner mehrere Möglichkeiten hat, die Bewegungen zu betrachten und zu verstehen, wodurch er die einzelnen räumlichen Informationen nicht mehr so schnell verarbeiten muss, um mit der Animation „Schritt zu halten“. Durch den repetitiven Charakter der Vorgänge, muss der Lerner ihre räumlichen Details auch nicht mehr in großem Umfang im Arbeitsgedächtnis speichern, er kann den Vergleich und die Integration der einzelnen Bewegungszustände mehr mit der externen Unterstützung der Animation durchführen (Lowe, 1999a). Die Animation hilft dem Lerner eventuell sogar dabei, die mentalen Rotationen und Transformationen zu vollziehen („Supplantationsfunktion“ der Animation) und kann damit einen Mangel an internem räumlichem Vorstellungsvermögen kompensieren (Salomon, 1979). Das Lernen mit dynamischen Darstellungen geht bei biologischen Bewegungen eher perzeptuell und automatisiert vonstatten, was sich auch darin andeutet, dass in der vorliegenden Studie das räumliche Vorstellungsvermögen kaum einen Einfluss auf die Lernergebnisse in der Animationsgruppe hatte. Bei höheren figuralen Fähigkeiten hingegen herrschte ein Gleichstand zwischen den Bedingungen bzw. war die parallele Bildgruppe sogar leicht deskriptiv im Vorteil im Vergleich zur Animationsgruppe. Das könnte mit darin begründet sein, dass das Inferieren der Zwischenzustände und die mentale Animation angemessene Anforderungen für Personen mit höheren räumlichen Fähigkeiten sind und ihren Kompetenzen somit entsprechen. Ein Lernvorgang mit Animationen könnte eine Unterforderung für solche Personen darstellen.

Die gefundenen Ergebnisse beziehen sich zunächst auf den Wertebereich an räumlichen Fähigkeiten, der in der Stichprobe vorlag und der sich auf einem durchschnittlichen Niveau befand ($M = 22.48$, $SD = 6.26$).

Interaktionen zwischen den Faktoren „Bedingung“ und „Merkfähigkeit“

Interaktionen zwischen den Bedingungen „Animation“ und „Key-Frames parallel“ mit der Merkfähigkeit von Personen

Ebenfalls gerichtete Hypothesen (5a – e) wurden für die Interaktionen zwischen den Faktoren „Bedingung“ und „Merkfähigkeit“ hinsichtlich der Kriterien „Nachtest gesamt“, „Animationsskala“, „Bildskala“, „Cognitive Load beim Lernen“ und „Cognitive Load beim Nachtest“ aufgestellt. Es wurde von signifikanten Interaktionen ausgegangen, die das Muster aufweisen sollten, dass der Lernvorteil durch Animationen bei geringerer Merkfähigkeit größer ist als bei höherer Merkfähigkeit, da die Inferenz der Zwischenzustände bei parallelen Key-Frames Arbeitsgedächtniskapazität beansprucht.

Keine der Hypothesen konnte durch die Rechnungen bestätigt werden, da die Interaktionen nicht signifikant wurden. Allerdings lagen alle Ergebnisse, deskriptiv gesehen, in der vorhergesagten Richtung, d.h., der Lernvorteil durch Animationen war bei geringerer Merkfähigkeit stärker ausgeprägt als bei höherer Merkfähigkeit. In beiden Versuchsgruppen hatten die Gedächtnisfähigkeiten der Probanden einen eher schwachen Einfluss auf die Lernergebnisse. Sowohl Personen mit geringerer als auch solche mit höherer Merkfähigkeit erreichten mit Animationen bei allen Kriterien deskriptiv bessere Ergebnisse als mit parallelen Key-Frames, beim „Cognitive Load beim Lernen“ waren die Unterschiede auch signifikant. Die Differenzen fielen hierbei bei geringerer Merkfähigkeit größer aus als bei höherer Merkfähigkeit.

Bei den Ergebnissen ist zu sehen, dass sich bei allen abhängigen Variablen konsistent das vorhergesagte Muster zeigte, die Interaktionen aber nicht signifikant wurden. Daher deutet sich nur in der Tendenz an, dass für das Lernen mit parallelen Key-Frames vergleichsweise höhere Gedächtnisfähigkeiten gebraucht werden, um bessere Lernresultate zu erzielen. Dies könnte daran liegen, dass bei parallelen Bildern internale Konstruktionsprozesse und mentale Animationen zusätzlich zum reinen Betrachten notwendig sind, um die Gangarten zu verstehen, was Arbeitsgedächtniskapazität beansprucht (Sims & Hegarty, 1997). Umgekehrt ist beim Lernen mit Animationen von biologischen Bewegungen Merkfähigkeit nicht so stark von Nöten, da es sich hierbei um repetitive Vorgänge handelt, wodurch die Animation an Flüchtigkeit verliert. Zudem läuft der Aufbau dynamischer perzeptueller Schemata bei Animationen vermutlich eher automatisiert ab.

Allerdings wurden die Interaktionen nicht signifikant, weshalb diese Interpretation nur sehr eingeschränkt vertreten werden darf.

Dass die Wechselwirkungen keine Signifikanz erreichten, zeigte sich unter anderem daran, dass die Merkfähigkeit einen unerwartet schwachen Einfluss in der parallelen Bildbedingung auf die Lernergebnisse hatte und dass Personen mit geringerer Merkfähigkeit mit statischen Key-Frames nur deskriptiv schlechter waren als mit Animationen. Probanden mit niedrigeren Gedächtnisfähigkeiten kamen somit vergleichsweise gut mit den parallelen Bildern zurecht, so dass keine signifikanten Interaktionen in der vorhergesagten Richtung entstanden.

Dies könnte darin begründet sein, dass bei den Key-Frames zwar Inferenzen auf die Zwischenzustände und mentale Animationen geleistet werden mussten (Hegarty et al., 2003; Hegarty, 1992), was durch die deskriptiven Interaktionsmuster angedeutet wurde, allerdings waren diese bei der relativ hohen Anzahl an Bildern anscheinend nicht übermäßig schwierig, so dass das Arbeitsgedächtnis nicht derart belastet wurde, dass die Interaktionen ein signifikantes Ausmaß annahmen. Ebenfalls erleichternd dürfte der Faktor gewirkt haben, dass Menschen generell relativ viel allgemeine Vorerfahrung mit biologischen Bewegungen haben, weshalb die internalen Konstruktionsprozesse wahrscheinlich einfacher gewesen waren und das Arbeitsgedächtnis nicht überfordert haben. Wenn nur wenige Features für die Lösung der Kategorisierungsaufgaben benötigt wurden, konnten Personen mit geringerer Gedächtniskapazität, die die mentale Animation und den Aufbau des dynamischen Schemas nur teilweise vollzogen haben, ebenfalls vergleichsweise gute Ergebnisse mit parallelen Key-Frames erzielen. Letztendlich könnte es eventuell einen zusätzlichen Lerneffekt durch den Nachtest gegeben haben, bei dem die einzelnen Gangarten im Rahmen der Aufgaben immer wieder gezeigt wurden, so dass das Arbeitsgedächtnis der Probanden zusätzlich entlastet wurde. Dies könnten Faktoren gewesen sein, weshalb sich die höheren Anforderungen an die Merkfähigkeit der Probanden durch parallele Key-Frames zwar andeuteten, aber zu keinen signifikanten Ergebnissen führten.

6 Integration und Gesamtdiskussion der Befunde

Es war bei der Gesamtfragestellung der Dissertation davon ausgegangen worden, dass ein differenzierterer Vergleich zwischen Animationen und statischen Bildern sinnvoll wäre, da sich in der bisherigen Forschung keine konsistente Überlegenheit einer dieser beiden Präsentationsformen gefunden hat (z.B. Rieber, 1990). Die aktuelle Annahme war, dass sich Animationen und statische Bilder für den Aufbau von kognitiven und perzeptuellen mentalen Repräsentationen und somit für die Lösung von Inferenz- und bildlichen Kategorisierungsaufgaben unterschiedlich gut eignen

Auf der kognitiven Ebene des Lernens, die im ersten Teil der ersten Studie der Dissertation untersucht wurde, wurde die Hypothese aufgestellt, dass mit statischen Bildern, die die Key-States eines Vorganges zeigen, Wissens- und Inferenzaufgaben besser gelöst werden als mit Animationen, was als Indikator dafür gewertet werden kann, dass mit Key-Frames dynamische mentale Modelle leichter aufgebaut werden. Die Key-Frames wurden hierbei in seriell schneller und in seriell langsamer Form präsentiert. Die Annahmen der ersten Teilstudie konnten nicht bestätigt werden, da die Ergebnisse der Mittelwertvergleiche nicht signifikant wurden. Deskriptiv gesehen, ergab sich ein gemischtes Bild. In einem Fall, hinsichtlich des Cognitive Loads, der beim Lernen empfunden wurde, war die Animationsbedingung signifikant besser als die seriell langsame Bildbedingung. Dies lag wahrscheinlich daran, dass bei seriell langsamen Key-Frames die Inferenzen zwischen den Bildern besonders schwierig sind und die Bildanordnung aus motivationaler Sicht eher monoton wirkt. Eine Animation wirkt motivational neuartig, und man muss den Vorgang beim Lernen vor allem wahrnehmen (Lowe, 2001; Hegarty et al., 2003). Durch die Studie haben sich die Annahmen der Theorie, dass durch statische Key-Frames dynamische mentale Modelle leichter aufgebaut werden (Schnotz & Lowe, 2008), erst einmal nicht bestätigt. Jedoch wäre es an der Stelle wahrscheinlich verfrüht, diese Überlegungen völlig zu verwerfen, da die Nichtsignifikanzen in den Ergebnissen möglicherweise durch die Thematik des Lernmaterials mit bedingt waren. Zudem haben auch Animationen inhärente Vorteile, wobei die Gestaltung der dynamischen Präsentation in der aktuellen Studie durch die mehrfachen Wiederholungen des Prozesses, durch das Standbild zu Beginn der Lerneinheit und durch die gleichmäßige Zeichnung der Animation, bei der nichts konzeptuell Irrelevantes hervorstach, vereinfacht war.

Auf der kognitiven Ebene wurde als zusätzliche Fragestellung innerhalb der ersten Teilstudie der Vergleich zwischen seriell schnellen und seriell langsamen Key-Frames behandelt. Auch hier ergaben sich keine signifikanten Unterschiede. Anscheinend hatte die Präsentationsgeschwindigkeit der Key-Frames keinen Einfluss auf die Lernergebnisse oder der Effekt wurde mit dem verwendeten Lernmaterial nicht adäquat widergespiegelt.

Auch im zweiten Teil der ersten Studie, in dem parallele mit seriell schnellen und seriell langsamen Key-Frames verglichen wurden, zeigten sich keine signifikanten Differenzen zwischen den Bedingungen. Lediglich bei den Inferierfähigkeiten ergab sich ein fast signifikanter Unterschied zwischen der parallelen und der seriell langsamen Bildbedingung, wobei letztere im Vorteil war. Auffällig an den Resultaten war allerdings, dass die parallele Bildgruppe bei sieben von acht Vergleichen schlechter war als die beiden seriellen Bildgruppen. In der parallelen Bildbedingung war in das Lernmaterial aus pragmatischen Gründen ein Interaktivitätsfeature eingebaut worden, das offenbar die Lernleistung mehr beeinflusst hat als vermutet und das den autonomen Charakter eines Lernvorganges mit simultanen Bildern noch verstärkt hat. Die höhere Selbststeuerung scheint eher zu Defiziten geführt zu haben, da durch Interaktivität auf kognitiver Ebene mehr Metakognition sowie Entscheidungsprozesse geleistet werden müssen (Schnotz, Böckheler & Grzondziel, 1999; Schnotz & Lowe, 2008). Zudem stellt ein solches Lernmaterial auch Anforderungen an die Motivation der Lerner. Die Probanden wirkten in der Pilotstudie beim Scrollen der simultanen Bilder eher unkonzentriert und schienen die Lernzeit nicht vollständig zu nutzen. Zusätzlich zeigten sich in der Hauptstudie größere bis signifikante Interaktionen zwischen der parallelen und der seriell langsamen Bildbedingung mit der motivationalen Variable „Herausforderung“. Personen, die die Aufgabe als Herausforderung für sich sahen, profitierten mehr von den parallelen Key-Frames, während Personen mit niedrigeren Herausforderungsscores besser mit den seriell langsamen Key-Frames zurechtkamen. Die Interaktivität in der parallelen Bildbedingung schien mit für deren deskriptive Nachteile in den Kriterien verantwortlich zu sein, da die Probanden die Vorteile dieser Materialgestaltung nicht vollständig nutzten. Entweder waren sie hierfür nicht motiviert genug oder sie überschätzten eventuell ihr eigenes Verständnis der Thematik.

Dieses Ergebnis, dass Interaktivität und somit ein selbstgesteuerter Lernvorgang nicht zu Vorteilen geführt, sondern eher das Gegenteil bewirkt hat, steht in Einklang mit weiteren Studien in der Literatur, in denen ähnliche Resultate gefunden wurden. In der Untersuchung von Schnotz, Böckheler und Grzondziel (1999) war die interaktive Animationsbedingung

der statischen Bildbedingung gegenüber unterlegen, wenn kooperativ gelernt werden sollte. Die Autoren erklärten dies damit, dass in der Animationsgruppe das Bedienen des Systems mit dem eigentlichen Lernvorgang konkurrierte. In der statischen Bildgruppe, die nicht interaktiv war, redeten die Teilnehmer mit der Zeit immer weniger über das System und konnten sich hingegen konzeptuell mit dem Lerninhalt beschäftigen. In der Studie von Bétrancourt und Realini (2005) wurde eine Animationsbedingung ohne Interaktivität mit zwei Animationsbedingungen jeweils verglichen, die in einem Fall mit partiellen und im anderen Fall mit vollständigen Kontrollmöglichkeiten ausgestattet waren. Die Versuchspersonen, die keine Kontrolle über das System hatten, erzielten die besten Ergebnisse im anschließenden Verstehenstest. Die Autoren vermuteten, dass die Interaktivität eventuell einen „Split attention-Effekt“ (Sweller et al., 1990) verursacht hat. Zudem wurde in dieser Studie eine signifikante Interaktion zwischen der Lernbedingung und der schulischen Ausrichtung der Teilnehmer gefunden. Personen, die keine naturwissenschaftliche Ausrichtung hatten, verstanden das Lernmaterial, das Aktionspotentiale im Nervensystem thematisierte, besser in der nicht-interaktiven Bedingung, während Personen mit naturwissenschaftlicher Ausrichtung mehr von der partiellen Kontrolle innerhalb der Animation profitierten.

Aus diesen Ergebnissen lässt sich schließen, dass Interaktivität beim Lernen anscheinend differenzierter betrachtet werden muss und in Zusammenhang mit kognitiven und motivationalen Eigenschaften der Lerner gesehen werden sollte.

In der zweiten Studie der Dissertation wurde die perzeptuelle Ebene des Lernens fokussiert, wobei hier angenommen wurde, dass Animationen besser für den Aufbau dynamischer perzeptueller Schemata und somit auch besser für die Lösung von bildlichen Kategorisierungsaufgaben geeignet sind als statische Bilder, die die Key-States der Bewegungen zeigen. Die Key-Frames wurden in dem Experiment parallel auf dem Bildschirm angeordnet präsentiert. Die Unterschiede zwischen den Bedingungen wiesen alle in die vorhergesagte Richtung, waren aber größtenteils nicht signifikant. Lediglich beim Cognitive Load, der beim Lernen empfunden wurde, fand sich ein hochsignifikantes Ergebnis: Die Teilnehmer, die die Gangarten mit Animationen betrachteten, bewerteten den Lernvorgang als viel weniger anstrengend als die Teilnehmer, die mit statischen Bildern konfrontiert waren.

Bei der zweiten Studie zeichnete sich somit ein deutlicheres Bild als bei der ersten Studie zu Gunsten der gemachten Annahmen ab. Durch die hochsignifikante Überlegenheit der

Animationen bezüglich der mentalen Belastung und durch ihre auch sonst deskriptiv besseren Ergebnisse, bestätigte sich die Theorie, dass Bewegungsmuster einfacher aus Animationen heraus extrahiert werden und dynamische perzeptuelle Schemata somit durch Animationen leichter aufgebaut werden, zumindest im Ansatz. Der Lernvorgang von biologischen Bewegungen mit Animationen ähnelt mehr der alltäglichen Wahrnehmung und profitiert somit mehr von der diesbezüglichen evolutionären Prädisposition des Menschen, während bei statischen Bildern Zustände zwischen den Frames kognitiv inferiert werden müssen (Schnotz & Lowe, 2008; Wade & Swanston, 2001; Hegarty et al., 2003). Allerdings bleibt zu betonen, dass die Mehrzahl der Unterschiede in der Untersuchung nicht signifikant war, weswegen diese Argumentation nur mit großen Einschränkungen vertreten werden kann. Die Kategorisierungsaufgaben innerhalb der aktuellen Studie waren sehr einfach für die Probanden gewesen, bei einer höheren Schwierigkeit der Aufgaben durch beispielsweise geringere Lern- und Testzeiten wären die Vorteile von Animationen möglicherweise noch stärker hervorgetreten. Die Interaktion zwischen dem Präsentationsformat des Lernmaterials und dem der Nachtestaufgaben war nicht signifikant, und der Vorteil durch Animationen war nur leicht größer auf der Animations- als auf der statischen Bildskala des Nachtests.

Zusammenfassend kann gesagt werden, dass die Mehrzahl der Haupteffekte zwischen den Bedingungen in den beiden Studien nicht signifikant waren. Dieses Ergebnis ist kongruent zu weiteren Studien aus der Literatur, bei denen die Unterschiede zwischen Animationen und statischen Bildern ebenfalls nicht signifikant wurden (z.B. Morrison & Tversky, 2001; Mayer et al., 2005). Auch im Review von Rieber (1990) wurde deutlich, dass viele Befunde zum Vergleich von Animationen und statischen Bildern keine Signifikanz aufwiesen. Beispielsweise fanden J. Caraballo und A. Caraballo (beide 1985) in ihren jeweiligen Studien keine Effekte zwischen einer Textbedingung und einer Textbedingung, die mit statischen Bildern versehen war sowie einer Textbedingung, die mit animierten Bildern angereichert war. Die Thematik des Lernmaterials war in den beiden Studien jeweils variiert (Physiologie des menschlichen Herzens und geometrische Flächenberechnung).

Jedoch wurden in allen Studien der Dissertation signifikante Interaktionen der Lernbedingungen mit Personenmerkmalen gefunden, die an der Stelle nochmals kurz resümiert werden:

Im ersten Teil der ersten Studie:

- Die Interaktion zwischen der Animations- und der seriell schnellen Bildbedingung mit dem Vorwissen der Teilnehmer wurde signifikant bei den Kriterien „Inferenzskala“ und „Cognitive Load beim Lernen“. Personen mit weniger Vorwissen lernten signifikant besser mit Animationen, während Personen mit höherem Vorwissen besser mit den seriell schnellen Key-Frames zurechtkamen. Die Inferenzen und die mentale Animation bei den Key-Frames schienen für Personen mit weniger Vorwissen sehr schwierig zu sein, die Animation lieferte ihnen hingegen vollständige Informationen über den Prozess (Lewalter, 1997).
- Ebenfalls signifikant wurde die Interaktion zwischen der seriell schnellen und der seriell langsamen Bildbedingung mit dem Vorwissen der Teilnehmer hinsichtlich des Cognitive Loads beim Lernen. Personen mit weniger Vorwissen lernten deskriptiv besser mit seriell langsamen Key-Frames, Personen mit höherem Vorwissen fühlten sich signifikant weniger belastet durch seriell schnelle Key-Frames. Seriell langsame Bilder liefern vor allem Informationen über Teilprozesse, die von Novizen für den Aufbau des mentalen Modelles benötigt werden (Mayer & Chandler, 2001), die für fortgeschrittene Lerner aber redundant sind.
- Fast signifikant wurden die Interaktionen zwischen der Animations- und der seriell schnellen Bildbedingung mit der Merkfähigkeit der Probanden bezüglich der mentalen Belastungen, die nach dem Lernen und nach dem Nachtest gemessen wurden. Personen mit weniger Merkfähigkeit verspürten deskriptiv mehr Cognitive Load in der Animationsgruppe, bei Personen mit höherer Merkfähigkeit war, teilweise signifikant, das Gegenteil der Fall. Beim Betrachten von Animationen strömen viele Informationen auf das Arbeitsgedächtnis des Lerners ein, die es überlasten können (Sweller et al., 1998). Zudem müssen bei einer Animation Vergleiche zwischen den Zuständen internal durchgeführt werden (Lowe, 1999a).

Im zweiten Teil der ersten Studie:

- Hier zeigten sich fast signifikante bzw. signifikante Interaktionen zwischen der parallelen und der seriell schnellen Bildbedingung mit dem Vorwissen der Teilnehmer bezüglich ihrer Inferierfähigkeiten und ihrer mentalen Belastung. Personen mit höherem Vorwissen lernten deskriptiv besser mit seriell schnellen Key-Frames, Personen mit niedrigerem Vorwissen profitierten mehr von parallelen Key-Frames. Zudem wurde auch die Interaktion zwischen der parallelen und der seriell langsamen Bildbedingung mit dem Vorwissen der Probanden bezüglich des Cognitive Loads beim Lernen signifikant. Fortgeschrittene Lerner kamen hier deskriptiv mit den parallelen Key-Frames besser zurecht, während für Personen mit weniger Vorwissen seriell langsame Bilder geeigneter waren. Die Interaktionen zwischen seriellen und parallelen Bildern mit dem Vorwissen der Teilnehmer wiesen somit keine klare Richtung auf, was eventuell durch die unterschiedlichen Präsentationsgeschwindigkeiten der Key-Frames in den beiden seriellen Bildbedingungen begründet war.

Bei den Interaktionen der Lernbedingungen mit dem Vorwissen der Probanden ist in den beiden Teilstudien außerdem zu sagen, dass der Vorwissensbereich in den Teilstichproben auf einem eher niedrigen Niveau lag und somit eingeschränkt war.

In der zweiten Studie:

- In dieser Studie fanden sich signifikante Interaktionen zwischen der Animations- und der parallelen Bildbedingung mit den räumlichen Fähigkeiten der Teilnehmer bei den Kriterien „Nachttest gesamt“ und „Bildskala“. Bei den Kriterien „Animationsskala“ und „Cognitive Load beim Nachttest“ wurden die Interaktionen fast signifikant. Der Lernvorteil durch Animationen war bei Personen mit niedrigeren räumlichen Fähigkeiten größer als bei Personen mit höheren räumlichen Fähigkeiten. Personen mit niedrigeren räumlichen Fähigkeiten hatten bei allen fünf abhängigen Variablen signifikant schlechtere Ergebnisse mit den parallelen Key-Frames. Bei den statischen Bildern mussten die Teilnehmer die Zustände zwischen den Frames inferieren, was insbesondere in räumlicher Hinsicht schwierig war, und sie mussten die Bewegungen

mental animieren, wofür räumliche Fähigkeiten benötigt wurden (Sims & Hegarty, 1997). Die Animationen hingegen könnten für Personen mit niedrigerem räumlichem Vorstellungsvermögen eine „Supplantationsfunktion“ (Salomon, 1979) gehabt haben und ihnen mentale Transformationen abgenommen haben.

Das letzte Ergebnis steht in Einklang mit der Studie von Blake (1977), in der Versuchspersonen die Bewegungsabfolgen von Schachzügen mit Hilfe von vollständig oder teilweise animierten oder mit Hilfe von statischen Darstellungen lernen sollten. Die animierten Versionen führten bei Personen mit niedrigerem räumlichem Vorstellungsvermögen zu signifikant besseren Ergebnissen als die statischen Bilder. Allerdings stehen diese Befunde im Gegensatz zu der Untersuchung von Hegarty et al. (2003), die keine signifikanten Interaktionen zwischen dem Instruktionsformat (Animation, mehrere statische Bilder, ein statisches Bild) und den räumlichen Fähigkeiten der Teilnehmer fanden. Das Thema des Lernmaterials dieser Studie war die Funktionsweise eines Spülkastens.

In den Studien der Dissertation hingen somit die Effekte der Präsentationsformen von individuellen Merkmalen, wie dem Vorwissen oder den räumlichen Fähigkeiten der Personen, mit ab. Solche Abhängigkeiten wurden schon im Artikel von Hegarty und Kriz (2008) berichtet und theoretisch erläutert. Die Befunde legen nahe, dass die differenzierten Fragestellungen der Studien innerhalb der Dissertation noch differenzierter formuliert werden müssten und dass Personenmerkmale in weiteren Studien noch stärker in die Analysen einbezogen werden sollten. Animationen und die verschiedenen Präsentationsarten der statischen Bilder hatten bei den aktuellen Ergebnissen teilweise positive Effekte für bestimmte Personengruppen gehabt. Grob gesprochen, deutete sich in der Tendenz an, dass sich Animationen bei niedrigeren Fähigkeiten besser eigneten und dass bei höheren Fähigkeiten der Einsatz von statischen Bildern leichte Vorteile brachte.

7 Methodische Kritik und Ausblick

Im ersten Teil der ersten Studie, in dem Animationen mit seriell schnellen und seriell langsamen Key-Frames sowie die beiden seriellen Bildbedingungen miteinander verglichen wurden, kann methodisch die Auswahl des Lernmaterials kritisiert werden. Das Thema des Lernmaterials „Signalübertragung an Synapsen“ war eher einfach für Studenten, die diese Inhalte in der gymnasialen Oberstufe gelernt hatten und die sich, nach eigener Aussage, an ihr Schulwissen während des Lernens erinnern konnten, obwohl sie im Vorwissenstest der Untersuchung eher schlechte Ergebnisse erzielten. Dieses aktualisierte Wissen könnte Unterschiede zwischen den Gruppen nivelliert haben. Daher würde es sich bei einer Wiederholung der Studie anbieten, ein anderes, schwierigeres Thema zu wählen, beispielsweise aus dem technischen oder physikalischen Bereich, oder die Untersuchung an Oberstufenschülern durchzuführen, die die Thematik noch nicht behandelt haben. Bezüglich des Lernmaterials zu Synapsen war weiterhin anzumerken, dass durch die mehrfache Wiederholung des Vorganges in der Animationsbedingung die Nachteile dieser dynamischen Darstellung abgemindert waren. Vielleicht wären bei einer kürzeren Lernzeit und bei weniger Wiederholungen des Prozesses innerhalb der Animationsbedingung eher größere Unterschiede zwischen den Gruppen erzeugt worden.

In der ersten Teilstudie waren die Haupteffekte größtenteils nicht signifikant, jedoch zeigten sich signifikante Interaktionen zwischen der Animations- und der seriell schnellen Bildbedingung mit dem Vorwissen der Probanden. Dies kann als Hinweis dafür gelten, dass „Aptitude-Treatment-Interaktionen“ in weiteren Studien zu Animationen und statischen Bildern eine größere Rolle spielen sollten. Auch für die pädagogische Praxis ist dieser Befund interessant, da der gewinnbringende Einsatz von Animationen und statischen Bildern auch von individuellen Voraussetzungen der Lerner abhängen könnte. Beispielsweise zeigte sich in der aktuellen Studie, dass für Personen mit geringerem Vorwissen Animationen besser geeignet sind als seriell schnelle Bilder.

Einschränkend ist hier allerdings noch hinzuzufügen, dass sich die aktuellen Ergebnisse zu den Interaktionen nur auf einen begrenzten, eher niedrigen Vorwissensbereich bezogen und zur Stabilisierung der Ergebnisse eine Stichprobe mit einem breiteren Vorwissensbereich oder ein Experten-Novizen-Design gewählt werden müsste.

Eines der signifikanten Resultate der Interaktionen basierte auf dem Cognitive Load der Probanden, der, in Anlehnung an die „Cognitive Load Rating Scale“ (Paas, 1992), mit Hilfe von zwei introspektiven Items ermittelt wurde, was eine eher geringe Skalenlänge darstellte.

Allerdings wird die „Cognitive Load Rating Scale“ bei Paas et al. (1994) als reliabel und sensitiv beschrieben, zumal es als schwierig erscheint, eine Alternative zur Messung mentaler Belastung zu finden, da beispielsweise physiologische Verfahren sehr aufwändig in der Durchführung sind.

Im zweiten Teil der ersten Studie wurden parallele mit seriellen Bildern verglichen. Bei den parallelen Bildern war aus pragmatischen Gründen ein einfaches Interaktivitätsfeature (Scrolleiste am rechten Bildrand) eingebaut worden, das in der Untersuchung die Lernleistung anscheinend mehr beeinflusst hat als angenommen. Daher sollte in weiteren Studien darauf geachtet werden, in simultanen Bildbedingungen die Frames möglichst alle so auf dem Screen anzuordnen, dass sie auch ohne Interaktivität sichtbar sind.

Das Interaktivitätsfeature hat eher zu einer Verminderung der Lernleistung geführt, da die parallele Bildgruppe in sieben von acht Vergleichen deskriptiv schlechter war als die beiden seriellen Bildgruppen. Eine mögliche Begründung für dieses Ergebnis basierte auf der Beobachtung aus der Pilotstudie, dass die Versuchspersonen der parallelen Bildgruppe die Lernzeit nicht voll zu nutzen schienen, die Bilder eher unsystematisch auf- und abscrollten und das Lernmaterial gegen Ende der Lernzeit nicht mehr bewusst zu betrachten schienen. „Eye tracking“-Studien, bei denen die Blickbewegungen und die Dauer der Fixationen der Probanden gemessen werden können, könnten dabei helfen, aufzuklären, ob diese Beobachtungen objektive Substanz haben. Dann könnte die Verwendung von Interaktivität auch einen negativen Effekt haben, der wahrscheinlich aber nur unter gewissen Umständen eintritt. Es zeigte sich nämlich zudem in der Studie, dass der Effekt von Interaktivität auch mit den motivationalen Voraussetzungen der Lerner in Zusammenhang stand. Es wurden größere bis signifikante Interaktionen zwischen der parallelen und der seriell langsamen Bildbedingung mit der Variable „Herausforderung“ gefunden, in der Richtung, dass Personen, die die Aufgabe sehr als Herausforderung für sich bewerteten, besser mit parallelen Bildern lernten, während Personen mit geringeren Herausforderungswerten besser mit den seriell langsamen Bildern zurechtkamen. Für die weitere Forschung wäre es interessant, die Wechselwirkungen zwischen interaktiven Präsentationsformen und der Motivation der Lerner noch stärker zu untersuchen. Dies hätte die praktische Relevanz, zu zeigen, dass Interaktivität, durch die ein aktiverer Lernvorgang erzeugt wird, nicht immer hilfreich ist, sondern dass ihre Effektivität von den individuellen Eigenschaften des Lerners mit abhängt.

Zudem ergaben sich auch in der zweiten Teilstudie signifikante Interaktionen zwischen den Lernbedingungen und dem Vorwissen der Teilnehmer, was für die weitere Forschung und für die Praxis ebenfalls impliziert, dass die Effektivität verschiedener statischer Bildformen individuell betrachtet werden muss.

Bei der zweiten Studie der Dissertation war methodisch auffällig, dass die Items des Nachttests sehr leicht für die Probanden waren und der Test daher vielleicht nicht optimal zwischen den Gruppen differenzieren konnte, so dass die Unterschiede zwischen den Bedingungen größtenteils nicht signifikant wurden.

Die geringe Schwierigkeit des Nachttests könnte an einer zu hohen Lernzeit sowie an einer zu hohen Testzeit der Items gelegen haben. Bei weiteren Studien sollte die Lernzeit von 10 s pro Pferdegangart und die Testzeiten der Items somit reduziert werden. Zudem hatte der Nachttest, nach Aussage einiger Teilnehmer in der Pilotstudie, für sie einen zusätzlichen Übungseffekt gehabt, wodurch die geringen Itemschwierigkeiten, zumindest bei den Aufgaben am Ende des Tests, mit erklärt werden könnten. Eventuell sollte ein Nachttest bei der Art von Aufgaben generell etwas kürzer gehalten werden. Man könnte auch das Testpferd stärker manipulieren, so dass es dem Lernpferd weniger ähnlich sieht, um die Schwierigkeiten der Aufgaben zu erhöhen und um eventuell stärkere Effekte zu erhalten. Eine Veränderung der Gangrichtung oder die Verwendung eines anderen Pferdes mit einem anderen Gangstil wären hierbei mögliche Optionen. Man könnte auch ein völlig anderes Lernmaterial verwenden, das schwierigere, weniger alltagsnahe Bewegungsformen thematisiert. Aus der geringen Schwierigkeit des Nachttests heraus leitete sich zudem die Vermutung ab, dass für die Lösung von bildlichen Kategorisierungsaufgaben eventuell kein vollständiges dynamisches perzeptuelles Schema benötigt wird, sondern dass nur einige seiner Features für die Bearbeitung herangezogen werden. In dem Fall wäre es vielleicht sinnvoll, einen anderen, schwierigeren Aufgabentyp zu wählen, um die Qualität des dynamischen Schemas zu messen. Beispielsweise könnte man die Probanden die Bewegungsabfolgen der Pferdegangarten aufzeichnen oder vervollständigen lassen, wozu sie nur in der Lage wären, wenn sie das Bewegungsmuster zu einem gewissen Grad verstanden hätten. Dann würde vielleicht der positive Effekt einer Animation beim Lernen von biologischen Bewegungen noch mehr heraustreten.

Es ergaben sich zwar auch bei dieser Studie keine bedeutsamen Leistungsunterschiede zwischen Animationen und parallelen Bildern, jedoch empfanden die Probanden beim Betrachten der Animationen zu den Pferdegangarten hochsignifikant weniger Cognitive

Load als die Probanden, die mit den statischen Bildern konfrontiert waren. Dies kann als ein Hinweis darauf gewertet werden, dass beim praktischen Erlernen von biologischen Bewegungen der Einsatz von Animationen hilfreich sein kann und dass der Lernvorgang dadurch erleichtert wird. Beispielsweise könnten Sportstudenten beim Erlernen verschiedener Schwimmstile oder Medizinstudenten beim Erlernen von Ganganomalien bei verschiedenen Krankheiten durch Animationen oder Videos effizienter unterstützt werden als durch stationäre Bilder, die in einem Lehrbuch enthalten sind.

In der zweiten Studie der Dissertation konnten teilweise signifikante Wechselwirkungen zwischen der Lernbedingung und den räumlichen Fähigkeiten von Probanden gefunden werden, was abermals für die Einbeziehung der individuellen Eigenschaften der Teilnehmer in den Vergleich von Animationen mit statischen Bildern spricht. Es konnten jedoch keine signifikanten Interaktionen zwischen der Animations- und der parallelen Bildbedingung mit der Wahrnehmungsvorerfahrung der Teilnehmer gefunden werden, was vielleicht auch daran gelegen hat, dass die Messung der Wahrnehmungsvorerfahrung in der aktuellen Studie nicht optimal war und dass man eher differenziertere Items zur Erfassung dieser Variable nehmen sollte, um eventuell vorhandene Interaktionen messen zu können.

Die verschiedenen Vergleiche zwischen Animationen und statischen Bildern und die Vergleiche zwischen den unterschiedlichen Präsentationsformen der statischen Bilder in den vorliegenden Studien haben gezeigt, dass die Haupteffekte zwischen den Bedingungen kaum signifikant wurden und dass die Bedeutsamkeit eher in den Interaktionen lag. Das bedeutet, dass in der weiteren Forschung zum Vergleich von Animationen mit statischen Bildern die Fragestellungen eventuell noch differenzierter betrachtet werden müssen. Die Effektivität der Präsentationsformen scheint von Personenmerkmalen, wie dem Vorwissen oder den räumlichen Fähigkeiten, mit abzuhängen. Diesbezüglich sind zwar schon einige Studien vorhanden (z.B. Blake, 1977), es könnten aber noch in weitere Designs Interaktionsanalysen eingebaut werden. Zudem könnten auch noch weitere Personenmerkmale auf moderierende Funktionen hin untersucht werden, beispielsweise, inwiefern eine Person überhaupt gerne am Computer lernt und modernen Präsentationsformen wie Animationen gegenüber positiv eingestellt ist. Eventuell könnte auch das Alter der Versuchspersonen eine Rolle spielen, da heutige Jugendliche und junge Erwachsene mit Computerspielen und multimedialen Darstellungen aufgewachsen sind und daher mit animierten Präsentationsformen besser zurechtkommen dürften als ältere Personen.

Da sich zudem auch bei einigen anderen Studien in der Literatur keine signifikanten Unterschiede zwischen Animationen und statischen Bildern finden ließen (z.B. Morrison & Tversky, 2001), wäre vielleicht auch eine fruchtbare Fragestellung, ob man diese beiden Darstellungsformen miteinander verbinden kann, um optimale Lernergebnisse zu erhalten. Beispielsweise könnte man Personen erst die Key-Frames eines Vorganges zeigen, damit sie die einzelnen Bestandteile des Prozesses verstehen können und damit ihre Aufmerksamkeit auf relevante Aspekte gelenkt wird. Die Präsentation einer Animation im Anschluss würde die Informationslücken zwischen den Frames schließen. In dieser Hinsicht wären auch noch andere Kombinationsmöglichkeiten zwischen Animationen und statischen Bildern denkbar, die man in weiteren Studien näher untersuchen könnte.

8 Literaturverzeichnis

- Atkinson, J. W. (1974). Motivational determinants of intellectual performance and cumulative achievement. In J. W. Atkinson & J. O. Raynor (Hrsg.), *Motivation and achievement* (S. 389-410). Washington, DC: Winston.
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York, NY: Holt, Rinehart & Winston.
- Ayres, P. & Paas, F. G. W. C. (2007). Can the Cognitive Load Approach Make Instructional Animations More Effective? *Applied Cognitive Psychology*, 21, 811-820.
- Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Clarendon Press.
- Baek, Y. & Layne, B. (1988). Color, graphics and animation in a computer-assisted learning tutorial lesson. *Journal of Computer-Based Instruction*, 15(4), 131-135.
- Barclay, C. D., Cutting, J. E. & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics*, 23, 145-152.
- Bayrhuber, H. & Kull, U. (Hrsg.). (1989). *Linder Biologie* (20., neubearbeitete Auflage). Stuttgart: J. B. Metzler.
- Bétrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 287-296). New York, NY: Cambridge University Press.
- Bétrancourt, M., Morrison, J. B. & Tversky, B. (2001). Les animations sont-elles vraiment plus efficaces? *Revue d'intelligence artificielle*, 14, 149-166.
- Bétrancourt, M. & Realini, N. (2005). *Faut-il vraiment laisser le déroulement d'une animation sous le contrôle de l'apprenant ?* Paper presented at the 11ème journée d'étude sur le traitement cognitif de systèmes d'informations complexes (JETCSIC), Nice (France).
- Blake, T. (1977). Motion in instructional media: some subject-display mode interactions. *Perceptual and Motor Skills*, 44, 975-985.
- Bogacz, S. & Trafton, J. G. (2005). Understanding dynamic and static displays: using images to reason dynamically. *Cognitive Systems Research*, 6, 312-319.
- Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (3. Auflage). Berlin: Springer.
- Boucheix, J.-M. & Schneider, E. (2009). Static and animated presentations in learning dynamic mechanical systems. *Learning and Instruction*, 19, 112-127.
- Brandl, A. (1982). *Modernes Reiten. Schritt – Trab – Galopp*. Zürich: BLV Verlagsgesellschaft.

- Bratfisch, O., Borg, G. & Dornic, S. (1972). *Perceived item-difficulty in three tests of intellectual performance capacity*. Report No. 29. Stockholm: Institute of Applied Psychology.
- Brown, J. (1983). Learning-by-doing revisited for electronic learning environments. In M. A. White (Hrsg.), *The future of electronic learning* (S. 13-32). Hillsdale, NJ: Erlbaum.
- Bryant, P. E. & Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature*, 232, 456-458.
- Caraballo, A. (1985). *An experimental study to investigate the effects of computer animation on the understanding and retention of selected levels of learning outcomes*. Unpublished doctoral dissertation, The Pennsylvania State University.
- Caraballo, J. (1985). *The effect of various visual display modes of selected educational objectives*. Unpublished doctoral dissertation, The Pennsylvania State University.
- Clark, R. (1983). Reconsidering research on learning from media. *Review of Educational Research*, 17, 92-101.
- De Kleer, J. (1975). *Qualitative and quantitative knowledge in classical mechanics*. Master thesis. Cambridge, MA: MIT Press.
- Diehl, J. M. & Arbinger, R. (2001). *Einführung in die Inferenzstatistik* (3. Auflage). Eschborn bei Frankfurt am Main: Verlag Dietmar Klotz.
- Dittrich, W. H., Troscianko, T., Lea, S. & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25, 727-738.
- Dwyer, F. (1978). *Strategies for improving visual learning*. State College, PA: Learning Services.
- Carney, R. N. & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14, 5-26.
- Catrambone, R. & Fleming Seay, A. (2002). Using animation to help students learn computer algorithms. *Human Factors*, 44(3), 495-511.
- Engelkamp, J. (1991). Bild und Ton aus Sicht der kognitiven Psychologie. *Medienpsychologie*, 3, 278-299.
- Freyd, J. (1983a). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics*, 33, 575-581.
- Freyd, J. (1983b). Representing the dynamics of a static form. *Memory and Cognition*, 11, 342-346.
- Gagné, R. (1985). *The conditions of learning* (4th edition). New York, NY: Holt, Rinehart & Winston.
- Garg, A. X., Norman, G. R. & Sperotable, L. (2001). How medical students learn spatial anatomy. *The Lancet*, 357, 363-364.

- Goldstein, A. G., Chance, J. E., Hoisington, M. & Buescher, K. (1982). Recognition memory for pictures: Dynamic vs. static stimuli. *Bulletin of the Psychonomic Society*, 20(1), 37-40.
- Grossman, E. D. & Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. *Vision Research*, 42, 1475-1482.
- Guttman, J., Levin, J. & Pressley, M. (1977). Pictures, partial pictures, and young children's oral prose learning. *Journal of Educational Psychology*, 69, 473-480.
- Hegarty, M. (1992). Mental Animation: Inferring Motion From Static Displays of Mechanical Systems. *Journal of Experimental Psychology: Memory and Cognition*, 18(5), 1084-1102.
- Hegarty, M. (2004). Mechanical reasoning as mental simulation. *TRENDS in Cognitive Sciences*, 8, 280-285.
- Hegarty, M. & Kriz, S. (2008). Effects of Knowledge and Spatial Ability on Learning from Animation. In R. K. Lowe & W. Schnotz (Hrsg.), *Learning with Animation* (S. 3-29). New York, NY: Cambridge University Press.
- Hegarty, M., Kriz, S. & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction*, 21, 325-360.
- Hegarty, M. & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Hrsg.), *Handbook of higher-level spatial cognition* (S. 121-169). New York, NY: Cambridge University Press.
- Heiser, J. & Tversky, B. (2002). *Diagrams and descriptions in acquiring complex systems. Proceedings of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hidrio, C. & Jamet, E. (2002). Compréhension d'un dispositif technique: Apports d'une illustration dynamique et des traitements multiples. *Psychologie Française*, 47, 61-67.
- Johansson, G. (1975). Visual motion perception. *Scientific American*, 232, 76-89.
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Cambridge University Press.
- Kalyuga, S., Ayres, P., Chandler, P. & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38(1), 23-31.
- Kintsch, W. (1998). *Comprehension. A paradigm for cognition*. Cambridge, MA: Cambridge University Press.
- Knauff, M. & Johnson-Laird, P. N. (2002). Visual imagery can impede reasoning. *Memory and Cognition*, 30, 363-371.
- Kosslyn, S. M. (1994). *Image and brain*. Cambridge, MA: MIT Press.

- Kosslyn, S. M., Ball, T. M. & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 47-60.
- Large, A., Beheshti, J., Breuleux, A. & Renaud, A. (1996). The effect of animation in enhancing descriptive and procedural texts in a multimedia learning environment. *Journal of the American Society for Information Science*, 47, 437-448.
- Larkin, J. H. (1983). The role of problem representation in physics. In D. Gentner & A. L. Stevens (Hrsg.), *Mental models* (S. 75-98). Hillsdale, NJ: Erlbaum.
- Larkin, J. H. & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65-99.
- Lewalter, D. (1997). *Lernen mit Bildern und Animationen*. Münster: Waxmann.
- Liepmann, D., Beauducel, A., Brocke, B. & Amthauer, R. (2007). *I-S-T 2000 R – Intelligenz-Struktur-Test 2000 R* (2., erweiterte und überarbeitete Auflage). Göttingen: Hogrefe.
- Lowe, R. K. (1993). Constructing a mental representation from an abstract technical diagram. *Learning and Instruction*, 3, 157-179.
- Lowe, R. K. (1996). Background knowledge and the construction of a situational representation from a diagram. *European Journal of Psychology of Education*, 11(4), 377-397.
- Lowe, R. K. (1999a). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, 14(2), 225-244.
- Lowe, R. K. (1999b). Domain-specific constraints on conceptual change in knowledge acquisition from diagrams. In W. Schnotz, S. Vosniadou & M. Carretero (Hrsg.), *New perspectives on conceptual change* (S. 223-245). Amsterdam: Elsevier.
- Lowe, R. K. (2001). Understanding Information Presented by Complex Animated Diagrams. In J. F. Rouet, J. J. Levonen & A. Biarreau (Hrsg.), *Multimedia Learning: Cognitive and Instructional Issues* (S. 65-74). London: Pergamon.
- Lowe, R. K. (2006). Changing perceptions of animated diagrams. In D. Barker-Plummer, R. Cox & N. Swoboda (Hrsg.), *Diagrammatic representations and inference* (S. 168-172). Berlin: Springer
- Lowe, R. K. (2008). Learning from Animation - Where to Look, When to Look. In R. K. Lowe & W. Schnotz (Hrsg.), *Learning with Animation* (S. 49-68). New York, NY: Cambridge University Press.
- Marcus, N., KhengJoo, A. L., Beng-Fei, K. & Ayres, P. (2006). *Animations with a trace lead to more effective learning than equivalent static graphics*. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.
- Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, 32, 1-19.
- Mayer, R. E. (2001). *Multimedia learning*. New York, NY: Cambridge University Press.

- Mayer, R. E. (2008). Research-Based Principles for Learning with Animation. In R. K. Lowe & W. Schnotz (Hrsg.), *Learning with Animation* (S. 30-48). New York, NY: Cambridge University Press.
- Mayer, R. E. & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84, 444-452.
- Mayer, R. E. & Chandler, P. (2001). When Learning is Just a Click Away: Does Simple User Interaction Foster Deeper Understanding of Multimedia Messages? *Journal of Educational Psychology*, 93(2), 390-397.
- Mayer, R. E., Hegarty, M., Mayer, S. & Campbell, J. (2005). When static media promote active learning: annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied*, 11, 256-265.
- Mayring, P. (2002). *Einführung in die qualitative Sozialforschung* (5. Auflage). Weinheim: Beltz Verlag.
- McCloud, S. (1994). *Understanding comics: The invisible art*. New York, NY: Harper Perennial.
- Morrison, J. B. & Tversky, B. (2001). The (in)effectiveness of animation in instruction. In J. Jacko & A. Sears (Hrsg.), *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems* (S. 377-378). Seattle, WA: ACM.
- Narayanan, N. H. & Hegarty, M. (1998). On designing comprehensible hypermedia manuals. *International Journal of Human-Computer Studies*, 48, 267-301.
- Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84, 429-434.
- Paas, F. G. W. C., van Gerven, P. W. M. & Wouters, P. (2007). Instructional Efficiency of Animation: Effects of Interactivity through Mental Reconstruction of Static Key Frames. *Applied Cognitive Psychology*, 21, 783-793.
- Paas, F. G. W. C. & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skill: A cognitive-load approach. *Journal of Educational Psychology*, 86, 122-133.
- Paas, F. G. W. C., van Merriënboer, J. J. G. & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, 79, 419-430.
- Paivio, A. (1978). A dual coding approach to perception and cognition. In H. L. Pick & E. Saltzman (Hrsg.), *Modes of perceiving and processing information* (S. 39-52). Hillsdale, NJ: Erlbaum.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, UK: Oxford University Press.
- Perez, E. C. & White, M. A. (1985). Student evaluation of motivational and learning attributes of microcomputer software. *Journal of Computer-Based Instruction*, 12, 39-43.

- Reed, S. (1985). Effect of computer graphics on improving estimates to algebra word problems. *Journal of Educational Psychology*, 77(3), 285-298.
- Reigeluth, C. M. & Stein, F. S. (1983). The elaboration theory of instruction. In C. M. Reigeluth (Hrsg.), *Instructional-design theories and models: An overview of their current status* (S. 335-381). Hillsdale, NJ: Erlbaum.
- Rheinberg, F., Vollmeyer, R. & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica*, 47, 57-66.
- Richter, T. (2007). Wie analysiert man Interaktionen von metrischen und kategorialen Prädiktoren? Nicht mit Median-Splits! *Zeitschrift für Medienpsychologie*, 19, 116-125.
- Rieber, L. P. (1989). The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. *Journal of Educational Computing Research*, 5, 431-444.
- Rieber, L. P. (1990). Animation in Computer-Based Instruction. *Educational Technology Research & Development*, 38(1), 77-86.
- Rieber, L. P. (1991). Animation, incidental learning and continuing motivation. *Journal of Educational Psychology*, 83, 318-328.
- Rieber, L. P., Boyce, M. J. & Assad, C. (1990). The effects of computer animation on adult learning and retrieval tasks. *Journal of Computer Based Instruction*, 17, 46-52.
- Rieber, L. P. & Kini, A. S. (1991). Theoretical foundations of instructional applications of computer-generated animated visuals. *Journal of Computer-Based Instruction*, 18, 83-88.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Hrsg.), *Cognition and categorization* (S. 27-48). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rosenzweig, M. R., Leiman, A. L. & Breedlove, S. M. (1999). *Biological Psychology. An Introduction to Behavioral, Cognitive and Clinical neuroscience* (2nd edition). Sunderland, MA: Sinauer Associates, Inc.
- Salomon, G. (1979). *Interaction of media, cognition and learning: An exploration of how symbolic forms cultivate mental skills and affect knowledge acquisition*. San Francisco, CA: Jossey-Bass.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403-428.
- Schnotz, W. (1994). *Aufbau von Wissensstrukturen. Untersuchungen zur Kohärenzbildung beim Wissenserwerb mit Texten*. Weinheim: Beltz Verlag.
- Schnotz, W. (2005). An Integrated Model of Text and Picture Comprehension. In R. E. Mayer (Hrsg.), *Cambridge Handbook of Multimedia Learning* (S. 49-69). New York, NY: Cambridge University Press.
- Schnotz, W. (2006). *Pädagogische Psychologie – Workbook*. Weinheim: Beltz Verlag.

- Schnotz, W. & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction, 13*, 141-156.
- Schnotz, W., Böckheler, J. & Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. *European Journal of Psychology of Education, 14*(2), 245-265.
- Schnotz, W. & Kürschner, C. (2007). A Reconsideration of Cognitive Load Theory. *Educational Psychological Review, 19*, 469-508.
- Schnotz, W. & Lowe, R. K. (2008). A unified view of learning from animated and static graphics. In R. K. Lowe & W. Schnotz (Hrsg.), *Learning with Animation. Research Implications for Design* (S. 304-356). New York, NY: Cambridge University Press.
- Schnotz, W. & Preuß, A. (1999). Task-dependent construction of mental models as a basis for conceptual change. In G. Rickheit & C. Habel (Hrsg.), *Mental models in discourse processing and reasoning* (S. 131-167). Oxford, UK: Elsevier Science Publishers.
- Schwartz, D. L. & Black, J. B. (1996a). Shuttling between depictive models and abstract rules: Induction and fall-back. *Cognitive Science, 20*, 457-497.
- Schwartz, D. L. & Black, J. B. (1996b). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology, 30*, 154-219.
- Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bulhore, E. T., Brammer, M. & David, A. S. (2000). The functional neuroanatomy of implicit-motion perception or "representational momentum." *Current Biology, 10*, 16-22.
- Seufert, T. (2003). *Kohärenzbildung beim Wissenserwerb mit multiplen Repräsentationen – Wirksamkeit von Kohärenzbildungshilfen*. Berlin: Logos.
- Shah, P. & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*, 4-27.
- Sims, V. K. & Hegarty, M. (1997). Mental animation in the visuospatial sketchpad: Evidence from dual-task studies. *Memory and Cognition, 25*(3), 321-332.
- Spangenberg, R. (1973). The motion variable in procedural learning. *AV Communication Review, 21*(4), 419-436.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.
- Sweller, J. (2005). Implications of Cognitive Load Theory for Multimedia Learning. In R. E. Mayer (Hrsg.), *Cambridge Handbook of Multimedia Learning* (S. 19-31). New York, NY: Cambridge University Press.
- Sweller, J., Chandler, P., Tierney, P. & Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *Journal of Experimental Psychology: General, 119*, 176-192.

- Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychological Review*, 10, 251-296.
- Thompson, S. V. & Riding, R. J. (1990). The effect of animated diagrams on the understanding of a mathematical demonstration in 1 to 14-years-old pupils. *British Journal of Educational Psychology*, 60, 93-98.
- Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, 2, 371-387.
- Tversky, B., Heiser, J., Mackenzie, R., Lozano, S. & Morrison, J. B. (2008). Enriching Animations. In R. K. Lowe & W. Schnotz (Hrsg.), *Learning with Animation. Research Implications for Design* (S. 263-285). New York, NY: Cambridge University Press.
- Tversky, B., Morrison, J. B. & Bétrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies*, 57, 247-262.
- Wade, N. J. & Swanston, M. T. (2001). *Visual Perception: An Introduction* (2nd edition). Hove, East Sussex: Psychology Press Ltd.
- Willows, D. (1978). A picture is not always worth a thousand words: Pictures as distractors in reading. *Journal of Educational Psychology*, 70(2), 255-262.
- Zacks, J. M. & Tversky, B. (2001). Event Structure in Perception and Conception. *Psychological Bulletin*, 127(1), 3-21.
- Zacks, J. M., Tversky, B. & Iyer, G. (2001). Perceiving, remembering and communicating structure in events. *Journal of Experimental Psychology: General*, 130, 29-58.

Internetquellen:

<http://www.geysir.com/de/islandpferde/gangarten>

9 Anhang

A1:	Studie 1 – Vorpilotstudie: Lernmaterialien.....s. beiliegende CD
A2:	Studie 1 – Vorpilotstudie: Messinstrumente.....358
B1:	Studie 1 – Pilotstudie: Lernmaterialien.....s. beiliegende CD
B2:	Studie 1 – Pilotstudie: Messinstrumente.....370
B3:	Studie 1 – Pilotstudie: Ratingsysteme.....s. beiliegende CD
C1:	Studie 1 – Hauptteilstudien 1 und 2: Lernmaterialien.....s. beiliegende CD
C2:	Studie 1 – Hauptteilstudien 1 und 2: Messinstrumente.....377
C3:	Studie 1 – Hauptteilstudien 1 und 2: Ratingsysteme.....s. beiliegende CD
D1:	Studie 2 – Pilotstudie: Lernmaterialien.....s. beiliegende CD
D2:	Studie 2 – Pilotstudie: Messinstrumente.....380
D3:	Studie 2 – Pilotstudie: Ratingsysteme.....s. beiliegende CD
E1:	Studie 2 – Hauptstudie: Lernmaterialien.....s. beiliegende CD
E2:	Studie 2 – Hauptstudie: Messinstrumente.....385
E3:	Studie 2 – Hauptstudie: Ratingsysteme.....s. beiliegende CD

Anhang A2 - 1: Soziodemographischer Fragebogen in der Vorpilotstudie**Angaben zur Person:**

Alter: Ich bin _____ Jahre alt.

Geschlecht: Ich bin männlich weiblich

Schulabschluß (bitte den bisher am höchsten erreichten nennen):

- | | | |
|---|-------------------------------------|--|
| <input type="checkbox"/> Hauptschulabschluß | <input type="checkbox"/> Fachabitur | <input type="checkbox"/> Fachhochschulabschluß |
| <input type="checkbox"/> Mittlere Reife | <input type="checkbox"/> Abitur | <input type="checkbox"/> Hochschulabschluß |

Beruf:

- Schüler Student Angestellter Beamter

Wenn Sie **Student** sind:

Welches Fach studieren _____ Sie?

In welchem Semester studieren Sie dieses Fach? _____

Anhang A2 - 2: Vorwissenstest zu Synapsen in der Vorpilotstudie

In den nachfolgenden Fragen möchten wir gerne wissen, wie gut Sie sich mit den Vorgängen an **chemischen** Synapsen auskennen. Wenn Sie die Antwort zu einer Frage **nicht wissen**, machen Sie dies bitte durch ein „?“ oder durch ein „/“ deutlich.

1. Welche Zellart ist in der Regel an der Bildung einer chemischen Synapse beteiligt?

2. Was findet grundlegend an einer Synapse statt (**1 Satz**)?

3. Was sind Ionen und welche beiden Arten von Ionen gibt es?

4. Wie nennt man allgemein einen Stoff, der Signale über den synaptischen Spalt hinweg überträgt?

5. Was passiert als Nächstes, wenn ein präsynaptisches Aktionspotential am synaptischen Endknopf ankommt?

6. Was ist die direkte Folge des Einstroms von Calciumionen im synaptischen Endknopf?

7. Wie heißt das Enzym, das Acetylcholin spaltet?

8. Wozu führt der Einstrom von Natriumionen an der postsynaptischen Membran einer Synapse?

9. Nennen Sie drei Gifte, die die Funktion einer Synapse beeinträchtigen.

10. Beschreiben Sie die Wirkung von **einem** der von Ihnen eben genannten Gifte an einer Synapse.

(Schreiben Sie bitte den Namen des zugehörigen Gifts nochmals dazu.)

11. Nennen Sie die wichtigsten Teilprozesse des Gesamtvorganges an einer erregenden Synapse in der richtigen Reihenfolge (**ca. 8 Teilprozesse**).

Anhang A2 - 3: Nachtest zu Synapsen in der Vorpilotstudie – Version 1

Es folgen nun einige Fragen zu den Vorgängen an der Synapse, die Sie auf dem Lernmaterial gesehen haben. Falls Sie eine Frage **nicht beantworten** können, machen Sie dies bitte wieder durch ein „?“ oder ein „ / „ deutlich. Wenn Sie eine Seite bearbeitet haben, blättern Sie bitte nicht mehr zurück.

1. Wodurch werden die Vorgänge an einer Synapse am Anfang ausgelöst?

2. Was geschieht an der nachsynaptischen Membran, wenn der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle andockt?

3. Beschreiben Sie die Folgen im Endknopf, wenn ein Mensch Calciummangel hat und nur wenig Calcium an den erregenden Synapsen vorhanden ist.

4. Curare (indianisches Pfeilgift) ist ein Stoff, der sich wie der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle setzt. Es öffnet diese aber nicht. Was ist die Folge, wenn Curare in hoher Konzentration im Spalt vorhanden ist?

5. Das Gift Botulinustoxin (vorhanden in verdorbenem Fisch) hemmt die Acetylcholinfreisetzung aus dem Endknopf. Warum kann dann kein nachsynaptisches Aktionspotential entstehen?

6. Was ist die direkte Folge der Öffnung der Calciumkanäle an der Synapse (**1 Satz**)?

7. Welchen Einfluß hat es auf die Erregungsweiterleitung, wenn man die Menge der Natriumionen im Spalt verringert?

8. Stärkere Reize werden im Nervensystem durch eine höhere Frequenz (=Häufigkeit) des Aktionspotentials ausgedrückt. Was bedeutet es für die Menge des aus dem Endknopf ausgeschütteten Transmitters, wenn sich die Häufigkeit des Aktionspotentials erhöht?

9. Welche Vorgänge liegen zwischen dem Ausschütten des Transmitters Acetylcholin in den Spalt und dem Einstrom von Natriumionen an der nachsynaptischen Membran?

10. Bei einer Form der Calciumkanalkrankheit sind die Calciumkanäle abnorm häufig geschlossen. Inwiefern beeinflusst dies die Erregungsweiterleitung, wenn ein Aktionspotential ankommt?

11. Bei einer anderen Form der Calciumkanalkrankheit befinden sich die Calciumkanäle besonders häufig in offenem Zustand. Inwiefern verändert sich die Erregungsweiterleitung?

12. Was geschieht mit dem Transmitter Acetylcholin am Ende des Übertragungsprozesses?

13. Das Fliegenpilzgift Muscarin setzt sich wie der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle und öffnet sie. Es wird aber im Gegensatz zu dem Transmitter im Spalt nicht abgebaut, sondern bleibt dort länger. Was passiert an der Synapse?

14. Atropin (Gift der Tollkirsche) dockt an die gleichen Rezeptoren an wie der Transmitter Acetylcholin, hat aber nicht dessen Wirkung. Was ist die Folge, wenn eine große Menge Atropin in den synaptischen Spalt gelangt?

15. Muscarin setzt sich wie der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle und öffnet sie. Es wird aber im Spalt nicht abgebaut.

Atropin dockt ebenfalls an die Rezeptoren der Natriumkanäle an, öffnet sie aber nicht.

Führt es zu einer Verschlimmerung oder Verminderung der Vergiftungserscheinungen, wenn die Nervengifte Muscarin und Atropin gleichzeitig an einer Synapse vorhanden sind? Begründen Sie.

16. Was folgt unmittelbar, nachdem die Bläschen auf die Zellmembran zugewandert sind?

17. Was bewirkt die Cholinesterase?

18. Das Aktionspotential der nachsynaptischen Nervenzelle wird durch eine positive Aufladung ihres Zellinnern ausgelöst. Welche Art von Ionenströmen müßten in die nachsynaptische Zelle erfolgen, um die Erregungsweiterleitung zu erschweren?

19. Das Insektenvertilgungsmittel E 605 hemmt die Aktivität der Cholinesterase. Was ist die Folge?

20. Wodurch werden die Calciumkanäle im Endknopf geöffnet?

21. Was ist die Folge für die Erregungsweiterleitung, wenn die Bläschen nicht mehr mit der vorsynaptischen Membran verschmelzen können? Begründen Sie kurz (**1 Satz**).

22. Was passiert, wenn das Immunsystem die Natriumkanäle in der nachsynaptischen Membran angreift und sich ihre Anzahl dadurch verringert?

23. Welche Prozesse laufen zwischen dem Ankommen eines Aktionspotentials am Endknopf und dem Verschmelzen der Bläschen mit der vorsynaptischen Zellmembran ab?

24. Bei der Alzheimer-Erkrankung produzieren die Nervenzellen zu wenig Acetylcholin. Warum verschreibt der Arzt hier häufig Cholinesterasehemmer?

25. Warum sollte man Cholinesterasehemmer nicht gesunden Menschen geben? Begründen Sie mittels der Vorgänge an der Synapse.

26. Wodurch werden die Natriumkanäle in der nachsynaptischen Membran geöffnet?

27. Was folgt direkt auf den Natriumionen - Einstrom an der nachsynaptischen Membran?

28. Welche Folge hat eine Erhöhung der Calciumkonzentration an der Synapse für die Transmitterausschüttung?

29. Das Gift der „Schwarzen Witwe“ hemmt unter anderem die Wiederaufnahme des Transmitters Acetylcholin in den Endknopf. Dieses wird aber durch seine Spaltung nach wie vor deaktiviert. Beschreiben Sie die Folge.

30. Welcher Prozeß geschieht, bevor die Bläschen auf die Zellmembran zuwandern?

Anhang A2 - 4: Nachtest zu Synapsen in der Vorpilotstudie – Version 2

Aus Platzgründen werden nur die Items dieses Nachtests genannt. Die formale Gestaltung des Tests war analog zu der ersten Version des Nachtests, daher sei für die Ansicht der Gestaltungsmerkmale auf den Anhang A2 – 3 verwiesen.

1. Wodurch werden die Vorgänge an einer Synapse anfangs ausgelöst?
2. Was geschieht an der nachsynaptischen Membran, wenn der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle andockt?
3. Beschreiben Sie die Folgen im Endknopf, wenn ein Mensch Calciummangel hat und nur wenig Calcium an den erregenden Synapsen vorhanden ist.
4. Die Entstehung eines Aktionspotentials hängt mit einer positiven Aufladung des Zellinneren der nachsynaptischen Zelle zusammen. Was passiert, wenn ein Transmitter anstatt Natriumkanäle (Na^+), Chloridkanäle (Cl^-) an der nachsynaptischen Membran öffnet und Chloridionen (Cl^-) in die Zelle einströmen?
5. Curare (indianisches Pfeilgift) ist ein Stoff, der sich an die Rezeptoren der Natriumkanäle setzt, ohne die Wirkung eines Transmitters zu haben. Was ist die Folge, wenn Curare in hoher Konzentration im Spalt vorhanden ist?
6. Was ist die direkte Folge der Öffnung der Calciumkanäle an der Synapse (**1 Satz**)?
7. Was geschieht mit dem Transmitter Acetylcholin am Ende des Übertragungsprozesses?
8. Was folgt direkt auf den Natriumionen - Einstrom an der nachsynaptischen Membran?
9. Das Gift Botulinustoxin (vorhanden in verdorbenem Fisch) hemmt die Acetylcholinfreisetzung aus dem Endknopf. Warum kann dann kein nachsynaptisches Aktionspotential entstehen?
10. Stärkere Reize werden im Nervensystem durch eine höhere Frequenz (=Häufigkeit) des Aktionspotentials ausgedrückt. Wie ist das Verhältnis zwischen der Frequenz des Aktionspotentials und der Menge des ausgeschütteten Transmitters?
11. Welche Folge hat eine Erhöhung der Calciumkonzentration an der Synapse für die Transmitterausschüttung?
12. Bei der Calciumkanalkrankheit sind die Calciumkanäle entweder abnorm häufig geschlossen oder abnorm häufig geöffnet. In welche beiden Richtungen können diese beiden Krankheitsfälle jeweils die Erregungsweiterleitung beeinflussen und warum?
13. Das Fliegenpilzgift Muscarin setzt sich wie der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle und öffnet sie. Es wird aber im Gegensatz zu dem Transmitter im Spalt nicht abgebaut, sondern bleibt dort länger. Was passiert an der Synapse?
14. Atropin (Gift der Tollkirsche) dockt an die gleichen Rezeptoren an wie der Transmitter Acetylcholin, hat aber nicht dessen Wirkung. Was ist die Folge, wenn eine große Menge Atropin in den Spalt gelangt?

15. Muscarin setzt sich wie der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle und öffnet sie. Es wird aber im Spalt nicht abgebaut.
Atropin dockt ebenfalls an die Rezeptoren der Natriumkanäle an, öffnet sie aber nicht.
Führt es zu einer Verschlimmerung oder Verminderung der Vergiftungserscheinungen, wenn die Nervengifte Muscarin und Atropin gleichzeitig an einer Synapse vorhanden sind? Begründen Sie.
16. Welcher Prozeß geschieht, bevor die Bläschen auf die Zellmembran zuwandern?
17. Was folgt unmittelbar, nachdem die Bläschen auf die Zellmembran zugewandert sind?
18. Was passiert an einer Synapse, bevor sich die Calciumkanäle öffnen?
19. Das Aktionspotential der nachsynaptischen Nervenzelle wird durch eine positive Aufladung ihres Zellinnern ausgelöst. Welche Art von Ionenströmen müßten an der nachsynaptischen Membran erfolgen, um die Erregungsweiterleitung zu erschweren?
20. Das Insektenvertilgungsmittel E 605 hemmt die Aktivität der Cholinesterase. Was ist die Folge?
21. Muscarin setzt sich wie der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle und öffnet sie. Es wird aber im Spalt nicht abgebaut.
Atropin dockt ebenfalls an die Rezeptoren der Natriumkanäle an, öffnet sie aber nicht.
Würden Sie bei einer Vergiftung durch E 605, das die Cholinesterase hemmt, eher Atropin oder Muscarin als lindernde Maßnahme geben? Begründen Sie.
22. Curare ist ein Gift, das sich an die Rezeptoren der Natriumkanäle setzt, aber bei genügend hoher Konzentration des Transmitters Acetylcholin verdrängt werden kann.
Könnte E 605, das die Aktivität der Cholinesterase hemmt, theoretisch ein Gegengift für Curare sein? Begründen Sie.
23. Wodurch werden die Natriumkanäle in der nachsynaptischen Membran geöffnet?
24. Welche Prozesse laufen zwischen dem Ankommen eines Aktionspotentials am Endknopf und dem Verschmelzen der Bläschen mit der vorsynaptischen Zellmembran ab?
25. Bei der Alzheimer-Erkrankung produzieren die Nervenzellen zu wenig Acetylcholin. Warum verschreibt der Arzt hier häufig Cholinesterasehemmer?
26. Warum sollte man Cholinesterasehemmer nicht gesunden Menschen geben? Begründen Sie mittels der Vorgänge an der Synapse.
27. Welche Vorgänge liegen zwischen dem Ausschütten des Transmitters Acetylcholin in den Spalt und dem Einstrom von Natriumionen an der nachsynaptischen Membran?
28. Das Gift der „Schwarzen Witwe“ erzeugt ständig offene Calciumkanäle in der vorsynaptischen Zellmembran. Zudem hemmt es die Wiederaufnahme des Acetylcholins in den Endknopf. Beschreiben Sie die Folgen dieser Wirkungen an der Synapse.
29. Bei der Krankheit „Myasthenia Gravis“ zerstört das Immunsystem einen Teil der Rezeptoren der Natriumkanäle in den Synapsen. Was sind die Folgen für das Acetylcholin und die Erregungsweiterleitung?
30. Was bewirkt die Cholinesterase (**1 Satz**)?

Anhang A2 - 5: Interviewleitfaden in der Vorpilotstudie

Zeiten stoppen für die Bearbeitung der Vortests und des Nachtests

1. Wie schätzt Du Dein Vorwissen ein?

2. Wie schwierig fandest Du es, den Vortest zu bearbeiten?

Irgendein konkretes Item zu einfach / zu schwierig?

3. Wie gut findest Du die Materialien? (Farbe, Art der Darstellung, Formen, etc.)

4. Was fehlt Deiner Meinung nach an den Materialien oder könnte man an ihnen besser machen?

5. Wie fandest Du die Länge der Lernzeit (zu lang / zu kurz)? Wie findest Du die Länge der Lernzeit für eine Person ohne Vorwissen?

6. Wie schwierig fandest Du es, den Nachtest zu bearbeiten?

7. Wie schwierig findest Du den Nachtest für eine Person ohne Vorwissen? Sind die Aufgaben für sie zu schwierig?

8. Findest Du, daß man die Items des Nachtests aus dem Lernmaterial heraus beantworten kann?

Wenn nein, welche nicht und warum?

Anm.: Zwischen den einzelnen Fragen wurden auf dem Originalbogen Freiräume zum Notieren der Antworten gelassen. Ansonsten entspricht die vorliegende formale Gestaltung der des Originals.

Anhang B2 - 1: Soziodemographischer Fragebogen in der Pilotstudie

Alter: Ich bin _____ Jahre alt.

Geschlecht: Ich bin männlich weiblich

Schulabschluss (bitte den bisher am höchsten erreichten nennen):

- Hauptschulabschluss Fachabitur Fachhochschulabschluss
 Mittlere Reife Abitur Hochschulabschluss

Beruf:

- Schüler Student Angestellter Beamter

Wenn Sie **Student** sind:

Welches Fach studieren Sie? _____

In welchem Semester studieren Sie dieses Fach? _____

Damit Ihre Angaben vertraulich bleiben, benötigen wir für die Studie einen von Ihnen erstellten Code.

Der Code setzt sich wie folgt zusammen:

1. Die ersten beiden Buchstaben des Vornamens Ihrer Mutter. (z.B. **DORIS**)
2. Die ersten beiden Buchstaben des Vornamens Ihres Vaters. (z.B. **PETER**)
3. Die ersten beiden Buchstaben Ihres Geburtsortes. (z.B. **KARLSRUHE**)
4. Der Tag Ihres Geburtstages als zweistellige Zahl (z.B. **08.9.1986**)

Der Code in diesem Beispiel wäre somit: **DOPEKA08**

Ihr Code:

--	--	--	--	--	--	--	--

Anhang B2 - 2: Vorwissenstest zu Synapsen in der Pilotstudie

Aus Platzgründen werden nur die Items dieses Vorwissenstests genannt. Die formale Gestaltung des Tests war analog zu der Version in der Vorpilotstudie, daher sei für die Ansicht der Gestaltungsmerkmale auf den Anhang A2 – 2 verwiesen.

1. Welche Zellart ist in der Regel an der Bildung einer chemischen Synapse beteiligt?
2. Was findet grundlegend an einer Synapse statt (**1 Satz**)?
3. Was sind Ionen und welche beiden Arten von Ionen gibt es?
4. Wie nennt man allgemein einen Stoff, der Signale über den synaptischen Spalt hinweg überträgt?
5. Was passiert als Nächstes, wenn ein vorsynaptisches Aktionspotential am synaptischen Endknopf ankommt (**1 Satz**)?
6. Was ist die direkte Folge des Einstroms von Calciumionen im synaptischen Endknopf?
7. Wie heißt das Enzym, das Acetylcholin spaltet?
8. Wozu führt der Einstrom von Natriumionen an der nachsynaptischen Membran einer Synapse?
9. Nennen Sie drei Gifte, die die Funktion einer Synapse beeinträchtigen.
10. Beschreiben Sie die Wirkung von **einem** der von Ihnen eben genannten Gifte an einer Synapse. (Schreiben Sie bitte den Namen des zugehörigen Gifts nochmals dazu.)
11. Nennen Sie die wichtigsten Teilprozesse des Gesamtvorganges an einer erregenden Synapse in der richtigen Reihenfolge (**ca. 8 Teilprozesse**).

Anhang B2 - 3: FAM (Rheinberg et al., 2001)

Sie werden gleich etwas zu den Vorgängen an einer Synapse am Computer lernen.

Nun wollen wir erst noch wissen, wie Ihre **momentane Einstellung** zu dieser Lernaufgabe ist. Dazu finden Sie auf dieser Seite Aussagen. Kreuzen Sie bitte jene Zahl an, die auf Sie am Besten passt.

	trifft nicht zu					trifft zu	
1. Ich mag solche Lernaufgaben.	1	2	3	4	5	6	7
2. Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein.	1	2	3	4	5	6	7
3. Wahrscheinlich werde ich die Aufgabe nicht schaffen.	1	2	3	4	5	6	7
4. Bei der Aufgabe mag ich die Rolle des Wissenschaftlers, der Zusammenhänge entdeckt.	1	2	3	4	5	6	7
5. Ich fühle mich unter Druck, bei der Aufgabe gut abschneiden zu müssen.	1	2	3	4	5	6	7
6. Die Aufgabe ist eine richtige Herausforderung für mich.	1	2	3	4	5	6	7
7. Nach dem Lesen der Instruktion erscheint mir die Aufgabe sehr interessant.	1	2	3	4	5	6	7
8. Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde.	1	2	3	4	5	6	7
9. Ich fürchte mich ein wenig davor, dass ich mich hier blamieren könnte.	1	2	3	4	5	6	7
10. Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen.	1	2	3	4	5	6	7
11. Bei Aufgaben wie dieser brauche ich keine Belohnung, sie machen mir auch so viel Spaß.	1	2	3	4	5	6	7
12. Es ist mir etwas peinlich, hier zu versagen.	1	2	3	4	5	6	7
13. Ich glaube, das kann jeder schaffen.	1	2	3	4	5	6	7
14. Ich glaube, ich schaffe diese Aufgabe nicht.	1	2	3	4	5	6	7
15. Wenn ich die Aufgabe schaffe, werde ich schon ein wenig stolz auf meine Tüchtigkeit sein.	1	2	3	4	5	6	7
16. Wenn ich an die Aufgabe denke, bin ich etwas beunruhigt.	1	2	3	4	5	6	7
17. Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten.	1	2	3	4	5	6	7
18. Die konkreten Leistungsanforderungen hier lähmen mich.	1	2	3	4	5	6	7

Anm.: Die Schriftgröße ist im Original im Testheft leicht größer

Anhang B2 - 4: Items zur Erfassung des Cognitive Loads

Nach der Lerneinheit:

Wie schwierig fanden Sie es, die Lernaufgabe zu bearbeiten?

sehr, sehr leicht	1	2	3	4	5	6	7	8	9	sehr, sehr schwierig
-------------------	---	---	---	---	---	---	---	---	---	----------------------

Wie stark haben Sie sich beim Bearbeiten der Lernaufgabe angestrengt?

sehr, sehr wenig	1	2	3	4	5	6	7	8	9	sehr, sehr stark
------------------	---	---	---	---	---	---	---	---	---	------------------

Nach dem Nachtest:

Wie schwierig fanden Sie es, die Fragen zur Synapse zu bearbeiten?

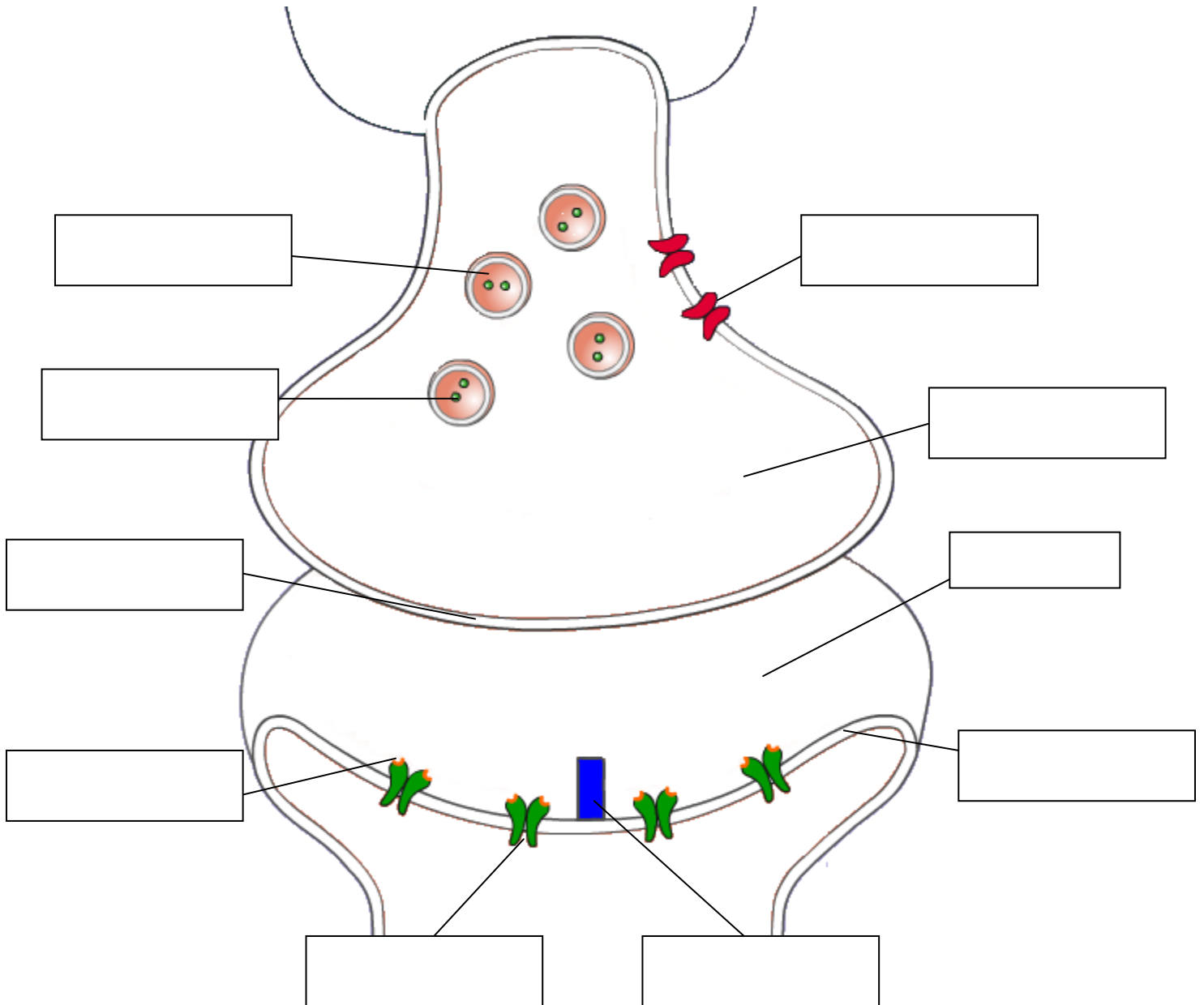
sehr, sehr leicht	1	2	3	4	5	6	7	8	9	sehr, sehr schwierig
-------------------	---	---	---	---	---	---	---	---	---	----------------------

Wie stark haben Sie sich beim Bearbeiten der Fragen zur Synapse angestrengt?

sehr, sehr wenig	1	2	3	4	5	6	7	8	9	sehr, sehr stark
------------------	---	---	---	---	---	---	---	---	---	------------------

Anhang B2 - 5: Test zur Erfassung des begrifflichen Wissens zu Synapsen

Beschriften Sie bitte mit den Begriffen, die Sie auf dem Standbild der Lerneinheit gesehen haben.



Anhang B2 - 6: Nachtest zu Synapsen in der Pilotstudie

Aus Platzgründen werden nur die Items dieses Nachtests genannt. Die formale Gestaltung des Tests war analog zu den beiden Versionen in der Vorpilotstudie, daher sei für die Ansicht der Gestaltungsmerkmale auf den Anhang A2 – 3 verwiesen.

1. Was ist die Aufgabe der Calciumionen im Übertragungsprozess?
2. Nach welchem Vorgang nimmt die Cholinesterase ihre Tätigkeit auf?
3. Das Gift Botulinustoxin (vorhanden in verdorbenem Fisch) hemmt die Acetylcholinfreisetzung aus dem Endknopf. Warum kann dann kein nachsynaptisches Aktionspotential entstehen?
4. Welchen Einfluss hat es auf die Erregungsweiterleitung, wenn man die Menge der Natriumionen im Spalt verringert?
5. Stärkere Reize werden im Nervensystem durch eine höhere Frequenz (=Häufigkeit) des Aktionspotentials ausgedrückt. Was bedeutet es für die Menge des aus dem Endknopf ausgeschütteten Transmitters, wenn sich die Häufigkeit des Aktionspotentials erhöht?
6. Nennen Sie die drei letzten Teilprozesse, mit denen der synaptische Übertragungsprozess endet, in der richtigen Reihenfolge.
7. Bei einer Form der Calciumkanalkrankheit sind die Calciumkanäle abnorm häufig geschlossen. Inwiefern beeinflusst dies die Erregungsweiterleitung, wenn ein Aktionspotential ankommt und warum?
8. Bei einer anderen Form der Calciumkanalkrankheit befinden sich die Calciumkanäle besonders häufig in offenem Zustand. Inwiefern verändert sich dann die Erregungsweiterleitung, wenn ein Aktionspotential ankommt und warum?
9. Wodurch werden die Natriumkanäle in der nachsynaptischen Membran geöffnet?
10. Welcher Prozess geschieht, bevor die Bläschen auf die Zellmembran zuwandern?
11. Das Fliegenpilzgift Muscarin setzt sich wie der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle und öffnet sie. Es wird aber im Gegensatz zu dem Transmitter im Spalt nicht abgebaut, sondern bleibt dort länger. Was passiert an der Synapse?
12. Atropin (Gift der Tollkirsche) dockt an die gleichen Rezeptoren an wie der Transmitter Acetylcholin, hat aber nicht dessen Wirkung. Was ist die Folge, wenn eine große Menge Atropin in den synaptischen Spalt gelangt?
13. Muscarin setzt sich wie der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle und öffnet sie. Es wird aber im Spalt nicht abgebaut. Atropin dockt ebenfalls an die Rezeptoren der Natriumkanäle an, öffnet sie aber nicht. Führt es zu einer Verschlimmerung oder Verminderung der Vergiftungserscheinungen, wenn die Nervengifte Muscarin und Atropin gleichzeitig an einer Synapse vorhanden sind? Begründen Sie.
14. Was geschieht mit dem Transmitter Acetylcholin am Ende des Übertragungsprozesses?

15. Welche Vorgänge liegen zwischen dem Ausschütten des Transmitters Acetylcholin in den Spalt und dem Einstrom von Natriumionen an der nachsynaptischen Membran?
16. Das Aktionspotential der nachsynaptischen Nervenzelle wird durch eine positive Aufladung ihres Zellinneren ausgelöst. Welche Art von Ionenströmen müssten in die nachsynaptische Zelle erfolgen, um die Erregungsweiterleitung zu erschweren?
17. Das Insektenvertilgungsmittel E 605 hemmt die Aktivität der Cholinesterase langfristig. Was sind die Folgen?
18. Curare ist ein Gift, das sich an die Rezeptoren der Natriumkanäle setzt, aber bei genügend hoher Konzentration des Transmitters Acetylcholin verdrängt werden kann.
Könnte ein Stoff, der die Aktivität der Cholinesterase **kurzzeitig** hemmt, ein Gegenmittel für Curare sein? Begründen Sie.
19. Welche Prozesse laufen zwischen dem Ankommen eines Aktionspotentials am Endknopf und dem Verschmelzen der Bläschen mit der vorsynaptischen Zellmembran ab (**3 Sätze**)?
20. Was ist die wichtige Funktion des Transmitters Acetylcholin im Übertragungsprozess?
21. Was ist die Folge für die Erregungsweiterleitung, wenn die Bläschen nicht mehr mit der vorsynaptischen Membran verschmelzen können? Begründen Sie kurz (**1 Satz**).
22. Bei der Krankheit „Myasthenia Gravis“ zerstört das Immunsystem einen Teil der Rezeptoren der Natriumkanäle in den Synapsen. Was sind die Folgen für das Acetylcholin und die Erregungsweiterleitung?
23. Bei der Alzheimer-Erkrankung produzieren die Nervenzellen zu wenig Acetylcholin. Warum verschreibt der Arzt hier häufig Cholinesterasehemmer?
24. Warum sollte man Cholinesterasehemmer nicht gesunden Menschen geben? Begründen Sie mittels der Vorgänge an der Synapse.
25. Die Entstehung eines Aktionspotentials hängt mit einer positiven Aufladung des Zellinneren der nachsynaptischen Zelle zusammen. Was passiert, wenn ein Transmitter anstatt Natriumkanäle (**Na⁺**), Chloridkanäle (**Cl⁻**) an der nachsynaptischen Membran öffnet und Chloridionen (**Cl⁻**) in die Zelle einströmen?
26. Welche Prozesse laufen ab, bevor sich die Calciumkanäle öffnen und nachdem sich die Calciumkanäle geöffnet haben? (**jeweils ein Prozess**)
27. Welche beiden Vorgänge passieren, nachdem die Bläschen mit der vorsynaptischen Membran verschmolzen sind?
28. Welche Folge hat eine Erhöhung der Calciumkonzentration an der Synapse für die Transmitterausschüttung?
29. Das Gift der „Schwarzen Witwe“ hemmt unter anderem die Wiederaufnahme des Transmitters Acetylcholin in den Endknopf. Dieses wird aber durch seine Spaltung nach wie vor deaktiviert. Beschreiben Sie die Folge.
30. Was geschieht an der nachsynaptischen Membran, wenn der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle andockt (**2 Sätze**)?

Anhang C2 - 2: Nachtest zu Synapsen in den Hauptteilstudien 1 und 2

Aus Platzgründen werden nur die Items dieses Nachtests genannt. Die formale Gestaltung des Tests war analog zu den Versionen aus den vorigen Studien, daher sei für die Ansicht der Gestaltungsmerkmale auf den Anhang A2 – 3 verwiesen.

1. Was ist die Aufgabe der Calciumionen im Übertragungsprozess?
2. Nach welchem Vorgang nimmt die Cholinesterase ihre Tätigkeit auf?
3. Das Gift Botulinustoxin (vorhanden in verdorbenem Fisch) hemmt die Acetylcholinfreisetzung aus dem Endknopf. Warum kann dann kein nachsynaptisches Aktionspotential entstehen?
4. Welchen Einfluss hat es auf die Erregungsweiterleitung, wenn man die Menge der Natriumionen im Spalt verringert?
5. Stärkere Reize werden im Nervensystem durch eine höhere Frequenz (=Häufigkeit) des Aktionspotentials ausgedrückt. Was bedeutet es für die Menge des aus dem Endknopf ausgeschütteten Transmitters, wenn sich die Häufigkeit des Aktionspotentials erhöht?
6. Nennen Sie die drei letzten Teilprozesse, mit denen der synaptische Übertragungsprozess endet, in der richtigen Reihenfolge.
7. Bei einer Form der Calciumkanalkrankheit sind die Calciumkanäle abnorm häufig geschlossen. Inwiefern beeinflusst dies die Erregungsweiterleitung, wenn ein Aktionspotential ankommt und warum?
8. Bei einer anderen Form der Calciumkanalkrankheit lassen sich die Calciumkanäle besonders leicht öffnen. Inwiefern verändert sich dann die Erregungsweiterleitung, wenn ein Aktionspotential ankommt und warum?
9. Wodurch werden die Natriumkanäle in der nachsynaptischen Membran geöffnet?
10. Welcher Prozess geschieht, bevor die Bläschen auf die vorsynaptische Zellmembran zuwandern?
11. Das Fliegenpilzgift Muscarin setzt sich wie der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle und öffnet sie. Es wird aber im Gegensatz zu dem Transmitter im Spalt nicht abgebaut, sondern bleibt dort länger. Was passiert an der Synapse?
12. Atropin (Gift der Tollkirsche) dockt an die gleichen Rezeptoren an wie der Transmitter Acetylcholin, hat aber nicht dessen Wirkung. Was ist die Folge, wenn eine große Menge Atropin in den synaptischen Spalt gelangt?
13. Muscarin setzt sich wie der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle und öffnet sie. Es wird aber im Spalt nicht abgebaut.
Atropin dockt ebenfalls an die Rezeptoren der Natriumkanäle an, öffnet sie aber nicht.
Führt es zu einer Verschlimmerung oder Verminderung der Vergiftungserscheinungen, wenn die Nervengifte Muscarin und Atropin gleichzeitig an einer Synapse vorhanden sind? Begründen Sie.
14. Was geschieht mit dem Transmitter Acetylcholin am Ende des Übertragungsprozesses?
15. Welche Vorgänge liegen zwischen dem Ausschütten des Transmitters Acetylcholin in den Spalt und dem Einstrom von Natriumionen an der nachsynaptischen Membran?

16. Das Aktionspotential der nachsynaptischen Nervenzelle wird durch eine positive Aufladung ihres Zellinneren ausgelöst. Welche Art von Ionenströmen müssten in die nachsynaptische Zelle erfolgen, um die Erregungsweiterleitung zu erschweren?
17. Das Insektenvertilgungsmittel E 605 hemmt die Aktivität der Cholinesterase langfristig. Was sind die Folgen?
18. Curare ist ein Gift, das sich an die Rezeptoren der Natriumkanäle setzt, aber bei genügend hoher Konzentration des Transmitters Acetylcholin verdrängt werden kann.
Könnte ein Stoff, der die Aktivität der Cholinesterase **kurzzeitig** hemmt, ein Gegenmittel für Curare sein? Begründen Sie.
19. Welche Prozesse laufen zwischen dem Ankommen eines Aktionspotentials am Endknopf und dem Verschmelzen der Bläschen mit der vorsynaptischen Zellmembran ab (**3 Sätze**)?
20. Was ist die wichtige Funktion des Transmitters Acetylcholin im Übertragungsprozess?
21. Was ist die Folge für die Erregungsweiterleitung, wenn die Bläschen nicht mehr mit der vorsynaptischen Membran verschmelzen können? Begründen Sie kurz (**1 Satz**).
22. Bei der Krankheit „Myasthenia Gravis“ zerstört das Immunsystem einen Teil der Rezeptoren der Natriumkanäle in den Synapsen. Was sind die Folgen für das Acetylcholin und die Erregungsweiterleitung?
23. Bei der Alzheimer-Erkrankung produzieren die Nervenzellen zu wenig Acetylcholin. Warum verschreibt der Arzt hier häufig Cholinesterasehemmer?
24. Warum sollte man Cholinesterasehemmer nicht gesunden Menschen geben? Begründen Sie mittels der Vorgänge an der Synapse.
25. Die Entstehung eines Aktionspotentials hängt mit einer positiven Aufladung des Zellinneren der nachsynaptischen Zelle zusammen. Was passiert, wenn ein Transmitter anstatt Natriumkanäle (**Na⁺**), Chloridkanäle (**Cl⁻**) an der nachsynaptischen Membran öffnet und Chloridionen (**Cl⁻**) in die Zelle einströmen?
26. Welche Prozesse laufen ab, bevor sich die Calciumkanäle öffnen und nachdem sich die Calciumkanäle geöffnet haben? (**jeweils ein Prozess**)
27. Welche beiden Vorgänge passieren, nachdem die Bläschen mit der vorsynaptischen Membran verschmolzen sind?
28. Welche Folge hat eine Erhöhung der Calciumkonzentration an der Synapse für die Transmitterausschüttung?
29. Das Gift der „Schwarzen Witwe“ hemmt unter anderem die Wiederaufnahme des Transmitters Acetylcholin in den Endknopf. Dieses wird aber durch seine Spaltung nach wie vor deaktiviert. Beschreiben Sie die Folge.
30. Was geschieht an der nachsynaptischen Membran, wenn der Transmitter Acetylcholin an die Rezeptoren der Natriumkanäle andockt (**2 Sätze**)?

Anhang D2 – 1: Test zur Erfassung der Wahrnehmungsvorerfahrungen mit Pferdebewegungen in der Pilotstudie

Mit den nachfolgenden Fragen würden wir gerne wissen, wie viel Erfahrung Sie mit Pferden und ihren Bewegungen haben. Kreuzen Sie bitte an.

	Ja	Nein
1. Haben oder hatten Sie früher häufiger mit Pferden zu tun?	<input type="checkbox"/>	<input type="checkbox"/>
2. Reiten Sie aktiv oder sind Sie früher mal aktiv geritten?	<input type="checkbox"/>	<input type="checkbox"/>
3. Sehen oder sahen Sie sich früher im Fernsehen häufiger Sportsendungen mit Pferden an?	<input type="checkbox"/>	<input type="checkbox"/>
4. Sehen oder sahen Sie sich früher im Fernsehen öfters Filme und Serien mit Pferden an?	<input type="checkbox"/>	<input type="checkbox"/>
5. Leben Sie in einer Gegend, in der es viele Pferde gibt oder haben Sie in einer solchen Gegend schon mal gelebt?	<input type="checkbox"/>	<input type="checkbox"/>
6. Gibt oder gab es in Ihrer Familie oder Ihrem Freundeskreis aktive Reiter?	<input type="checkbox"/>	<input type="checkbox"/>
7. Lesen Sie häufiger Bücher über Pferde oder haben Sie dies in der Vergangenheit öfters getan?	<input type="checkbox"/>	<input type="checkbox"/>
8. Haben Sie schon einmal etwas über die Gangarten von Pferden gelernt?	<input type="checkbox"/>	<input type="checkbox"/>
9. Gehen oder gingen Sie früher häufiger zu Pferderennen oder Reitturnieren?	<input type="checkbox"/>	<input type="checkbox"/>
10. Schauen Sie sich häufiger Bilder mit Pferden an oder haben Sie dies früher mal getan?	<input type="checkbox"/>	<input type="checkbox"/>
11. Gehen oder gingen Sie früher öfters in einen Reitstall?	<input type="checkbox"/>	<input type="checkbox"/>

Anhang D2 – 2: Vorwissenstest zu Pferdegangarten in der Pilotstudie

Nun möchten wir gerne wissen, wie gut Sie sich mit Pferdegangarten auskennen. Wenn Sie die Antwort zu einer Frage **nicht wissen**, machen Sie dies bitte durch ein „?“ oder durch ein „/“ deutlich.

1. Nennen Sie drei Pferdegangarten.

2. Beschreiben Sie **eine** der von Ihnen eben genannten Pferdegangarten.

(Schreiben Sie bitte den Namen der betreffenden Gangart nochmals dazu.)

Anhang D2 – 3: Items zur Erfassung des Cognitive Loads

Nach der Lerneinheit:

Wie schwierig fanden Sie es, die Lernaufgabe zu bearbeiten?

sehr, sehr leicht	1	2	3	4	5	6	7	8	9	sehr, sehr schwierig
-------------------	---	---	---	---	---	---	---	---	---	----------------------

Wie stark haben Sie sich beim Bearbeiten der Lernaufgabe angestrengt?

sehr, sehr wenig	1	2	3	4	5	6	7	8	9	sehr, sehr stark
------------------	---	---	---	---	---	---	---	---	---	------------------

Nach dem Nachtest:

Wie schwierig fanden Sie es, die Aufgaben zu den Pferdegangarten zu bearbeiten?

sehr, sehr leicht	1	2	3	4	5	6	7	8	9	sehr, sehr schwierig
-------------------	---	---	---	---	---	---	---	---	---	----------------------

Wie stark haben Sie sich beim Bearbeiten der Aufgaben zu den Pferdegangarten angestrengt?

sehr, sehr wenig	1	2	3	4	5	6	7	8	9	sehr, sehr stark
------------------	---	---	---	---	---	---	---	---	---	------------------

Anhang D2 – 4: Nachtest zu Pferdegangarten in der Pilotstudie

Da die Items des Nachtests den Probanden computergestützt nach der Betrachtung des Lernmaterials präsentiert wurden, sei für ihre Ansicht auf Anhang D1 auf der Begleit-CD verwiesen. Der Nachtest folgt jeweils direkt auf das Lernmaterial in den Lerndateien.

Für die Gestaltung des Antwortbogens, der den Probanden in Papierform vorgelegt wurde, werden hier einige Beispielitems gezeigt:

Aufgabe 1

- Schritt
 - Trab
 - Galopp
 - Tölt
-

Aufgabe 2

- Tölt
 - Schritt
 - Trab
 - Galopp
-

Aufgabe 3

- Galopp
 - Tölt
 - Schritt
 - Trab
-

usw. bis Aufgabe 32

Anhang D2 – 5: Interviewleitfaden in der Pilotstudie

Zeit für Gesamtdurchlauf:

1. An was haben Sie die Pferdegangarten im Nachtest jeweils erkannt?

2. Wie fanden Sie den Versuch insgesamt? z.B. Ablauf, Dauer

3. Hatten Sie an irgendeiner Stelle besondere Schwierigkeiten?

4. Was könnte man Ihrer Meinung nach an dem Versuch besser machen?

Anhang E2 – 1: Test zur Erfassung der Wahrnehmungsvorerfahrungen mit Pferdebewegungen in der Hauptstudie

Aus Platzgründen werden nur die Items dieses Tests genannt. Die formale Gestaltung des Tests war analog zu der Version in der Pilotstudie, daher sei für die Ansicht der Gestaltungsmerkmale auf den Anhang D2 – 1 verwiesen.

1. Haben oder hatten Sie früher häufiger mit Pferden zu tun?
2. Reiten Sie aktiv oder sind Sie früher mal aktiv geritten?
3. Sehen oder sahen Sie sich früher im Fernsehen häufiger Sportsendungen mit Pferden an?
4. Sehen oder sahen Sie sich früher im Fernsehen öfters Filme und Serien mit Pferden an?
5. Leben Sie in einer Gegend, in der es viele Pferde gibt oder haben Sie in einer solchen Gegend schon mal gelebt?
6. Lesen Sie häufiger Bücher über Pferde oder haben Sie dies in der Vergangenheit öfters getan?
7. Haben Sie schon einmal etwas über die Gangarten von Pferden gelernt?
8. Gehen oder gingen Sie früher häufiger zu Pferderennen oder Reitturnieren?
9. Schauen Sie sich häufiger Bilder mit Pferden an oder haben Sie dies früher mal getan?
10. Gehen oder gingen Sie früher öfters in einen Reitstall?

Anhang E2 – 2: Nachtest zu Pferdegangarten in der Hauptstudie

Da die Items des Nachtests den Probanden computergestützt nach der Betrachtung des Lernmaterials präsentiert wurden, sei für ihre Ansicht auf Anhang E1 auf der Begleit-CD verwiesen. Der Nachtest folgt jeweils direkt auf das Lernmaterial in den Lerndateien.

Die Gestaltung des Antwortbogens in der Hauptstudie war analog zu der in der Pilotstudie, weshalb für die Ansicht von den Gestaltungsmerkmalen des Antwortbogens auf den Anhang D2 – 4 verwiesen sei.

Tabellenverzeichnis

Kapitel 4.1:

4.1.1	Beschreibung der drei Versuchsbedingungen der Vorpilotstudie.....	94
4.1.2	Trennschärfen und Schwierigkeiten der Items des Vorwissenstests in der Vorpilotstudie.....	104
4.1.3	Trennschärfen und Schwierigkeiten der Items der Wissensskala des Nachtests (Version 1) in der Vorpilotstudie.....	105
4.1.4	Trennschärfen und Schwierigkeiten der Items der Wissensskala des Nachtests (Version 2) in der Vorpilotstudie.....	106
4.1.5	Trennschärfen und Schwierigkeiten der Items der Inferenzskala des Nachtests (Version 1) in der Vorpilotstudie.....	107
4.1.6	Trennschärfen und Schwierigkeiten der Items der Inferenzskala des Nachtests (Version 2) in der Vorpilotstudie.....	108
4.1.7	Deskriptive Statistiken der Variable „Vorwissenstest“ bei den Gruppen 1, 2 und 3 der Vorpilotstudie.....	110
4.1.8	Deskriptive Statistiken der Variable „Nachwissensskala“ bei den Gruppen 1, 2 und 3 der Vorpilotstudie.....	110
4.1.9	Deskriptive Statistiken der Variable „Inferenzskala“ bei den Gruppen 1, 2 und 3 der Vorpilotstudie.....	111
4.1.10	Deskriptive Statistiken der Variable „Bearbeitungszeit des Vorwissenstests“ bei den Gruppen 1, 2 und 3 der Vorpilotstudie.....	112
4.1.11	Deskriptive Statistiken der Variable „Bearbeitungszeit des Nachtests“ bei den Gruppen 1, 2 und 3 der Vorpilotstudie.....	112

Kapitel 4.2:

4.2.1 Trennschärfen und Schwierigkeiten der Items des Vorwissenstests in der Pilotstudie.....	144
4.2.2 Trennschärfen und Schwierigkeiten der Items der Beschriftungsaufgabe zum begrifflichen Wissen in der Pilotstudie.....	145
4.2.3 Trennschärfen und Schwierigkeiten der Items der Nachwissensskala in der Pilotstudie.....	146
4.2.4 Trennschärfen und Schwierigkeiten der Items der Inferenzskala in der Pilotstudie.	147
4.2.5 Reliabilitäten der Subskalen des FAM in der Pilotstudie.....	148
4.2.6 Deskriptive Statistiken der abhängigen Variablen für die Teilstichprobe „Animation“ mit „Bild seriell schnell“ und „Bild seriell langsam“.....	149
4.2.7 Deskriptive Statistiken der Kontrollvariablen für die Teilstichprobe „Animation“ mit „Bild seriell schnell“ und „Bild seriell langsam“.....	150
4.2.8 Deskriptive Statistiken der abhängigen Variablen für die Teilstichprobe „Bild parallel“ mit „Bild seriell schnell“ und „Bild seriell langsam“.....	151
4.2.9 Deskriptive Statistiken der Kontrollvariablen für die Teilstichprobe „Bild parallel“ mit „Bild seriell schnell“ und „Bild seriell langsam“.....	151
4.2.10 Mittelwerte und Standardabweichungen der Bedingungen in der AV „Nachwissensskala“.....	153
4.2.11 Mittelwerte und Standardabweichungen der Bedingungen in der AV „Inferenzskala“.....	154
4.2.12 Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load bei der Lernaufgabe“.....	155
4.2.13 Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load beim Nachttest“.....	155
4.2.14 Mittelwerte und Standardabweichungen der Bedingungen in der AV „Nachwissensskala“.....	157

4.2.15 Mittelwerte und Standardabweichungen der Bedingungen in der AV „Inferenzskala“	158
4.2.16 Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load bei der Lernaufgabe“	159
4.2.17 Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load beim Nachttest“	159

Kapitel 4.3.1:

4.3.1.1 Gütekriterien des Vorwissenstests, der Wissens- und der Inferenzskala sowie der Begriffsaufgabe in der Hauptstudie	182
4.3.1.2 Reliabilitäten der Subskalen des FAM in der Hauptstudie	182
4.3.1.3 Deskriptive Statistiken der abhängigen Variablen in der ersten Hauptteilstudie	184
4.3.1.4 Deskriptive Statistiken der Kontrollvariablen in der ersten Hauptteilstudie	184
4.3.1.5 Mittelwerte und Standardabweichungen der Bedingungen in der AV „Nachwissensskala“	187
4.3.1.6 Adjustierte Mittelwerte und Standardfehler der Bedingungen in der AV „Inferenzskala“	188
4.3.1.7 Adjustierte Mittelwerte und Standardfehler der Bedingungen in der AV „CL Lern“	189
4.3.1.8 Mittelwerte und Standardabweichungen der Bedingungen in der AV „CL Nachttest“	190
4.3.1.9 Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild seriell schnell x Vorwissen“ bei der AV „Inferenzskala“	192
4.3.1.10 Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild seriell schnell x Vorwissen“ bei der AV „CL Lern“	193

4.3.1.11	Regressionsanalyse für die Bestimmung der Interaktion „Bild seriell schnell vs. Bild seriell langsam x Vorwissen“ bei der AV „Inferenzskala“196
4.3.1.12	Regressionsanalyse für die Bestimmung der Interaktion „Bild seriell schnell vs. Bild seriell langsam x Vorwissen“ bei der AV „CL Lern“198
4.3.1.13	Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild seriell schnell x Räumliche Fähigkeiten“ bei der AV „Inferenzskala“200
4.3.1.14	Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild seriell schnell x Merkfähigkeit“ bei der AV „CL Lern“204
4.3.1.15	Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild seriell schnell x Merkfähigkeit“ bei der AV „CL Nachtest“206

Kapitel 4.3.2:

4.3.2.1	Deskriptive Statistiken der abhängigen Variablen in der zweiten Hauptteilstudie.....227
4.3.2.2	Deskriptive Statistiken der Kontrollvariablen in der zweiten Hauptteilstudie.....228
4.3.2.3	Mittelwerte und Standardabweichungen der Bedingungen in der AV „Nachwissensskala“230
4.3.2.4	Mittelwerte und Standardabweichungen der Bedingungen in der AV „Inferenzskala“230
4.3.2.5	Adjustierte Mittelwerte und Standardfehler der Bedingungen in der AV „CL Lern“231
4.3.2.6	Mittelwerte und Standardabweichungen der Bedingungen in der AV „CL Nachtest“232
4.3.2.7	Regressionsanalyse für die Bestimmung der Interaktion „Bild parallel vs. Bild seriell schnell x Vorwissen“ bei der AV „Inferenzskala“233

4.3.2.8	Regressionsanalyse für die Bestimmung der Interaktion „Bild parallel vs. Bild seriell schnell x Vorwissen“ bei der AV „CL Lern“	235
4.3.2.9	Regressionsanalyse für die Bestimmung der Interaktion „Bild parallel vs. Bild seriell langsam x Vorwissen“ bei der AV „CL Lern“	237

Kapitel 5.1:

5.1.1	Trennschärfen und Schwierigkeiten der Items des Wahrnehmungsvorerfahrungstests in der Pilotstudie.....	272
5.1.2	Trennschärfen und Schwierigkeiten der Items des Vorwissenstests in der Pilotstudie.....	273
5.1.3	Trennschärfen und Schwierigkeiten der Items des Nachtests (gesamt) in der Pilotstudie.....	274
5.1.4	Trennschärfen und Schwierigkeiten der Items der Animationsskala in der Pilotstudie.....	276
5.1.5	Trennschärfen und Schwierigkeiten der Items der Bildskala in der Pilotstudie.....	277
5.1.6	Reliabilitäten der Subskalen des FAM in der Pilotstudie.....	279
5.1.7	Deskriptive Statistiken der abhängigen Variablen in der Pilotstudie.....	280
5.1.8	Deskriptive Statistiken der Kontrollvariablen in der Pilotstudie.....	280
5.1.9	Mittelwerte und Standardabweichungen der Bedingungen in der AV „Nachtest gesamt“	282
5.1.10	Mittelwerte und Standardabweichungen der Bedingungen in der AV „Animationsskala“	283
5.1.11	Mittelwerte und Standardabweichungen der Bedingungen in der AV „Bildskala“ ..	283
5.1.12	Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load bei der Lernaufgabe“	284

5.1.13 Mittelwerte und Standardabweichungen der Bedingungen in der AV „Cognitive Load beim Nachttest“	284
5.1.14 Zellenmittelwerte und Standardfehler in der zweifaktoriellen ANOVA mit Messwiederholung	285

Kapitel 5.2:

5.2.1 Gütekriterien des Vorerfahrungs- und des Vorwissenstests, des Nachttests sowie der Animations- und der Bildskala in der Hauptstudie	305
5.2.2 Reliabilitäten der Subskalen des FAM in der Hauptstudie	305
5.2.3 Deskriptive Statistiken der abhängigen Variablen in der Hauptstudie	307
5.2.4 Deskriptive Statistiken der Kontrollvariablen in der Hauptstudie	307
5.2.5 Adjustierte Mittelwerte und Standardfehler der Bedingungen in der AV „Nachttest gesamt“	309
5.2.6 Mittelwerte und Standardabweichungen der Bedingungen in der AV „Animationsskala“	310
5.2.7 Adjustierte Mittelwerte und Standardfehler der Bedingungen in der AV „Bildskala“	310
5.2.8 Mittelwerte und Standardabweichungen der Bedingungen in der AV „CL Lern“	311
5.2.9 Mittelwerte und Standardabweichungen der Bedingungen in der AV „CL Nachttest“	311
5.2.10 Zellenmittelwerte und Standardfehler in der zweifaktoriellen ANCOVA mit Messwiederholung	312
5.2.11 Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild parallel x Räumliche Fähigkeiten“ bei der AV „Nachttest gesamt“	316
5.2.12 Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild parallel x Räumliche Fähigkeiten“ bei der AV „Animationsskala“	317

5.2.13 Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild parallel x Räumliche Fähigkeiten“ bei der AV „Bildskala“	318
5.2.14 Regressionsanalyse für die Bestimmung der Interaktion „Animation vs. Bild parallel x Räumliche Fähigkeiten“ bei der AV „CL Nachtest“	320

Abbildungsverzeichnis

Kapitel 2:

- 2.1 Das integrierte Modell des Text- und Bildverstehens (angelehnt an die Darstellung in Schnotz, 2005).....23

Kapitel 4.1:

- 4.1.1 Beispielbild für die Bildgestaltung der Synapse in der Vorpilotstudie.....96

Kapitel 4.2:

- 4.2.1 Darstellung des beschrifteten Standbildes.....131
4.2.2 Beispielbild für die graphische Gestaltung der Synapse.....132

Kapitel 4.3.1:

- 4.3.1.1 Regressionsgeraden des Vorwissens in den Gruppen „Animation“ und „Bild seriell schnell“ (AV: Inferenzskala).....192
4.3.1.2 Regressionsgeraden des Vorwissens in den Gruppen „Animation“ und „Bild seriell schnell“ (AV: CL Lern).....194
4.3.1.3 Regressionsgeraden des Vorwissens in den Gruppen „Bild seriell schnell“ und „Bild seriell langsam“ (AV: Inferenzskala).....197
4.3.1.4 Regressionsgeraden des Vorwissens in den Gruppen „Bild seriell schnell“ und „Bild seriell langsam“ (AV: CL Lern).....199
4.3.1.5 Regressionsgeraden der räumlichen Fähigkeiten in den Gruppen „Animation“ und „Bild seriell schnell“ (AV: Inferenzskala).....201
4.3.1.6 Regressionsgeraden der Merkfähigkeit in den Gruppen „Animation“ und „Bild seriell schnell“ (AV: CL Lern).....205

4.3.1.7	Regressionsgeraden der Merkfähigkeit in den Gruppen „Animation“ und „Bild seriell schnell“ (AV: CL Nachtest).....	206
---------	---	-----

Kapitel 4.3.2:

4.3.2.1	Regressionsgeraden des Vorwissens in den Gruppen „Bild parallel“ und „Bild seriell schnell“ (AV: Inferenzskala).....	234
4.3.2.2	Regressionsgeraden des Vorwissens in den Gruppen „Bild parallel“ und „Bild seriell schnell“ (AV: CL Lern).....	235
4.3.2.3	Regressionsgeraden des Vorwissens in den Gruppen „Bild parallel“ und „Bild seriell langsam“ (AV: CL Lern).....	238

Kapitel 5.1:

5.1.1	Beispielbild für die Animation der Gangart „Schritt“.....	257
5.1.2	Beispielbild für die Animation der Gangart „Trab“.....	257
5.1.3	Beispielbild für die Animation der Gangart „Galopp“.....	258
5.1.4	Beispielbild für die Animation der Gangart „Tölt“.....	258
5.1.5	Anordnung der Key-Frames im Lernmaterial am Beispiel „Schritt“.....	259
5.1.6	Beispiel für ein Bilditem mit vier Key-Frames anhand der Gangart „Galopp“.....	265
5.1.7	Die Itemabfolge des Nachtests in der Pilotstudie.....	266
5.1.8	Graphische Veranschaulichung der Interaktion zwischen dem Format der Lernbedingung und dem Format der Nachtestskalen in der Pilotstudie.....	286

Kapitel 5.2:

5.2.1	Graphische Veranschaulichung der Interaktion zwischen dem Format der Lernbedingung und dem Format der Nachtestskalen in der Hauptstudie.....	313
5.2.2	Regressionsgeraden der räumlichen Fähigkeiten in den Gruppen „Animation“ und „Bild parallel“ (AV: Nachtest gesamt).....	317

5.2.3	Regressionsgeraden der räumlichen Fähigkeiten in den Gruppen „Animation“ und „Bild parallel“ (AV: Animationsskala).....	318
5.2.4	Regressionsgeraden der räumlichen Fähigkeiten in den Gruppen „Animation“ und „Bild parallel“ (AV: Bildskala).....	319
5.2.5	Regressionsgeraden der räumlichen Fähigkeiten in den Gruppen „Animation“ und „Bild parallel“ (AV: CL Nachttest).....	321

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die Dissertation selbst angefertigt und alle von mir benutzten Hilfsmittel in der Arbeit angegeben habe.

Ich habe die Dissertation bislang weder als Gesamtschrift noch in Auszügen als Prüfungsarbeit für eine staatliche oder eine andere wissenschaftliche Prüfung eingereicht.

Zudem erkläre ich, dass ich weder diese noch eine andere Abhandlung bisher bei einer anderen Hochschule als Dissertation eingereicht habe.

Landau, den

Inga Wagner

Lebenslauf

Persönliche Angaben:

Name: Inga Wagner
Geburtsdatum: 25.05.1981
Geburtsort: Pirmasens
Familienstand: ledig
Email: wagneri@uni-landau.de

Bildungswegdegang:

1991 – 2000 Hugo Ball Gymnasium, Pirmasens; Abschluss: Abitur
2001 – 2008 Studium der Psychologie an der Universität Koblenz-Landau;
Abschluss: Diplom
Thema der Diplomarbeit: Gemischtgeschlechtliche Freundschaften im
mittleren Kindesalter

Beruflicher Werdegang nach dem Studium:

05/2008 – 11/2008 Tätigkeit als Wissenschaftliche Mitarbeiterin in der Arbeitseinheit
„Entwicklungspsychologie“: „Projekt VERA – Vergleichsarbeiten in
der Grundschule“, unter Leitung von Herrn Prof. Dr. Hosenfeld, an der
Universität Koblenz-Landau
Seit 11/2008 Promotionsstipendiatin in der Arbeitseinheit „Allgemeine und
Pädagogische Psychologie“, unter Leitung von Herrn Prof. Dr.
Schnotz, an der Universität Koblenz-Landau