

UNIVERSITÄT DUISBURG-ESSEN  
FACHBEREICH BILDUNGSWISSENSCHAFTEN  
LEHRSTUHL FÜR LEHR-LERNPSYCHOLOGIE

Lernen mit dynamischen Visualisierungen:  
Metaanalyse und experimentelle Untersuchungen  
zu einem naturwissenschaftlichen Lerninhalt

Dissertation zur Erlangung des Doktorgrades  
der Philosophischen Fakultät  
der Universität Duisburg-Essen

vorgelegt von  
Tim Niclas Höffler  
April 2007

Tag der Disputation:  
07. September 2007

Gutachter:  
Prof. Dr. Detlev Leutner  
Prof. Dr. Rolf Plötzner

# Inhaltsverzeichnis

<b>Zusammenfassung</b> .....	<b>5</b>
<b>Abstract</b> .....	<b>7</b>
<b>Danksagung und Widmung</b> .....	<b>8</b>
<b>1 Einführung</b> .....	<b>9</b>
1.1 Einleitung .....	9
1.2 Struktur dieser Arbeit.....	10
<b>2 Theoretischer Hintergrund</b> .....	<b>11</b>
2.1 Der Einfluss von Standbildern auf den Lernerfolg .....	11
2.2 Der Einfluss von Animationen auf den Lernerfolg .....	18
2.3 Animation versus Standbilder.....	20
2.4 Beeinflussende Moderatoreffekte.....	23
<b>3 Metaanalyse zur Wirksamkeit von Animationen und Standbildern</b> .....	<b>26</b>
3.1 Einführung.....	26
3.2 Methode .....	27
3.2.1 Auffindung und Selektion geeigneter Studien.....	27
3.2.2 Kodierung .....	28
3.2.3 Effektstärkenbestimmung und -analyse.....	29
3.3 Ergebnisse .....	32
3.3.1 Studiencharakteristika .....	32
3.3.2 Overall-Effekte .....	32
3.3.3 Einfluss von Moderatoren .....	35
3.3.4 Vertiefende Analyse zweiter und dritter Dimension .....	41
3.4 Diskussion.....	47
3.4.1 Overall-Effekte von Animationen gegenüber Standbildern.....	48
3.4.2 Repräsentational oder dekorativ?.....	49
3.4.3 Deklarativ oder prozedural?.....	50
3.4.4 Computer- oder videobasiert – Grad des Realismus?.....	51
3.4.5 Eigenschaften von Standbildern .....	52
3.4.6 Limitationen dieser Metaanalyse .....	53
3.4.7 Fazit.....	54
<b>4 Erste experimentelle Studie (Pilotstudie)</b> .....	<b>56</b>
4.1 Einführung.....	56
4.2 Fragestellung und inhaltliche Hypothesen.....	57
4.3 Methode .....	58

4.3.1	Variablen.....	58
4.3.2	Materialien .....	61
4.3.3	Räumlichkeiten und Hardware.....	65
4.3.4	Untersuchungsablauf.....	65
4.3.5	Probanden .....	66
4.4	Ergebnisse .....	66
4.4.1	Hypothesen 1 und 2: Animationen versus Standbilder .....	67
4.4.2	Hypothese 3: Die Rolle des themenspezifischen Vorwissens .....	68
4.4.3	Hypothese 4: Die Rolle der Spatial Ability .....	70
4.4.4	Hypothese 5: Die Rolle des Cognitive Load .....	74
4.4.5	Weitere Berechnungen .....	75
4.5	Diskussion .....	77
<b>5</b>	<b>Zweite experimentelle Studie .....</b>	<b>80</b>
5.1	Einführung.....	80
5.2	Zusätzliche theoretische Erörterungen.....	81
5.3	Fragestellung und inhaltliche Hypothesen.....	82
5.4	Methode .....	84
5.4.1	Variablen.....	84
5.4.2	Materialien .....	85
5.4.3	Räumlichkeiten und Hardware.....	88
5.4.4	Untersuchungsablauf.....	88
5.4.5	Probanden .....	89
5.5	Ergebnisse .....	89
5.5.1	Hypothesen 1a, 1b und 2: Animationen versus Standbilder .....	89
5.5.2	Hypothese 3: Die Rolle des themenspezifischen Vorwissens .....	92
5.5.3	Hypothese 4: Die Rolle der Spatial Ability .....	95
5.5.4	Weitere Berechnungen .....	101
5.5.5	Follow-Up.....	104
5.6	Diskussion .....	105
<b>6</b>	<b>Dritte experimentelle Studie .....</b>	<b>109</b>
6.1	Einführung.....	109
6.2	Zusätzliche theoretische Erörterungen.....	109
6.3	Fragestellung und inhaltliche Hypothesen.....	110
6.4	Methode .....	112
6.4.1	Variablen.....	112
6.4.2	Materialien .....	113
6.4.3	Räumlichkeiten und Hardware.....	114

6.4.4	Untersuchungsablauf.....	114
6.4.5	Probanden .....	115
6.5	Ergebnisse .....	116
6.5.1	Hypothesen 1a, 1b und 2: Animationen versus Standbilder .....	116
6.5.2	Hypothese 3: Die Rolle des themenspezifischen Vorwissens .....	119
6.5.3	Hypothese 4: Die Rolle der Spatial Ability .....	123
6.5.4	Weitere Berechnungen .....	125
6.6	Diskussion.....	128
<b>7</b>	<b>Vergleich der drei experimentellen Studien.....</b>	<b>130</b>
7.1	Einführung.....	130
7.2	Ergebnisse .....	130
7.3	Diskussion.....	135
<b>8</b>	<b>Zusammenfassende Diskussion.....</b>	<b>137</b>
8.1	Zusammenfassung der Ergebnisse.....	137
8.2	Theoretische Implikationen .....	138
8.3	Praktische Implikationen .....	141
8.4	Fachdidaktische Implikationen .....	142
8.5	Abschließende Bemerkungen und Ausblick .....	143
<b>9</b>	<b>Literatur.....</b>	<b>145</b>
<b>10</b>	<b>Anhang.....</b>	<b>158</b>
	Anhang A: Tabelle der in die Metaanalyse aufgenommenen Studien.....	159
	Anhang B: Vortest der ersten und dritten empirischen Studie.....	166
	Anhang C: Nachtest der ersten bis dritten empirischen Studie .....	170
	Anhang D: Vortest der zweiten empirischen Studie .....	176
	Anhang E: Paper-Folding-Test.....	180
	Anhang F: Card-Rotation-Test.....	183
	Anhang G: Abschrift des Audiokommentars der Lernumgebungen .....	186

## Zusammenfassung

Die Frage nach Lernvorteilen von dynamischen Visualisierungen gegenüber Standbildern ist nach wie vor umstritten. Während auf der einen Seite computerbasierte Animationen eine realistischere Abbildung komplexer Prozesse und Prozeduren versprechen, die Vorlage für ein exaktes mentales Modell liefern können, verweisen auf der anderen Seite Forschungsergebnisse der Instruktionspsychologie auf die erhöhte kognitive Belastung, die mit bewegten Bildern einhergehen und so die Lernenden überfordern kann (Tversky, Morrison & Bétrancourt, 2002). Mit Hilfe einer Metaanalyse fasst vorliegende Dissertation zunächst den aktuellen Stand der Forschung zum Lernen mit Animationen und Standbildern zusammen. Neben einer statistisch signifikanten durchschnittlichen Überlegenheit von Animationen in den publizierten Studien ( $d=0,40$ ) wird darüber hinaus eine Reihe von Moderatorvariablen identifiziert, die Einfluss auf das Lernen mit Visualisierungen haben. Dazu gehört unter anderem die instruktionale Rolle der Animation, d.h. ob die Animation dekorativ oder repräsentational (Carney & Levin, 2002) verwendet wird, der Grad des Realismus der Animation sowie die Funktion erläuternden zusätzlichen Textes.

Ausgehend von den Ergebnissen der Metaanalyse sowie zusätzlicher theoretischer Überlegungen werden sodann die Ergebnisse dreier experimenteller Studien zum Lernen mit Animationen und Standbildern vorgestellt. Mittels verschiedener Versionen einer eigens entwickelten computerbasierten Lernumgebung mit chemischer Thematik wird der Einfluss der Anzahl von Standbildern untersucht, wenn diese mit einer Animation verglichen werden. Es zeigt sich, dass die Animation einer Serie von vier Standbildern überlegen ist, wohingegen zumindest Hinweise dafür gefunden werden können, dass die Steigerung von vier auf elf Standbilder sowie die Ermöglichung minimaler Interaktivität diesen Unterschied reduziert.

Ein besonderer Schwerpunkt der Arbeit liegt auf dem Einfluss lernerspezifischer Faktoren wie denen des themenspezifischen Vorwissens, des *Cognitive Load* sowie insbesondere des räumlichen Vorstellungsvermögens. Diesbezüglich zeigt sich ein kompensierender „Aptitude-Treatment“-Effekt: Lernende mit hohem räumlichen Vorstellungsvermögen können die Nachteile, die eine Standbildversion gegenüber der Animation hinsichtlich des Lernerfolgs verursacht, kompensieren und mit Standbildern genauso gut lernen wie mit Animationen. Lernende mit niedrigem

räumlichen Vorstellungsvermögen profitieren hingegen besonders von einer Animation. Um dieses Ergebnis weiter zu untersuchen, wird auf unterschiedliche Facetten des räumlichen Vorstellungsvermögens fokussiert. Es zeigt sich, dass der beobachtete Effekt nur für den Bereich der *Spatial Visualization* (Carroll, 1993) gilt.

Die Ergebnisse der Arbeit haben theoretische, praktische und fachdidaktische Implikationen, die abschließend diskutiert werden.

## Abstract

The question of whether dynamic visualizations or static pictures are superior in learning is still a matter of discussion. On the one hand, computer-based animations may be able to give a more realistic picture of complex processes and procedures and therefore provide the basis for an accurate mental model. On the other hand, instructional psychologists suggest that the transient character of dynamic pictures may increase cognitive load, which can overtax the learner (Tversky, Morrison & Bétrancourt, 2002). Via a statistical meta-analysis this dissertational thesis summarizes published studies concerning learning with animations and static pictures. In addition to the main finding of a clear superiority of animations over static pictures ( $d=.40$ ), several moderator variables can be identified: These include the instructional role of animations, i.e. whether the animation is used to serve a decorative or representational function (Carney & Levin, 2002); the level of realism; and the role of additional illustrating text.

Based on the meta-analytic findings as well as on further theoretical consideration three experimental studies on learning with animations and static pictures were conducted. Using different versions of a self-developed computer-based learning environment dealing with a topic in chemistry, the effect of the number of still pictures when compared to an animation is investigated. Results indicate that the animation is clearly superior to a series of four static pictures, while there is some evidence that this difference diminishes when compared to a series of eleven static pictures and when minimal interactivity is introduced.

A specific focus of the present thesis is on the influence of learner characteristics, such as prior knowledge, cognitive load, and, in particular, spatial ability. A compensating effect of spatial ability is identified: Learners with high spatial ability are able to compensate for the disadvantages of the static version and therefore learn as much as with the animation. For learners with low spatial ability, however, the animation is especially beneficial. To further explore these results, different facets of spatial ability are examined. It can be shown that the observed effect is only valid for the area of *spatial visualization* (Carroll, 1993).

The results of the present dissertational thesis have theoretical, practical, and educational implications which are discussed conclusively.

## Danksagung und Widmung

Gewidmet meinem Großvater Hans Nack, der leider kurz vor Vollendung dieser Arbeit verstarb. Er hätte auch dieses Werk sicher stolz in die lange Reihe von Zeugnissen und Arbeiten seiner Kinder und Enkel gestellt.

Zu Dank verpflichtet bin ich, wie das immer bei Arbeiten ist, die sich über einen längeren Zeitraum erstrecken und in die viel Mühe investiert wird, einer ganzen Reihe von Menschen. Allen voran sind sicherlich meine Betreuer an der Universität Duisburg-Essen, Herr Prof. Dr. Detlev Leutner und Frau Prof. Dr. Elke Sumfleth zu nennen; dies nicht nur wegen der wohldosierten Betreuung, sondern auch, zusammen mit Herrn Prof. Dr. Hans Fischer, Herrn Prof. Dr. Klaus Klemm sowie Frau Prof. Dr. Angela Sandmann, für die Einrichtung und Betreuung der Forschergruppe und des Graduiertenkollegs nwu-essen, das einen exzellenten Rahmen für interdisziplinäre und fruchtbare Forschung bereitstellt.

Herrn Prof. Dr. Rolf Plötzner aus Freiburg danke ich herzlich, dass er sich dazu bereit erklärt hat, als externer Gutachter meiner Dissertation aufzutreten.

Ohne die vielen Studierenden und Schüler und Schülerinnen hätte diese Dissertation zu keinen Ergebnissen geführt, also möchte ich sowohl ihnen als auch den Dozenten und Lehrern danken, die mir ihre Kurse freundlicherweise für meine Testungen zur Verfügung gestellt haben: Dr. Christian Henke, Silke Klos, Dr. Ulrich Läer, Michael Langer, Dr. Andrea Maier-Richter, Dr. Jasmin Neuroth, Dr. Stefan Rumann, Holger Stawitz, Dr. Maik Walpuski. Darüber hinaus vielen Dank an Mathias Ständer für erfolgreiche Vermittlungstätigkeiten von Lehrerkollegen.

Unter keinen Umständen zu unterschätzen sind die motivationssteigernden Wirkungen einer netten Bürobesezung; hier hatte ich mit Regina Hübinger, Silke Klos und Julia Wadouh besonderes Glück. Wir fanden zumeist die Waage zwischen Arbeit und Ablenkung, Stress und Spaß. Auch allen anderen Mitgliedern des Graduiertenkollegs möchte ich danken; es war eine wirklich schöne Zeit! Besonders hervorheben möchte ich Ina Glemnitz, Willi Liegener, Thomas Reyer und Tina Thillmann, die regen Anteil an einer ausgeglichenen Work-Life-Balance hatten.

Schließlich bleibt noch, meiner Frau Claudia Blechert zu danken, die mir in den vergangenen Jahren wie immer am meisten geholfen hat, trotz zuweilen sehr weit, zuweilen ganz nah erscheinender 513 Kilometer Distanz.



# 1 Einführung

## 1.1 Einleitung

Noch immer sind die Wirksamkeit und der sinnvolle Einsatz von dynamischen Animationen zur Steigerung des Lernerfolgs unter Instruktionspsychologen und Didaktikern umstritten. Während noch vor wenigen Jahren die didaktischen Möglichkeiten von Animationen unbegrenzt erschienen, hat mittlerweile eine skeptischere Ansicht die Oberhand gewonnen, die, beeinflusst besonders von der Arbeit von Tversky, Morrisson & Bétrancourt (2002), auf die zusätzliche kognitive Belastung verweist, die dynamische Visualisierungen induzieren. Standbilder, so die Auffassung, könnten zumeist zu mindestens ebenso guten Lernerfolgen führen.

Allerdings ruht dieser Konsens noch auf einem relativ unsicheren, weil empirisch wenig abgesicherten Standbein. Zudem dürfte die überwiegende Mehrzahl professioneller Entwickler und Designer von Lernumgebungen konträrer Meinung sein, die besonders auf die motivierende Wirkung von Animationen verweist.

Ziel dieser Arbeit ist es deshalb, mittels einer Metaanalyse zunächst den tatsächlich aktuellen – und nicht den „gefühlten“ – Forschungsstand zum Lernen mit Animationen versus Standbildern zu erheben und abzubilden. Darüber hinaus sollen womöglich entscheidende Moderatorvariablen identifiziert und mit einer Reihe von empirischen Studien abgesichert werden. Zu den in dieser Dissertation besonders untersuchten Faktoren zählt in erster Linie die Anzahl der dargebotenen Standbilder – kann eine Serie von Standbildern durch schlichtes Hinzufügen weiterer Zwischenschritte an das Lernerfolgsniveau einer äquivalenten Animation angepasst werden? Weiterhin wird die Rolle des räumlichen Vorstellungsvermögens in den Fokus gestellt – dienen besondere Fähigkeiten in diesem Bereich eher als zusätzlicher Verbesserer beim Lernen mit dynamischen Visualisierungen, oder kompensieren sie stattdessen mangelnde Dynamik bei Standbildern? Daneben wird eine Reihe weiterer Größen in Betracht gezogen, die ihren Einfluss beim Lernen mit computerbasierten Lernumgebungen geltend machen könnten, wie etwa kognitive Belastung, Vorwissen, Realistik und einige mehr.

Insgesamt wird mit dieser Arbeit angestrebt, einen Beitrag zu der Diskussion um die Frage „Animation oder Standbilder?“ zu liefern, der zum einen in zusammenfassender Art und Weise so objektiv wie möglich den aktuellen Stand der Forschung abbildet, zum anderen aber auch weitere Erklärungsmöglichkeiten für sich

gegenseitig scheinbar widersprechende Ergebnisse der Vergangenheit anbietet und mögliche moderierende Variablen zur Diskussion stellt.

## 1.2 Struktur dieser Arbeit

Der Schwerpunkt dieser Dissertation liegt auf dem Vergleich dynamischer und nichtdynamischer Visualisierungen. Dazu wurden neben einer theoretisch orientierten Metaanalyse insgesamt drei Experimente durchgeführt. Dementsprechend gliedert sich diese Arbeit in vier Hauptteile, in denen nacheinander auf die auch in dieser Reihenfolge bearbeiteten Untersuchungen eingegangen wird, jeweils inklusive Methodik, Ergebnissen und Einzeldiskussion. Dazu folgt noch ein weiterer Part, der die drei Experimente zusammenführt und übergreifende Ergebnisse diskutiert.

Da sich alle Untersuchungen auf dieselbe Thematik beziehen und aufeinander aufbauen, ist ein gemeinsamer theoretischer Hintergrund als eigenständiges Kapitel vorgeschaltet.

Zum Abschluss folgen eine Zusammenfassung aller Untersuchungsergebnisse sowie eine allgemeine Diskussion zu den theoretischen, praktischen und fachdidaktischen Implikationen der Ergebnisse vorliegender Dissertation.

## 2 Theoretischer Hintergrund

### 2.1 Der Einfluss von Standbildern auf den Lernerfolg

Sicherlich die Mehrzahl heutiger Studien zum Lernen mit Texten und Bildern hat Mayers „Cognitive Theory of Multimedia Learning“ (Mayer, 1997; 2001) zur Grundlage. In dieser wird der multimedial Lernende als Konstrukteur des eigenen Wissens angesehen, der aktiv relevante Informationen, seien sie nun piktorial oder verbal dargeboten, auswählt und miteinander verknüpft.

Den Ausgangspunkt von Mayers Theorie bildet Paivios „Dual Coding Theory“ (Clark & Paivio, 1991; Paivio, 1986), welche annimmt, dass zwei verschiedene kognitive Systeme zur Informationsverarbeitung existieren: Im verbalen System werden Informationen sequentieller Natur wie z.B. geschriebener oder gesprochener Text weitergeleitet und verarbeitet, wohingegen das nonverbale System für räumliche Informationen und Bilder zuständig ist.

Die Vorstellung zweier kognitiver Systeme stimmt mit allgemeinspsychologischen Vorstellungen zum Arbeitsgedächtnis nach Baddeley (1986; 1999) überein, der Indizien für die (funktionale) Existenz einer *phonological loop*, die für verbale Informationen zuständig ist, sowie eines *visuo-spatial sketch pad*, das visuelle und räumliche Informationen verarbeitet, gefunden hat. Darüber hinaus nimmt er eine Zentrale Exekutive an, die beide Subsysteme kontrolliert und mit Langzeitwissen verknüpft. Deren Arbeitsweise ist jedoch deutlich schwerer zu ergründen als die der untergeordneten visuellen und verbalen Systeme und ebenfalls noch nicht abschließend geklärt. Paivios bzw. Mayers Ansatz unterscheidet sich von dem Baddeleys vorrangig darin, dass bei Baddeley *phonological loop* und *visuo-spatial sketch pad* tatsächlich für das jeweilige Informationsformat so zuständig sind, wie es präsentiert und aufgenommen wird, also auditiv oder visuell. Demgegenüber verfolgt Paivio (1986) einen „stimulus-as-presented“-Ansatz (Mayer, 2005a, S. 34), in dem der verbale Kanal verbalen Input und der nonverbale Kanal nonverbale Informationen verarbeitet, unabhängig von der Art der sensorischen Aufnahmemodalität, der verbale Kanal also z.B. sowohl gesprochenen *als auch* geschriebenen Text, der nonverbale Kanal Bilder.

Mayer (2001; 2005c) versucht, mit seinem Modell einen Kompromiss zu finden, indem geschriebener Text zwar zuerst mittels der Augen, also über das visuelle, nonverbale System, aufgenommen und verarbeitet wird („at least initially“; Mayer &

Moreno, 2002), jedoch schnell in das verbale System überführt wird. Der Grund hierfür ist, so die Vorstellung, dass geschriebener Text im Allgemeinen nicht in Form eines Bildes, sondern als sozusagen mental gesprochene Lautfolge in derselben Funktion wie gesprochener Text weiter transportiert werde:

*On-screen text may initially be processed in the visual channel because it is presented to the eyes, but an experienced reader may be able to mentally convert images into sounds that are processed through the auditory channel. (Mayer, 2001, S. 48)*

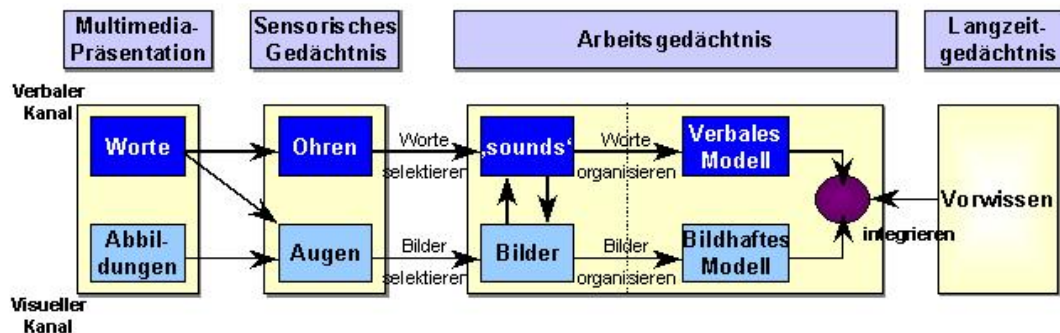
Effizientes Lernen gelingt nach Paivios Modellvorstellung dann am besten, wenn der zu lernende Gegenstand sowohl in Form geschriebenen oder gesprochenen Textes als auch bildhaft kodiert wird. Dadurch kann er über beide kognitiven Systeme, das visuelle und das verbale, verarbeitet werden, und beide mentalen Repräsentationen können miteinander verknüpft werden, womit auch eine leichtere visuelle und verbale Abrufbarkeit gewährleistet ist.

Angewendet auf das multimediale Lernen erweitert Mayer nun Paivios Theorie mittels Wittrocks generativer Lerntheorie (1974, 1989). Wie Wittrock geht Mayer von einem mehrstufigen Prozess des Lernens aus: Nach der Übertragung in das sensorische Gedächtnis erfolgt eine *Selektion* bedeutungstragender Elemente aus der gesamten Information heraus – und zwar, nach Paivio, in beiden Repräsentationskanälen parallel –, so dass eine oberflächliche visuelle bzw. verbale Repräsentation der Information im Arbeitsgedächtnis gebildet wird. Mayer (1997) bezeichnet diese als Bild- bzw. Textbasis.

Sodann werden diese Basen zu mentalen Modellen (Mayer & Moreno, 2002) organisiert, worunter Mayer kohärente visuelle bzw. verbale Repräsentationen von Systemen versteht. Erst zu diesem Zeitpunkt der *Organisation* setzt Verständnis ein; vorher konnte lediglich von einem oberflächlichen Einprägen etwa von Schlüsselbegriffen oder Bildanordnungen ausgegangen werden.

Schließlich, um höchstmögliche Behaltens- und Abrufmöglichkeiten zu gewährleisten, werden beide mentalen Modelle miteinander verknüpft und aufeinander bezogen; zudem wird nun gegebenenfalls bereits vorhandenes Vorwissen aus dem Langzeitgedächtnis *integriert*. Jetzt ist es dem Lernenden möglich, z.B. die in einem Text beschriebenen und in Bildern veranschaulichten Abläufe eines komplexen Prozesses miteinander zu vergleichen und zu verknüpfen sowie kausale Zusammenhänge zu verstehen.

In Abbildung 2-1 ist der gesamte Lernablauf nach Mayers Kognitiver Theorie des Multimedialen Lernens dargestellt (nach Mayer, 2001):



*Abbildung 2-1: Die Cognitive Theory of Multimedia Learning, eigene Darstellung, adaptiert nach Mayer, 2001, S.44.*

In vielen Punkten ähnlich, in entscheidenden Ideen aber ausdifferenzierter, stellt sich ein konkurrierendes Modell von Schnotz zum Verstehen von Text und Bildern dar (Schnotz, 2005; Schnotz & Bannert, 1999; 2003; Schnotz, Seufert & Bannert, 2001): Das „Integrative Model of Text and Picture Comprehension“ unterscheidet wesentlich deutlicher zwischen den Kanälen der Text- und Bildverarbeitung, die als nicht in dem Maße parallel angenommen werden wie in Mayers Modell. Es wird von der Prämisse ausgegangen, dass Texte und Bilder auf unterschiedlichen Repräsentationsprinzipien beruhen – während ein Text eine „deskriptionale Repräsentation“ darstellt, also aus komplexen Symbolen, die einen Sachverhalt beschreiben, besteht, ist ein Bild oder ein Diagramm eine „depiktionale Repräsentation“, die aus ikonischen Zeichen besteht und „mit dem Bezeichneten aufgrund einer Ähnlichkeits- oder Analogierelation durch gemeinsame Strukturmerkmale verknüpft“ ist (Schnotz et al., 2001). Diese Unterscheidung lässt sich, so Schnotz, auch auf mentale Repräsentationen übertragen. Sensorische Kanäle auf der perzeptuellen Ebene werden somit klar von repräsentierenden Kanälen auf der kognitiven Ebene abgegrenzt. Während in Mayers Theorie sensorische Modalität und Repräsentationsformat mittels eines auditiv-verbale und eines visuell-piktoralen Kanals zusammengeführt werden, geht Schnotz' integrativer Ansatz davon aus, dass piktoriale Information nicht zwingend mit visueller Modalität verknüpft ist, sondern auch durch andere sensorische Modalitäten aufgenommen werden kann. So wird zum Beispiel beim Hören des charakteristischen Rufs eines Vogels das akustische Muster durch den auditiven Kanal geleitet, was zu einer

perzeptuellen Repräsentation im auditiven Arbeitsgedächtnis führt. Bei genügend Vorwissen kann die ausgewählte Information dann über einen visuellen Filter im piktorialen Kanal weiter verarbeitet werden und zu einem mentalen Modell des entsprechenden Vogels führen (Schnotz, 2005, S. 58). Das „Integrative Model of Text and Picture Comprehension“ ist in Abbildung 2-2 dargestellt.

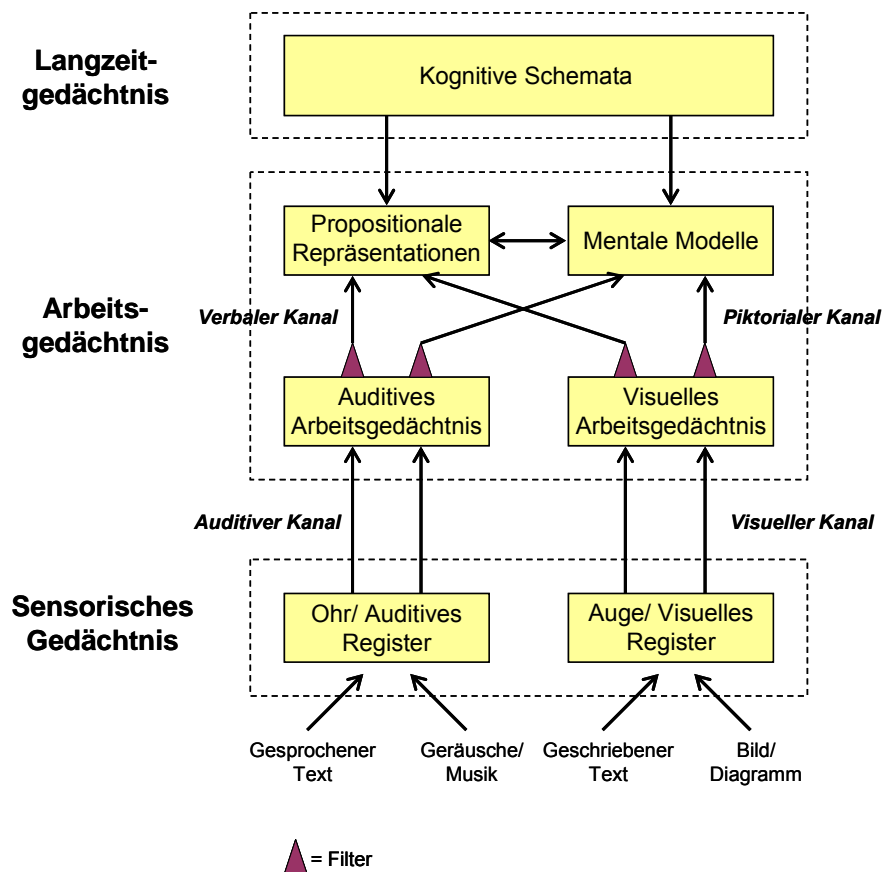


Abbildung 2-2: Das Integrative Model of Text and Picture Comprehension, eigene Darstellung, adaptiert nach Schnotz, 2005, S.57.

### Prinzipien multimedialen Lernens

Mayer (2001; 2005c) hat im Laufe der Zeit eine Reihe von Prinzipien zusammengetragen, teilweise auf eigene Arbeiten zurückgehend, teilweise sich auf Erkenntnisse anderer Autoren stützend, die sich als förderlich bei der Gestaltung multimedialer Lernumgebungen erwiesen haben. Er unterscheidet zwischen Grund- und erweiterten Prinzipien, wobei letztere zumeist Spezialfälle betreffen und besondere Verfeinerungen wie „worked-out examples“ (Renkl, 2005) oder „guided discovery learning“ (de Jong, 2005) darstellen, auf die hier nicht näher eingegangen werden soll. Die Grundprinzipien multimedialen Lernens sind folgende:

- *Das Multimediaprinzip:* Das Multimediaprinzip folgt direkt aus der *Dual-coding*-Theorie (vgl. Abbildung 2-1) und besagt, dass Menschen besser mittels Texten und Bildern lernen als mit Texten allein. Dieses Prinzip ist sicherlich intuitiv nachvollziehbar und wurde laut Schnotz (2005) bereits im 17. Jahrhundert von Comenius (1658) als sinnvoll postuliert und angewandt. Mittlerweile wird angenommen, dass auf diese Weise deshalb mehr gelernt wird, weil beide Kanäle der Informationsverarbeitung – der visuelle wie der verbale – gleichzeitig angesprochen und referentielle Verknüpfungen gebildet werden (Fletcher & Tobias, 2005). Eine Anzahl von Studien – viele davon in Mayer (2001) zusammengefasst – konnte diesen Effekt empirisch nachweisen. Allerdings weisen Schnotz & Bannert (2003) darauf hin, dass dem Lerngegenstand unangemessene Bilder negative Effekte auf das Lernen haben können – und damit das Hinzufügen von Bildern nicht, wie Paivios (1986) *Dual-coding-theory* impliziert, generell lernförderlich ist.
- *Das Redundanzprinzip:* Im Gegensatz zum Multimediaprinzip erscheint das Redundanzprinzip eher kontraintuitiv. Es postuliert, dass redundante Lernmaterialien eher weggelassen denn hinzugefügt werden sollten – sei es in Form zusätzlicher Präsentationsformen (dieselbe Information etwa als geschriebener Text und als Audiokommentar), sei es in Form einer ausführlicheren Präsentation statt einer Zusammenfassung (Mayer, 2001, bezeichnet diese Form noch als *Kohärenzprinzip*). Kontraintuitiv ist dieses Prinzip deshalb, weil im Allgemeinen die Wiederholung ein lernförderliches Element darstellt – jedoch bietet die *Cognitive Load Theory* (eine ausführliche Darstellung dieser Theorie folgt im weiteren Verlauf dieses Kapitels) eine plausible Erklärung für negative Effekte redundanter Information an: Gerade neue Informationen müssen ausschließlich im Arbeitsgedächtnis verarbeitet werden und dürfen deshalb nicht zu ausführlich oder doppelt präsentiert werden, um keine Überlastung des Arbeitsspeichers hervorzurufen (Sweller, 2005). Trotzdem gilt: „Revision is not a redundant activity“ (Sweller, 2005, S. 166) – Wiederholungen sind oftmals wichtig, erfordern jedoch keinen Wechsel der Präsentationsform oder unnötig ausführliche Präsentationen. Es bleibt noch festzuhalten, dass Schnotz (2005) dieses Prinzip noch erweitert (und gleichzeitig dem

Multimediaprinzip widerspricht), indem er für Lernende mit hohem Vorwissen einen *General Redundancy Effect* annimmt, der sogar die gleichzeitige Darstellung eines Sachverhalts in zwei Modalitäten (verbal/visuell) als unter Umständen schädlich postuliert. Bei genügend Vorwissen stellt das Ansprechen eines zweiten Kanals redundante Information dar, die unnötigerweise kognitive Ressourcen beansprucht. Dieser Effekt wird auch als *Expertise Reversal Effect* (Kalyuga, Chandler & Sweller, 2000) bezeichnet und im späteren Verlauf dieses Kapitels noch ausführlicher behandelt werden.

- *Das Split-Attention-Prinzip*: Ayres & Sweller (2005) definieren das Auftreten eines instruktionalen *Split-Attention*-Effektes als eine Situation, in der Lernenden abverlangt wird, ihre Aufmerksamkeit zwischen verschiedenen, räumlich oder zeitlich voneinander getrennten, aber für das Verständnis essentiellen Informationsquellen zu teilen und diese mental zu integrieren. Die Notwendigkeit, voneinander getrennte Informationen zusammenzuführen, erhöht die kognitive Belastung und kann zu einer Erniedrigung des Lernerfolgs im Vergleich zu Bedingungen führen, in denen die Informationen zeitgleich und/oder räumlich nah beieinander präsentiert werden. Tarmizi & Sweller (1988) konnten als erste zeigen, dass Lernende bessere Lernerfolge vorwiesen, wenn in *worked examples* (Renkl, 2005) Lösungen in Textform und dazugehörige Diagramme integriert waren, statt untereinander dargestellt zu werden. Eine Vielzahl weiterer Studien konnte diesen räumlichen *Split-Attention*-Effekt replizieren. Gleiches gilt für das zeitliche *Split-Attention*-Prinzip: Eine gleichzeitige Darstellung zusammengehöriger, relevanter Informationen ist sinnvoller als eine zeitlich gestaffelte (Ayres & Sweller, 2005).

Mayer (2001; 2005b) bezeichnet diese beiden Phänomene als *spatial* und *temporal contiguity principles* und sieht sie als Spezialfälle des *Split-Attention*-Prinzips an. Es bleiben jedoch noch einige Einschränkungen des Prinzips festzuhalten. Zum einen ist es unabdingbar, dass die verschiedenen Informationsquellen tatsächlich aufeinander aufbauende und nicht redundante Fakten präsentieren (sonst griffe das Redundanzprinzip, siehe oben), zum anderen gilt das Prinzip nur bei schwierigem, komplexen Material, denn sehr einfaches Material



überfordert in der Regel auch bei nicht optimaler Darstellung nicht das Arbeitsgedächtnis.

- *Das Modalitätsprinzip:* Das Modalitätsprinzip besagt, dass Erläuterungen in multimedialen Lernumgebungen eher als verbale Erklärung statt als geschriebener Text auf dem Bildschirm dargeboten werden sollten, wenn sie, dem Multimediaprinzip folgend, mit Bildern kombiniert werden (Mayer & Moreno, 2002). Die Begründung für die höhere Lernwirksamkeit der verbalen Erklärungen im Gegensatz zu geschriebenen Texten erfolgt abermals mit Hilfe der *Dual-coding*-Theorie: Geschriebener Text muss – zumindest zu Anfang – ebenso wie die gleichzeitig dargebotenen Bilder über den visuellen Kanal aufgenommen und verarbeitet werden. Verbale Erläuterungen hingegen nutzen von Beginn an den unbelasteten auditiven Kanal und erniedrigen so, wenn sie keine redundanten Informationen darstellen, die kognitive Belastung. Weiterhin ist die Aufmerksamkeit des Lernenden, wenn Text und Bild gleichzeitig auf dem Bildschirm zu sehen sind, immer nur auf einen bestimmten Teil gerichtet, lediglich eine Informationsquelle kann jeweils beachtet und bearbeitet werden. Damit stellt dieses Prinzip im Grunde einen Spezialfall des *Split-Attention*-Effektes dar (Low & Sweller, 2005).

Der Modalitätseffekt ist insgesamt gut empirisch belegt (z.B. Brünken, Plass & Leutner, 2004; Ginns, 2005), es gibt allerdings auch einige Autoren, die den Effekt nicht replizieren konnten (z.B. Tabbers, 2002). Zudem weist Schnotz (2005) darauf hin, dass gerade bei komplexeren Lerninhalten gesprochener Text einen entscheidenden Nachteil gegenüber dem geschriebenen Pendant hat: Er ist „fleeting“ (S. 65), schreitet also kontinuierlich voran, ohne dass der Lerner darüber Kontrolle hätte und ggf. einzelne Teile wiederholen könnte, wie das bei geschriebenem Text der Fall wäre. Im Falle schwer verständlicher Lerninhalte kombiniert mit statischen Bildern und nicht limitierter Lernzeit empfiehlt Schnotz (2005) deshalb also, eher geschriebenen Text zu verwenden.

- *Das Kohärenzprinzip:* Das Kohärenzprinzip befasst sich mit der Rolle zusätzlicher, für das Verständnis des Lerngegenstands nicht unbedingt notwendiger Elemente, durch die die Lernumgebung ergänzt wird. Als Beispiele führen Mayer & Moreno (2002) etwa interessante

Hintergrundinformationen oder begleitende Musik oder Soundeffekte an. Zwar wäre es anzunehmen, dass solcherlei Elemente eine Motivationssteigerung beim Lernenden zur Folge haben, jedoch zeigte eine Reihe von Autoren (Harp & Mayer, 1997; Moreno & Mayer, 2001; Renninger, Hidi & Krapp, 1992), dass diese stattdessen sogar zu einer Leistungsminderung führen können. Erklärbar sind die Ergebnisse mit Hilfe der *Cognitive Load Theory* (siehe unten): Je mehr zusätzliche, aber unnötige Elemente eine Lernumgebung enthält, desto mehr *Extraneous Cognitive Load* wird erzeugt, desto wahrscheinlicher entsteht eine kognitive Überlastung beim Lernenden. Das Kohärenzprinzip besagt also, dass nicht unbedingt notwendige Elemente der Lernumgebung im Zweifelsfall eliminiert werden sollten. Wie bereits bemerkt, stellt das Kohärenzprinzip (Mayer, 2001) damit nach neuerer Auffassung (Sweller, 2005) im Grunde nur einen Spezialfall des Redundanzprinzips dar ist somit ironischerweise mittlerweile redundant. Aufgrund der immer noch vorkommenden Nutzung dieses Begriffs (z.B. Mayer, 2005b) wurde dieses Prinzip der Vollständigkeit halber hier erwähnt.

## 2.2 Der Einfluss von Animationen auf den Lernerfolg

In den Kernaussagen von Mayers Theorie wird noch nicht prinzipiell zwischen stehenden Bildern und Animationen unterschieden. Eine Animation lässt sich definieren als eine Serie schnell wechselnder Bilder auf dem Computerbildschirm, die dem Betrachter die Vorstellung einer Bewegung vermittelt (Rieber & Kini, 1991). Sie soll dem Zweck dienen, sich die dargestellten Abläufe genau vorstellen zu können und so eine adäquate mentale Repräsentation des Vorgangs zu bilden. Nach dem Supplantationskonzept von Salomon (1979) sollte diese Eigenschaft der Animation, dynamische Abläufe darzustellen, die möglicherweise mangelhafte Fähigkeit der Person, Bewegungen zu imaginieren, ersetzen („supplantieren“). Die Animation dient dann als externes Modell für die interne Vorstellung („a ‚ready-made‘ explicit dynamic representation of the situation“; Schnotz & Lowe, 2003, S. 120).

Zwar werden Bewegungen nach gängiger Vorstellung (Schnotz, 2006; Squire & Kandel, 1999) im Langzeitgedächtnis vermutlich nicht als bewegte mentale Modelle gespeichert, sondern lediglich in Form der wichtigsten Schlüsselbilder – die dann

möglicherweise verwendet werden, um die Bewegung mental zu simulieren (Narayanan & Hegarty, 2002) –, jedoch dürfte eine exakte bewegte Darstellung von Prozessen die Bildung eines solchen Modells unterstützen. Dies gilt insbesondere für abstrakte, schwer nachzuvollziehende Sachverhalte (Weiss, Knowlton & Morrison, 2002). Besonders hervorzuheben als möglicherweise sogar „einzig haltbarer Grund“ (Weiss et al. 2002, S. 469) zur Nutzung von Animationen ist deren Eigenschaft, Bewegungsabläufe darstellen zu können. Schwartz (2003) fügt dem noch die Darstellbarkeit von Transformationen von einem Zustand in einen anderen hinzu.

Eine Reihe von Autoren (z.B. Michas & Berry, 2000; Rieber, 1991; Spotts & Dwyer, 1996; Szabo & Pookay, 1996) konnte in ihren Untersuchungen einen positiven Einfluss von Animationen auf den Lernerfolg belegen. Mayer & Moreno (2002) zeigten darüber hinaus die prinzipielle Übertragbarkeit der Kognitiven Theorie multimedialen Lernens vom Lernen mit Text und Bild auf das Lernen mit Text und Animationen auf. Bétrancourt (2005) resümiert einige Prinzipien, die aus der aktuellen Forschung zum Lernen mit Animationen folgen und noch über die Prinzipien multimedialen Lernens (s.o.) hinausgehen:

- *Das Apprehension-Prinzip:* Es ist vonnöten, dass das gewählte Design der Elemente einer Animation unmittelbar verständlich ist und den gängigen Konventionen folgt (Tversky et al., 2002). In diesem Zusammenhang folgt das Prinzip weiterhin dem Redundanzprinzip (s.o.), indem es speziell zusätzliche, aber für das Verständnis unnötige kosmetische Verschönerungen sowie die Darstellung in 3D verneint.
- *Das Kongruenzprinzip:* Veränderungen in der Animation sollten eher der Bildung eines funktionalen mentalen Modells dienen als auf eine hundertprozentig korrekte Darstellung der Realität Wert legen, d.h. zum Beispiel sollten Ereignisse, die eigentlich gleichzeitig geschehen, sich aber gegenseitig bedingen, nacheinander dargestellt werden, um die Kausalitätsprinzipien sinnvoll zu illustrieren und die Bildung eines korrekten mentalen Modells zu begünstigen.
- *Das Interaktivitätsprinzip:* Zwar kann es leicht geschehen, dass zusätzliche Interaktivität Lernende durch eine Vielzahl zusätzlicher Möglichkeiten überlastet (Nerdel, 2003), jedoch scheinen gewisse zusätzliche Funktionen wie „Weiter“ (Mayer & Chandler, 2001) oder übliche Menübuttons wie an einem Videorecorder (Schwan, Garsoffky & Hesse, 2000) den Lernerfolg

von Animationen zu unterstützen. Die Möglichkeit zur eigenen Einteilung der Informationsaufnahme sowie die Möglichkeit zur Wiederholung dürften dem Effekt des „Vorbeirauschens“ der Animation entgegenwirken.

- *Das Aufmerksamkeitssteuerungsprinzip:* Lowe (2003) konnte zeigen, dass mit Animationen Lernende häufig besser graphisch auffällige Charakteristika erinnern denn thematisch wichtige Vorgänge, da besonders Novizen nicht zwischen wichtigen und unwichtigen Details unterscheiden können. Es scheint daher gerade in Animationen geraten, wichtige Elemente optisch hervorzuheben. Dieses Prinzip wird auch als „cueing“ (Rieber, 1994) oder „signaling“ (Bétrancourt, 2005; Mautone & Mayer, 2001) bezeichnet.

### 2.3 Animation versus Standbilder

Salomons (1979) Konzept lässt zunächst einmal die Vermutung zu, dass Animationen bezüglich des Lernerfolgs Lernumgebungen mit Standbildern überlegen sein sollten, da in diesem Falle die Animation als äußeres Medium den fehlenden inneren Prozess der Imagination dynamischer Abläufe ersetzt (Urhahne, Prenzel, von Davier, Senkbeil & Bleschke, 2000). Als weiterer Grund hierfür ist anzuführen, dass die Animation dem Lernenden die Last der Imagination und Deutung weitgehend abnimmt und so eine niedrigere kognitive Belastung darstellen könnte. Bei statischen Bildern müssen mitunter erst mehr oder minder abstrakte Bewegungssymbole wie Pfeile etc. gedeutet und verschiedene Informationen miteinander verbunden werden. Diese könnten jedoch leicht missinterpretiert werden, wohingegen die dargestellten Bewegungen der Animation eindeutiger sind. Besonders die Darstellung eines zeitlichen Ablaufs fällt mit den begrenzten Aussagemöglichkeiten der Bewegungsindikatoren statischer Bilder naturgemäß deutlich schwerer (Lewalter, 1997, S. 105).

Die Animation hingegen regt, so Rieber & Kini (1991), eine automatisierte Fähigkeit zur Herstellung einer (scheinbaren) Bewegung an und setzt so Kapazitäten des Arbeitsgedächtnisses frei. Diese frei gewordenen Ressourcen können wiederum zusätzlich zum Bemühen um Verständnis eingesetzt werden.

Dagegen lässt sich allerdings argumentieren, dass eine Animation im Gegensatz zu Standbildern keine permanenten, sondern „transiente“, flüchtige Informationen

liefert (z.B. Ainsworth & van Labeke, 2004), was die zeitlichen Limits des Arbeitsgedächtnisses nach Baddeley (1986; 1992) sehr strapazieren kann. Die *Cognitive Load Theory* (Chandler, 2004; Chandler & Sweller, 1991; Sweller, 1994) nimmt in diesem Zusammenhang verschiedene Arten kognitiver Belastung an, auf die Animationen oder Standbilder unterschiedliche Auswirkungen haben dürften.

### *Die Cognitive Load Theory*

Der *Cognitive Load Theory* (Chandler, 2004; Chandler & Sweller, 1991; Sweller, 1994) folgend existieren drei verschiedene Typen kognitiver Belastung des Arbeitsgedächtnisses: *Intrinsic*, *Extraneous* und *Germane Cognitive Load*. Dabei wird intrinsische Belastung als bestimmt durch die Menge der „interagierenden“ Elemente in einem zu lernenden Bereich angesehen und stellt somit ein direktes Maß der Schwierigkeit eines Bereichs dar. Es herrscht im Allgemeinen Konsens darüber, dass der *Intrinsic Load* vom Designer einer Lernumgebung nicht beeinflusst werden kann<sup>1</sup>. Hingegen sind *Extraneous* und *Germane Load* sehr stark manipulierbar. Zur Erreichung höchsten Lernerfolges sollte versucht werden, den *Extraneous Load* zu minimieren und gleichzeitig den *Germane Load* unter Berücksichtigung der Kapazitätsgrenzen des Arbeitsgedächtnisses zu optimieren. Ersterer wird definiert als unnötig implizierte kognitive Belastung, die durch ineffizientes instruktionales Design hervorgerufen wird. Dies kann etwa redundanten Text oder eine ungünstige, einen *Split-Attention*-Effekt hervorrufende räumliche Aufteilung der Lernumgebung beinhalten. Als Antagonist hierzu kann der *Germane Load* bezeichnet werden. Dieser bezeichnet die vom Lernenden aufgewandten kognitiven Ressourcen, die zur Schemabildung, d.h. zur Zusammenfassung und Organisation einer Vielzahl einzelner Elemente eines Sachverhaltes, benötigt werden. Gebildete Schemata werden, so die Modellvorstellung, zusammenhängend im Langzeitgedächtnis abgelegt und reduzieren so die kognitive Belastung. Eine optimal gestaltete

---

<sup>1</sup> Dies ist in jüngster Zeit insofern umstritten, als eine Reihe von Autoren (z.B. Gerjets, Scheiter & Catrambone, 2004; Kester, Paas & van Merriënboer, 2006) die Meinung vertreten, den *Intrinsic Load* durch Segmentierung des Lerngegenstands reduzieren zu können. Da jedoch die gesamte intrinsische Belastung über die Zeit trotzdem gleich bleibt, widerspricht diese Ansicht tendenziell der ursprünglichen Theorie (Chandler & Sweller, 1991).

Lernumgebung erleichtert die Schemakonstruktion und weist eine hohe *Germane Cognitive Load* bei gleichzeitig niedriger *Extraneous Load* auf. Dies ist insbesondere bei hochkomplexen Lerninhalten notwendig, da in solchen Fällen ein hoher *Intrinsic Load* vorliegt und zuviel *Extraneous Load* die Kapazität des Arbeitsgedächtnisses überlasten könnte, es also zu einem *Overload* käme, bei dem keine Arbeitsgedächtniskapazität mehr verfügbar wäre, um *Germane Load* zu induzieren und kognitive Verarbeitungsprozesse zu ermöglichen. Ist die durch einen Lerngegenstand hervorgerufene intrinsische Belastung niedrig, muss auch eine erhöhte extrinsische Ladung nicht zu niedrigerer Lernleistung führen:

*If the total cognitive load is not excessive due to a relatively low intrinsic cognitive load, then a high extraneous cognitive load may be irrelevant because students are readily able to handle low element interactivity material with almost any form of presentation (Sweller, 1994, S. 308).*

Es ließe sich also argumentieren, dass eine Animation im Gegensatz zu Standbildern aufgrund ihres schnell fortschreitenden Charakters einen höheren *Extraneous Cognitive Load* impliziert und so, bei entsprechender hoher intrinsischer Belastung, leichter die Kapazität des Arbeitsgedächtnisses überlastet. Dies stünde in Übereinstimmung zu Ergebnissen von Mayer, Hegarty, Mayer & Campbell (2005), die in vier von acht Vergleichen eine signifikante Überlegenheit (in den anderen Fällen keinen Unterschied) von papierbasierten Standbildern gegenüber computerbasierten Animationen hinsichtlich des Lernerfolgs fanden und auch die *Cognitive Load* als eine mögliche Erklärung heranziehen. Jedoch erscheint in diesem Falle die Vergleichbarkeit der Bedingungen (papierbasiert versus computerbasiert) zumindest zweifelhaft.

Weiterhin deutet vieles darauf hin, dass die lern- bzw. verstehensförderliche Wirkung von Standbildern neben dem Zusammenwirken mit Text auch durch Hinweisreize (Park, 1998), durch Beschränkung auf Schlüsselbilder (Catrambone & Seay, 2002; Hegarty, Kriz & Cate, 2003) oder durch besondere Realistik (Michas & Berry, 2000) entscheidend erhöht werden kann. Tversky e al. (2002) weisen zusammenfassend darauf hin, dass Animationen nicht grundsätzlich überlegen seien und oft deutlich überschätzt würden; häufig sei es nämlich in Primäruntersuchungen der Fall gewesen, dass Animationen und Standbilder nicht dasselbe Maß an Information beinhaltet hätten und deshalb nicht vergleichbar gewesen seien. Schließlich zeigten Brünken, Gall und Leutner (2001) als weiteren Negativpunkt von

Animationen, dass mit Animationen und dazugehörigen Texten Lernende ihre Aufmerksamkeit deutlich stärker den Animationen zu Lasten der Texte zuwandten, was eine niedrigere Lösungsquote textbasierter Aufgaben zur Folge hatte.

## 2.4 Beeinflussende Moderatoreffekte

Auch eine Reihe von inter- und intraindividuellen Moderatoren kann Einfluss darauf haben, ob eher Standbilder oder Animationen bei dem jeweiligen Probanden zu höherer Lernleistung führen. So wird z.B. die *Spatial Ability*, das räumliche Vorstellungsvermögen, dafür verantwortlich gemacht, dass nicht jeder von Animationen gleichermaßen profitiert (Blake, 1977; Hays, 1996; Large, Beheshti, Breuleux & Renaud, 1996; Yang, Andre & Greenbowe, 2003). Mit Brünken, Steinbacher & Leutner (2000) wird hier als ein zentraler Aspekt des räumlichen Vorstellungsvermögens die Fähigkeit verstanden, „Objekte in zwei oder drei Dimensionen mental rotieren oder falten und sich die Lage- oder Formveränderungen derart manipulierter Objekte vorstellen zu können“ (ebd., S. 38).

Hegarty (2005, S. 457) fasst zusammen, dass das räumliche Vorstellungsvermögen wohl in jener Hinsicht eine Rolle beim Lernen mit Animationen spiele, dass Lernende mit hoher *Spatial Ability* eher vom Lernen mit Animationen profitierten als solche mit niedrigem räumlichen Vorstellungsvermögen (Isaak & Just, 1995). Mayer & Sims (1994) bezeichnen dies als *ability-as-enhancer*-Hypothese.

Intuitiv nahe liegender wäre es jedoch, wenn gerade Lernende mit niedriger *Spatial Ability* durch Animationen geholfen werden würde, ein adäquates mentales Modell des dargestellten Prozesses zu entwickeln bzw. direkt zu übernehmen; ein solches aus Standbildern zu generieren sollte an und für sich ungleich schwerer und deshalb eher von Lernenden mit hoher *Spatial Ability* zu bewältigen sein (*ability-as-compensator*-Hypothese, Mayer & Sims, 1994). Dies stünde im Einklang mit der Supplantationstheorie von Salomon (1979), wonach eine fehlende Fähigkeit des Lernenden (hier: räumliches Vorstellungsvermögen) durch die Vorführung eines dynamischen Prozesses ersetzt wird (Urhahne et al., 2000).

Eine weitere Einfluss nehmende Variable stellt das themenspezifische Vorwissen dar, da bei hohem Vorwissen weniger Kapazität des Arbeitsgedächtnisses für die Inhalte zugunsten des Bemühens um Verständnis der dargestellten Bewegung aufgebracht werden sollte (ChanLin, 2001; Höffler, 2003; Nerdel, 2003; Szabo &

Poohkay, 1996). Lernende mit hohem Vorwissen konzentrieren sich eher auf die tatsächlich relevante Information, um ein adäquates mentales Modell zu generieren (Schnotz, Picard & Hron, 1993), während Novizen leicht von perzeptorisch besonders auffälligen Reizen abgelenkt werden (Lowe, 2003): „Prior knowledge can determine the ease with which learners can perceive and interpret visual representations in working memory“ (Cook, 2006, S. 1073).

### *Der „Expertise Reversal“-Effekt*

In diesem Zusammenhang spielt der „*Expertise Reversal*“-Effekt eine Rolle, der in erster Linie auf Kalyuga zurückgeht (Kalyuga, Ayres, Chandler & Sweller, 2003; Kalyuga, Chandler & Sweller, 1998; 2000). Kalyuga konnte zeigen, dass bestimmte instruktionale Techniken, die sich zur Erleichterung des Lernens bewährt haben, für Lernende mit hohem Vorwissen („Experten“) nicht nur wirkungslos bleiben, sondern sogar negative Effekte zeigen können. Zu diesen bewährten Techniken gehört die Beachtung einer Reihe der oben aufgeführten Multimediaprinzipien wie z.B. des *Redundancy*- und des *Modality*-Prinzips. Erklärt wird dieser Effekt mittels der *Cognitive Load Theory*: Experten weisen in dem zu lernenden Bereich im Gegensatz zu Novizen bereits zuvor gebildete und gespeicherte Schemata auf, die mit der jeweiligen Aufgabe unmittelbar assoziiert werden können. Dadurch können Experten in der Regel schneller und besser eine Aufgabe lösen, da das jeweilige Schema ihnen ausreichende Hilfestellung und Anleitung (*guidance*) zur Bewältigung bietet. Novizen hingegen müssen ausweichen auf „problem-solving search strategies that are cognitively inefficient because they impose a heavy working memory Load“ (Kalyuga et al., 2003, S. 24). Instruktionale Hilfestellungen wie die Beachtung der Multimediaprinzipien oder z.B. die Bereitstellung von *worked-out-examples* (Renkl, 2005) zielen auf eine Anleitung zum Umgang mit dem entsprechenden Lerngegenstand, setzen also genau an der Stelle an, an der es Novizen mangelt – sie unterstützen die Schemabildung. So hilfreich dies für Novizen ist, so leicht kann es Experten passieren, dass die instruktionalen Hilfestellungen mit den bestehenden Schemata kollidieren und so versucht werden muss, redundante Informationen miteinander zu verknüpfen. Dies wiederum verursacht eine zusätzliche kognitive Belastung des Arbeitsgedächtnisses und kann zu einem *Cognitive Overload* führen, was eine reduzierte Lernleistung nach sich zieht.



Relativ häufig wurde der Unterschied in der Lernleistung bei prozeduralen, eher auf die Problemlösefähigkeit fokussierenden Aufgaben und deklarativen, reines Faktenwissen betreffenden Aufgaben untersucht (zur Unterscheidung von Lernen und Problemlösen vgl. Wirth, 2004). Die gängige Argumentation diese Frage betreffend lautet, dass besonders das tiefere Verständnis von der Verwendung von Animationen profitieren würde (Mayer & Moreno, 2002), Faktenwissen hingegen auch mit Standbildern ausreichend gefördert würde. Verschiedene Untersuchungsergebnisse liefern jedoch ein sehr heterogenes Bild (u.a. Catrambone & Seay, 2002; ChanLin, 1998; 2001; Large et al., 1996; Nerdel, 2003; Wright, Milroy & Lickorish, 1999; Yang et al., 2003). Die Rolle von eher prozedural-motorischem Lernen ist daneben noch weitgehend ungeklärt und nur teilweise untersucht (z.B. Spangenberg, 1973).

Daneben erscheinen einige Moderatoren denkbar, die bisher noch nicht im Fokus der Untersuchungen standen, nichtsdestotrotz aber eine Rolle spielen könnten. So könnte es etwa einen Unterschied machen, ob die in der Animation dargestellte Bewegung zentrales Element des Lerngegenstands ist oder nur benutzt wird, um einen Sachverhalt zu verdeutlichen (vgl. Rieber, 1990). Auch könnte es sein, dass durchaus Unterschiede zwischen computer- und videobasierten Animationen hinsichtlich der Lernleistung existieren, allein schon wegen der höheren Realitätsnähe, aber möglicherweise auch darüber hinausgehend. Weiterhin mag es der Fall sein, dass Animationen in bestimmten Wissensdomänen nützlicher sind als in anderen (etwa in der Physik, wie z.B. von Mayer, 2001, immer wieder verwendet, oder in der Biologie, vgl. Nerdel, 2003).

Dieser Überblick macht bereits deutlich, dass für Animationen wie auch für Standbilder eine Fülle von guten Argumenten bereitstehen, weshalb die eine oder andere Art überlegen sein könnte. Und so breit die Argumentationsmöglichkeiten sind, so vielfältig scheinen auch die Ergebnisse der Untersuchungen zur Thematik zu sein. Darum ist die (erstmalige) Durchführung einer Metaanalyse sinnvoll, um zum einen mögliche *Overall*-Effekte der einen oder anderen Version aufzudecken und zum anderen in quantitativer Weise Moderatoren des Lernerfolgs unter den verschiedenen Bedingungen zu identifizieren.

## 3 Metaanalyse zur Wirksamkeit von Animationen und Standbildern

### 3.1 Einführung

Während die Nutzung von Bildern zur Veranschaulichung von Sachverhalten ein seit langem gängiges Mittel ist, drängen in den letzten Jahren auch immer stärker computergenerierte Animationen in die Didaktik vor. Sowohl in Form spezieller Lernprogramme als auch als Zwischensequenzen in Lehrfilmen kann ihnen mehr und mehr begegnet werden. Dabei ist es in der wissenschaftlichen Diskussion höchst umstritten (vgl. Kapitel 2.3), ob Animationen gegenüber Standbildern überhaupt einen Lernvorteil erbringen. Es herrscht der Eindruck, jeder Untersuchung, die die Animationen im Vorteil sieht, stünde mindestens eine entgegen, die das Gegenteil indiziert.

Dabei werden verschiedenste Untersuchungsbedingungen eingeführt, die für die eine oder die andere Version möglicherweise moderierende Einflüsse aufweist. Erleichtern z.B. spezielle Hinweisreize („cues“) in den Standbildern das Erlernen von Sachverhalten (u.a. Park, 1998; Lewalter, 1997), wirken sich Animationen nur bei Aufgaben zum Problemlösen positiv aus (Large, Beheshti, Breuleux & Renaud, 1996), spielt das Vorwissen (ChanLin, 2001; Nerdel, 2003; Höffler, 2003; Szabo & Poohkay, 1996) oder die *Spatial Ability* (Blake, 1977; Hays, 1996; Yang, Andre & Greenbowe, 2003) eine Rolle? Sollte eine Animation so gestaltet sein, dass thematisch Relevantes auch graphisch besonders auffällig ist (Lowe, 2003)? Kann die Wirkung von Standbildern dadurch verbessert werden, dass sie die entscheidenden Schlüsselbilder des Bewegungsablaufs zeigen (Catrambone & Seay, 2002)? Eventuell ist auch der inhaltliche Wissensbereich oder die Komplexität der Animation der entscheidende Faktor.

Es bleibt festzustellen, dass es noch keinen „common sense“ über diese Frage gibt und dass auch verschiedene Reviews (u.a. Large, 1996; Tversky et al., 2002) keinen umfassenden Überblick oder gar eine abschließende Lösung anbieten konnten.

Es scheint also dringend notwendig, die vorliegende Literatur zum Vergleich von Animationen und Standbildern zu sichten, zu ordnen und aufzuarbeiten. Die hier zu berichtende Studie versucht genau dies, indem sie mittels einer Metaanalyse

zunächst mögliche „Overall-Effekte“ identifizieren möchte und anschließend bestrebt ist, aus der Vielzahl von verschiedenen Untersuchungsbedingungen relevante Moderatoren für den erhöhten Lernerfolg bei Animationen oder Standbildern zu finden.

## 3.2 Methode

In dieser Studie wurde auf die Methode der Metaanalyse (Glass, McGaw & Smith, 1981; Hedges & Olkin, 1985) zurückgegriffen. Diese hat das Ziel, die Ergebnisse einer Vielzahl von Studien zusammenzufassen und sowohl allgemeine Effekte aufzudecken als auch mögliche moderierende Faktoren zu identifizieren. Im vorliegenden Fall sollten mögliche *Overall*-Effekte der Wirksamkeit von Animationen oder Standbildern hinsichtlich des Lernerfolgs untersucht und zudem Faktoren gefunden werden, die einen solchen möglichen Vorteil weiter ausbauen oder schmälern können.

Eine Metaanalyse gliedert sich traditionell in drei Stufen: (1) Auffindung und Selektion geeigneter Studien, (2) Kodierung interessierender Eigenschaften, (3) Effektstärkenbestimmung und -analyse zur Untersuchung der Beziehungen zwischen Studieneigenschaften und -effekten.

### 3.2.1 Auffindung und Selektion geeigneter Studien

Zur Identifikation geeigneter Studien wurden zunächst die einschlägigen computerisierten Datenbanken SCI und SSCI (1993-2004), ERIC (1966-2004) sowie PsycInfo (1887-2004) und Psyn dex (1977-2004) durchsucht. Hierzu wurde eine Vielzahl von Kombinationen aus den Suchbegriffen „Animation“, „dynamic picture“, „dynamic image“, „still picture“, „still image“, „static picture“, „static image“, „motion“, „steps“, „simulation“ etc. verwendet. Die sich ergebende sehr große Menge an Treffern wurde daraufhin anhand der Abstracts, soweit sie verfügbar waren, überblickt und nach geeigneten Studien selektiert. Im Zweifelsfall wurde der gesamte Artikel beschafft und ggf. nach Lektüre aussortiert. Zusätzlich zu den Treffern aus den Datenbanken lieferten die Literaturverzeichnisse und Querverweise der bereits gefundenen Artikel Hinweise auf weitere Studien. Als besonders hilfreich erwiesen sich in diesem Zusammenhang die Reviews von Alesandrini (1982), Bétrancourt &

Tversky (2000), Large (1996), Milheim (1993), Park & Hopkins (1993) sowie Tversky et al. (2002).

Um dem „Veröffentlichungsbias“ ein Stück weit entgegenzuwirken, war auch das Bemühen vorhanden, nicht veröffentlichte Dissertationen, Diplomarbeiten oder Konferenzabstracts zu beschaffen und aufzunehmen. Dies glückte jedoch nur in geringem Umfang, nicht zuletzt wegen der häufig unzureichenden Methodik in solchen Studien.

Die Suche lieferte zunächst 57 Artikel zur näheren Analyse. In die Metaanalyse übernommen wurde ein Artikel dann, wenn a) ein Vergleich einer Animationen beinhaltenden Lernumgebung mit Standbildern vorgenommen wurde, b) kein Mix aus Standbildern und Animationen in einer Bedingung vorlag, c) keine oder nur minimale Interaktivität<sup>2</sup> vom Lernenden gefordert war und d) augenscheinlich beide Bedingungen inhaltlich vergleichbar waren.

Auf diese Weise fielen 25 der 57 Artikel wieder heraus, weil sie eines oder mehrere der obigen Kriterien nicht erfüllten. In fünf Fällen wurden weiterhin Studien ausgeschlossen, weil sie methodische Mindeststandards nicht erfüllten (wobei dieses Kriterium, ganz im etablierten Glass'schen Sinne, vgl. Bangert-Drowns, 1986, liberal gehandhabt wurde) oder nicht genügend bzw. überhaupt keine statistischen Angaben machten. Bei einer Studie schließlich (ChanLin, 2000) wurde festgestellt, dass sie dasselbe statistische Material einer anderen Studie nochmals aufarbeitete. Im Endeffekt konnten so 26 Studien in die Metaanalyse übernommen werden.

### **3.2.2 Kodierung**

Zwar untersuchten alle gefundenen Studien den Unterschied im Lernerfolg zwischen Standbildern und Animationen, die wenigsten jedoch beschränkten sich darauf, und viele setzten unterschiedliche Schwerpunkte. Teilweise war der hier interessierende Unterschied auch lediglich ein Nebenprodukt der Untersuchung. So war es geraten, eine Vielzahl von Faktoren zu kodieren, um mögliche moderierende Effekte identifizieren zu können. Einige Variablen mussten jedoch nach einer

---

<sup>2</sup> „Minimale Interaktivität“ ist hier nach Lawless & Brown (1997) so definiert, dass der Lernende einem vorgegebenen Informationspfad folgen muss und sich höchstens entscheiden kann, bestimmte Teile nochmals anzusehen oder vorwärts zu springen. Vollständige Autonomie z.B. durch Wählen aus einem Menü oder Einstellen verschiedener Parameter ist somit ausgeschlossen.

versuchsweisen Kodierung an einer kleinen Stichprobe wieder herausgenommen werden, da es sich als unmöglich herausstellte, sie reliabel zu kodieren.

So blieben letztendlich neben statistischen Angaben der jeweiligen Studien zwölf Eigenschaften der Studien, die in die Kodierung gingen.

Zwei Variablen bezogen sich auf die Art der Animation; zum einen wurde unterschieden, ob es sich um computer- oder videobasierte Animationen handelte, da es durchaus umstritten ist, ob beide Möglichkeiten gleich zu behandeln sind. Mit dieser Fragestellung verknüpft, jedoch weiter differenzierend ist die zweite Variable, die den Realitätsgrad der Animation auf vier Stufen (schematisch, eher einfach, eher realistisch, photo-realistisch) bewertete.

Zwei weitere Variablen beleuchteten die Eigenschaften der Standbildversionen; so könnte es einen Unterschied ausmachen, ob einerseits zusätzlicher Text vorhanden ist oder ob andererseits Bewegungsindikatoren („cues“) die Standbilder vereinfachen sollen.

Auch der Bereich der zu lernenden Aufgabe wurde auf mehreren Ebenen kodiert: Eine Variable bildete die zugrunde liegende Wissensdomäne (z.B. Biologie, Mathematik, Militär etc.) ab, eine weitere sollte verschlüsseln, ob es sich eher um einen zu lernenden Prozess oder eine Prozedur handelt, da eine Prozedur möglicherweise besser in animierter, weil genauerer Form gelernt werden kann. Eine andere Hypothese sollte mit der Kodierung überprüft werden, ob die Bewegung ein zentraler Prozess des zu lernenden Bereichs ist oder ob die Animation lediglich als Mittel zum Zweck benutzt wird.

Weil von Tversky et al. (2002) angemerkt wurde, dass die beiden Versionen sich oftmals in der aufgewendeten Zeit unterschieden und damit nicht vergleichbar seien, wurde, soweit angegeben, der jeweilige Zeitaufwand ebenfalls verzeichnet.

Schließlich wurden substanzielle Studienmerkmale wie das Jahr der Veröffentlichung sowie die Art und Größe der Stichproben aufgenommen.

### **3.2.3 Effektstärkenbestimmung und -analyse**

Als Maß der Effektstärke wurde Cohens  $d$  in der nach Hedges & Olkin (1985) modifizierten Form benutzt, in der die beiden Gruppenmittelwerte subtrahiert und durch eine gepoolte Streuung dividiert werden.

Wenn Mittelwerte und/oder Streuungen nicht angegeben waren, wurden Formeln zur Berechnung aus t- und F-Werten (Glass, McGaw & Smith, 1981; Rustenbach,

2003) oder zur Schätzung aus  $\chi^2$ -Werten (Cohen, 1966; Fricke & Treinies, 1985) verwendet.

Da die Untersuchung des Animationseffektes in vielen Fällen nicht das Hauptziel der Studien darstellte, wurden deren Ergebnisse auch oft nicht explizit über alle Versuchsbedingungen dargestellt, sondern in Form vieler einzelner Mittelwertsunterschiede der jeweiligen Bedingungen. Oftmals wäre es so inhaltlich nicht zu vertreten gewesen, lediglich eine Bedingung und deren zugrunde liegende Effektstärke auszuwählen. Deshalb wurde auf die Methode von Glass et al. (1981) zurückgegriffen, sämtliche Effektstärken in die Metaanalyse zu übernehmen und jede Versuchsbedingung als eigenständige Untersuchung zu behandeln. Dies führte im vorliegenden Fall dazu, dass 26 Untersuchungen 76 Effektstärken lieferten.

Da ein solches Vorgehen unter Umständen zu sehr niedrigen Stichprobengrößen pro Zelle und damit einer Verzerrung der Ergebnisse führen kann, war eine Korrektur der Effektstärken für kleine Stichproben von Nöten, wie von Hedges & Olkin (1985) empfohlen. Deren Formel korrigiert die Effektstärke durch  $d \approx (1-3/(4N-9))g$ , wobei  $g$  die unkorrigierte Effektstärke und  $N$  die Gesamtstichprobengröße darstellt.

Darüber hinaus sollte der Kritik an der Glass'schen Methode, dass auf diese Weise manche Studien (und damit mögliche Fehler und Mängel in der Validität) überrepräsentiert würden, dadurch Rechnung getragen werden, dass neben einer ungewichteten (wenngleich korrigierten) Analyse der Effektstärken, die mit konventionellen statistischen Mitteln wie der Varianzanalyse durchgeführt wurde, eine zusätzliche gewichtete Analyse erfolgen sollte.

Dieses in der metaanalytischen Praxis gebräuchliche Vorgehen gibt Studien mit größerer Stichprobengröße größeres Gewicht in der Auswertung, dabei voraussetzend, dass größere Studien bessere Schätzer der Gesamtpopulation darstellen (Bangert-Drowns, Hurley & Wilkinson, 2004). Auf diese Weise bekommt keine Studie nur aufgrund einer großen Anzahl von gelieferten Effektstärken ein unverhältnismäßig großes Gewicht, da jede einzelne Versuchsbedingung dementsprechend weniger Versuchspersonen enthält<sup>3</sup>. Erreicht wird dieser Effekt

---

<sup>3</sup> In manchen, in der Tabelle in Anhang A gesondert gekennzeichneten Fällen, war es nötig, darüber hinaus eine zusätzliche Adjustierung der Stichprobengröße je nach Menge der integrierten Einzelvergleiche aus dieser Studie vorzunehmen. Diese erfolgte, um einer möglichen Abhängigkeit verschiedener integrierter Variablen aus einer Studie mit denselben Probanden entgegenzuwirken und folgt einer Empfehlung von Hedges & Olkin (1985, S. 206), um so einen Ausschluss einer Anzahl

durch Gewichtung mit der Inversen des Standardschätzfehlers (Hedges & Olkin, 1985). Ein anschließender Test auf Homogenität mittels der Q-Statistik, die eine annähernde  $\chi^2$ -Verteilung mit  $k-1$  Freiheitsgraden ( $k$ = Anzahl der Effektstärken) aufweist und Teil des „kategorialen Modells fester Effekte“ (Rustenbach, 2003) der Metaanalyse ist, liefert Aufschluss darüber, ob die Varianz der Effektstärken auf mehr als nur den Standardschätzfehler zurückzuführen ist, also Heterogenität vorliegt (Hedges, 1994).

Ist das der Fall, werden im Anschluss die Einflüsse möglicher Moderatoren durch Varianzzerlegung untersucht, d.h. es werden die Varianzen zwischen einzelnen Kategorien ( $Q_{\text{between}}$ ) mit denen innerhalb der Kategorien ( $Q_{\text{within}}$ ) verglichen und auf Signifikanz geprüft. Konservative<sup>4</sup> Scheffé-korrigierte post-hoc-Kontrasttests werden benutzt, um bei mehr als zwei Subkategorien aufzudecken, welche von ihnen sich signifikant unterscheiden.

Gegebenenfalls werden in einem weiteren Schritt heterogene Unterkategorien weiter nach den Stufen anderer Unterkategorien unterteilt. Dieses Vorgehen, in etwa vergleichbar mit der Analyse von Interaktionseffekten bei traditioneller Varianzanalyse, wird beendet, wenn homogene Kategorien gefunden sind oder diese sehr klein werden (Hedges & Olkin, 1985).

Wie bereits erwähnt, wurde das kategoriale Modell fester Effekte für die Metaanalyse gewählt, das teilweise als weniger zuverlässig im Gegensatz zu einem Modell zufälliger Effekte diskutiert wird, da es in der Regel engere Vertrauensintervalle liefert (z.B. Hunter & Schmidt, 2000). Jedoch ist das kategoriale Modell nach wie vor die am häufigsten angewandte Methode (Schulze, 2004, S. 82), der zudem zugeschrieben wird, aus ihr qualitativ höherwertige Schlüsse ziehen zu können:

*The random effects models put fewer restrictions on the variation that may be observed in effect sizes or correlations. The price for the reduction in the number of restrictions, of course, is that the inferences made are weaker. In random effects models, studies do not*

---

von informationstragenden Einzelvergleichen zu umgehen. Hierzu wurde das jeweilige N des Einzelvergleichs durch die Anzahl der verschiedenen Einzelvergleiche mit denselben Probanden geteilt. Auf diese Weise ging kein Proband numerisch mehrfach in die Metaanalyse ein, selbst wenn er z.B. verschiedene abhängige Variablen bearbeitet hatte oder eine Kontrollgruppe mit verschiedenen Versuchsgruppen verglichen wurde.

<sup>4</sup> im Vergleich zur Korrektur nach Bonferroni

*replicate findings in any strong way. [...] Qualitatively stronger conclusions usually can be obtained by analysis under fixed effects models. (Hedges & Olkin, 1985, S. 248)*

### 3.3 Ergebnisse

#### 3.3.1 Studiencharakteristika

In die Metaanalyse gingen aus 26 verschiedenen Studien 76 unterscheidbare Vergleiche der Lernleistung bei Standbildern und Animationen ein. Diese wurden durch Effektstärken abgebildet.

Die Studien erstrecken sich über einen Zeitraum von 1973 bis 2003, allerdings stammen nur drei Untersuchungen aus den 1970er Jahren, in denen die Möglichkeiten zur Erstellung computerbasierter Animationen noch sehr stark eingeschränkt waren. Somit vergleichen denn auch nur sechs von 26 Studien Videoclips mit Standbildern, alle anderen benutzen computerbasierte Animationen.

Inhaltlich ist eine sehr breite Spannweite zu verzeichnen: Von der Gewitterentstehung über die Photosynthese und Programmieren mit BASIC bis hin zum Zusammenbau eines Maschinengewehrs oder dem Bandagieren einer Hand erstrecken sich die Themengebiete der einbezogenen Studien. Die Stichprobengrößen, in 17 von 26 Untersuchungen (65%) aus Studierenden, in acht Fällen aus Schülern (31%) rekrutiert, variierten in den Einzeluntersuchungen von N=21 bis N=263 mit einem Mittelwert von 55 und einem Median von 40.

In der Tabelle in Anhang A sind alle eingegangenen Studien mit ausgewählten Merkmalen verzeichnet.

#### 3.3.2 Overall-Effekte

Falls es keine Unterschiede machen sollte, ob mit Standbildern oder Animationen gelernt wird, sollten in etwa gleich viele Studien jeweils die Überlegenheit der einen oder anderen Bedingung aufzeigen. Tatsächlich gingen in die Metaanalyse 21 Vergleiche ein, in der Animationen signifikant überlegen waren, aber nur zwei, in denen signifikante Ergebnisse zugunsten der Standbilder gefunden wurden. Trotzdem muss dies nicht per se für die Überlegenheit von Animationen sprechen, da auch eine Verzerrung durch den „publication bias“ in Betracht zu ziehen ist. Dieser besagt, dass ein gefundenes Ergebnis eventuell aufgrund einer Verzerrung durch



vordringliche Veröffentlichung signifikanter Studien entstanden ist – während Studien mit nichtsignifikanten Ergebnissen häufig nicht veröffentlicht werden und somit auch nur schwerlich in eine Metaanalyse eingehen können.

Dass dies in vorliegender Analyse der Fall sein könnte, dem widersprechen zunächst per Augenschein die 53 Bedingungen, in denen kein signifikanter Unterschied festgestellt werden konnte, es konnten also auch diverse Studien mit nichtsignifikanten Ergebnissen publiziert werden. Weiterhin wurde das so genannte *fail-safe-N* berechnet (Rosenthal, 1979, adaptiert von Orwin, 1983). Dieser Parameter gibt an, wie viele Studien mit nichtsignifikanten Ergebnissen, die nicht in die Analyse eingegangen sind, es geben müsste, um den gefundenen Effekt zum Verschwinden zu bringen. Er beträgt in vorliegender Untersuchung 2964. Nach Rosenthals Faustformel (1991), nach der die Plausibilität einer Publikationsverzerrung dann ausreichend unwahrscheinlich ist, wenn der Wert des *fail-safe-N* das Fünffache der eingegangenen Studienzahl plus 10 übersteigt (hier: 390), erscheint ein *publication bias* ausreichend unplausibel.

Betrachtet man die jeweiligen Effektstärken, so kann festgestellt werden, dass 55 von 76 Bedingungen (72,4%) einen positiven Wert aufweisen und damit eine Überlegenheit der Animation indizieren.

Die mittlere ungewichtete, nach Hedges & Olkin (1985) korrigierte Effektstärke  $d$  beträgt 0,89 Standardabweichungen, mit einem 95%-Vertrauensintervall von 0,35 bis 1,43. Da die Spannweite der gefundenen Effektstärken mit von -2,34 bis 18,21<sup>5</sup> aber sehr groß ist (vgl. auch Abbildung 3-1), stellt der Median der Effektstärken eine weniger von Extremwerten beeinflusste Größe dar. Diese beträgt 0,63 Standardabweichungen, was nach der Definition von Cohen immer noch einem mittleren Effekt entspricht.

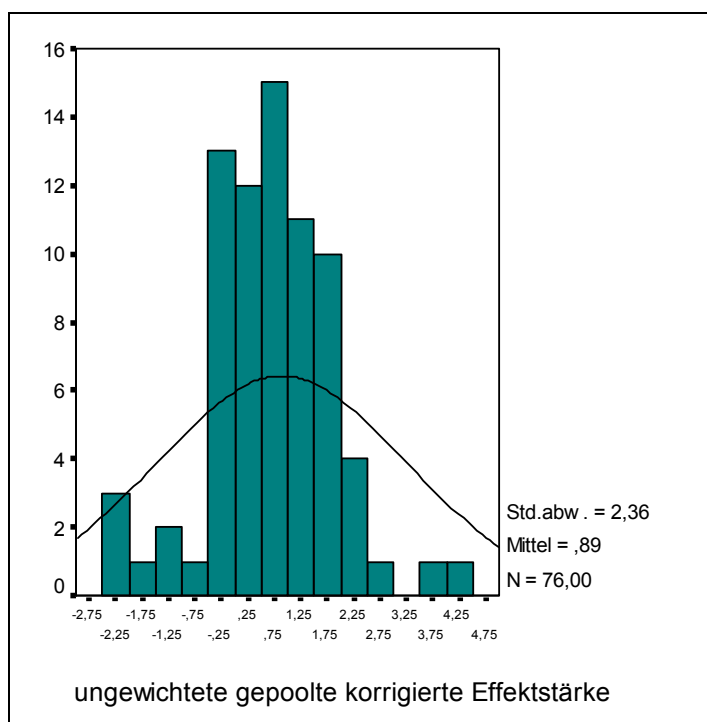
Gleiches gilt für die gewichtete Effektstärke mit einem  $d$  von 0,40 Standardabweichungen (Vertrauensintervall von 0,31 bis 0,48). Die gewichtete Effektstärke stellt normalerweise die verlässlichste und in Metaanalysen am häufigsten verwendete Maßzahl dar. Deshalb wird sich die Darstellung der weiteren

---

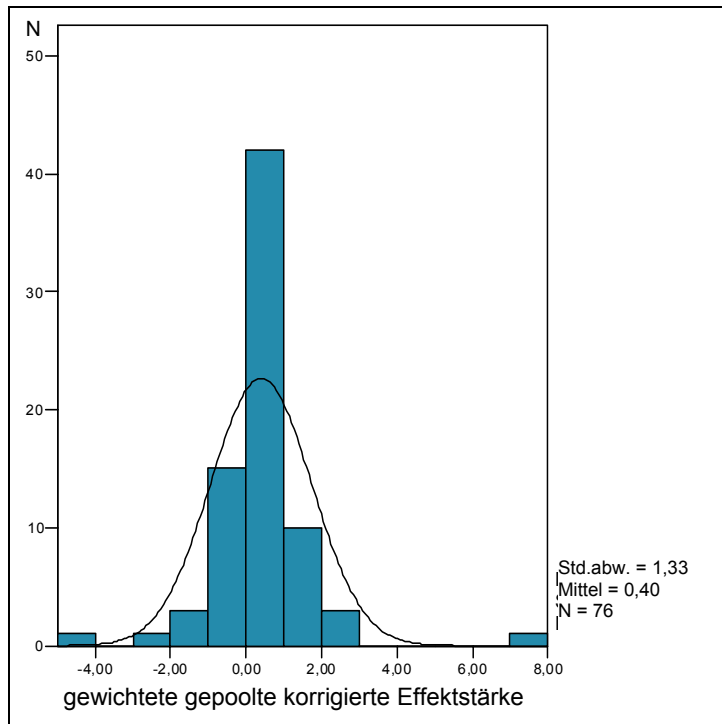
<sup>5</sup> Hier fällt besonders die als geradezu gigantisch groß zu bezeichnende Effektstärke von 18,21 auf. Diese ergibt sich aus einer Untersuchung von Rieber (1991). Im Allgemeinen wäre in einem solchen Fall ein Ausschluss dieses Extremwertes zu erwägen; da jedoch nach der Gewichtung (vgl. im Folgenden) dieser Wert auf moderate 0,54 fällt, konnte davon abgesehen werden.

Ergebnisse auch weitgehend auf die Analyse gewichteter Effektstärken konzentrieren.

Abbildung 3-1 zeigt die Verteilung der ungewichteten Effektstärken in einem Histogramm, während Abbildung 3-2 die Verteilung der gewichteten Effektstärken illustriert. Es zeigt sich jeweils, dass zwar die Mittelwerte der Verteilungen in positiver Richtung verschoben sind, die Ergebnisse aber eine große Streubreite zeigen, so dass es denkbar erscheint, dass mit diesen Verteilungen nicht eine einzelne Population repräsentiert wird, deren Werte nur aufgrund des Standardschätzfehlers um den Erwartungswert streuen, sondern dass die Varianz durch unterschiedliche Eigenschaften der jeweiligen Untersuchungen zustande kommt.



**Abbildung 3-1:** Histogramm der Verteilung der ungewichteten Effektstärke. Ein Outlier bei 18,21 ist nicht sichtbar.



**Abbildung 3-2:** Histogramm der Verteilung der gewichteten Effektstärke.

Ein Test auf Homogenität (Hedges & Olkin, 1985) liefert für diese Annahme jedenfalls Hinweise;  $Q_{\text{total}}=550,57$ ,  $df=75$ ,  $p<0,001$ , was eine deutliche Heterogenität der Verteilung indiziert. Dies macht eine genauere Untersuchung der Moderatoren notwendig.

### 3.3.3 Einfluss von Moderatoren

Eine Reihe möglicher Moderatorvariablen wurde untersucht. Hierzu zählten insbesondere die Art der Lernleistung, der Grad der Realistik der Animation, ob die Animation eher dekorative oder repräsentationale Funktion hatte, ob die Animation computer- oder videobasiert war, welcher inhaltliche Bereich zugrunde lag sowie ob die Visualisierungen mit Text und/oder Hinweisreizen (*cues*) versehen waren. Es ist anzumerken, dass daneben noch weitere Moderatoren avisiert waren, diese jedoch aufgrund mangelnder Dokumentation in den untersuchten Studien nicht in ausreichender Form durchgängig kodiert werden konnten. So war zum Beispiel in lediglich sieben Studien angegeben, wie viel Zeit für das Bearbeiten der einzelnen Lernumgebungen zugestanden wurde.

Tabelle 3-2 gibt die gewichteten Effektstärken, Vertrauensintervalle sowie die Stichprobengrößen für die einzelnen Bedingungen der Moderatorvariablen an. Zudem sind signifikante Unterschiede nach der von Hedges & Olkin (1985) empfohlenen Q-Statistik vermerkt ( $Q_{\text{between}}$ - und  $Q_{\text{within}}$ -Werte). Wenn eine Variable mehr als zwei Kategorien aufwies, wurden zusätzlich post-hoc -Kontraste berechnet, um die Interpretation signifikanter *between*-Effekte zu ermöglichen.

**Tabelle 3-2:** Gewichtete mittlere Effektstärken sowie die zugehörigen  $F$ -/ $t$ -Werte bzw.  $Q_{between}$ - und  $Q_{within}$ -Werte für die einzelnen Ausprägungen der potentiellen Moderatorvariablen.

Moderatorvariable	Anzahl der Effektstärken	Gewichtete mittlere Effektstärke $d$	95% Konfidenzintervall für $d$	$Q_{between}$	$Q_{within}$
<i>Inhaltsbereich</i>				23,53***	
Biologie	12	0,29	0,06 – 0,52		5,63
Physik	39	0,23	0,09 – 0,37		11,28
Chemie	7	0,77	0,57 – 0,97		25,14***
Mathematik	5	0,22	-0,02 – 0,46		3,70
Militaria	3	0,88	0,50 – 1,26		9,48**
andere	10	0,45	0,20 – 0,70		14,19
<i>Art der Lernleistung</i>				18,08***	
prozedural-motorisch	5	1,06	0,71 – 1,41		22,07***
deklarativ	40	0,43	0,31 – 0,55		44,27
prozedural-problemlösend	31	0,27	0,14 – 0,40		15,21
<i>Art der Animation</i>				7,80**	
computerbasiert	64	0,33	0,23 – 0,43		42,23
videobasiert	12	0,67	0,48 – 0,86		43,44***
<i>Realismusgrad der Animation</i>				12,89**	
schematisch	24	0,31	0,13 – 0,49		10,96
eher einfach	11	0,39	0,24 – 0,54		19,61*
eher realistisch	22	0,19	0,01 – 0,37		3,58
photorealistisch	12	0,67	0,48 – 0,86		43,44***
<i>Rolle der Animation</i>				30,90***	
Dekorative Animation	21	0,11	-0,02 – 0,24		3,71
Repräsentationale Animation	55	0,60	0,49 – 0,71		99,70***
<i>Erläuternder Text</i>				5,63*	
Enthalten	59	0,35	0,26 – 0,44		49,12
Nicht enthalten	17	0,62	0,42 – 0,82		33,69*
<i>Cues in Standbildern</i>				1,78	
Enthalten	52	0,34	0,23 – 0,45		34,92
Nicht enthalten	13	0,47	0,29 – 0,65		23,49*

**Anmerkung:**  $Q_{between}$  = Unterschiedseffekt zwischen den Gruppen;  $Q_{within}$  = Heterogenität innerhalb der Gruppe

\* =  $p < 0,05$ ; \*\* =  $p < 0,01$ ; \*\*\* =  $p < 0,001$

Wie sich zeigt, ergab die gewichtete Analyse eine Reihe signifikanter Effekte, die im Folgenden näher beleuchtet werden sollen.

### *Inhaltsbereich*

Diese potentielle Moderatorvariable weist eine eher grobe Unterteilung nach den inhaltlichen Bereichen auf, die für die Lernumgebungen gewählt wurden (Abbildung 3-3-B). Hierbei sind eindeutige Zuweisungen teilweise leider schwierig zu treffen, was möglicherweise der Grund für – in zwei von sechs Fällen – hochsignifikant heterogene Kategorien ist. Daneben ist jedoch auch ein  $Q_{\text{between}}$ -Effekt, also ein signifikanter Unterschied zwischen den Subkategorien, zu beobachten, der mit einem Wert von  $Q_{\text{between}}(5) = 23,53$  den kritischen Wert eines 0,1%-Signifikanzniveaus übertrifft. Spitzenreiter hinsichtlich der gewichteten mittleren Effektstärke sind die Bereiche der Chemie ( $d=0,77$ ) und des Militärs ( $d=0,88$ ), wobei beide aber nur auf sehr geringen Gruppengrößen beruhen ( $N=7$  bzw.  $N=3$ ). Die Bereiche der Physik und der Biologie sind beide durch deutlich größere Stichproben repräsentiert, unterscheiden sich mit  $d=0,23$  und  $d=0,29$  jedoch nicht signifikant voneinander.

### *Art der Lernleistung*

Es unterschied sich von Untersuchung zu Untersuchung, welcher Aspekt des Lernens im Fokus stand (Abbildung 3-3-G): In den meisten Fällen ( $N=40$ ) ging es um deklaratives Wissen, das mit Hilfe der jeweiligen Lernumgebungen erworben werden sollte. Häufig ( $N=31$ ) interessierte jedoch auch die Fähigkeit zu einem tieferen Verständnis, das z.B. durch die Anwendung erlernter Regeln ermöglicht, Probleme zu lösen. Diese Fähigkeit wird hier im Folgenden „prozedural-problemlösend“ genannt, um sie von einer dritten Lernfacette abzuheben, dem „prozedural-motorischen“ Lernen. Diese Unterscheidung wird deshalb getroffen, weil in der Literatur der Begriff des prozeduralen Lernens, der normalerweise im Gegensatz zum deklarativen Lernen steht, unklar ist und teilweise als das Erlernen einer bestimmten motorischen Tätigkeit, teilweise als Fähigkeit zum Problemlösen definiert wird.

„Prozedural-motorisches“ Lernen, in vorliegender Metaanalyse in fünf Fällen getestet, kann z.B. die erworbene Fertigkeit sein, ein Maschinengewehr zusammenzubauen (Spangenberg, 1973).

Dieses zeigte dann auch die mit Abstand höchste mittlere gewichtete Effektstärke mit  $d=1,06$  (95% Vertrauensintervall von 0,71 bis 1,41). Wenn deklaratives Wissen abgefragt wurde, ergab sich eine mittlere Effektstärke von  $d=0,43$  (95% VI: 0,31 – 0,55), bei prozedural-problemlösenden Wissen war es lediglich  $d=0,27$  (95% VI: 0,14 – 0,40).

Dieser Unterschied wurde bei der gewichteten varianzanalytischen Auswertung deutlich (Tabelle 3-2):  $Q_{\text{between}}(2)=18,08$ ,  $p<0,001$ . Der Kontrast zwischen prozedural-motorischem und deklarativem Wissen ist ebenso statistisch signifikant ( $Q_{\text{contrast}}(1)=11,42$ ,  $p<0,001$ ) wie der Einzelvergleich zwischen prozedural-motorischem und prozedural-problemlösendem Wissen ( $Q_{\text{contrast}}(1)=17,57$ ,  $p<0,001$ ), nicht jedoch derjenige zwischen deklarativem und problemlösendem Wissen. Zudem weist einer der drei  $Q_{\text{within}}$ -Werte hochsignifikante Abweichungen auf, die eine Heterogenität dieser Subkategorie indiziert (Tabelle 3-2). Dieses Problem wird in der späteren Auswertung noch aufgegriffen werden.

### *Probanden*

Die Variable, die die Art der Probanden (Schüler, Studierende, Rekruten, Erwachsene) unterschied, zeigte keinerlei signifikante Unterschiede ( $Q_{\text{between}}(3)=6,34$ ,  $p=0,10$ ) zwischen den Kategorien.

### *Art der Animation*

Ob die Animation computer- oder videobasiert ist, spielte hinsichtlich der mittleren Effektstärke eine deutliche Rolle: So war die gewichtete mittlere Effektstärke von Untersuchungen, in denen videobasierte Animationen mit statischen Bildern verglichen wurden, mit  $d=0,67$  (95% VI 0,48 bis 0,86;  $N=12$ ) deutlich größer als die von computerbasierten Animationen ( $d=0,33$  (95% VI 0,23 bis 0,43,  $N=64$ ), vgl. Abbildung 3-3-E. Dementsprechend wurde auch der Vergleich zwischen den Kategorien signifikant:  $Q_{\text{between}}(1)=7,80$ ,  $p<0,01$ .

Allerdings gilt dies auch für den Heterogenitätsindikator  $Q_{\text{within}}$  der Kategorie der videobasierten Animationen ( $p < 0,001$ ), was die Effektstärkendifferenz weniger deutlich auf formale Unterschiede zwischen den beiden Kategorien zurückführen lässt (vgl. Tabelle 3-2).

Auch diese Kategorie wird in der vertiefenden Analyse (siehe unten) aufgrund des Heterogenitätsproblems genauer untersucht werden.

### *Grad der Realistik der Animation*

Der Grad der Realistik der Animation wurde auf vier Stufen beurteilt, von einfachen, schematischen Darstellungen bis hin zu photorealistischen Videos (Abbildung 3-3-C). Es zeigten sich deutliche Unterschiede zwischen den Subkategorien hinsichtlich der gewichteten mittleren Effektstärke ( $Q_{\text{between}}(3) = 13,83$ ,  $p < 0,01$ ).

Nicht überraschend, da mit der Kategorie der videobasierten Animationen (siehe oben) übereinstimmend, wies die höchste Realistikstufe ein  $d$  von 0,67 auf. Diese photorealistischen Videos waren computerbasierten Animationen der Stufe 3 („eher realistisch“) im paarweisen Vergleich signifikant überlegen ( $Q_{\text{contrast}}(1) = 24,92$ ,  $p < 0,001$ ). Innerhalb der computerbasierten Lernumgebungen schienen eher solche Animationen für höhere Effektstärken ( $d = 0,31$  bzw.  $0,39$  vs.  $0,19$ ) zu stehen, die weniger realistisch und damit weniger detailgetreu waren. Jedoch konnten die paarweisen Kontraste keine signifikanten Unterschiede nachweisen (z.B. Stufe 2 vs. Stufe 3:  $Q_{\text{contrast}}(1) = 2,74$ ,  $p = 0,10$ ). Die  $Q_{\text{within}}$ -Werte erreichten in zwei von vier Fällen signifikante Höhen (vgl. Tabelle 3-2).

### *Rolle der Animation*

Es stellte sich heraus, dass es eine wichtige Rolle spielt, ob die Bewegung, die von der Animation dargestellt wird, ein zentraler Lerngegenstand ist, also *repräsentational* ist, oder lediglich *dekorativ* zur Veranschaulichung einer anderen Thematik verwandt wird (Bezeichnungen adaptiert nach Carney & Levin, 2002). Im repräsentationalen Falle sind Animationen den Standbildern wesentlich deutlicher überlegen ( $d = 0,60$ ; 95% VI 0,49 bis 0,71) als bei dekorativen Animationen ( $d = 0,11$ ; 95% VI -0,02 bis 0,24), bei denen aufgrund des großen Vertrauensintervalls, das Null



einschließt, nicht mehr von einer Überlegenheit der Animationen zu sprechen ist (siehe auch Tabelle 3-2 sowie Abbildung 3-3-F).

Der Unterschied zwischen den beiden Kategorien wird auch statistisch signifikant:  $Q_{\text{between}}(1)=30,90$ ,  $p<0,001$ . Dieser Moderator wird auch in der vertiefenden Analyse noch eine Rolle spielen.

### *Erklärender Text*

Nicht grundsätzlich waren die Visualisierungen mit zusätzlichem erläuternden Text versehen; in N=20 von 76 Fällen war dies nicht der Fall (Abbildung 3-3-H). Es zeigt sich, dass in diesem Fall die Standbilder den Animationen wesentlich deutlicher unterlegen sind ( $Q_{\text{between}}(1)=5,63$ ,  $p<0,05$ ). Es wird eine mittlere gewichtete Effektstärke von  $d=0,62$  (95% VI 0,42 bis 0,82) erreicht – gegenüber lediglich  $d=0,35$  bei enthaltenem Text (95% VI 0,26 bis 0,44). Da allerdings auch in diesem Fall eine der Kategorien signifikant heterogen ist (vgl. Tabelle 3-2), wird noch eine vertiefende Analyse vorzunehmen sein (siehe unten).

### *Cues in den Standbildern*

Auch „Cues“, besondere Hinweisreize wie Pfeile, Hervorhebungen etc., waren nicht in jeder Lernumgebung mit Standbildern enthalten. Entgegen der Erwartungen macht dies keinerlei Unterschiede hinsichtlich der gewichteten mittleren Effektstärke: Waren Cues enthalten, ergibt sich eine Effektstärke von  $d=0,34$ , war dies nicht der Fall, ist die Effektstärke nur unwesentlich höher mit  $d=0,47$  (Abbildung 3-3-D). Dieser Abstand ist nicht statistisch signifikant,  $Q_{\text{between}}(1)=1,78$ ,  $p=0,18$ , beide Kategorien weisen heterogene Effektstärken auf (Tabelle 3-2).

### **3.3.4 Vertiefende Analyse zweiter und dritter Dimension**

Zwar konnten einige signifikante Unterschiede zwischen verschiedenen Unterkategorien potentieller Moderatorvariablen ausgemacht werden, es bestand jedoch, wie erwähnt, häufiger das Problem, dass auch *innerhalb* der Unterkategorien signifikante Varianzen festgestellt wurden, die Kategorien also heterogen waren.

Tritt ein solcher Fall ein, so empfehlen Hedges & Olkin (1985), die betroffenen Kategorien zu identifizieren und nach weiteren Dimensionen aufzuschlüsseln, um so unter Umständen homogene Kategorien zu erzeugen, die sich im Idealfall signifikant voneinander unterscheiden.

*The classes obtained by one partitioning may be partitioned again according to a second classification dimension, and in turn partitioned according to other dimensions to yield finer and finer groupings. [...] Proceed [...] until an acceptable level of within-class fit is obtained or the classification dimensions are exhausted. (Hedges & Olkin, 1985, S. 157/58)*

Dieses Verfahren wurde im vorliegenden Fall angewandt, indem eine Vielzahl unterschiedlicher Kombinationen (Paare und Tripel) der verschiedenen bereits kodierten Variablen gebildet wurde und so die Unterkategorien immer feiner wurden. Oftmals wurden dabei die Subkategorien jedoch so klein, dass sie statistisch nicht mehr sinnvoll auszuwerten waren. Effektstärken von Kategorien mit  $N < 2$  werden deshalb in der weitergehenden Analyse nicht beachtet.

Da nun grundsätzlich mehr als zwei Bedingungen miteinander verglichen wurden, wurden an einigen Stellen zusätzliche Kontrasttests zum Einzelvergleich eingesetzt, um signifikante Effekte zwischen den Klassen sinnvoll interpretieren zu können. Dieses Verfahren ist nach Hedges & Olkin (1985) bei Signifikanz der Quadratsumme zwischen den Kategorien ( $Q_{\text{between}}$ ) gerechtfertigt.

Im Folgenden sollen nun zwei Kombinationen zweiter Dimension, d.h. Effektstärkenvergleiche, in denen zwei Moderatorvariablen miteinander gekreuzt wurden, sowie zwei Kombinationen dritter Dimension, bei denen noch eine weitere Variable hinzukam, dargestellt werden. Diese ausgewählten Kombinationen (siehe auch Abbildung 3-3) sind diejenigen, bei denen sich über die bisher aufgeführten Ergebnisse hinaus aus inhaltlicher und statistischer Sicht besonders aussagekräftige Resultate ergeben haben. Alle anderen Kombinationen wurden zwar berechnet, werden hier aber nicht weiter aufgeführt.

Erste Klassifizierungsdimension

Zweite Klassifizierungsdimension

Dritte Klassifizierungsdimension

**(A) Overall-Effekte:**

	Animationen
d	.40
N	76

$Q_{total}(75) = 550,57, p < .0001$

**(B) Inhaltsbereich:**

	Bio- logie	Physik	Chemie
d	.29	.23	.77*
N	12	39	7

	Mathe- matik	Militär	andere
d	.22	.88*	.45
N	5	3	10

$Q_{between}(5) = 23,53, p < .001$

**(C) Grad der Realistik:**

	1	2	3	4
d	.31	.39*	.19	.67*
N	24	11	22	12

$Q_{between}(3) = 13,83, p < .01$

**(D) Cues in Standbildern:**

	Ja	Nein
d	.34	.47*
N	52	13

$Q_{between}(1) = 1,78, p = .18$

**(E) Art der Animation:**

	Video- basiert	Computer- basiert
d	.67*	.33
N	12	64

$Q_{between}(1) = 7,80, p < .01$

**(F) Rolle der Animation:**

	Dekorativ	Reprä- sentativ
d	.11	.60*
N	21	55

$Q_{between}(1) = 30,90, p < .001$

**(G) Art der Lernleistung:**

	Proze- dural- motorisch	Dekla- rativ	Problem- lösend
d	1.06*	.43	.27
N	5	40	31

$Q_{between}(2) = 18,08, p < .001$

**(H) Erklärender Text:**

	Ja	Nein
d	.35	.62*
N	59	17

$Q_{between}(1) = 5,63, p < .05$

**(J) Art der Animation & Rolle der Animation:**

	Video- basiert	Computer- basiert
Deko- rativ	d .00	.12
N	1	20
Reprä- sentativ	d .93*	.49
N	11	44

$Q_{between}(2) = 38,28, p < .001$

**(K) Art der Lernleistung & Rolle der Animation:**

	Proze- dural- motorisch	Dekla- rativ	Problem- lösend
Deko- rativ	d -	.05	.19
N	0	14	7
Reprä- sentativ	d 1.06*	.75*	.32
N	5	26	24

$Q_{between}(4) = 52,40, p < .001$

**(L) Art der Animation & Rolle der Animation & Text:**

	Video- basiert	Computer- basiert	
Kein Text	Dekorativ	d .00	--
	N	1	0
Text	Dekorativ	d --	.12
	N	0	20
Kein Text	Reprä- sentativ	d .95*	.78
	N	8	8
Text	Reprä- sentativ	d .88*	.47
	N	3	36

$Q_{between}(4) = 39,16, p < .001$

**(M) Art der Lernleistung & Rolle der Animation & Text:**

	Proze- dural- motorisch	Dekla- rativ	Problem- lösend	
Kein Text	Dekorativ	d --	.00	--
	N	0	1	0
Text	Dekorativ	d --	.06	.19
	N	0	13	7
Kein Text	Reprä- sentativ	d 2.02*	1.48*	.39
	N	2	8	6
Text	Reprä- sentativ	d .88*	.61	.29
	N	3	18	18

$Q_{between}(7) = 68,30, p < .001$

Abbildung 3-3: Darstellung der Moderatoranalyse erster, zweiter und dritter Dimension. Heterogene Effektstärken sind mit einem Sternchen markiert. In die Analyse aufgenommene Effektstärken sind fett gedruckt.

### *Art der Animation sowie Rolle der Animation*

In den Einzelanalysen wurde bereits festgestellt, dass zum einen Untersuchungen mit repräsentationalen Animationen höhere mittlere gewichtete Effektstärken aufweisen als solche mit dekorativen, zum anderen videobasierte Animationen computerbasierten überlegen sind. In der Kombination beider Variablen, also einer 2x2-Analyse, verstärkt sich dieser Trend: Die Unterschiede zwischen den Subkategorien werden signifikant:  $Q_{\text{between}}(2)=38,28$ ,  $p<0,001$  (vgl. Abbildung 3-3-J). Die Extremkategorie – aus repräsentationalen, videobasierten Animationen bestehend – weist nun eine nochmals erhöhte mittlere Effektstärke von  $d=0,93$  auf, ist allerdings nach wie vor nicht homogen ( $Q_{\text{within}}(10)=57,74$ ,  $p<0,001$ ). Dekorative videobasierte Animationen kommen in dieser Untersuchung hingegen nur einmal vor, diese Zelle ist daher statistisch nicht sinnvoll auswertbar.

Der Unterschied zwischen videobasierten und computerbasierten Animationen ist aber mit der Aufspaltung in Studien mit repräsentationalen und dekorativen Animationen noch deutlicher geworden: Repräsentationale *computerbasierte* Animationen erreichen lediglich eine mittlere gewichtete Effektstärke von  $d=0,49$  (und stellen mit  $Q_{\text{within}}(43)=45,25$ ,  $p=0,38$  eine homogene Subkategorie dar). Der Unterschied zur Kategorie der Untersuchungen mit *videobasierten* repräsentationalen Animationen ( $d=0,93$ ) ist bei Berechnung eines a-priori-Einzelvergleichs damit signifikant ( $Q_{\text{contrast}}(1)=11,25$ ,  $p<0,001$ ).

Damit liegen die Untersuchungen mit repräsentationalen computerbasierten Animationen aber noch vor der Kategorie der Untersuchungen mit *dekorativen* computerbasierten Untersuchungen, bei denen eine mittlere Effektstärke von  $d=0,12$  ( $Q_{\text{within}}(19)=4,13$ ,  $p=0,99$ ) erreicht wurde. Der Unterschied zwischen diesen beiden Kategorien ist ebenfalls signifikant ( $Q_{\text{contrast}}(1)=14,36$ ,  $p<0,001$ ), womit nun das Ergebnis der Analyse erster Dimension des Vergleichs zwischen dekorativen und repräsentationalen Animationen (vgl. oben) mit homogenen Subkategorien zufrieden stellend repliziert wurde.

### *Art der Lernleistung sowie Rolle der Animation*

Werden die Variablen „Art der Lernleistung“ (prozedural-motorisch, deklarativ und prozedural-problemlösend) und „Rolle der Animation“ (dekorative Animation vs.

repräsentationale Animation) kombiniert (Abbildung 3-3-K), verstärken sich die sich bereits in den Einzelanalysen abzeichnenden Trends weiter – repräsentationale Animationen sind dekorativen überlegen, ebenso wie Animationen bei prozedural-motorischen Aufgaben mehr Lernerfolg erbringen als bei deklarativen oder prozedural-problemlösenden.

So weisen Bedingungen, die *repräsentationale* Animationen mit deklarativer Wissensabfrage kombinieren, mit  $d=0,75$  ( $Q_{\text{within}}(25)=57,10$ ,  $p<0,001$ ) eine deutlich höhere mittlere gewichtete Effektstärke auf als *dekorative* Animationen mit deklarativer Wissensabfrage ( $d=0,05$ ;  $Q_{\text{within}}(13)=2,60$ ,  $p=0,99$ ). Dieser Unterschied wird im Einzelvergleich signifikant:  $Q_{\text{contrast}}(1)=33,30$ ,  $p<0,001$ . Neu ist, dass erstgenannte Kategorie (repräsentationale Animationen mit *deklarativer* Wissensabfrage) nun auch der Kategorie repräsentationaler Animationen mit *prozedural-problemlösenden* Aufgaben ( $d=0,32$ ;  $Q_{\text{within}}(23)=12,06$ ,  $p=0,97$ ) statistisch signifikant überlegen ist ( $Q_{\text{contrast}}(1)=13,01$ ,  $p<0,001$ ). Dekorative Animationen mit prozedural-problemlösenden Anforderungen schließlich weisen eine relativ niedrige mittlere Effektstärke von  $d=0,19$  ( $Q_{\text{within}}(6)=3,62$ ,  $p=0,73$ ) auf, unterscheiden sich aber nicht signifikant von einer der angrenzenden Subkategorien (vgl. Abbildung 3-3-K).

Da keine Lernumgebung, die *prozedural-motorisches* Wissen abfragt, die Animation in dekorativer Funktion verwendet, bringt auf dieser Dimension die weitere Aufschlüsselung keinen weiteren Erkenntnisgewinn. Der Einzelvergleich zwischen repräsentationalen Animationen, die entweder *prozedural-motorisches* ( $d=1,06$ ,  $Q_{\text{within}}(4)=22,07$ ,  $p<0,001$ ) oder *deklaratives* Wissen ( $d=0,75$ , siehe oben) abfragen, zeigt nun keine statistisch signifikante Differenz mehr:  $Q_{\text{contrast}}(1)=14,22$ ,  $p<0,001$ .

Der Vollständigkeit halber sei noch erwähnt, dass sich die einzelnen kombinierten Kategorien, wie die gewichtete varianzanalytische Auswertung ergibt, voneinander hochsignifikant unterscheiden ( $Q_{\text{between}}(4)=52,40$ ,  $p<0,001$ ).

*Fokussierung auf Standbilder als Kontrollbedingung: Art der Animation, Rolle der Animation sowie erklärender Text:*

Führt man zusätzlich die Variable ein, die codiert, ob die mit den Animationen verglichenen Standbilder zusätzlichen erläuternden Text enthalten, erhält man mittels einer dreidimensionalen Analyse mit den Variablen „Art der Animation“ und „Rolle

der Animation“ weiteren Einblick in die Unterschiede zwischen statischen und dynamischen Visualisierungen aus einer anderen Perspektive.

Wie Abbildung 3-3-L zu entnehmen ist, führt dies zu fünf Subkategorien, die sich signifikant voneinander unterscheiden ( $Q_{\text{between}}(4)=39,16$ ,  $p<0,001$ ). Es zeigt sich, dass ein Großteil der Untersuchungen mit repräsentationalen *videobasierten* Visualisierungen *ohne* zusätzlichen Text auskamen und so auf eine mittlere gewichtete Effektstärke von 0,95 ( $Q_{\text{within}}(7)=39,17$ ,  $p<0,001$ ) kommen. Wird diese Subkategorie mit repräsentationalen *computerbasierten* Visualisierungen ohne zusätzlichen Text verglichen ( $d=0,78$ ,  $Q_{\text{within}}(7)=6,74$ ,  $p=0,46$ ), was die fairste und akkurateste Vergleichskategorie bildet, zeigt sich nicht länger ein signifikanter Unterschied:  $Q_{\text{contrast}}(1)=0,31$ ,  $p=0,58$ .

Ebenso unterscheiden sich repräsentationale *computerbasierte* Animationen *mit* zusätzlichem Text (mittlere gewichtete Effektstärke:  $d=0,47$ ;  $Q_{\text{within}}(35)=37,83$ ,  $p=0,34$ ) nicht signifikant von repräsentationalen *videobasierten* Animationen *mit* zusätzlichem Text ( $d=0,88$ ;  $Q_{\text{within}}(2)=9,46$ ,  $p=0,25$ ). Wohl aber ist eine Signifikanz ( $Q_{\text{contrast}}(1)=12,51$ ,  $p<0,001$ ) zu verzeichnen zwischen *repräsentationalen* und *dekorativen* computerbasierten Animationen mit zusätzlichem Text, wobei letztere mit einer mittleren Effektstärke von 0,12 ( $Q_{\text{within}}(19)=4,13$ ,  $p=0,99$ ) den niedrigsten Wert aufweist.

*Fokussierung auf Standbilder als Kontrollbedingung: Art der Lernleistung, Rolle der Animation sowie erklärender Text:*

Auch die Variablen „Art der Lernleistung“ und „Rolle der Animation“ können sinnvoll mit der Variable „Text in den Standbildern“ kombiniert werden, um Schlüsse hinsichtlich der Ursprünge der Effektstärkenunterschiede zu ziehen (Abbildung 3-3-M).

Richtet man das Augenmerk auf den Vergleich der verschiedenen Dimensionen der Art der Lernleistung (prozedural-motorisch, deklarativ, prozedural-problemlösend, vgl. Abbildung 3-3-M), so können die in der Analyse zweiter Dimension noch beobachteten Unterschiede (s.o.) nur noch teilweise festgestellt werden: Lediglich repräsentationale Visualisierungen ohne zusätzlichen Text, die *deklaratives* Wissen erfragen, weisen noch eine statistisch signifikant höhere mittlere Effektstärke ( $d=1,48$ ;  $Q_{\text{within}}(7)=39,96$ ,  $p<0,001$ ) auf als ihr Gegenpart bei *prozedural-*

*problemlösendem* Wissen ( $d=0,39$ ;  $Q_{\text{within}}(5)=5,51$ ,  $p=0,36$ ):  $Q_{\text{contrast}}(1)=16,70$ ,  $p<0,001$ . Dabei ist jedoch zu beachten, dass eine der Subkategorien nach wie vor heterogen ist. Ebenso, jedoch mit homogenen Subkategorien, sind repräsentationale Visualisierungen *mit* zusätzlichem Text, die *deklaratives* Wissen erfragen ( $d=0,61$ ;  $Q_{\text{within}}(17)=25,72$ ,  $p=0,08$ ), solchen überlegen, die *prozedural-problemlösendes* Wissen erfordern ( $d=0,29$ ;  $Q_{\text{within}}(17)=6,68$ ,  $p=0,99$ ):  $Q_{\text{contrast}}(1)=5,65$ ,  $p<0,05$ .

Vergleichbare Relationen zwischen deklarativem und prozedural-problemlösendem Wissen (vgl. Abbildung 3-3-M) können ansonsten nicht mehr festgestellt werden. Für *dekorative* Visualisierungen mit zusätzlichem Text ( $d=0,06$ ;  $Q_{\text{within}}(12)=4,22$ ,  $p=0,98$  vs.  $d=0,19$ ;  $Q_{\text{within}}(6)=3,62$ ,  $p=0,73$ ):  $Q_{\text{contrast}}(1)=0,34$ ,  $p=0,56$ ) gilt dieser Unterschied nicht. Jedoch sind alle Subkategorien homogen.

Weiterhin unterscheiden sich repräsentationale Animationen mit *prozedural-motorischen* Aufgaben (mit Text:  $d=0,88$ ; ohne Text:  $d=2,02$ ) von keiner vergleichbaren deklarativen oder prozedural-problemlösenden Kategorie. Zudem sind beide Kategorien heterogen ( $Q_{\text{within}}(2)=9,46$ ,  $p<0,01$  bzw.  $Q_{\text{within}}(1)=5,05$ ,  $p<0,05$ ).

Fokussierend auf Unterschiede zwischen Visualisierungen mit oder ohne zusätzlichem erläuternden Text kann auf der Ebene deklarativen Wissens ein signifikanter Unterschied festgestellt werden: Repräsentationale Visualisierungen *ohne* Text sind mit einer mittleren Effektstärke von 1,48 statistisch repräsentationalen Visualisierungen *mit* Text ( $d=0,61$ ) überlegen;  $Q_{\text{contrast}}(1)=14,86$ ,  $p<0,001$ .

Auf der Ebene *prozedural-problemlösender* sowie *prozedural-motorischer* Aufgaben lassen sich hingegen keine vergleichbaren signifikanten Unterschiede finden, auch wenn dies von der Höhe der Werte her hätte erwartet werden können ( $d=2,02$  vs.  $0,88$ ;  $d=0,39$  vs.  $0,29$ ).

Der Gesamtunterschied zwischen allen Subkategorien bleibt jedoch weiterhin signifikant ( $Q_{\text{between}}(7)=68,30$ ,  $p<0,001$ ).

### 3.4 Diskussion

Die entscheidenden Befunde, die die Metaanalyse hinsichtlich des Unterschiedes des Lerneffekts von Animationen und Standbildern liefert, lassen sich unter fünf Punkte subsumieren:

### 3.4.1 Overall-Effekte von Animationen gegenüber Standbildern

Ungeachtet der verschiedenen Moderatoren ergab sich, über alle Einzelvergleiche betrachtet, ein deutlicher Vorteil von Animationen gegenüber Standbildern: Während eine negative mittlere Effektstärke die Überlegenheit von Standbildern gezeigt hätte, indiziert die gefundene gewichtete mittlere Effektstärke von 0,40 (95% VI 0,31 bis 0,48) einen Effekt zugunsten der Animationen. Nach Cohen (1977), der Effektgrößen von 0,20, 0,50 und 0,80 als kleine, mittlere und große Effekte bezeichnet, kann hier also zumindest von einem kleinen bis mittleren Effekt gesprochen werden. Nach der Definition von Tallmadge (1977) können selbst Effektstärken von 0,25 bis 0,33 als „educational significant“ betrachtet werden, womit im vorliegenden Fall sogar die untere Grenze des 95% Vertrauensintervalls (0,31) noch einen im pädagogischen Umfeld praktisch bedeutsamen Wert darstellt. Betrachtete man die ungewichtete Effektstärke, die aus oben genannten Gründen aber die Gefahr der Verzerrung in sich trägt, könnte man bei der gefundenen mittleren Effektstärke von 0,89 Standardabweichungen sogar von einem großen Effekt sprechen.

Nachdem das Potential von Animationen noch vor einigen Jahren als sehr groß eingeschätzt wurde, geht der gegenwärtige *common sense* unter kognitionspsychologisch orientierten Forschern eher dahin, dass zumindest nichtinteraktive Animationen hinsichtlich der Lernsteigerung oft mehr Nach- als Vorteile bringen und Standbildern, außer unter bestimmten Bedingungen, tendenziell nicht überlegen sind („Computer animation is not a panacea in itself“; Bétrancourt & Tversky, 2000).

Damit stehen die Befunde zunächst in einem deutlichen Widerspruch zu der aktuell gängigen Meinung, denn insgesamt gesehen zeigt sich ein deutlicher Effekt zugunsten der Animationen. Da ein *publication bias* durch Ermittlung des *fail-safe-N* weitgehend ausgeschlossen werden kann und zudem mit gewichteten Effektstärken operiert wurde, müssen andere als methodische Gründe für diesen Unterschied gesucht werden.

So erscheint es besonders im Licht der nachfolgenden Moderatorenanalyse wahrscheinlich, dass Animationen nur unter bestimmten Bedingungen ihre Vorteile gegenüber Standbildern entfalten können – die dann unter idealen Voraussetzungen besonders stark zum Tragen kommen und in sehr hohen Effektstärken resultieren. Diese Folgerung geht aus der Tatsache hervor, dass eine Vielzahl der Primärstudien



keine signifikante Überlegenheit der Animation belegen konnte und trotzdem eine mittlere Effektstärke von 0,40 resultierte. Im Einklang mit einer Reihe von Autoren, die in Reviews Voraussetzungen formulierten, unter denen Animationen besonders wirksam sein sollten (u.a. Mayer & Moreno, 2002; Milheim, 1993; Tversky et al., 2002; Weiss et al., 2002), zielte auch die im Folgenden diskutierte Moderatorenanalyse auf die Identifikation solcher Bedingungen ab.

### 3.4.2 Repräsentational oder dekorativ?

Eine Moderatorvariable, auf die die Metaanalyse Hinweise liefert, ist die der Frage, ob die Bewegung zentrales Element des Lerngegenstandes ist. Sehr deutlich zeigt sich, dass repräsentationale Animationen Standbildern in größerem Maße überlegen sind, als dass dies dekorative Animationen sind (mittlere Effektstärke von 0,60 gegenüber 0,11). Die Anwendung von Animationen scheint also besonders dann sinnvoll zu sein, wenn die Bewegung, die dargestellt wird, auch tatsächlich der zu lernende Sachverhalt ist. Beispiele hierfür sind die Untersuchungen von Michas & Berry (2000), in der der genaue Bewegungsablauf, wie eine Hand zu bandagieren ist, vermittelt wurde, oder von Yang & Andre (2003), in der unter anderem die Elektronenbewegungen in einer Taschenlampenbatterie zu erlernen waren.

Eine Erklärung hierfür könnte sein, dass die Animation es im repräsentationalen Fall erleichtert, ein mentales Modell der zu erlernenden Bewegung zu erstellen, da sie bereits ein Muster für dieses anbietet:

*An animation is likely to be useful when the learning material entails motion, trajectory or change over time so that the animation helps to build a mental model of the dynamics (Bétrancourt & Tversky, 2000).*

Wenn Animationen hingegen rein dekorativ verwendet werden, wird die Bildung eines funktionalen „bewegten“ mentalen Modells eher nicht gefördert. Das Vertrauensintervall (-0,02 bis 0,24) schließt in diesem Fall den Wert Null mit ein, so dass keine signifikante Überlegenheit von Animationen gegenüber Standbildern festzustellen ist. Ein Beispiel hierfür sind die Visualisierungen bei Lai (2000), in dessen Lernumgebung Konzepte des Programmierens mit QuickBASIC mit Hilfe einer animierten Zeichentrickfigur vermittelt werden sollten – die Animationen erbrachten keinen zusätzlichen Vorteil. Man könnte mutmaßen, dass die gezeigte Bewegung auf dem Bildschirm im Gegenteil sogar von dem eigentlichen Lerngegenstand ablenkt. Nach der *Cognitive Load Theory* (Chandler & Sweller,

1991) würde durch die Animation zusätzliche *Extraneous Cognitive Load* produziert, der das Arbeitsgedächtnis des Lernenden unnötig belastet. Mit anderen Worten: Die dekorative Animation offeriert lediglich unnötige Einzelheiten, so genannte *seductive details* (Harp & Mayer, 1998).

### 3.4.3 Deklarativ oder prozedural?

Die Analyse zeigt, dass Animationen in einigen Fällen dann größere Lernerfolge erbrachten, wenn eher deklaratives statt prozedural-problemlösendes Wissen gefordert wurde, also eher reine Fakten abgefragt statt Verständnisaufgaben gestellt wurden. In anderen Fällen waren Animationen, die problemlösendes Wissen abfragten, solchen mit deklarativen Aufgaben zumindest nicht überlegen.

Beides erscheint überraschend, da von Animationen im Allgemeinen eher die Möglichkeit erwartet wird, tieferes Verständnis für einen Prozess zu gewinnen, während für das Memorieren von Fakten auch Standbilder ausreichen sollten. So geben Weiss et al. (2002) u.a. für „facts requiring memorization“, „principles/rules requiring a simple explanation“ die Empfehlung: “Your CBI does not need animation to effectively teach the subject matter” (S. 474).

Die Metaanalyse weist aber konträre Ergebnisse auf. Zwar sind Animationen Standbildern auch beim Erwerben von Verständnis (im Sinne von prozedural-problemlösendem Wissen) überlegen ( $d=0,27$ , 95% VI 0,14 bis 0,40), dieser Vorteil ist aber noch deutlich stärker beim Faktenwissen ( $d=0,43$ , 95% VI 0,31 bis 0,55) feststellbar und wird in einer Reihe von Einzelvergleichen – wenngleich nicht grundsätzlich – statistisch signifikant. Es mag sein, dass die Möglichkeit, mit Hilfe von Animationen leichter ein mentales Modell zu generieren, beim Zugriff auf Informationen mehr Vorteile bringt als bei der Anwendung dieses Wissens zur Problemlösung. Dies erscheint insofern nachvollziehbar, dass Informationsabruf eine notwendige, aber nicht hinreichende Vorbedingung für das Problemlösen ist.

Diese Frage bedarf in jedem Fall weiterer Forschung; es bleibt festzuhalten, dass die Faustregel „A well-designed animated display can be expected to facilitate [...] transfer rather than learning“ (Bétrancourt & Tversky, 2000, S. 326) zumindest überdenkenswert ist.

Weiterhin zeigten sich hohe Lernerfolge bei prozedural-motorischen Aufgaben; wie die detaillierteren Auswertungen zeigen, war deren scheinbare Überlegenheit gegenüber Vergleichen mit deklarativen Aufgaben ( $d=1,06$  versus 0,43) aber

hauptsächlich der Tatsache zu schulden, dass die gesamte Kategorie der motorischen Einzelvergleiche aus repräsentationalen, videobasierten Animationen bestand. Wird dieser Kategorie eine vergleichbare Kategorie der repräsentationalen Animationen entgegengestellt, bei denen deklaratives Wissen gefordert wird, verschwinden die Vorteile der prozedural-motorischen Kategorie. Dies ist insofern erstaunlich, als solcherlei Prozeduren, besonders, wenn sie direkt einen Vorgang darstellen, den Menschen ausführen und den es sich anzueignen gilt, zunächst allgemein leichter zu erlernen sein sollten, da dieser Vorgang eher vom nichtdeklarativen Gedächtnis gesteuert wird und nach gewisser Übung zu reflexartigen statt reflektiven Kenntnissen führt (Squire & Kandel, 1999). Zudem deuten neuere Untersuchungen darauf hin, dass gerade in solchen Fällen von Animationen in besonderem Fall profitiert werden könnte (Marcus, Wong, Ayres, Paas & Sweller, 2007).

#### **3.4.4 Computer- oder videobasiert – Grad des Realismus?**

Hinsichtlich der Frage, ob Animationen eher computer- oder videobasiert erstellt werden sollten – und bezüglich der unmittelbar daraus resultierenden Frage, wie realistisch eine Animation sein sollte –, ergibt sich ein gemischtes Bild. Zwar scheinen videobasierte Animationen, und damit diejenigen mit höchstem Realitätsgrad, zunächst den computerbasierten überlegen zu sein, jedoch relativiert sich dieser Vorteil (von  $d=0,67$  vs.  $d=0,33$ ) bei der Zerlegung in Subkategorien und Einbeziehung statistischer Interaktionen mit anderen Variablen weiter und weiter; die Bedingung „repräsentationale, videobasierte Visualisierungen ohne Text“ ( $d=0,95$ , 95% VI 0,67 bis 1,23), der acht von zwölf Videos angehören, hat bei Gegensetzung zu ihrem computerbasierten Pendant keinerlei statistisch signifikanten Vorteil mehr ( $d=0,78$ , 95% VI 0,25 bis 1,31). Die scheinbare Überlegenheit der hochrealistischen Bedingungen dürfte also auf andere Faktoren wie die Konfundierung mit der An- oder Abwesenheit zusätzlichen erläuternden Textes zurückzuführen sein. Dazu passt auch, dass bei der Einteilung der Bedingungen in vier Realismusstufen die des zweithöchsten Realismusgrades (Stufe 3), also die der hochrealistischen *computerbasierten* Animationen, die niedrigste mittlere gewichtete Effektstärke aufweist ( $d=0,19$ , 95% VI 0,01 bis 0,37).

Animationen mit eher niedrigem Realismus resultieren also nicht in niedrigeren Lernerfolgen, was auch mit den Forderungen Tverskys et al. (2002) in Einklang steht: „Animations should lean toward the schematic and away from the realistic“ (S. 258). Auf diese Weise nämlich sollte es Lernenden leichter fallen, sich auf die wesentlichen inhaltlichen Bereiche zu konzentrieren, sich also weniger z.B. von graphischen Details ablenken zu lassen („sufficiently complex to convey the important information within it, yet simple enough to be easily understood“; Milheim, 1993, S. 173). Bereits Rieber (1994) forderte dementsprechend Animationen, die den Lernenden ausdrücklich auf den entscheidenden Lerngegenstand hinweisen. Er benutzte dazu in seinen Untersuchungen (Rieber, 1990; 1991) eine „chunking“-Strategie, die die Animationen mehrfach unterteilte und so den Lernenden die neuen Informationen Schritt für Schritt präsentierte. Zudem benutzte Rieber Animationen mit sehr niedrigem Realitätsgrad – bezeichnenderweise stammt ein Großteil der Einzelvergleiche mit Realitätsstufe 1 in dieser Metaanalyse aus den Untersuchungen Riebers.

Rieber (1994) verwies aber auch darauf, dass seiner Überzeugung nach die „chunking“-Strategie nur eine von vielen Möglichkeiten des „Cueings“ in Animationen darstellt, die zum Erfolg führen sollten (S. 176). Dies könnte z.B., nach Tversky et al. (2002), Erklärungen, Pfeile oder grafische Hervorhebungen umfassen.

Letztendlich mag es wieder der entscheidende Punkt zu sein, die kognitive Belastung des Lernenden zu reduzieren und zu diesem Zweck möglichst viele störende, den *Extraneous Cognitive Load* (Sweller, 1994) erhöhende Elemente zu entfernen.

### **3.4.5 Eigenschaften von Standbildern**

Laut dieser Metaanalyse sind Animationen Standbildern hinsichtlich des Lernerfolgs im Allgemeinen weit überlegen. Wie aber müssen Standbilder gestaltet sein, um diesen Rückstand wenigstens teilweise auszugleichen?

Es hat sich einmal mehr gezeigt, dass begleitende, erläuternde Texte für einen hinreichenden Lernerfolg unabdingbar sind: War kein Text vorhanden, wie bei 20 Einzelvergleichen der Fall, stieg der Vorteil der Animationen deutlich an ( $d=0,62$  vs.  $0,35$ ), in einigen Subkategorien sogar auf Werte größer als eine Standardabweichung.

Dieses wenig überraschende Resultat befindet sich in Übereinstimmung mit Mayers "Generative Theory of Multimedia Learning" (Mayer, 1997, 2001), die davon ausgeht, dass erfolgreiches Lernen dann stattfinden kann, wenn beide Kanäle der Informationsverarbeitung – visuell wie verbal – angesprochen und referentielle Verknüpfungen zwischen beiden erstellt werden. Mayers *Multimedia Principle* (Mayer & Gallini, 1990) besagt, dass statische Bilder in Verbindung mit Text bessere Lernwirksamkeit aufweisen als Text allein. In Relation hierzu zeigen die Ergebnisse der Metaanalyse nun (indirekt, durch den Vergleich mit Animationen), dass statische Bilder in Verbindung mit Text bessere Lernwirksamkeit aufweisen als statische Bilder allein.

Unerwarteter ist jedoch, dass Cues in den Standbildern, Hinweisreize in Form von Pfeilen etc., offensichtlich keinen zusätzlichen Lerngewinn erbringen, nachdem beide Bedingungen ähnlich abschnitten ( $d=0,34$  vs.  $0,47$ ). Auch eine (hier nicht dargestellte) Unterteilung in weitere Subkategorien brachte keinerlei Veränderung.

So sinnvoll Cues in Verbindung mit Animationen nach allgemeiner Meinung auch sein mögen (s.o.) – in Verbindung mit Standbildern sind Standbilder Animationen nicht weniger unterlegen als ohne zusätzliche Cues. Dies könnte daran liegen, dass Standbilder im Gegensatz zu Animationen es nicht nötig haben, auf deren besonders wichtige Bestandteile gesondert hinzuweisen, da der Betrachter genügend Zeit hat, diese selbständig herauszusuchen. Bei nicht oder nur minimal interaktiven Animationen hingegen besteht die Gefahr, dass schon im nächsten Moment ein anderer Aspekt entscheidend ist und der erste nicht mehr zu sehen, was das Arbeitsgedächtnis zusätzlich belastet (Ainsworth & Van Labeke, 2004):

*When one views an animation or video, one views one frame at a time, and once the animation or video has advanced beyond a given frame, it is no longer available to the viewer. This places heavy demands on working memory if information presented earlier in the animation or video must be integrated with information that is presented later. (Hegarty, 2004)*

### **3.4.6 Limitationen dieser Metaanalyse**

Die spezifischen Charakteristika einer Metaanalyse sollten bei der Bewertung der hier berichteten Ergebnisse in Erinnerung behalten werden. So ist zum Beispiel zu beachten, dass der Bereich der Studien, die die hier vorliegende Metaanalyse umfasst, limitiert ist auf genau jene Untersuchungen, die Standbilder mit Animationen

vergleichen, bei denen die Animationen nicht oder nur minimal interaktiv sind (was einen weitgehend „fairen“ Vergleich ermöglicht), bei denen beide Bedingungen zumindest ansatzweise zu vergleichen sind und die nicht zuletzt auch gewisse methodische Mindeststandards einhalten. Auch müssen die Ergebnisse mit denen der anderen Studien vergleichbar sein und zumindest grundlegende statistische Angaben machen. Daher sind eine Reihe von Untersuchungen, die sich zwar auch mit diesem Feld beschäftigen, aber nicht all diese notwendigen Kriterien einhalten, in der Metaanalyse nicht enthalten, auch wenn sie zum Teil sehr interessante Ergebnisse enthalten – als nur ein Beispiel seien hier etwa die Arbeiten von Lowe (z.B. 1999; 2003) genannt.

Zudem ist nochmals auf das Problem der Heterogenität der Subkategorien zu verweisen, das zwar ein häufiges Charakteristikum von gewichteten Metaanalysen ist (z.B. LaFrance, Hecht & Paluck, 2003), nichtsdestotrotz die statistische Aussagekraft und Interpretierbarkeit der dargestellten Ergebnisse aber mindert.

Schließlich erhebt die Analyse nicht den Anspruch, sämtliche möglichen Moderatoren erfasst zu haben. Im Gegenteil sind sicherlich noch eine große Anzahl weiterer, den Erfolg von Animationen mitbestimmender Variablen denkbar: so z.B. lernerabhängige Faktoren wie dessen themenspezifisches Vorwissen (z.B. ChanLin, 2001), *Spatial Ability* (z.B. Yang et al., 2003) oder Motivation (z.B. Höffler, 2003), die Anzahl der vorgegebenen Schlüsselbilder (z.B. Hegarty, 1992), die Zeit bei der Bearbeitung (Tversky et al., 2002) oder natürlich auch die Interaktivität mancher Animationen (z.B. Nerdel, 2003; Plötzner & Lowe, 2004; Schnotz, Bockheler & Grzondziel, 1998; Schuh, Gerjets & Scheiter, 2005). Diese und andere Faktoren waren in die Metaanalyse, weil jeweils nur von einem kleinen Teil der Primärstudien erfasst und untersucht, nicht integrierbar.

### **3.4.7 Fazit**

Die vorliegende Metaanalyse stellt den Versuch dar, in ein nach wie vor relativ schwer zu überblickendes und unübersichtliches Forschungsfeld, in dem verschiedene Ansichten zur Wirksamkeit und zum Nutzen von Animationen miteinander konkurrieren, eine gewisse Ordnung und Objektivität hineinzubringen.

Dabei konnte eine durchaus überraschende klare Überlegenheit von Animationen gegenüber Standbildern aufgezeigt werden, die sich, wie bereits von verschiedenen

anderen Autoren gefunden bzw. vermutet, besonders unter spezifischen Bedingungen herauskristallisierte.

So gibt es nun weitere Evidenz für die Nützlichkeit des Einsatzes von Animationen, wenn die dargestellte Bewegung auch tatsächlich der Lerngegenstand ist (und nicht nur dekorativ verwendet wird). Ein weiteres Mal konnte gezeigt werden, dass Standbilder immer zusammen mit begleitendem Textmaterial präsentiert werden sollten – nicht notwendigerweise jedoch mit zusätzlichen Hinweisreizen (Cues) – wenn dies auch jeweils sicher grundsätzlich vom Lerninhalt abhängig ist. Für den angemessenen Grad der Realistik von Animationen konnten lediglich mehrdeutige Hinweise gefunden werden, schließlich scheinen Animationen auch und gerade für den Erwerb deklarativen Wissens geeignet.

Es erscheint sinnvoll, diese Erkenntnisse mit weiteren Primärstudien zu validieren. Trotzdem kann auf Basis dieser Analyse behauptet werden, dass Animationen, in bestimmten Bereichen, unter bestimmten Voraussetzungen, selbst ohne Einbeziehung von Interaktivität nach wie vor ein großes Potential zur Lernerleichterung haben. Animationen scheinen besser zu sein als ihr Ruf, erfordern aber, um effektiv zu sein, einen durchdachten und sorgfältigen Einsatz sowie ein Design, das auf forschungsbasierten Theorien des Lernens und der Instruktion beruht.

## 4 Erste experimentelle Studie (Pilotstudie)

### 4.1 Einführung

Die Ergebnisse der Metaanalyse warfen eine ganze Reihe von möglichen Fragestellungen auf, die einer näheren experimentellen Untersuchung würdig gewesen wären. Sollte zum Beispiel der Inhaltsbereich von Lernumgebungen systematisch variiert werden, um Vorteile von Animationen in bestimmten Domänen auszumachen? Lassen sich Unterschiede bei verschiedenen Graden der Realistik der Visualisierungen (Tversky et al., 2002) finden? Ist „Cueing“ (Lowe, 2003; Rieber, 1994), wenn schon nicht bei Standbildern im Vergleich mit Animationen, wofür die Metaanalyse Hinweise erbrachte, so doch vielleicht bei Animationen eine sinnvolle Möglichkeit zur Steigerung des Lernerfolgs?

Die Wahl fiel schließlich auf eine andere Frage: Könnte es möglich sein, dass Animationen zwar, wie in der Metaanalyse gefunden, zunächst einmal einer Serie von Standbildern überlegen sind, dass sich dieser Unterschied aber aufhebt, wenn die *Anzahl* der Standbilder deutlich angehoben wird (vgl. Hegarty et al., 2003)? Schließlich ist eine Animation ebenfalls nur eine Serie schnell dargebotener Standbilder (Rieber & Kini, 1991). Daneben erscheinen die Effekte potentiell moderierender Variablen besonders interessant, wie etwa das Vorwissen, der *Cognitive Load* oder das räumliche Vorstellungsvermögen. Die Beantwortung dieser Fragen könnte eine Erklärung für die große Heterogenität der Befunde zum Lernen mit Animationen und Standbildern liefern, die in der Metaanalyse festgestellt wurde.

Im Folgenden wird nun eine Serie von drei Experimenten dargestellt, die das Ziel hat, anhand verschiedener Versionen einer selbst entwickelten Lernumgebung im Bereich der Chemie mögliche Effekte der Variation der Anzahl der Standbilder sowie potentieller Moderatorvariablen nachzuweisen. Die in diesem Kapitel beschriebene Pilotstudie dient dabei zunächst einmal, neben dem Ziel der Erprobung der Materialien, dazu, eine erste Bestätigung für die Hypothese zu bekommen, dass Animationen Standbildern im Allgemeinen hinsichtlich des Lernerfolgs überlegen sind. Erst in einem nachfolgenden Schritt wird dann eine weitere Lernumgebungsversion eingeführt, die aus deutlich mehr Standbildern besteht.



## 4.2 Fragestellung und inhaltliche Hypothesen

Untersuchungsgegenstand dieser ersten experimentellen Untersuchung ist die unterschiedliche Wirkung von Animationen und Serien von Standbildern im Bereich der Chemie hinsichtlich ihrer Lernwirksamkeit. Hierbei wird unterschieden zwischen fakten- und verständnisorientierten Lernaufgaben. Zudem ist von Interesse, ob lernerabhängige Variablen eine moderierende Rolle spielen, wie etwa das themenspezifische Vorwissen des Lernenden oder dessen räumlich-visuelle Begabung. Auch die Rolle der kognitiven Belastung wird untersucht.

Abgeleitet besonders aus den Ergebnissen der Metaanalyse, die im vorherigen Kapitel geschildert wurden, ergibt sich folgende erste inhaltliche Hypothese:

***H1: Mit einer animierten Lernumgebung werden bessere Lernleistungen erzielt als mit einer aus statischen Bildern bestehenden Lernumgebung.***

Spezifischer auf die Art der Lernleistung hin ausgerichtet zeigen gängige Literatur und Metaanalyse konträre Ergebnisse: Während bisherige Untersuchungen eher davon ausgingen, dass Animationen besonders für den Erwerb von tieferem Verständnis sinnvoll sind, deutete die Metaanalyse auf einen ausgeprägten Vorteil von Animationen auch gerade beim Faktenwissen hin. Im Vertrauen auf die Allgemeingültigkeit der Ergebnisse der Metaanalyse lautet dementsprechend die ergänzende zweite Hypothese:

***H2: Die Überlegenheit der Animation zeigt sich sowohl bei Aufgaben, die Faktenwissen erfragen, als auch bei solchen, für die tieferes Verständnis erforderlich ist.***

Hinsichtlich des themenspezifischen Vorwissens ist zu bemerken, dass allgemein davon ausgegangen wird (Dean & Enemoh, 1983; Lewalter, 1997, S. 65 ff.; Mayer, 1997), dass besonders Lernende mit niedrigem Vorwissen von Animationen profitieren, während denjenigen mit hohem Vorwissen Standbilder genügen und sich Animationen sogar als lernhinderlich erweisen können. Ausgehend von der Voraussetzung, dass sich sowohl Lernende mit hohem als auch mit niedrigem Vorwissen in der Stichprobe befinden, ergibt sich als dritte Hypothese:

***H3: Die lernförderliche Wirkung von Animationen lässt sich besonders für Lernende mit niedrigem themenspezifischen Vorwissen nachweisen. Lerner***

***mit hohem Vorwissen erzielen mit stehenden Bildern mindestens genauso gute Ergebnisse wie mit Animationen.***

Sollte die räumliche Vorstellungskraft (*Spatial Ability*) eine moderierende Rolle spielen, so liefert die Literatur (Hegarty, 2005; Issak & Just, 1995; Large et al., 1996; Mayer & Sims, 1994; Yang et al., 2003) dafür eher widersprüchliche Hinweise: Kann eine Animation Defizite im räumlichen Vorstellungsvermögen ausgleichen oder überfordert sie gerade solche Lernenden mit solcherlei Defiziten? Plausibler anzunehmen erscheint, der Supplantationstheorie von Salomon (1979) folgend, ein Interaktionseffekt der ersten Variante:

***H4: Lernende mit niedrigem räumlichen Vorstellungsvermögen profitieren von Animationen im Vergleich zu Standbildern in höherem Maße als Lernende mit hohem räumlichen Vorstellungsvermögen.***

Schließlich ist bezüglich des *Cognitive Load* davon auszugehen, dass die Animation aufgrund ihres schnell fortlaufenden Charakters eine höhere kognitive Belastung des Arbeitsgedächtnisses darstellt als statische Bilder – wobei die Animation jedoch nicht zu einer kognitiven *Überlastung* führen darf, da dies wiederum den Lernerfolg negativ beeinträchtigte und damit Auswirkungen auf die Gültigkeit der ersten Hypothese hätte:

***H5: Die Animation führt bei den Lernenden zu einer höheren kognitiven Belastung als Standbilder.***

## 4.3 Methode

### 4.3.1 Variablen

#### *Unabhängige Variable (UV)*

Als Unabhängige Variable in dieser Untersuchung dient die *Art der Visualisierung*, also ob dem Lernenden eine Animation oder eine Serie von vier Standbildern präsentiert wird.

### *Abhängige Variablen (AV)*

Die Abhängige Variable stellt in erster Linie der *Lernerfolg* dar, gemessen anhand der erzielten Punkte in einem Wissensnachtest. Dieser lässt sich differenzieren in offene und Multiple-Choice-Fragen, ferner in Fragen nach Faktenwissen und tiefer gehendem Verständnis. Faktenwissen wird allgemein als die Kenntnis von Begriffen und Definitionen sowie von Tatsachen und Eigenschaften definiert (Lewalter, 1997); es handelt sich also lediglich um eine Erfassung der Oberflächenstruktur der Informationen. Verständnis hingegen umfasst darüber hinaus die Fähigkeit, gelernte Bedeutungen und Definitionen zu übertragen sowie Schlussfolgerungen zu treffen und Interpretationen vorzunehmen (Metfessel, Michael & Kirsner, 1969). Lewalter (1997) definiert Verständnis als die Anwendung des Wissens in Aufgaben, die in direktem Zusammenhang mit den Lerninhalten stehen.

Weiterhin lässt sich der Wissensnachtest differenzieren in Aufgaben, die spezifisch anhand der Animation bzw. der Standbilder zu beantworten sind und solche, für die Informationen aus der umgebenden Lernumgebung, in die die Animation bzw. die Standbilder eingebettet sind, ausreichend sind. Besonders letztere Unterscheidung ist entscheidend, da für die Differenzierung des Lernerfolgs zwischen Animationen und Standbildern natürlich nur die Aufgaben entscheidend sind, für deren Beantwortung beide Gruppen der UV unterschiedliche Bedingungen zur Verfügung gestellt bekommen.

Zudem wird das Maß des *Cognitive Load*, gemessen mit zwei von Paas (1992) adaptierten Items, als AV erhoben, auch wenn die kognitive Belastung möglicherweise lediglich moderierende oder mediierende Funktion (vgl. Baron & Kenny, 1986) hat.

### *Moderatorvariablen*

Als (mutmaßlich) moderierende Variablen sind zum einen die *Höhe des themenspezifischen Vorwissens* zu nennen und zum anderen das Ausmaß räumlich-visueller Fähigkeiten, der *Spatial Ability*, gemessen mit Hilfe eines Paper-Folding-Tests von Ekstrom, French, Harman & Dermen (1976), der in der Multimediaforschung sehr häufig eingesetzt wird (u.a. Barnea & Dori, 1999; Mayer &

Sims, 1994; Moreno & Mayer, 1999) und besonders wegen seiner hohen Komplexität geschätzt wird (vgl. Hegarty & Waller, 2004).

### *Störvariablen*

Eine klassische potentielle Störvariable beim Einsatz von computerbasierten Lernumgebungen (z.B. Lewalter, 1997; Nerdel, 2003; Urhahne, 2002) ist die Vertrautheit mit dem Medium Computer. Während noch vor einigen Jahren davon auszugehen war, dass nicht jeder Proband zwangsläufig im Umgang mit Hypertexten, Animationen und allgemein mit der Bedienung eines Computers geschult ist, so dass eine computerbasierte Untersuchung immer auch eine Kontrolle dieser Variable beinhalten musste (z.B. Nerdel, 2003), ist dies heutzutage nicht mehr der Fall. Gerade im Falle der hier rekrutierten Probanden (vgl. unten) – Studierende des zweiten bis vierten Semesters –, aber auch allgemein bei Schülern und Studierenden, sind Basiskenntnisse im Computerumgang mittlerweile gängiger Standard. Die OECD (2006) führt hierzu an, dass „nahezu alle 15-jährigen Schüler in den OECD-Ländern [...] über etwas Erfahrung im Umgang mit Computern“ (S. 2) verfügen. Da zur Bearbeitung der hier verwandten Lernumgebung lediglich Basiskenntnisse (wie etwa das Bedienen einer Computermouse, das Klicken auf „Weiter“-Buttons) nötig sind, kann davon ausgegangen werden, dass diese Variable keinen störenden Einfluss hat.

Eine wichtige, in Vergleichen von dynamischen und nichtdynamischen Visualisierungen häufig nicht beachtete Störvariable ist der Zeitfaktor. Es stellt eine wichtige Vorbedingung dar, dass gewährleistet wird, dass zum Lernen mit beiden Versuchsbedingungen gleich viel Zeit eingeräumt wird. Wegen ihres spezifischen Charakters werden Standbilder in der Regel deutlich länger betrachtet als eine Animation mit ihrem dynamischen Ablauf; wird dieser Unterschied im Design nicht berücksichtigt, ist zwar eine hohe Alltagsvalidität gegeben, auf der anderen Seite jedoch können beide Bedingungen nicht wirklich fair miteinander verglichen werden, wenn die Standbilder deutlich länger betrachtet werden als die Animation. Um diesem Problem entgegenzuwirken, wurde bei vorliegender Studie die Zeit, in der Animationen oder Standbilder gezeigt wurden, exakt gleich gehalten und mit der verbalen Erläuterung synchronisiert. Die Lernenden können die Standbilder nicht nach Wunsch länger betrachten, die Bilder wechseln stattdessen im Stile einer Diashow und sind daher mit der Animation vergleichbarer.

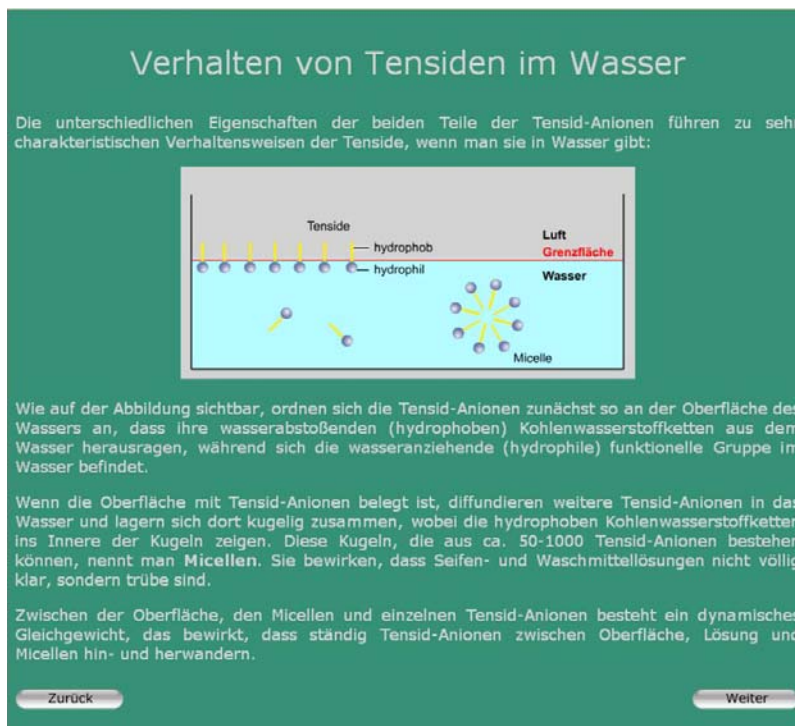
### 4.3.2 Materialien

#### *Computergestützte Lernumgebungen*

Wie bereits deutlich gemacht, basiert diese Studie zum Großteil auf einer computergestützten Lernumgebung. Diese wurde speziell für diese und die Folgeuntersuchungen mit den Autorenprogrammen *Macromedia Flash MX Professional 2004* sowie *Macromedia Dreamweaver MX 2004* erstellt und existiert in zwei Versionen, einer Standbild- und einer Animationsversion.

Die Thematik dieser Lernumgebungen ist der Vorgang der Schmutzablösung beim Waschen durch Tenside. Dabei stimmen beide Versionen zunächst in den Grundzügen der formalen Gestaltung und Bedienung überein: Schriftart und -größe, Farbgebung und Rahmengrößen etc. sind gleich, auch die Bedienelemente zur Navigation innerhalb der multimedialen Lernumgebung zwischen einzelnen Seiten wie z.B. „Weiter“-Buttons sind identisch.

Tatsächlich sind die ersten Seiten beider Lernumgebungen, die einen Einstieg in die Thematik vermitteln sollen und eine Mischung aus Texten und stehenden Bildern darstellen, sogar exakt dieselben (vgl. als Beispiel Abbildung 4-1), ebenso wie die letzte Seite.



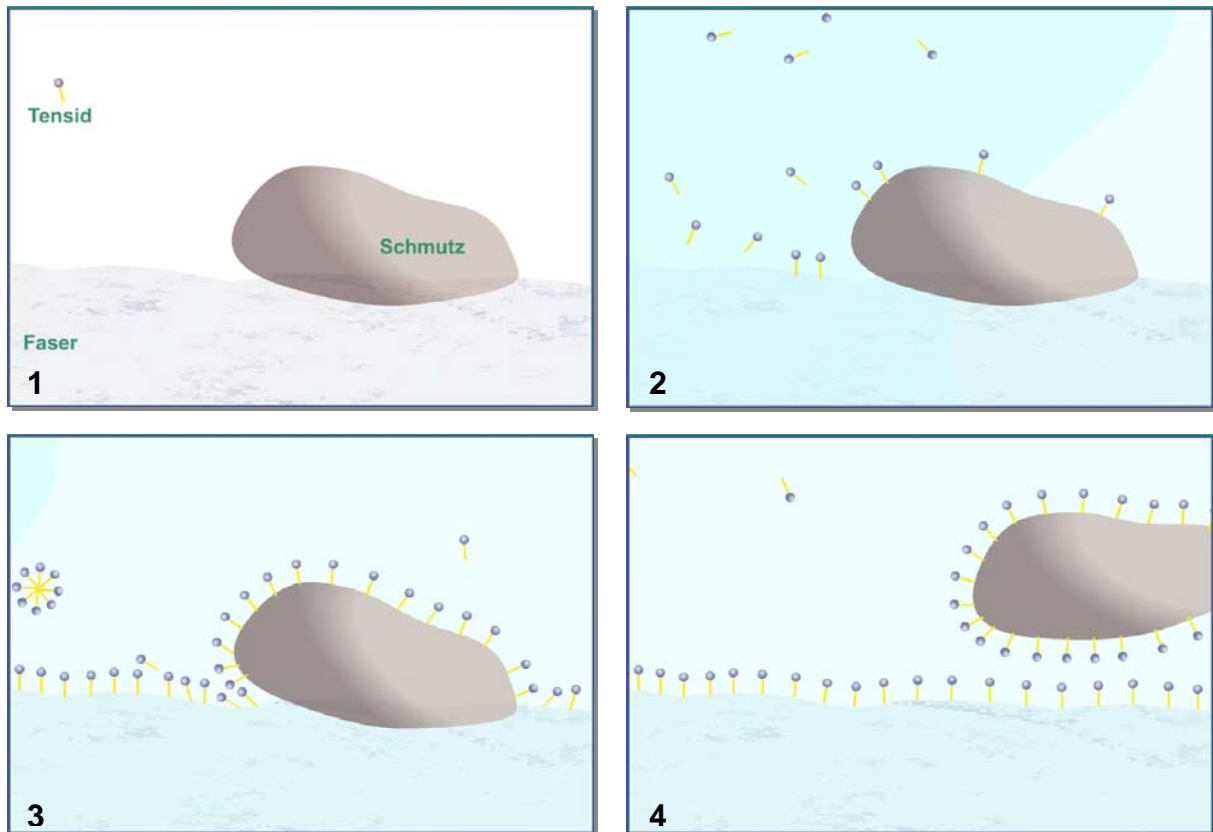
**Abbildung 4-1:** Ein Beispiel aus den in die Thematik einführenden Anfangsseiten beider Lernumgebungsversionen.

Dazwischen befindet sich die Seite mit der Animation bzw. der Serie von Standbildern. Auch diese sehen zunächst gleich aus, beide beginnen mit demselben Standbild (Abbildung 4-2), das die Elemente der Visualisierung benennt.



*Abbildung 4-2: Die Startseite der Animation bzw. der Serie von Standbildern. Klicken auf den „Start“-Button startet die Lernsequenz.*

Während nach Klicken auf den Start-Button bei der Animationsversion eine 73 Sekunden dauernde Animation folgt, werden in der Standbildversion in demselben Zeitraum vier Standbilder, Screenshots aus der Animation, nacheinander eingeblendet. Diese wurden nach inhaltlichen Gesichtspunkten ausgewählt und stellen jeweilige Schlüsselmomente des Ablaufs der Schmutzablösung dar, die zum Verständnis des Vorgangs essentiell sind. Abbildung 4-3 stellt die Serie in ihrer Gesamtheit dar.



*Abbildung 4-3: Die vier verwendeten Standbilder, hier gemeinsam dargestellt. Die Nummerierungen geben die Reihenfolge der Darbietung an und waren nicht Teil der Lernumgebung.*

Der jeweilige Zeitpunkt des Wechsels zum nächsten Standbild erfolgt in Abstimmung mit dem hinterlegten Audiokommentar, der die Vorgänge auf dem Schirm erläutert (eine Abschrift des Kommentars ist in Anhang G zu finden). Dieser ist in der Animationsversion ebenfalls enthalten; die Probanden haben zu diesem Zweck Kopfhörer aufgesetzt.

Die Nutzung des Audiokommentars anstelle etwa von *on-screen* Text erfolgt in Übereinstimmung mit dem Modalitätsprinzip Mayers (2001; vgl. Kapitel 2.1).

### *Fragebögen*

Neben dem Computermaterial erhalten die Probanden sowohl vor als auch nach dessen Bearbeitung einen Fragebogen (vgl. Anhänge B und C), die im Folgenden als Vor- und Nachtest bezeichnet werden.

### *Vortest*

In dem ersten Fragebogen (Anhang B), der vor der Bearbeitung der Lernsoftware auszufüllen ist, werden zum einen Informationen zum schulischen und universitären Werdegang der Versuchspersonen abgefragt. Dies soll einem Screening zur Einschätzung der Vorbildung bezüglich der behandelten Materie dienen. Die Fragen beziehen sich z.B. auf die Abiturnote, die letzte Chemienote oder das Hauptfach des Studiums. Auch das Geschlecht sowie das Alter werden erhoben. Die Fragebögen werden anonymisiert, indem die Probanden angewiesen werden, mit Hilfe der Initialen ihres Vaters und dem Geburtstag ihrer Mutter (vierstellig, TTMM) ein individuelles Kennzeichen zu erstellen. Zudem sollen die Probanden die an ihrem Arbeitsplatz angebrachte Nummer eintragen, um ihnen auf diese Weise Treatments zuordnen zu können.

Zur Erfassung des themenspezifischen Vorwissens werden vier offene Fragen gestellt, die sowohl Wissen des Weiteren allgemeinen Themenbereichs erfragen (z.B. „Erläutern Sie den Begriff ‚Oberflächenspannung‘!“) als auch sehr spezifisch auf Vorwissen zu der Thematik der Animation bzw. der Standbilder zielen („Wie funktioniert die Schmutzablösung beim Wäschewaschen?“).

Zusätzlich enthält der Vortest eine Version des Paper-Folding-Tests nach Ekstrom et al., 1976 (vgl. Anhang E) zur Messung der *Spatial Ability*. Dieser besteht aus einer Seite Instruktion und einer Seite mit zehn Aufgaben ein- und desselben Typs: Die schrittweise Faltung eines Papiers in zwei bis vier Schritten ist dargestellt, an dessen Ende an einer Stelle ein Loch gestanzt wird. Daneben sind fünf Möglichkeiten mit unterschiedlichen Positionen und unterschiedlicher Anzahl von Löchern für das wieder aufgefaltete Papier dargestellt. Von diesen ist nur eine korrekt (*single-choice*), die von den Probanden anzukreuzen ist. Der Paper-Folding-Test ist ein Speed-Test, d.h. die Probanden haben nur drei Minuten Zeit zur Bearbeitung und beginnen dementsprechend mit der Bearbeitung gleichzeitig auf ein Zeichen des Versuchsleiters hin.

### *Nachtest*

Der Nachtest (Anhang C) besteht aus zwei verschiedenen Fragebögen, da die Beantwortung späterer Fragen die Teilnehmer mit Wissen versorgen könnte, um ihre Lösung von Aufgaben zu Beginn nochmals zu revidieren oder zu ergänzen. Daher erhalten die Probanden den zweiten Teil des Nachtests erst im Austausch gegen den



ersten Teil, der wiederum unmittelbar nach der Beschäftigung mit der computergestützten Lernumgebung ausgeteilt wird.

Der erste Teil enthält drei offene Aufgaben zum soeben bearbeiteten Inhalt, davon eine, die tiefgehendes Verständnis erfordert. Außerdem sollen nochmals Kennzeichen und Nummer des Arbeitsplatzes eingetragen werden, weiterhin wird anhand zweier Items, einer Adaption der Skalen von Paas (1992), die subjektiv erlebte kognitive Belastung der Probanden während der Testung erhoben.

Im zweiten Teil finden sich weitere 17 Fragen zum Inhalt der Lernumgebung. Davon weisen sieben Aufgaben ein offenes Frageformat auf, neun Fragen sind im *Single-Choice-Format*, d.h. die Probanden haben die richtige aus vier Antwortmöglichkeiten anzukreuzen. Eine Aufgabe schließlich erfordert die Beschriftung einer Abbildung mit vorgegeben Begriffen.

### **4.3.3 Räumlichkeiten und Hardware**

Die Probanden arbeiten alle an Laptops neuerer Generation, die eine angemessene Auflösung und Darstellungsgeschwindigkeit aufweisen. Jeder von ihnen trägt zum Zeitpunkt der Bearbeitung der Lernumgebung einen Kopfhörer, über den er den Audiokommentar störungsfrei aufnehmen kann.

Die Räume in den Gebäuden der Chemiedidaktik der Universität Duisburg-Essen, die für die Untersuchung genutzt werden, weisen eine ausreichende Größe auf, so dass die bis zu 15 Teilnehmer beider Testungen weit genug voneinander getrennt sitzen, um nicht den Bildschirm des Nachbarn einsehen zu können.

### **4.3.4 Untersuchungsablauf**

Eine Untersuchung dauerte etwa 50 Minuten. Nach einer kurzen Begrüßung, Einführung und Aufklärung über die Anonymität der Untersuchung bekamen die Probanden den ersten Fragebogen (Anhang B) mit der Auflage zugeteilt, ihn zügig, aber trotzdem so genau wie möglich zu beantworten. Dies nahm in etwa 15 Minuten in Anspruch. Dabei wurde der Paper-Folding-Test (Anhang E), der Bestandteil des ersten Fragebogens war, aufgrund seines Speed-Test-Charakters zum Schluss zusammen durchgeführt.

Nach der Rückgabe der Fragebögen begann, gleichzeitig für alle Probanden, die Phase der individuellen Bearbeitung der jeweiligen Lernsoftware. Die Probanden

wurden dabei angewiesen, still und für sich zu arbeiten und die Kopfhörer aufzusetzen. Für eventuell auftauchende technische Probleme stand der Versuchsleiter zur Verfügung, für fachliche bzw. den Programmablauf betreffende Fragen jedoch ausdrücklich nicht. Nachdem alle Teilnehmer die kurze Software in etwa zehn Minuten bearbeitet hatten, wurden die Rechner ausgeschaltet und der erste Teil des Posttests verteilt. Dieser sollte ebenfalls einzeln bearbeitet werden. Nach dessen Beendigung erhielt jeder Proband im Austausch den zweiten Teil des Nachtests.

Sobald jeder diesen ebenfalls abgeschlossen hatte, bedankte sich der Versuchsleiter für die Teilnahme und bot noch eine Frage- und Diskussionsrunde zum Hintergrund der Untersuchung an.

Die beiden Untersuchungen dieser Studie fanden am 20. Juni sowie 21. Juli 2005 statt.

#### **4.3.5 Probanden**

Insgesamt nahmen an dieser (Pilot-)Studie 26 Probanden aus zwei verschiedenen Chemiekursen der Universität Duisburg-Essen teil. Fast alle (24) waren weiblich, das Durchschnittsalter betrug 23 Jahre. Es handelte sich ausnahmslos um Studierende des Lehramts Chemie bzw. Naturwissenschaften, die sich zumeist, in 18 Fällen, im zweiten Fachsemester befanden, höchstens im vierten. Dies lässt, ebenso wie die Angabe, dass nur ein sehr kleiner Teil der Studierenden in der Schule Chemie länger als bis zur elften Klasse belegt hatte (Durchschnitt: 10,24 Schuljahre), den vorläufigen Schluss zu, dass sich das mittlere Vorwissen wohl eher auf einem niedrigen Niveau bewegt.

### **4.4 Ergebnisse**

In diesem Kapitel werden die statistischen Ergebnisse dieser Studie dargestellt. Fast alle Berechnungen wurden mit dem Statistikprogramm SPSS 12.02 (2004) durchgeführt, einzelne Berechnungen der Effektstärke mit dem Programm GPOWER von Faul & Erdfelder (1992) mit Hilfe des Maßes der gepoolten Streuung. Die Darstellung der Ergebnisse erfolgt anhand der zuvor aufgestellten Hypothesen.

Von den ursprünglich 26 Versuchspersonen musste bei der Berechnung eine Probandin ausgeschlossen werden, weil diese durch Schauen bei ihrer Nachbarin beide Testversionen gesehen hatte.

#### 4.4.1 Hypothesen 1 und 2: Animationen versus Standbilder

Wie auch aus Abbildung 4-4 und Tabelle 4-1 ersichtlich, kann die erste Hypothese bestätigt werden – im Mittel ist die Lernumgebung, die eine Animation enthält, derjenigen mit vier statischen Bildern hinsichtlich des Lernerfolgs statistisch signifikant überlegen ( $t(19)=1,79$ ,  $p<0,05$  (einseitig),  $d=0,73$ ).

Hypothese 2 bestätigt sich hingegen nicht. Wie sich herausstellt, ist der Vorteil der Animationen in besonderem Maße auf die Verständnisaufgaben zurückzuführen, nur hier zeigt sich ein einseitig signifikanter Vorteil ( $t(23)=1,74$ ,  $p<0,05$ ,  $d=0,70$ ). Im Falle der reinen Faktenwissensaufgaben verfehlt der Unterschied zwischen Animation und Standbildern das einseitige 5%-Signifikanzniveau ( $t(23)=1,50$ ,  $p=0,07$ ,  $d=0,60$ ).

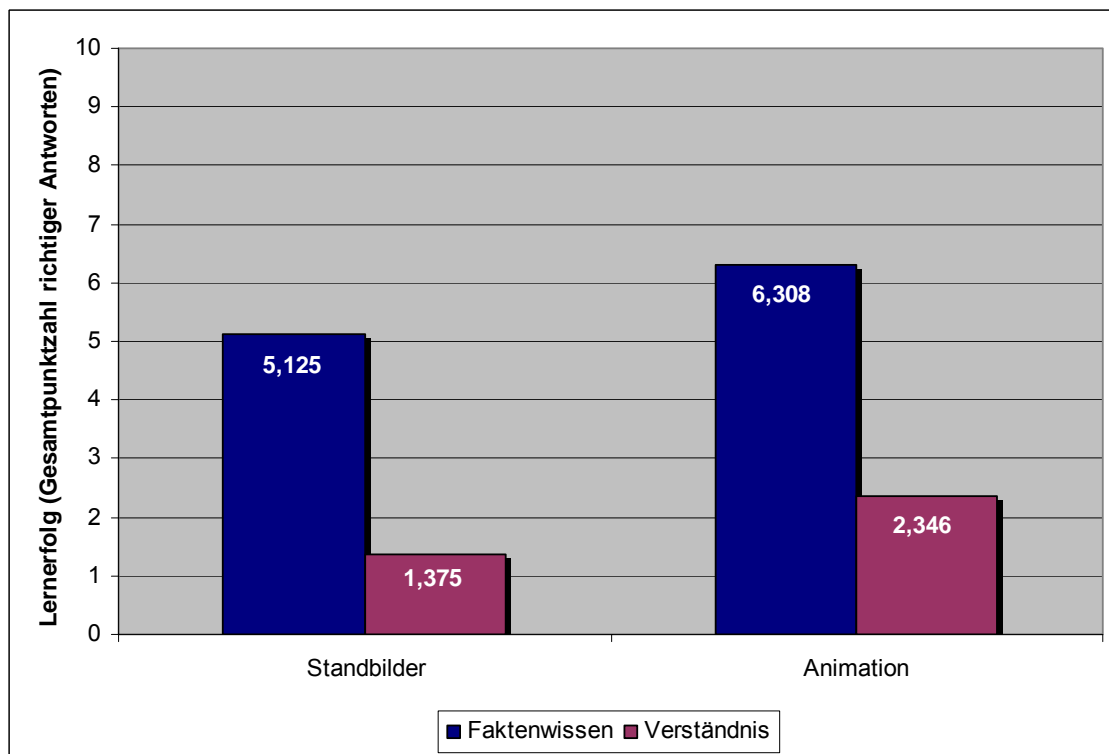


Abbildung 4-4: Mittelwerte der erzielten Lernleistungen der Probanden mit Standbild- oder Animationsversion, aufgeteilt nach Faktenwissen und Verständnisaufgaben.

*Tabelle 4-1: Ergebnisse des Vergleichs von Animation und Standbildern. Angegeben sind neben deskriptiven Kennwerten die Ergebnisse des (einseitigen) t-Tests inklusive Effektstärke d.*

<b>Abhängige Variable</b>	<b>Gestaltungsweise</b>	<b>N</b>	<b>M</b>	<b>SD</b>	<b>df</b>	<b>t<sub>emp</sub></b>	<b>p*</b>	<b>d</b>
Lernerfolg gesamt	Standbilder	12	6,500	3,484	19**	1,792	0,041	0,73
	Animation	13	8,654	2,375				
Lernerfolg Faktenwissen	Standbilder	12	5,125	2,312	23	1,499	0,074	0,60
	Animation	13	6,308	1,588				
Lernerfolg Verständnis	Standbilder	12	1,375	1,351	23	1,739	0,048	0,70
	Animation	13	2,346	1,434				

\* einseitiger Test

\*\* Die niedrigere Anzahl von Freiheitsgraden (df) ergibt sich aus der Nichthomogenität der Varianzen, was die Heranziehung eines speziellen Tests nötig machte.

#### 4.4.2 Hypothese 3: Die Rolle des themenspezifischen Vorwissens

Dass das themenspezifische Vorwissen eine moderierende Rolle spielt, bestätigt sich nicht. Es können keinerlei Effekte dieser Art nachgewiesen werden.

Hierzu werden zunächst die Probanden mittels eines Mediansplits ( $Md=1,5$ ) in hohes und niedriges Vorwissen eingeteilt. Eine alternative Vorgehensweise, die „hohes Vorwissen“ als dann gegeben definiert, wenn mindestens 50% der möglichen Punkte erreicht werden, musste verworfen werden, da sich ansonsten zu ungleiche Gruppengrößen ( $N=18$  vs.  $N=7$ ) ergeben hätten. Insgesamt ist das Vorwissen mit einem Mittelwert von 1,82 und einem Median von 1,5 bei einem möglichen Maximalscore von 8 also wie erwartet recht niedrig. Die Berechnung eines t-Tests mit dem Vorwissensscore als abhängiger Variable und der unabhängigen Variable „Animation oder Standbilder“ zur Absicherung, falls sich per Zufall systematische Unterschiede im Vorwissen zwischen den Gruppen ergeben haben könnten, ergibt kein statistisch signifikantes Ergebnis:  $M_{Animation}=2,23$ ,  $SD_{Animation}=1,80$ ;  $M_{Standbilder}=1,38$ ,  $SD_{Standbilder}=1,30$ ;  $t(23)=1,35$ ,  $p=0,19$ ,  $d=0,54$ . Allerdings ist anzumerken, dass in diesem Falle, da eine Bestätigung anstelle eines Verwerfens der Nullhypothese angestrebt wird, üblicherweise ein höheres Signifikanzniveau angenommen wird, zum Beispiel 20 Prozent. Dieses wird knapp erreicht, so dass die Voraussetzung der Gleichheit des Vorwissens nicht mit letzter Gewissheit gegeben ist.

Im Anschluss an die Aufteilung der Probanden via Mediansplit wurde sodann eine zweifaktorielle Varianzanalyse mit dem Lernerfolg als abhängiger Variable berechnet. Hierbei zeigen sich keine Interaktionseffekte, die den erwarteten moderierenden Effekt des Vorwissens bestätigt hätten (vgl. Tabelle 4-2), auch wenn Abbildung 4-5, das Profildiagramm, Gegenteiliges suggerieren mag. Es kann ebenfalls kein Haupteffekt für das Vorwissen gefunden werden, ein höheres Vorwissen erbringt also keinen signifikanten Vorteil für den Lernerfolg im Posttest.

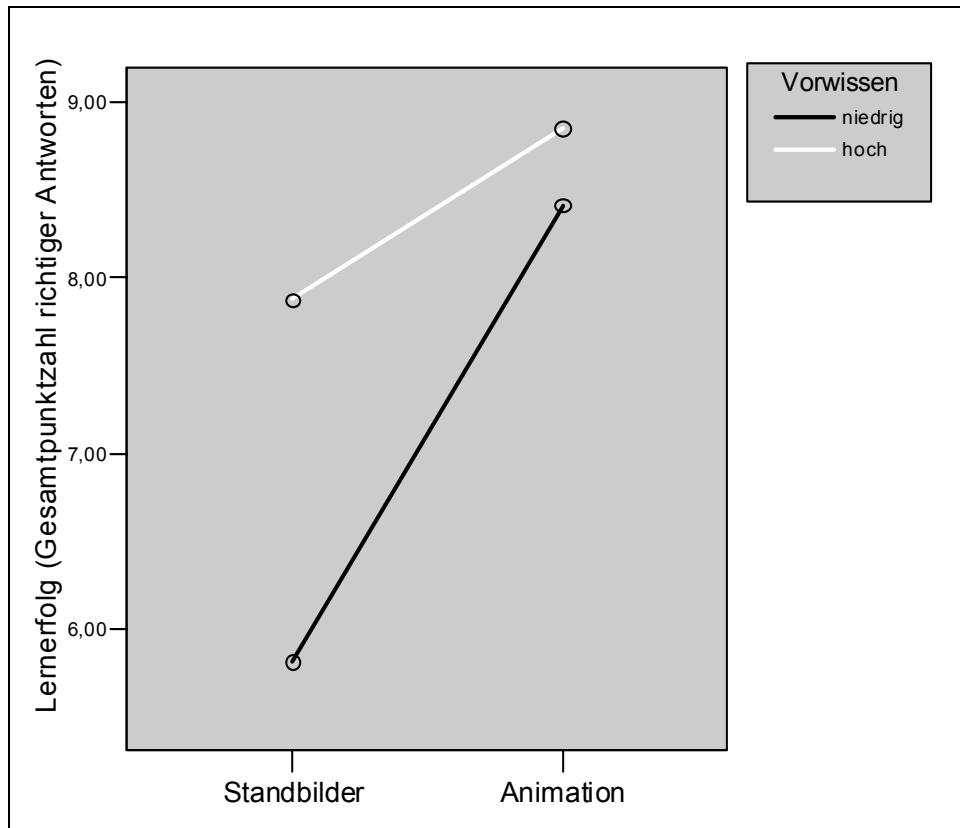
Verzichtet man auf die Durchführung des Mediansplit, der zwar den Vorteil der direkten Vergleichbarkeit zweier Gruppen mit sich bringt, dafür jedoch mit einem gewissen Informationsverlust durch die Dichotomisierung einer intervallskalierten Variablen einhergeht (vgl. Leutner & Rammsayer, 1995), und bezieht man das Vorwissen als Kovariate in die Berechnungen mit ein, ergeben sich ähnliche Ergebnisse – die Interaktion des Vorwissens und der Visualisierungsart ergibt bei sequentieller Varianzzerlegung keinerlei statistisch signifikante Resultate (Lernzuwachs gesamt:  $F(1,21) < 1$ ,  $p = 0,54$ ,  $MSE = 8,27$ ,  $\eta^2 = 0,02$ ).

*Tabelle 4-2: Ergebnisse der Varianzanalysen für Hypothese 3.*

<b>Abhängige Variable</b>	<b>Varianzquelle</b>	<b>SS*</b>	<b>df</b>	<b>MS</b>	<b>F<sub>emp</sub></b>	<b>p</b>	<b><math>\eta^2</math></b>
<b>Lernzuwachs gesamt</b>	<i>Haupteffekt Vorwissen**</i>	15,21	1	15,21	1,69	0,21	0,07
	<i>Haupteffekt Visualisierungsart</i>	21,86	1	21,86	2,43	0,13	0,10
	<i>Interaktionseffekt</i>	3,84	1	3,84	0,43	0,52	0,02
	<i>Fehler</i>	189,22	21	9,01	---	---	---
<b>Lernzuwachs Faktenwissen</b>	<i>Haupteffekt Vorwissen**</i>	14,22	1	14,22	3,82	0,06	0,15
	<i>Haupteffekt Visualisierungsart</i>	4,95	1	4,95	1,33	0,26	0,06
	<i>Interaktionseffekt</i>	0,62	1	,621	0,17	0,69	0,01
	<i>Fehler</i>	78,27	21	3,727	---	---	---
<b>Lernzuwachs Verständnis</b>	<i>Haupteffekt Vorwissen**</i>	0,02	1	0,02	0,01	0,93	0,00
	<i>Haupteffekt Visualisierungsart</i>	6,01	1	6,01	2,92	0,10	0,12
	<i>Interaktionseffekt</i>	1,38	1	1,38	0,67	0,42	0,03
	<i>Fehler</i>	43,24	21	2,06	---	---	---

\* sequentielle Varianzzerlegung

\*\* Mediansplit



**Abbildung 4-5:** Lernerfolg gesamt in Abhängigkeit vom Vorwissen und der Art der Visualisierung. Ein Interaktionseffekt mit dem Vorwissen wird zwar suggeriert, aber nicht statistisch signifikant.

#### 4.4.3 Hypothese 4: Die Rolle der Spatial Ability

Hypothese 4 geht davon aus, dass Lernende mit niedrigem räumlichen Vorstellungsvermögen stärker von Animationen im Vergleich zu Standbildern profitieren als Lernende mit hohem räumlichen Vorstellungsvermögen. Mit anderen Worten wurde ein Interaktionseffekt dergestalt erwartet, dass es für Animationen keinerlei Unterschiede machen sollte, ob der Lerner hohe oder niedrige räumliche Vorstellungskraft besitzt, bei der Präsentation von Standbildern hingegen Lernende mit niedriger *Spatial Ability* signifikant weniger lernen sollten.

Diese Hypothese kann bestätigt werden. Sowohl nach Durchführung eines Mediansplits als auch bei Einbeziehung der Gesamtpunktzahl im Paper-Folding-Test (Ekstrom et al., 1976) als Kovariate wird die Interaktion der *Spatial Ability* mit dem Lernergebnis statistisch signifikant, sofern die Vorbildung in Form der Abiturnote statistisch kontrolliert wird.

Im Falle des Vorgehens mittels Mediansplit ergeben sich zwei Gruppen mit den Mittelwerten 2,17 (SD 1,85) für niedriges und 6,69 (SD 1,32) für hohes räumliches Vorstellungsvermögen bei einem möglichen Maximalscore von 10. Tabelle 4-3 zeigt die statistischen Kennwerte beider Gruppen bezüglich des Lernerfolgs.

*Tabelle 4-3: Deskriptive Kennwerte für die Subgruppen räumlichen Vorstellungsvermögens nach der Aufteilung per Mediansplit für die AV Lernerfolg gesamt.*

<b>Visualisierungsart</b>	<b>Räumliches Vorstellungsvermögen</b>	<b>N</b>	<b>M</b>	<b>SD</b>
Standbilder	niedrig	4	4,00	1,78
	hoch	7	8,21	3,52
	gesamt	11	6,68	3,59
Animation	niedrig	7	8,21	2,10
	hoch	6	9,17	2,77
	gesamt	13	8,65	2,38
Gesamt	niedrig	11	6,68	2,85
	hoch	13	8,65	3,11
	gesamt	24*	7,75	3,09

\* Anmerkung: Die auf N=24 reduzierte Gesamtzahl erklärt sich in diesem Falle durch einen fehlenden Wert in der Kovariate „Abiturnote“.

Die Berechnung einer zweifaktoriellen Varianzanalyse mit der Abiturnote als Kovariate und sequentieller Varianzzerlegung erbringt sodann eine statistisch signifikante Interaktion ( $F(1,19)=4,40$ ,  $p<0,05$ ,  $MSE=6,69$ ,  $\eta^2=0,19$ ; vgl. Tabelle 4-4 sowie Abbildung 4-5) der erwarteten Art. Zudem wird der Haupteffekt für das räumliche Vorstellungsvermögen einseitig signifikant ( $F(1,19)=3,86$ ,  $p(\text{einseitig})=0,03$ ,  $MSE=6,69$ ,  $\eta^2=0,17$ ; vgl. Tabelle 4-4 sowie Abbildung 4-6) – Lernende mit hohem räumlichen Vorstellungsvermögen zeigen höhere Lernerfolge als solche mit niedrigem räumlichen Vorstellungsvermögen. Ebenso zeigt sich ein einseitiger statistisch signifikanter Haupteffekt für die Visualisierungsart ( $F(1,19)=3,09$ ,  $p(\text{einseitig})=0,05$ ,  $MSE=6,69$ ,  $\eta^2=0,14$ ).

Tabelle 4-4: Ergebnisse der Kovarianzanalyse für den Effekt der Spatial Ability.

Abhängige Variable	Varianzquelle	SS*	df	MS	$F_{emp}$	p	$\eta^2$
Lernzuwachs gesamt	Haupteffekt Abiturnote	17,08	1	17,08	2,55	0,13	0,12
	Haupteffekt Spatial Ability**	25,81	1	25,81	3,86	0,06	0,17
	Haupteffekt Visualisierungsart	20,67	1	20,67	3,09	0,10	0,14
	Interaktionseffekt	29,42	1	29,42	4,40	0,05	0,19
	Fehler	127,03	19	6,69	---	---	---

\* sequentielle Varianzzerlegung

\*\* Mediansplit

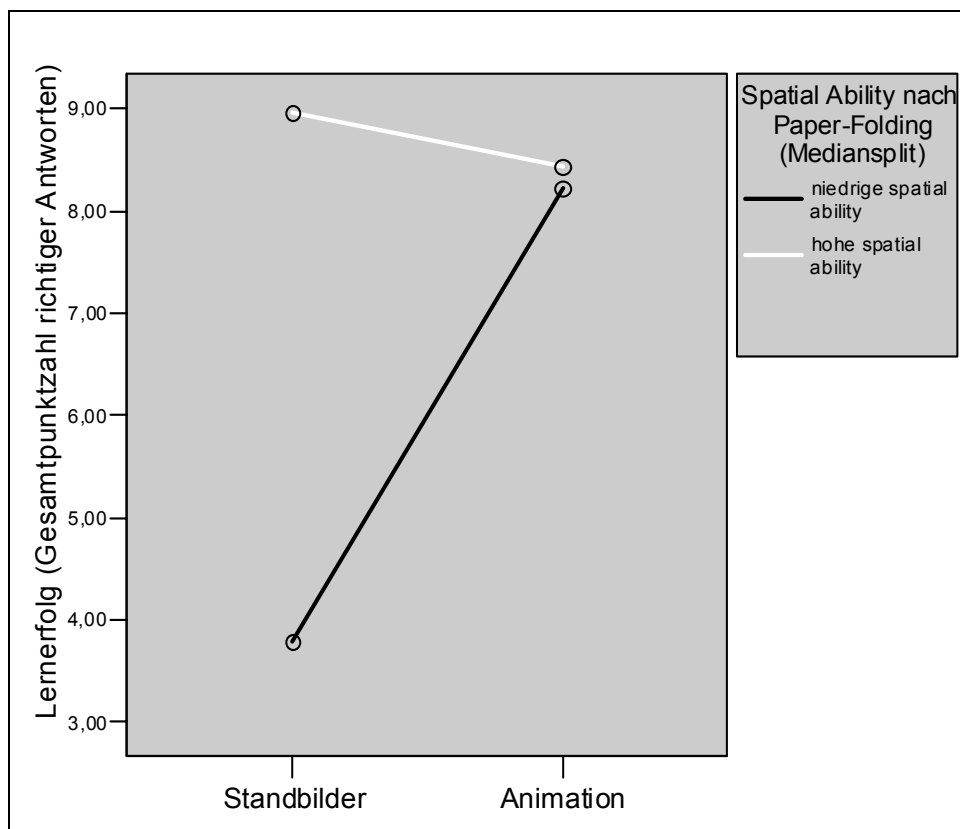


Abbildung 4-6: Lernerfolg in Abhängigkeit vom räumlichen Vorstellungsvermögen und der Art der Visualisierung unter Kontrolle der Abiturnote. Deutlich wird der Interaktionseffekt.

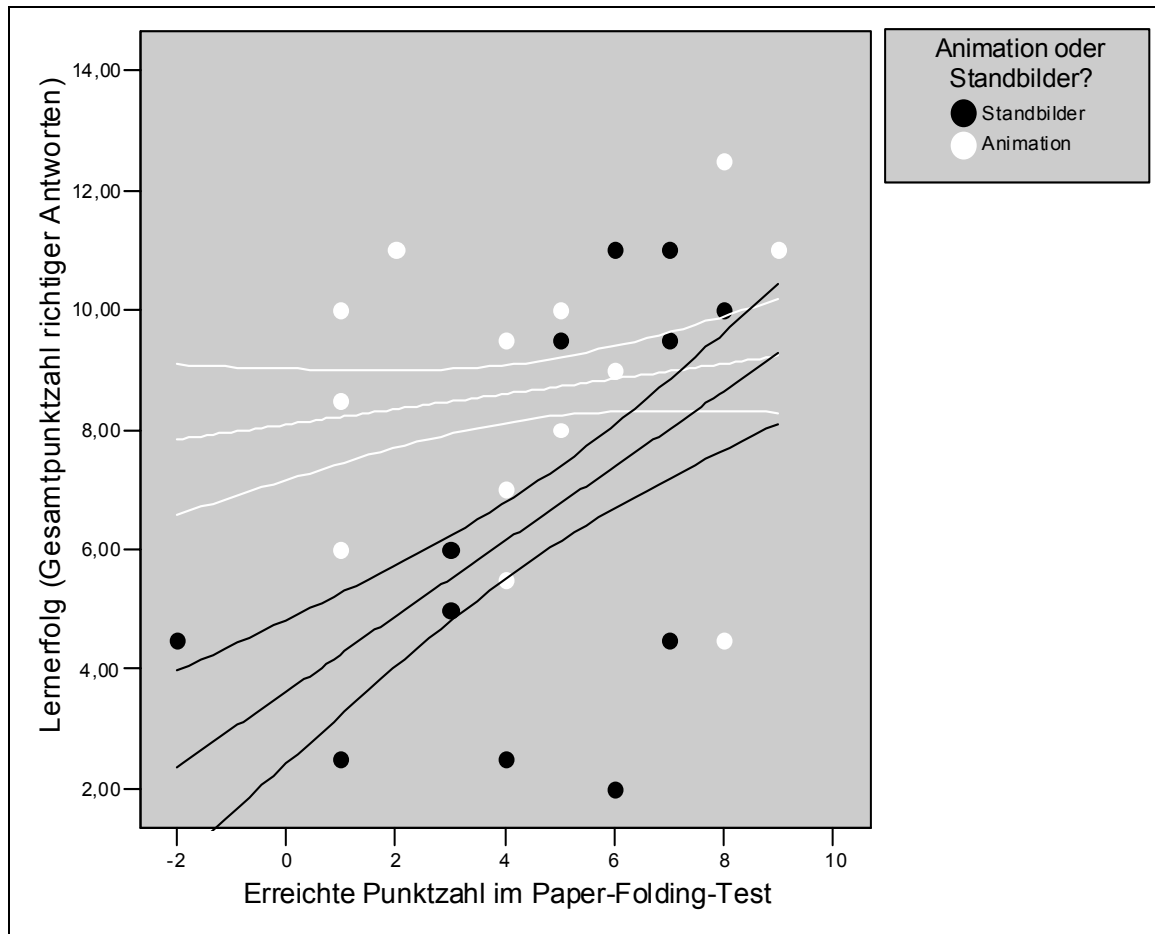


Die alternative Berechnungsart, in der das Ergebnis des Paper-Folding-Tests direkt als Kovariate in die varianzanalytischen Berechnungen mit eingeht, statt zur Einteilung in zwei Gruppen zu dienen, erbringt bei zusätzlicher Kontrolle der Abiturnote sowie sequentieller Varianzzerlegung vergleichbare Ergebnisse: Insbesondere wird auch der ATI-Effekt signifikant ( $F(1,19)=4,63$ ,  $p=0,045$ ,  $MSE=7,01$ ,  $\eta^2=0,20$ ; vgl. Tabelle 4-5 sowie Abbildung 4-7). Nicht signifikant bei dieser Berechnungsart wird jedoch der Haupteffekt für *Spatial Ability*.

*Tabelle 4-5: Ergebnisse der Kovarianzanalyse für den Effekt der Spatial Ability bei Nichtanwendung eines Mediansplits.*

<b>Abhängige Variable</b>	<b>Varianzquelle</b>	<b>SS*</b>	<b>df</b>	<b>MS</b>	<b><math>F_{emp}</math></b>	<b>p</b>	<b><math>\eta^2</math></b>
<b>Lernzuwachs gesamt</b>	<i>Haupteffekt Abiturnote</i>	17,08	1	17,08	2,44	0,14	0,11
	<i>Haupteffekt Spatial Ability</i>	19,29	1	19,29	2,75	0,11	0,13
	<i>Haupteffekt Visualisierungsart</i>	18,10	1	18,10	2,58	0,12	0,12
	<i>Interaktionseffekt</i>	32,41	1	32,41	4,63	0,05	0,20
	<i>Fehler</i>	133,12	19	7,01	---	---	---

\* sequentielle Varianzzerlegung



*Abbildung 4-7: Streudiagramm mit Lernerfolg als Y-Achse und Punktzahl im Paper-Folding-Test als X-Achse in Abhängigkeit von der jeweiligen Visualisierungsart inklusive zugehöriger Regressionslinien (mit 75%-Vertrauensintervall).*

Im Falle eines versuchsweisen Ausschlusses des in Abbildung 4-7 sichtbaren Outliers rechts unten (8;4,5) aus der Berechnung bleibt nicht nur der Interaktionseffekt einseitig signifikant ( $F(1,18)=3,06$ ,  $p(\text{einseitig})=0,049$ ,  $MSE=6,06$ ,  $\eta^2=0,15$ ), sondern sind auch die Haupteffekte sowohl für das räumliche Vorstellungsvermögen ( $F(1,18)=5,22$ ,  $p=0,035$ ,  $MSE=6,06$ ,  $\eta^2=0,225$ ) als auch für die Visualisierungsart – Animation oder Standbilder – ( $F(1,18)=5,04$ ,  $p=0,037$ ,  $MSE=6,06$ ,  $\eta^2=0,22$ ) statistisch signifikant.

#### 4.4.4 Hypothese 5: Die Rolle des Cognitive Load

Aufgrund des schnell fortschreitenden Charakters einer Animation im Gegensatz zu Standbildern hätte hier eine höhere selbstberichtete kognitive Belastung vermutet werden können. Es zeigen sich jedoch keinerlei Unterschiede zwischen den

Bedingungen hinsichtlich des *Cognitive Load*; die Hypothese kann nicht bestätigt werden.

Insgesamt ist die selbstberichtete kognitive Belastung (beide verwendeten Items wurden addiert) relativ niedrig ( $M=6,4$ ;  $MD=6$ ;  $SD=1,58$ ; bei einem möglichen Maximalscore von 14; vgl. Tabelle 4-6), beide Gruppen unterscheiden sich nicht:  $t(23)=0,20$ ,  $p=0,84$ ,  $d=0,08$ . Auch Interaktionseffekte mit anderen Variablen, hier nicht weiter im Detail dargestellt, können varianzanalytisch nicht beobachtet werden.

*Tabelle 4-6: Deskriptive Kennwerte für die Items der Cognitive Load-Skala, abhängig von der Visualisierungsart.*

<b>Cognitive Load-Item</b>	<b>Visualisierungsart</b>	<b>N</b>	<b>M</b>	<b>SD</b>
Wie leicht oder schwer war das Lernmaterial zu verstehen?	Standbilder	12	2,75	0,87
	Animation	13	2,85	0,99
Bei der Bearbeitung des Lernmaterials war meine mentale Anstrengung...	Standbilder	12	3,58	0,90
	Animation	13	3,62	0,87
Gesamt	Standbilder	12	6,33	1,56
	Animation	13	6,46	1,66

#### 4.4.5 Weitere Berechnungen

Der Vollständigkeit halber ist in Tabelle 4-7 die Korrelationsmatrix für die drei abhängigen Variablen und andere relevante Variablen angegeben. Signifikante Korrelationen sind gesondert gekennzeichnet; in diesem Fall sind dies zum Beispiel nicht weiter überraschende lineare Zusammenhänge von Vorwissen und Lernerfolg sowie zwischen den beiden Items der *Cognitive Load*-Skala. Eher unerwartet ist hingegen die statistisch signifikante Korrelation des Lernerfolgs im Verständnis mit dem Ergebnis im Paper-Folding-Test, was möglicherweise auf eine Korrelation mit allgemeiner Intelligenz hindeutet – dies ist jedoch reine Spekulation.

**Tabelle 4-7:** Mittelwerte, Standardabweichungen und Matrix der Produktmomentkorrelationen relevanter Variablen.

	M	SD		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lernerfolg gesamt (1)	7,62	3,10	r	1									
			N	25									
Lernerfolg Fakten- wissen (2)	5,74	2,02	r	0,92**	1								
			N	25	25								
Lernerfolg Verständnis (3)	1,88	1,45	r	0,85**	0,58**	1							
			N	25	25	25							
Abiturnote (4)	2,65	0,54	r	-0,28	-0,25	-0,25	1						
			N	24	24	24	24						
Chemienote (5)	8,62	2,62	r	-0,25	-0,39	0,01	-0,49*	1					
			N	21	21	21	21	21					
Vorwissen (6)	1,82	1,61	r	0,41*	0,49*	0,19	-0,38	-0,43	1				
			N	25	25	25	24	21	25				
Paper- Folding-Test (7)	4,52	2,79	r	0,34	0,20	0,44*	0,08	-0,15	-0,18	1			
			N	25	25	25	24	21	25	25			
CL 1 (8)	2,80	0,91	r	-0,18	-0,28	0,00	0,28	0,10	-0,22	-0,14	1		
			N	25	25	25	24	21	25	25	25		
CL 2 (9)	3,60	0,87	r	-0,21	-0,29	-0,04	0,36	0,20	-0,16	-0,14	0,58**	1	
			N	25	25	25	24	21	25	25	25	25	
CL gesamt (10)	6,40	1,58	r	-0,22	-0,32	-0,02	0,36	0,16	-0,22	-0,15	0,90**	0,88**	1
			N	25	25	25	24	21	25	25	25	25	25

\* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

\*\* Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Die Reliabilitäten der eingesetzten Skalen sind durchweg zufriedenstellend: Die selbstentwickelte Skala zum Vorwissen, aus vier Items bestehend, zeigte dabei noch die niedrigste Reliabilität (Cronbachs Alpha = 0,68), die jedoch durch den Verzicht auf eines der Items<sup>6</sup> auf einen guten Wert von  $\alpha=0,80$  gesteigert werden könnte. Der Lernerfolg verzeichnet als Gesamtskala ein Alpha von ebenfalls 0,80; jedoch wurden für obige Berechnungen grundsätzlich lediglich die sieben Items verwendet, die sich explizit auf den Bereich der Visualisierung (Animation oder Standbilder) beziehen

<sup>6</sup> Dabei handelt es sich um die vierte Frage, die den Begriff der Oberflächenspannung erläutert wissen will. Tatsächlich fällt diese Frage etwas aus dem Rahmen, da die anderen Fragen des Vortests eher auf den später dargestellten Prozess abzielen, nicht auf Hintergrundwissen.

und nicht auf die einbettende Lernumgebung. Diese erreichen ein ebenfalls befriedigendes Cronbachs Alpha von 0,71. Schließlich beträgt die Reliabilität der Skala zur *Cognitive Load*, aus zwei Items bestehend,  $\alpha=0,73$ .

#### 4.5 Diskussion

Diese erste experimentelle (Pilot-)Studie erbrachte zum großen Teil die erwarteten Ergebnisse. Die eingesetzten Lernumgebungen führten zu Lernleistungen, die wie erwartet bei der Animationsversion statistisch signifikant höher ausfielen als bei der Standbildversion. Diese Annahme beruhte auf den Ergebnissen der vorhergehenden Metaanalyse und wurde nun bestätigt. Jedoch erfordert die bei einer Pilotstudie naturgemäß recht kleine Stichprobe selbstverständlich eine Replikation dieses Ergebnisses, wie es auch von vornherein geplant war. Der Nachweis der prinzipiellen Überlegenheit der Animation gegenüber den (vier) Standbildern von Schlüsselementen des Prozesses legitimiert zugleich die Einleitung des nächsten Schrittes, der Einführung einer weiteren Lernumgebungsversion, in der der Prozess durch eine deutlich größere Anzahl von Standbildern repräsentiert wird.

In Übereinstimmung mit gängiger Auffassung (Mayer & Moreno, 2002), nur eingeschränkt jedoch mit den Ergebnissen der Metaanalyse vereinbar, ist das Ergebnis zu interpretieren, dass die Vorteile der Animation gegenüber den Standbildern zum großen Teil auf Verständnisaufgaben zurückzuführen sind, die Ergebnisse zum Faktenwissen sich hingegen nicht signifikant unterschieden. Natürlich erscheint die Annahme sinnvoll, dass Animationen durch ihre exaktere Wiedergabemöglichkeit eines Prozesses besonders tieferes Verständnis dieses Prozesses anregen können, während für reines Faktenwissen Standbilder ausreichen sollten – genau dies scheint diese Studie zu implizieren. Jedoch (vgl. Kapitel 3) konnte die Metaanalyse einen solchen spezifischen Vorteil nicht belegen; Animationen waren Standbildern gleichermaßen beim Faktenwissen wie auch beim tieferen Verständnis überlegen. Es erscheint allerdings plausibel, dass dieser Unterschied in besonders hohem Maße kontextabhängig ist. Es ist gut denkbar, dass in bestimmten Kontexten Standbilder vollkommen ausreichen können, um tieferes Verständnis beim Lernenden anzuregen, wohingegen in anderen Kontexten reine Fakten mit Hilfe von Animationen besser vermittelt werden können. Die Rolle des

Kontexts zu kontrollieren durch z.B. eine hohe Variation desselben ist leider im Rahmen der hier vorliegenden Dissertation nicht möglich gewesen, da es einen zu hohen Aufwand bei der Entwicklung der Lernumgebungen bedeutet und zudem den Fokus der Arbeit deutlich verschoben hätte.

Aus diesem Grunde bleibt auch bei der Fragestellung nach der Rolle der *Spatial Ability* ein gewisser Unsicherheitsfaktor hinsichtlich der Verallgemeinbarkeit der Ergebnisse. Nichtsdestotrotz indiziert das Ergebnis zum räumlichen Vorstellungsvermögen eine kompensierende Funktion der *Spatial Ability* beim Lernen mit Standbildern versus Animationen der verwendeten Art und damit mutmaßlich eine maßgebliche Rolle bei der mentalen Modellbildung aus externen Visualisierungen: Die Nachteile einer weniger geeigneten Visualisierungsform (in diesem Falle die Serie von Standbildern) können durch ein erhöhtes Maß räumlichen Vorstellungsvermögens kompensiert werden. Umgekehrt gilt, dass der Einsatz von Animationen mangelndes räumliches Vorstellungsvermögen ausgleichen und zu einem ebenso guten Lernergebnis führen kann. Dieses hypothesenkonforme Ergebnis erscheint schlüssig, auch wenn es teilweise der Literatur widerspricht (z.B. Isaak & Just, 1995; Mayer, 2001). Es bleibt selbstverständlich noch, das Ergebnis zu replizieren und diesem damit eine größere Aussagekraft zu verschaffen.

Hinsichtlich des themenspezifischen Vorwissens konnte kein moderierender Effekt nachgewiesen werden, obwohl der *Expertise Reversal Effect* bereits gut dokumentiert ist (Kalyuga et al., 1998; 2000; 2003). Hierzu ist zu bemerken, dass bei der Planung des Experiments zwei einander widersprüchliche Forschungsinteressen hätten miteinander in Einklang gebracht werden müssen: Zum einen wäre für den Nachweis des *Expertise Reversal Effect* eine hinsichtlich ihres Vorwissens möglichst heterogene Stichprobe nötig gewesen, um tatsächliche „Experten“ und „Novizen“ zu ermitteln. Zum anderen jedoch ist es sinnvoll, Probanden mit möglichst geringem Vorwissen zu rekrutieren, damit bei dem relativ einfachen Lerninhalt, der hier zum Einsatz kam, überhaupt noch Lerneffekte nachweisbar sind. Das zweitgenannte Interesse erschien nach sorgfältiger Abwägung schließlich gewichtiger, da der Unterschied zwischen Animationen und Standbildern die Primärhypothese und der *Expertise Reversal Effect* bereits hinreichend dokumentiert ist.

Schließlich ist zum Effekt von *Cognitive Load* (Hypothese 5: Animationen sollten zu einer höheren kognitiven Belastung führen als Standbilder) zu bemerken, dass

der Lerngegenstand offensichtlich recht einfach war und damit der *Intrinsic Load* sehr niedrig, was wohl hauptsächlich durch die gemessenen Selbstauskünfte der Lernenden widerspiegelt werden dürfte. Da sich keinerlei Unterschiede zwischen den Visualisierungsarten hinsichtlich des *Cognitive Load* ergaben, wurden wohl auch keine Effekte hinsichtlich eines *Extraneous Load* gefunden, auch wenn zu erwarten gewesen wäre, dass zum Verstehen des Inhalts der schnell fortschreitende Charakter der Animation eine höhere kognitive Herausforderung darstellt als Standbilder. Natürlich ist dies reine Spekulation, da es bei dem hier eingesetzten Testinstrument zur Erfassung des *Cognitive Load*– wie auch bisher bei keinem sonstigen Messinstrument –nicht möglich ist, zwischen den einzelnen theoretischen Facetten des *Cognitive Load* zu unterscheiden. Somit ist zwar theoretisch ableitbar, dass zwischen den Versionen Unterschiede hinsichtlich der *Cognitive Load*-Facetten bestehen sollten; es ist bisher aber per Messung noch nicht möglich, diese Unterschiede empirisch abzusichern.

## 5 Zweite experimentelle Studie

### 5.1 Einführung

Die zweite experimentelle Studie baut direkt auf der oben beschriebenen Pilotstudie auf. Die erfolgreich erprobten Materialien werden nun ohne weitreichende Änderungen an einer größeren Stichprobe getestet, um die gefundenen Ergebnisse zu replizieren und die Interpretationen zu überprüfen. Zwar erwiesen sich das Lernmaterial tendenziell als relativ niedrig kognitiv belastend und somit möglicherweise als zu leicht, trotzdem wurde entschieden, keine Veränderungen vorzunehmen, und zwar aus folgenden Gründen: Zum einen wird die Stichprobe dieses Mal aus Schülern rekrutiert, um so zusätzlich die allgemeine Übertragbarkeit auf das Setting der Schule nachzuweisen und eine Unterrichtsrelevanz und breitere Gültigkeit der Hypothesen zeigen zu können. Dies bedeutet aber auch, dass in betracht gezogen werden muss, dass die Schülerinnen und Schüler evtl. ein noch niedrigeres Vorwissensniveau haben und möglicherweise weniger schnell und nachhaltig aus den computerbasierten Lernumgebungen lernen. Zum anderen ist als Grund für die Nichtveränderung der Materialien anzuführen, dass das Ausmaß der kognitiven Belastung der Probanden nur ein Nebenaspekt dieser Reihe experimenteller Untersuchungen ist. Im Vordergrund steht die Frage der Variation der Anzahl von Standbildern im Vergleich zu einer Animation sowie der Aspekt der *Spatial Ability*. Da eine zu anspruchsvolle Lernumgebung das Auffinden solcher Effekte unter Umständen verhindern könnte, wird eine insgesamt niedrige kognitive Belastung als in Kauf zu nehmender Nachteil hingenommen. Deshalb wird auch von einer weiteren Berücksichtigung dieser Variable in einer separaten Hypothese abgesehen, aber natürlich weiterhin mit erhoben.

Zusätzlich wird nun wie bereits vor Durchführung der ersten experimentellen Studie geplant eine dritte Versuchsbedingung eingeführt: Neben den Visualisierungsarten „Animation“ und „4 Standbilder“ existiert nun die Bedingung „11 Standbilder“, die den Ablauf des dargestellten Prozesses noch weiter untergliedert und deutlich mehr Zwischenschritte darstellt als die vier Standbilder. Auf diese Weise, so die Voraussage, sollten ebenso gute Lernergebnisse wie mit der Animation gefunden werden können. Zusätzlich wird die *Spatial Ability* mit einem weiteren Test erhoben, um den moderierenden Effekt des räumlichen



Vorstellungsvermögens unter Umständen weiter eingrenzend auf unterschiedliche Subdimensionen beziehen zu können.

## 5.2 Zusätzliche theoretische Erörterungen

*Spatial Ability* wurde in vorliegender Arbeit zunächst über den Paper-Folding-Test (Ekstrom et al., 1976) operationalisiert. Eliot & Smith (1983) weisen jedoch darauf hin, dass *Spatial Ability* bereits in so einer Fülle von unterschiedlichen Weisen definiert worden ist, dass es schwierig sei, sich jeweils im Klaren darüber zu sein, welche Bedeutung im Speziellen gemeint sei. Tatsächlich konnte Carroll (1993) in einer Reanalyse vorangegangener Faktorenanalysen fünf verschiedene Faktoren identifizieren, aus denen sich räumliches Vorstellungsvermögen zusammensetzen scheint:

- *Spatial Visualization*: Dieser Faktor fasst die Fähigkeiten zusammen, räumliche Formen zu verstehen, zu kodieren und mental zu manipulieren (Carroll, 1993, S. 309), was eine intensive Beteiligung exekutiver Kontrollprozesse nötig macht, die die in Tests wie dem Paper-Folding-Test gestellten Probleme strukturieren, analysieren und die daraufhin ausgearbeiteten Lösungsstrategien überwachen (Marshalek, Lohman & Snow, 1983, S. 124).
- *Spatial Relations*: Zwar sind Tests, die auf diesem Faktor laden (wie z.B. der Card-Rotation-Test; Ekstrom et al., 1976) *Spatial Visualization*-Tests dahingehend ähnlich, dass für deren Lösung ebenfalls mentale Transformationen nötig sind, als hervorstechendes Kennzeichen fordern sie jedoch Manipulationen (zumeist Rotationen) von 2D-Objekten in einem einzelnen Schritt in möglichst kurzer Zeit (Miyake, Friedman, Rettinger, Shah & Hegarty, 2001).
- *Closure Speed*: Dieser Faktor beinhaltet die Fähigkeit, Figuren zu erkennen, die in irgendeiner Weise erschwert wahrnehmbar sind, wie etwa durch einen ablenkenden oder tarnenden Hintergrund.
- *Closure Flexibility*: Der einzige Unterschied dieses Faktors zu dem vorhergehenden ist, dass der Proband im Vorhinein weiß, welches Muster er zu identifizieren hat. D.h. zusätzlich ist bei Tests, die auf diesem Faktor

hoch laden (z.B. der Hidden-Figures Test; Ekstrom et al. 1976), ein gewisses Maß an visuell-räumlicher Speichermöglichkeit vonnöten.

- *Perceptual Speed*: Bei diesem Faktor geht es darum, möglichst schnell einfache Figuren wahrzunehmen und mit anderen zu vergleichen, um auf diese Weise gleiche Objekte zu identifizieren. Dabei ist eine mentale Rotation ausdrücklich nicht notwendig.

Von diesen fünf Faktoren erscheinen lediglich die ersten beiden als höhere kognitive Herausforderungen, die auch tatsächlich beim Lernen mit dynamischen oder nichtdynamischen Visualisierungen eine Rolle spielen könnten. Mit Einschränkungen könnte eventuell auch der Faktor *Perceptual Speed* dazu zählen, wenn es wichtig ist, gleich aussehende Objekte als gleich zu klassifizieren. Mustererkennung vor unklarem Hintergrund dürfte, zumindest für diese Untersuchung, keine notwendige Fähigkeit sein.

Dass jedoch *Spatial Visualization*, messbar mit dem Paper-Folding-Test, eine wichtige Fähigkeit bei der Bearbeitung von Animationen ist, ist eine naheliegende Hypothese, die durch die Pilotstudie bereits gestützt wurde. Ähnliches könnte für *Spatial Relations* gelten – diesem soll in folgender Untersuchung auf den Grund gegangen werden.

### 5.3 Fragestellung und inhaltliche Hypothesen

Die Hypothesen der ersten Studie bedürfen für die zweite Untersuchung keiner wesentlichen Änderung, auch wenn sich nicht alle Hypothesen im ersten Anlauf bestätigen ließen. Weiterhin steht die unterschiedliche Wirkung dynamischer und nichtdynamischer Visualisierungen im Bereich der Chemie hinsichtlich ihrer Lernwirksamkeit im Vordergrund, während zudem auf verschiedene lernerabhängige Variablen hinsichtlich ihrer Moderatorwirkung fokussiert wird.

Lediglich im Hinblick auf die dritte, neu eingeführte Versuchsbedingung müssen die inhaltlichen Hypothesen angepasst werden:

***H1a: Mit der animierten Lernumgebung werden bessere Lernleistungen erzielt als mit der aus 4 statischen Bildern bestehenden Lernumgebung.***

***H1b: Es bestehen hinsichtlich der Lernleistung keine Unterschiede zwischen der Animation und der mit 11 Standbildern ausgestatteten Lernumgebung. Diese ist wiederum der 4-Standbilder-Version überlegen.***

Trotz gegenteiliger Ergebnisse der Pilotstudie wird die zweite Hypothese, den Unterschied hinsichtlich Faktenwissens und tieferen Verständnisses betreffend, aufgrund der Ergebnisse der Metaanalyse vorerst beibehalten:

***H2: Die Überlegenheit der Animation zeigt sich sowohl bei Aufgaben, die Faktenwissen erfragen, als auch bei solchen, für die tieferes Verständnis erforderlich ist.***

Die dritte Hypothese bezüglich des themenspezifischen Vorwissens verändert sich nicht und wird nur für die zusätzliche Versuchsbedingung angepasst:

***H3: Die lernförderliche Wirkung von Animationen lässt sich besonders für Lernende mit niedrigem themenspezifischen Vorwissen nachweisen. Lerner mit hohem Vorwissen erzielen mit 4 stehenden Bildern mindestens genauso gute Ergebnisse wie mit Animationen oder 11 Standbildern.***

Die Pilotstudie lieferte einen moderierenden ATI-Effekt der *Spatial Ability* im Sinne einer kompensierenden Funktion: Mangelndes räumliches Vorstellungsvermögen kann durch Vorlage einer Animation ausgeglichen werden, bzw. offensichtlich nicht optimale Visualisierungen (4 Standbilder) können mit erhöhtem räumlichen Vorstellungsvermögen kompensiert werden. Dieses Ergebnis soll repliziert werden, so dass die vierte Hypothese in Bezug auf die 4-Standbilderversion natürlich gleich bleibt. Was die 11-Standbilderversion betrifft, so gilt es, diametrale Möglichkeiten zu berücksichtigen: Zum einen besteht diese Version natürlich nach wie vor aus Standbildern, unterscheiden sich von ihrem Charakter her also nicht wesentlich von der 4-Standbilderversion – es ließe sich somit argumentieren, dass ein erhöhtes räumliches Vorstellungsvermögen hier im gleichen Maße nötig ist, um adäquaten Lernerfolg zu erzielen. Auf der anderen Seite besteht die Vorstellung, dass gerade die erhöhte Frequenz an Standbildern (wie auch Animationen durch einen – extrem – hochfrequenten Wechsel von Standbildern zustande kommen) der entscheidende Faktor sein könnte, um Defizite im räumlichen Vorstellungsvermögen kompensieren

zu können und die Bildung einer mentalen Animation (Hegarty et al., 2003) anzuregen. Die erweiterte vierte Hypothese lautet deswegen:

***H4: Lernende mit niedrigem räumlichen Vorstellungsvermögen profitieren von Animationen im Vergleich zu 4 Standbildern in höherem Maße als Lernende mit hohem räumlichen Vorstellungsvermögen. Ebenso profitieren Lernende mit niedrigem räumlichen Vorstellungsvermögen von 11 Standbildern im Vergleich zu 4 Standbildern in höherem Maße als Lernende mit hohem räumlichen Vorstellungsvermögen.***

Bei dieser vierten Hypothese ist jedoch zu beachten, dass sie sich, nach den Ergebnissen der Pilotstudie, bisher lediglich auf ein Teilkonstrukt des räumlichen Vorstellungsvermögens bezieht, die *Spatial Visualization*. Es bleibt abzuwarten, ob ein ähnlicher Effekt auch für den Bereich der *Spatial Relations* gilt.

Ausgehend von den Ergebnissen in der vorherigen Studie, die nur ein sehr geringes Niveau des selbstberichteten *Cognitive Load* fand, wird in dieser Studie keine separate Hypothese zu dieser Variable mehr verfolgt.

## 5.4 Methode

### 5.4.1 Variablen

#### *Unabhängige Variable (UV)*

Als Unabhängige Variable in dieser Untersuchung dient weiterhin die *Art der Visualisierung*, also ob dem Lernenden eine Animation, eine Serie von vier oder elf Standbildern präsentiert wird.

#### *Abhängige Variablen (AV)*

Die Abhängige Variable stellt weiterhin der *Lernerfolg* dar, gemessen anhand der erzielten Punkte im Nachtest. Dieser wird weiterhin differenziert in offene und Multiple-Choice-Fragen, ferner in Fragen nach Faktenwissen und tiefer gehendem Verständnis.

Der *Cognitive Load* wird nach wie vor mit zwei adaptierten Items von Paas (1992) erhoben, auch wenn der angewandte Test zwar gute Reliabilitätswerte erzielte, nicht

jedoch Unterschiede zwischen den Visualisierungsarten finden konnte. Andere Messmethoden als die Selbstauskunft, wie etwa die sogenannte *Dual-Task-Methode* (Brünken, Steinbacher, Plass & Leutner, 2002), würden einen zu hohen Aufwand erfordern, der nicht durch die Rolle des *Cognitive Load* als lediglich am Rande interessierender Untersuchungsaspekt gerechtfertigt wird.

#### *Moderatorvariablen*

Mutmaßliche Moderatorvariablen bleiben weiterhin die *Höhe des themenspezifischen Vorwissens* (erhoben mit denselben vier Items aus der Pilotstudie) sowie die *Spatial Ability*, gemessen mit Hilfe sowohl des bereits eingesetzten Paper-Folding-Tests von Ekstrom et al. (1976) als auch mittels des Card-Rotation-Tests derselben Autoren. Zwar sollen beide Tests das räumliche Vorstellungsvermögen erfassen, es ist jedoch möglich, dass unterschiedliche Facetten desselben Konstrukts getestet werden. Darum erscheint ein zweiter Test sinnvoll.

#### *Störvariablen*

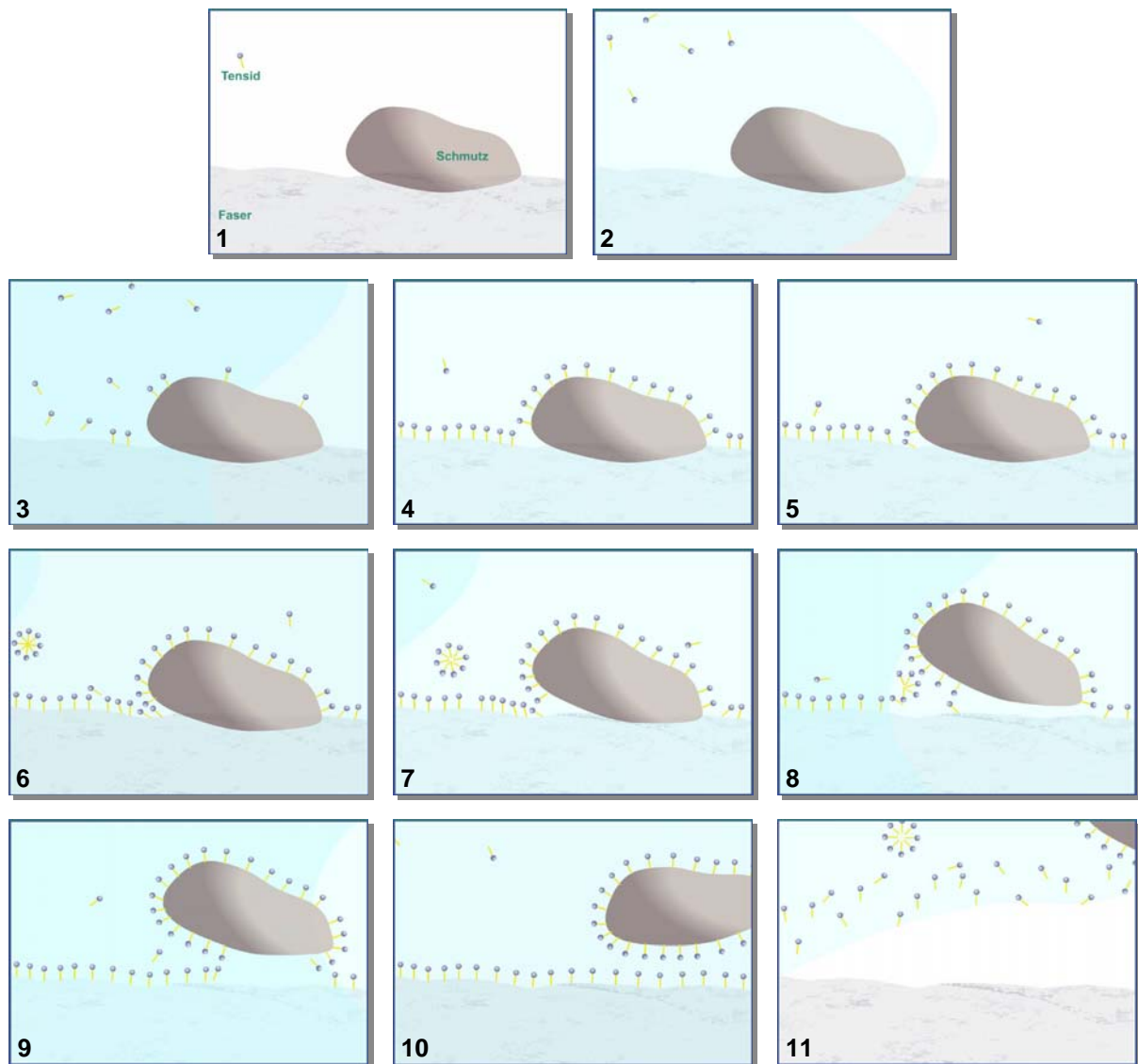
Da das Design der Studie weitgehend gleich geblieben ist, ergeben sich keine zusätzlichen potentiellen Störvariablen als bei der Pilotstudie (vgl. Kapitel 4.3.4).

### **5.4.2 Materialien**

#### *Computergestützte Lernumgebungen*

Es werden dieselben computergestützten Lernumgebungen verwendet wie in der Pilotstudie. Zusätzlich zu den beiden bereits existierenden Versionen, die den Vorgang der Schmutzablösung beim Waschen durch Tenside mittels derselben Lernumgebung darstellen, sich jedoch an entscheidender Stelle entweder durch eine Serie von vier Standbildern oder eine Animation auszeichnen, wird eine dritte Version eingeführt. Diese stimmt in den Grundzügen der formalen Gestaltung und Bedienung mit den anderen Versionen überein, ebenso im Inhalt der umgebenden Lernumgebung. An die Stelle der Animation bzw. der vier Standbilder tritt nun aber eine Serie von elf Standbildern, unterlegt mit derselben verbalen Erläuterung (vgl. Anhang G) wie bei den anderen Versionen und in derselben Länge (73 Sekunden).

Die zusätzlichen Bilder, die als Screenshots aus der Animation generiert wurden, unterteilen die vier Phasen des Prozesses der Schmutzablösung weiter und stellen zwar inhaltlich sinnvolle, jedoch zum Verständnis des Prozesses nicht zwingend notwendige Visualisierungen dar. Durch die Begrenzung der Länge der Darstellung auf die der Animation wird jetzt jedes Bild nur noch im Mittel 6,6 Sekunden gezeigt (Min: 4,1 sec, Max: 15,9 sec), während dies im Falle der vier Standbilder noch 18,3 Sekunden waren. Abbildung 5-1 zeigt die gesamte Serie der ausgewählten Standbilder.



*Abbildung 5-1: Die Serie der elf verwendeten Standbilder, hier gemeinsam dargestellt. Die Nummerierungen geben die Reihenfolge der Darbietung an und waren nicht Teil der Lernumgebung.*

### *Fragebögen*

Die Fragebögen der Pilotstudie werden für diese Untersuchung aufgrund insgesamt befriedigender Reliabilitäten nur geringfügig im Hinblick auf formale Kriterien (andere Probanden, daher andere Anrede etc.) geändert sowie durch einen weiteren Test zum räumlichen Vorstellungsvermögen ergänzt (vgl. Anhänge D und F).

### *Vortest*

Da für diese zweite Studie eine Rekrutierung von Probanden im schulischen Bereich vorgesehen ist, muss dieser auf Studierende angelegte Fragebogen im Teil zum biographischen Werdegang entsprechend angepasst werden. So wird nun nicht mehr nach der Abiturnote gefragt, sondern nach den letzten in Chemie und Biologie erworbenen Zeugnisnoten. Fragen zum Studium entfallen natürlich.

Die vier Fragen zur Erfassung des themenspezifischen Vorwissens werden beibehalten, auch wenn die vierte Frage im Pilottest zu einer deutlichen Erniedrigung der Reliabilität geführt hat. Im Sinne einer Erfassung eines breiteren Vorwissensstandes erscheint der Einbezug auch dieses Items jedoch gerechtfertigt.

Zusätzlich zum Paper-Folding-Test nach Ekstrom et al., 1976 (Anhang E) enthält der Vortest eine weitere Messung der *Spatial Ability*, den Card-Rotation-Test aus derselben Testbatterie (Anhang F). So soll eine zusätzliche Facette des räumlichen Vorstellungsvermögens (der *Spatial Relations*-Faktor) erhoben werden.

Der Card-Rotation-Test besteht aus einer Seite Instruktion und einer Seite mit zehn Aufgaben desselben Typs: Eine geometrische Beispielfigur wird jeweils in acht Variationen dargestellt. Für jede Variation muss der Proband entscheiden, ob es sich um die gleiche Figur handelt, die lediglich in der Ebene gedreht wurde, oder ob sie zusätzlich gespiegelt wurde, es also nicht mehr die „gleiche“ Figur ist. Das Ziel ist es, in drei Minuten so viele Aufgaben wie möglich richtig zu lösen – Fehler geben Punktabzug.

### *Nachtest*

Der Nachtest (Anhang C) wird im Vergleich zur Pilotstudie nicht verändert.

### **5.4.3 Räumlichkeiten und Hardware**

Die Probanden arbeiten alle an Computern neuerer Generation, die eine angemessene Auflösung und Darstellungsgeschwindigkeit aufweisen. Dazu werden die vorhandenen Computerräume der teilnehmenden Schulen sowie zusätzliche Laptops verwendet. Jeder der Teilnehmer trägt zum Zeitpunkt der Bearbeitung der Lernumgebung einen Kopfhörer, über den er den Audiokommentar störungsfrei aufnehmen kann.

### **5.4.4 Untersuchungsablauf**

Eine Untersuchung dauerte etwa 50 Minuten, der Ablauf verlief weitgehend analog zu dem der Pilotstudie. Nach einer kurzen Begrüßung, Einführung und Aufklärung über die Anonymität der Untersuchung bekamen die Probanden den ersten Fragebogen (Anhang D) mit der Auflage zugeteilt, ihn zügig, aber trotzdem so genau wie möglich zu beantworten. Dies nahm in etwa 15 Minuten in Anspruch. Dabei wurde der Paper-Folding-Test (Anhang E), der Bestandteil des ersten Fragebogens war, aufgrund seines Speed-Test-Charakters gleichzeitig mit allen Teilnehmern durchgeführt, woraufhin anschließend der Card-Rotation-Test (Anhang F) im selben Stil folgte.

Nach der Rückgabe der Fragebögen begann, gleichzeitig für alle Probanden, die Phase der individuellen Bearbeitung der jeweiligen Lernsoftware. Die Probanden wurden dabei angewiesen, still und für sich zu arbeiten und die Kopfhörer aufzusetzen. Für eventuell auftauchende technische Probleme stand der Versuchsleiter zur Verfügung, für fachliche bzw. den Programmablauf betreffende Fragen jedoch ausdrücklich nicht. Nachdem alle Teilnehmer die kurze Software in ca. zehn Minuten bearbeitet hatten, wurden die Rechner ausgeschaltet und der erste Teil des Posttests verteilt. Dieser sollte ebenfalls einzeln bearbeitet werden. Nach dessen Beendigung erhielt jeder Proband im Austausch den zweiten Teil des Nachtests.

Sobald jeder Proband diesen ebenfalls abgeschlossen hatte, bedankte sich der Versuchsleiter für die Teilnahme und bot noch eine Frage- und Diskussionsrunde zum Hintergrund der Untersuchung an.

Die drei Untersuchungen dieser Studie fanden am 22. Februar (im Gymnasium Baesweiler) sowie 22. März 2006 (in der Gesamtschule Kerpen) statt.



### 5.4.5 Probanden

An dieser Studie nahmen insgesamt 62 Probanden aus drei verschiedenen 11. Klassen zweier Schulen aus Nordrhein-Westfalen teil (Gymnasium Baesweiler sowie Gesamtschule Kerpen). Das Durchschnittsalter betrug 17 Jahre, 40,6% der Teilnehmer waren weiblich. Über zwei Drittel waren Gymnasiasten, die anderen Gesamtschüler. Die Durchschnittsnote in Chemie betrug 2,84, in Biologie 2,55.

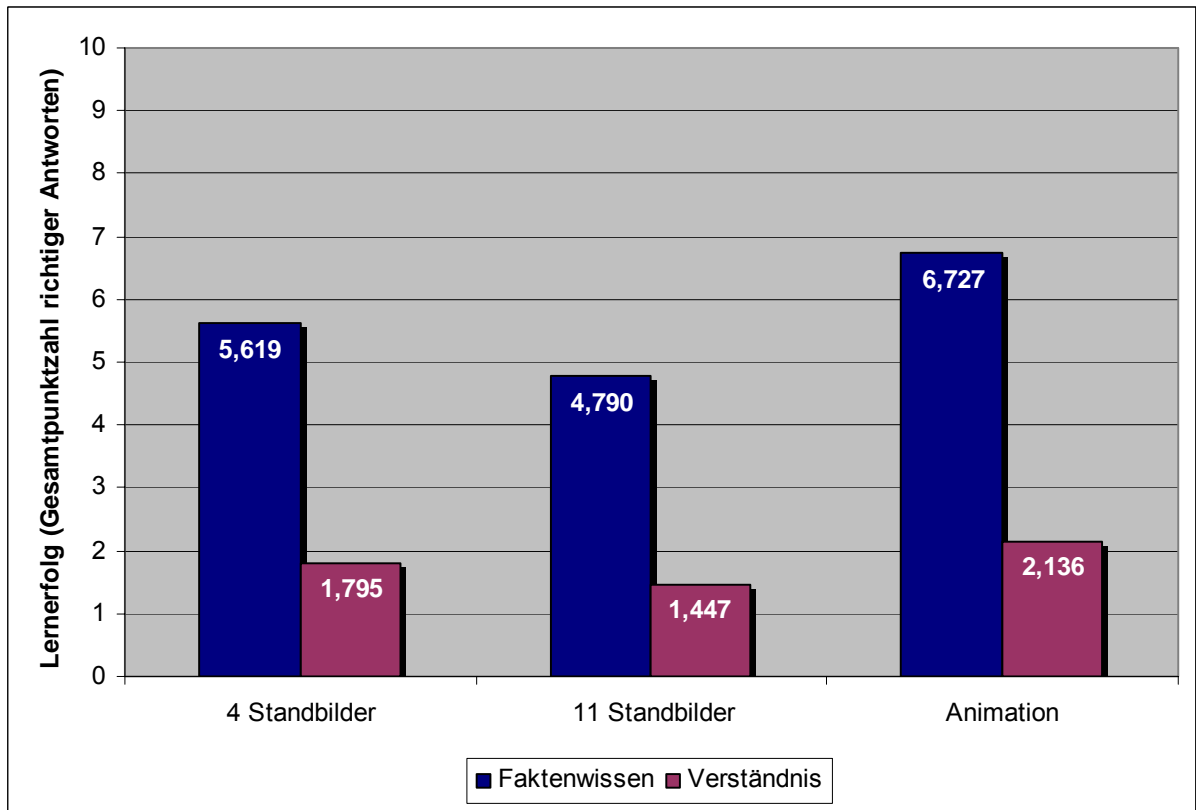
## 5.5 Ergebnisse

In diesem Kapitel werden die statistischen Ergebnisse dieser Studie dargestellt. Fast alle Kalkulationen wurden mit dem Statistikprogramm SPSS 12.02 (2004) durchgeführt, einzelne Berechnungen der Effektstärke mit dem Programm GPOWER von Faul & Erdfelder (1992) mit Hilfe des Maßes der gepoolten Streuung. Die Darstellung der Ergebnisse erfolgt anhand der zuvor aufgestellten Hypothesen.

### 5.5.1 Hypothesen 1a, 1b und 2: Animationen versus Standbilder

Hypothese 1a dient der Replikation der Pilotstudie und drückt aus, dass die Animationsversion insgesamt der 4-Standbilderversion überlegen sein sollte, wohingegen sich Hypothese 1b mit der Rolle der neu eingeführten 11-Standbilderversion beschäftigt.

Die Bestätigung der Hypothesen glückt nur teilweise. Insgesamt verfehlt der Vergleich der drei Gruppen untereinander das 5%-Signifikanzniveau nur knapp ( $F(2,59)=3,06$ ,  $p=0,055$ ,  $MSE=11,63$ ,  $\eta^2=0,09$ ; vgl. Tabelle 5-2). Der Lernerfolg ist aber wie erwartet bei der Animationsversion deskriptiv höher als bei der 4-Standbilderversion (vgl. Abbildung 5-2, Tabelle 5-1), unter Einbeziehung des Vorwissens als Kovariate wird dieser Unterschied auch in der erwarteten Richtung statistisch signifikant:  $F(1,39)=3,68$ ,  $p(\text{einseitig})=0,032$ ,  $MSE=10,62$ ,  $\eta^2=0,09$ . Hypothese 1a kann also als bestätigt angesehen werden.



*Abbildung 5-2: Mittelwerte der erzielten Lernleistungen der Probanden mit Standbilder- oder Animationsversion, aufgeteilt nach Faktenwissen und Verständnisaufgaben.*

*Tabelle 5-1: Deskriptive Kennwerte des Lernerfolgs, aufgeteilt nach Visualisierungsart.*

Abhängige Variable	Visualisierungsart	N	M	SD
Lernerfolg gesamt	4 Standbilder	21	7,41	4,01
	11 Standbilder	19	6,24	3,30
	Animation	22	8,86	2,83
Lernerfolg Faktenwissen	4 Standbilder	21	5,62	2,78
	11 Standbilder	19	4,79	2,84
	Animation	22	6,73	1,84
Lernerfolg Verständnis	4 Standbilder	21	1,80	1,60
	11 Standbilder	19	1,45	1,09
	Animation	22	2,14	1,43

*Tabelle 5-2: Ergebnisse des varianzanalytischen Vergleichs aller drei Gruppen ohne Einbeziehung weiterer Kovariaten.*

<b>Abhängige Variable</b>	<b>Varianzquelle</b>	<b>SS</b>	<b>df</b>	<b>MS</b>	<b>F<sub>emp</sub></b>	<b>p</b>	<b>η<sup>2</sup></b>
<b>Lernzuwachs gesamt</b>	<i>Haupteffekt UV</i>	71,10	2	35,55	3,06	.055	0,09
	<i>Fehler</i>	686,06	59	11,63	---	---	---
<b>Lernzuwachs Faktenwissen</b>	<i>Haupteffekt UV</i>	38,90	2	19,45	3,10	.052	0,10
	<i>Fehler</i>	369,97	59	6,27	---	---	---
<b>Lernzuwachs Verständnis</b>	<i>Haupteffekt UV</i>	4,85	2	2,42	1,24	.30	0,04
	<i>Fehler</i>	115,45	59	1,96	---	---	---

Hinsichtlich der Rolle der 11-Standbilderversion lautet Hypothese 1b, dass diese Bedingung sich bezüglich des Lernerfolges nicht von der Animationsversion unterscheiden sollte, wohl aber von der 4-Standbilderversion. Das Gegenteil ist der Fall. Rein deskriptiv ist die neue Version sogar noch der Version mit 4 Standbildern unterlegen (Tabelle 5-1), statistisch unterschieden sich beide aber nicht, wie die Berechnung des Kontrasts ergibt ( $t(59) = -1,09$ ,  $p = 0,28$ ,  $d = 0,32$ ). Stattdessen sind die elf Standbilder der Animation, entgegen der Erwartung, signifikant unterlegen:  $t(59) = 2,46$ ,  $p < 0,05$ ,  $d = 0,86$ .

Die zweite Hypothese beschäftigt sich mit der Differenzierung zwischen Faktenwissen und tieferem Verständnis. Ausgehend von den Ergebnissen der Metaanalyse wurde erwartet, dass die Animation sowohl bei Aufgaben, die tieferes Verständnis in die Materie erfordern, als auch im Bereich des Faktenwissens zu besseren Lernergebnissen als mit (vier) Standbildern führen sollte. Bereits die Varianzanalyse (Tabelle 5-2) deutet an, dass sich diese Hypothese wohl nicht bestätigen wird: Unterschiede scheinen stärker im Bereich des Faktenwissens als im Verständnis zu finden zu sein. Der a-priori-Kontrasttest bestätigt diese Vermutung ansatzweise – der Einzelvergleich 4 Standbilder-Animation verfehlt bei den Verständnisaufgaben das Signifikanzniveau ( $t(59) = 0,80$ ,  $p = 0,43$ ,  $d = 0,22$ ) ebenso wie im Bereich des Faktenwissens ( $t(59) = 1,45$ ,  $p = 0,15$ ,  $d = 0,47$ ); lediglich im Vergleich der Animation mit der Gesamtheit *beider* Standbildversionen, inhaltlich durchaus zu begründen, ist beim Faktenwissen ein signifikanter Unterschied festzustellen ( $t(59) = 2,29$ ,  $p < 0,05$ ,  $d = 0,59$ ).

### 5.5.2 Hypothese 3: Die Rolle des themenspezifischen Vorwissens

Trotz Nichtauffindens in der Pilotstudie wurde die Hypothese aufrecht erhalten, dass themenspezifisches Vorwissen eine moderierende Rolle dergestalt einnehme, dass Lernende mit niedrigem Vorwissen eher von Animationen profitierten, Experten hingegen auch von Standbildern.

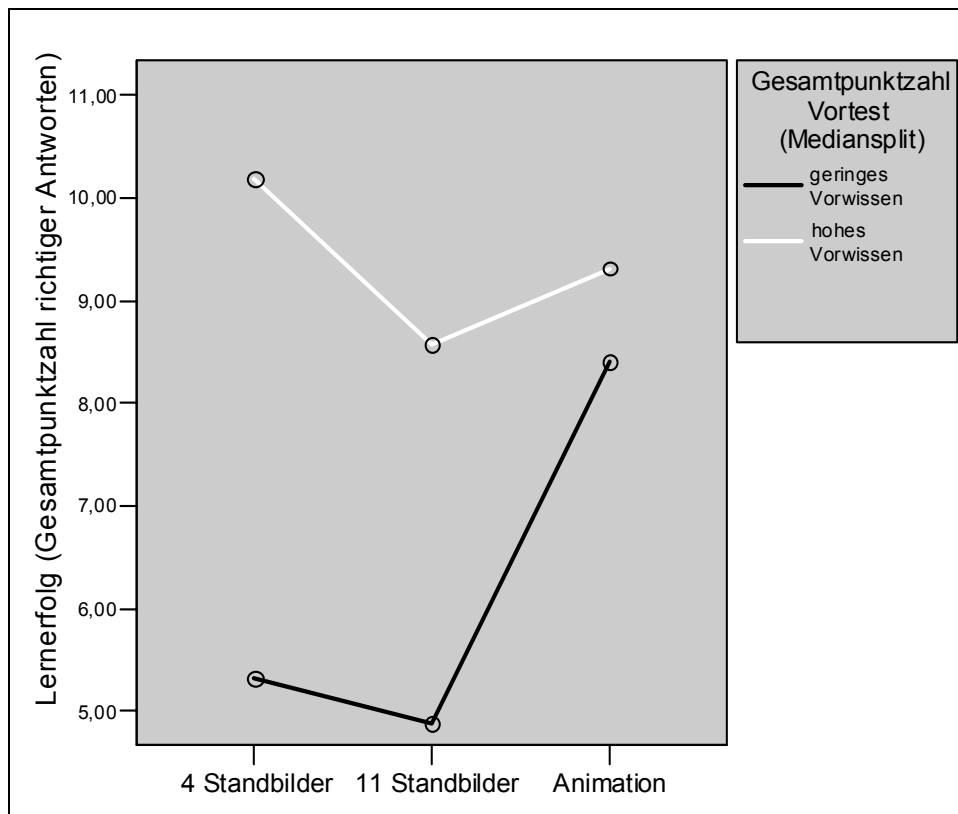
Zunächst ist hierzu festzustellen, dass das Vorwissen insgesamt wiederum sehr niedrig ausgeprägt ist, mit einem Mittelwert von 1,33 (SD 0,93) und einem Median von 1 bei einem möglichen Maximum von 8. Dies erschwert natürlich eine Einteilung in „Experten“ und „Novizen“ anhand eines Mediansplits, da in diesem Fall auch die „Experten“ maximal 50% der Aufgaben vollständig lösten. Somit sind die Befunde dieses Kapitels auch nur eingeschränkt interpretierbar.

Zur Abklärung der notwendigen Voraussetzungen für die weitere Analyse erbringt eine einfaktorielle Varianzanalyse das erwartete Ergebnis, dass die Randomisierung der Probanden insofern geglückt ist, dass es keine systematischen Unterschiede im Vorwissen zwischen den drei Versuchsbedingungen gibt; die drei Mittelwerte unterscheiden sich nicht signifikant voneinander ( $F(2,59) < 1$ ,  $p = 0,56$ ,  $MSE = 0,51$ ,  $\eta^2 = 0,02$ ).

Die nachfolgende zweifaktorielle ANOVA zum Nachweis eines möglichen ATI-Effektes von Vorwissen und Visualisierungsart (Animation, vier oder elf Standbilder) findet kein statistisch signifikantes Ergebnis; zwar deuten die deskriptiven Werte (vgl. Tabelle 5-3 und Abbildung 5-3) auf eine Interaktion hin, das Signifikanzniveau von 5% wird jedoch – relativ knapp – verfehlt ( $F(2,56) = 2,45$ ,  $p = 0,095$ ,  $MSE = 21,91$ ,  $\eta^2 = 0,08$ ; vgl. Tabelle 5-4).

**Tabelle 5-3:** Deskriptive Kennwerte für die Subgruppen themenspezifischen Vorwissens nach der Aufteilung per Mediansplit.

Visualisierungsart	Vorwissen	N	M	SD
4 Standbilder	niedrig	12	5,33	3,51
	hoch	9	10,19	2,85
	gesamt	21	7,41	4,01
11 Standbilder	niedrig	12	4,88	2,50
	hoch	7	8,57	3,34
	gesamt	19	6,24	3,30
Animation	niedrig	11	8,41	2,70
	hoch	11	9,32	3,02
	gesamt	22	8,86	2,83
Gesamt	niedrig	35	6,14	3,26
	hoch	27	9,42	3,00
	gesamt	62	7,57	3,52



**Abbildung 5-3:** Interaktionsdiagramm des Lernerfolgs der Gruppen mit geringem und hohem Vorwissen in Abhängigkeit der Visualisierungsart.

Nicht weiter erstaunlich ist der signifikante Haupteffekt (siehe auch Tabelle 5-4) für das Vorwissen – Lernende mit hohem Vorwissen haben signifikant höhere Nachtestergebnisse als Lernende mit niedrigem Vorwissen ( $F(1,56)=18,28$ ,  $p<0,001$ ,  $MSE=8,93$ ,  $\eta^2=0,25$ ).

*Tabelle 5-4: Ergebnisse der zweifaktoriellen Varianzanalyse für die Einflüsse der Visualisierungsart und des themenspezifischen Vorwissens.*

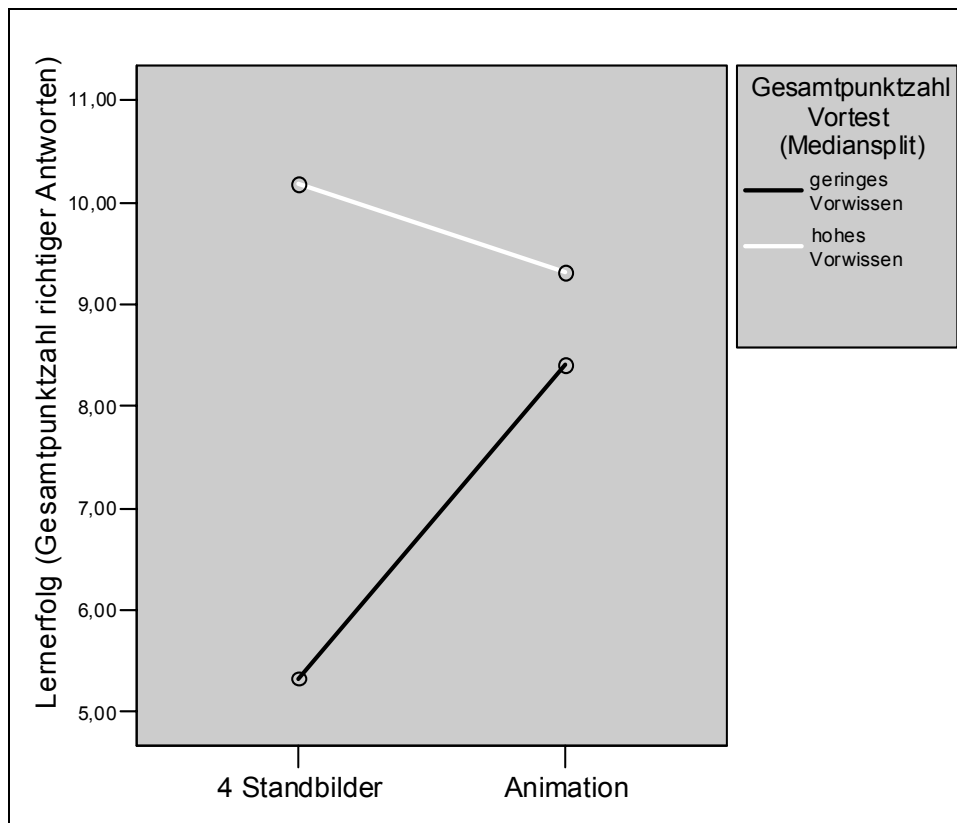
<b>Abhängige Variable</b>	<b>Varianzquelle</b>	<b>SS*</b>	<b>df</b>	<b>MS</b>	<b>F<sub>emp</sub></b>	<b>p</b>	<b><math>\eta^2</math></b>
<b>Lernzuwachs gesamt</b>	<i>Haupteffekt Vorwissen**</i>	163,18	1	163,18	18,28	0,01	0,25
	<i>Haupteffekt Visualisierungsart</i>	50,31	2	25,15	2,82	0,07	0,09
	<i>Interaktionseffekt</i>	43,81	2	21,91	2,45	0,10	0,08
	<i>Fehler</i>	499,86	56	8,93	---	---	---

\* *Sequentielle Varianzzerlegung*

\*\* *Mediansplit*

Betrachtet man nur die Animations- und die 4-Standbilderversion – wie bereits das vorherige Kapitel deutlich macht, verfehlt die 11-Standbilderversion die gesteckten Ziele bei weitem –, so wird der sich andeutende Interaktionseffekt signifikant ( $F(1,39)=4,43$ ,  $p<0,05$ ,  $MSE=9,35$ ,  $\eta^2=0,10$ ), verdeutlicht auch durch Abbildung 5-4.

Es bleibt zum Abschluss festzuhalten, dass bei Verzicht auf Durchführung des Mediansplit und bei alternativer Berechnung durch Einbeziehung des Vorwissens als Kovariate in eine ANCOVA mit sequentieller Varianzzerlegung die Interaktion von Vorwissen und Visualisierungsart (Animation, vier oder elf Standbilder) das Signifikanzniveau verfehlt:  $F(2,56)=1,67$ ,  $p=0,197$ ,  $MSE=9,84$ ,  $\eta^2=0,06$ , ebenso unter Ausschluss der 11-Standbilderversion.



*Abbildung 5-4: Profildiagramm des Lernerfolgs der Gruppen mit geringem und hohem Vorwissen bei Nichtberücksichtigung der 11-Standbilderversion. Der Interaktionseffekt wird deutlich.*

### 5.5.3 Hypothese 4: Die Rolle der Spatial Ability

In der Pilotstudie konnten Hinweise für einen kompensierenden Effekt der *Spatial Ability* in der Form gefunden werden, dass Lernende mit niedrigem räumlichen Vorstellungsvermögen stärker von Animationen im Vergleich zu Standbildern profitieren als Lernende mit hohem räumlichen Vorstellungsvermögen.

Dieser ATI-Effekt kann mit Einschränkungen repliziert werden. Um die Vergleichbarkeit zur Pilotstudie einerseits zu gewährleisten und um andererseits die Rolle der neu eingeführten 11-Standbilderversion speziell testen zu können, wird auf Einzelvergleiche zwischen jeweils zwei Visualisierungsarten zurückgegriffen.

Im Falle der 4-Standbilderversion vs. Animation kann beim Testen des räumlichen Vorstellungsvermögens mittels des Paper-Folding-Tests (analog zur Pilotstudie) zwar kein Interaktionseffekt ohne Einbeziehung weiterer Kovariaten (bei der Berechnung via sequentieller Varianzzerlegung) ermittelt werden – die statistische Signifikanz wird mit  $F(1,39) < 1$ ,  $p = 0,39$ ,  $MSE = 11,67$ ,  $\eta^2 = 0,02$  verfehlt. Werden als Kovariaten

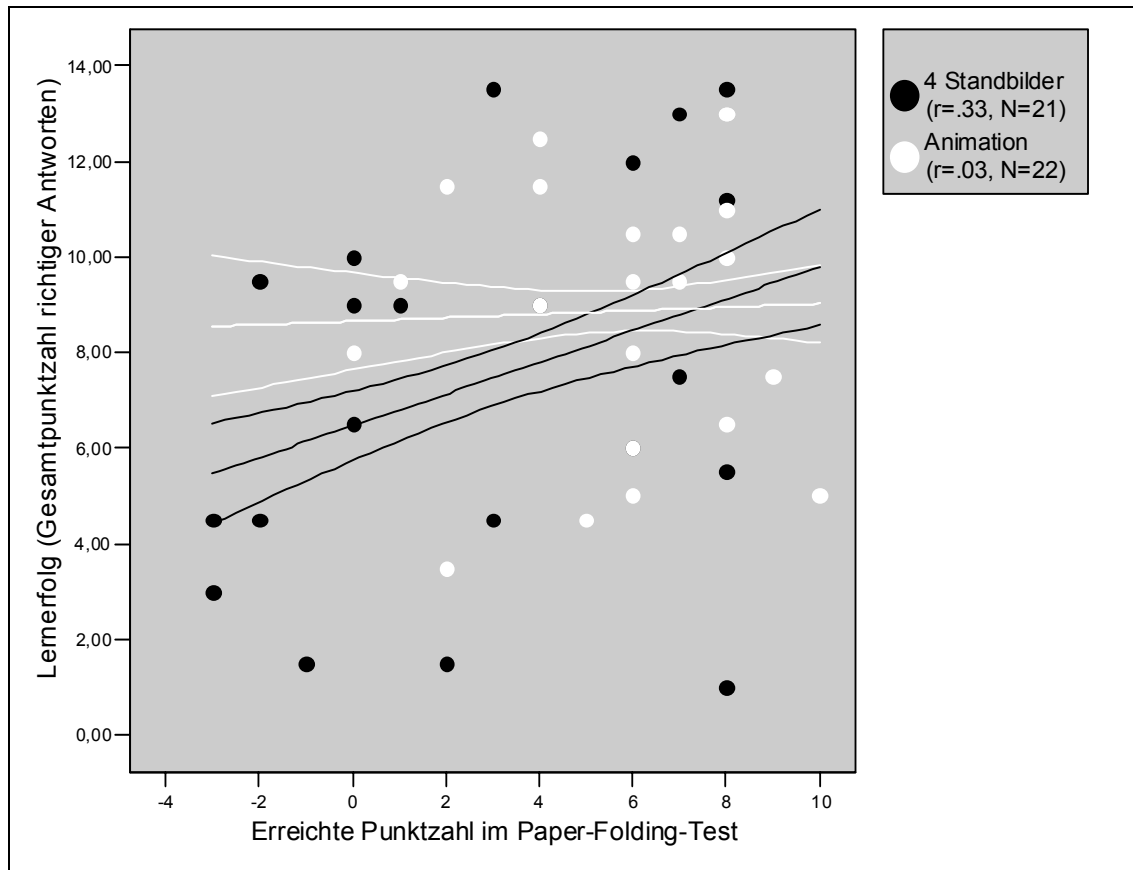
jedoch die letzte Schulnote in Chemie als zusätzliches Maß des Vorwissens sowie das Geschlecht in die Berechnung aufgenommen, so wird die Interaktion zwischen *Spatial Ability* und Visualisierungsart wiederum signifikant ( $F(1,36)=6,62$ ,  $p<0,05$ ,  $MSE=9,46$ ,  $\eta^2=0,16$ ; vgl. Tabelle 5-5). Abbildung 5-5 visualisiert die Interaktion durch ein Streudiagramm mit dazugehörigen Regressionslinien. Darüber hinaus zeigt sich ein signifikanter Haupteffekt der *Spatial Ability* ( $F(1,36)=4,12$ ,  $p<0,05$ ,  $MSE=9,46$ ,  $\eta^2=0,10$ ) – je höher das räumliche Vorstellungsvermögen, desto mehr wird mit Visualisierungen, gleich welcher Art, gelernt.

*Tabelle 5-5: Ergebnisse der Kovarianzanalyse zum Effekt der Spatial Ability (als Kovariate) beim Vergleich zwischen 4 Standbilder- und Animationsversion. Grau unterlegt ist der hypothesenkonforme Interaktionseffekt zwischen Spatial Ability und Visualisierungsart.*

<b>Abhängige Variable</b>	<b>Varianzquelle</b>	<b>SS*</b>	<b>df</b>	<b>MS</b>	<b>F<sub>emp</sub></b>	<b>p</b>	<b><math>\eta^2</math></b>
<b>Lernzuwachs gesamt</b>	<i>Haupteffekt Geschlecht</i>	7,62	1	7,62	0,81	0,375	0,02
	<i>Haupteffekt Spatial Ability</i>	39,30	1	39,30	4,15	0,049	0,10
	<i>Haupteffekt Chemienote</i>	55,63	1	55,63	5,88	0,020	0,14
	<i>Interaktionseffekt Spatial Ability - Chemienote</i>	5,72	1	5,72	0,60	0,442	0,02
	<i>Haupteffekt Visualisierungsart</i>	1,45	1	1,45	0,15	0,698	0,004
	<i>Interaktionseffekt Spatial Ability - Visualisierungsart</i>	62,62	1	62,62	6,62	0,014	0,16
	<i>Fehler</i>	340,62	36	9,46	---	---	---

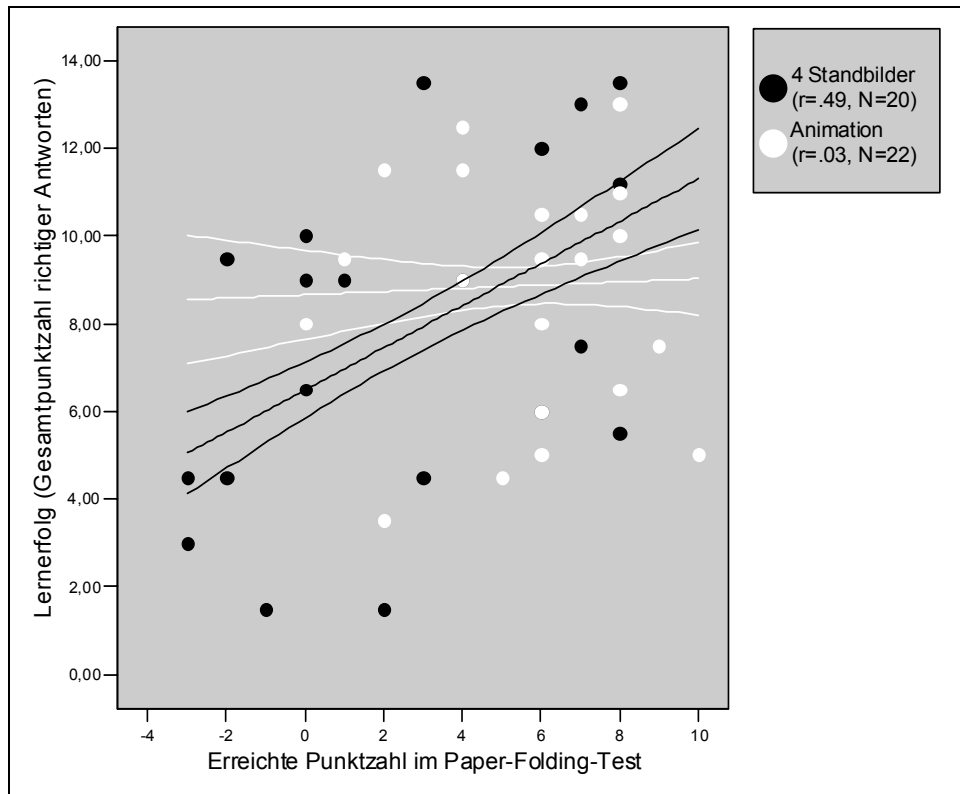
\* Sequentielle Varianzzerlegung





**Abbildung 5-5:** Streudiagramm mit Lernerfolg als Y-Achse und Punktzahl im Paper-Folding-Test als X-Achse in Abhängigkeit von der jeweiligen Visualisierungsart inklusive zugehöriger Regressionslinien (mit 75%-Vertrauensintervall). Die 11-Standbilderversion ist nicht einbezogen.

Eliminiert man versuchsweise den Outlier bei 8;1 (vgl. Abbildung 5-5 rechts unten), wird das Bild (Abbildung 5-6) noch deutlicher. Die Interaktion bleibt signifikant, auch wenn der Effekt etwas schwächer ist ( $F(1,36)=5,19$ ,  $p<0,05$ ,  $MSE=8,62$ ,  $\eta^2=0,13$ , vgl. Tabelle 5-6).



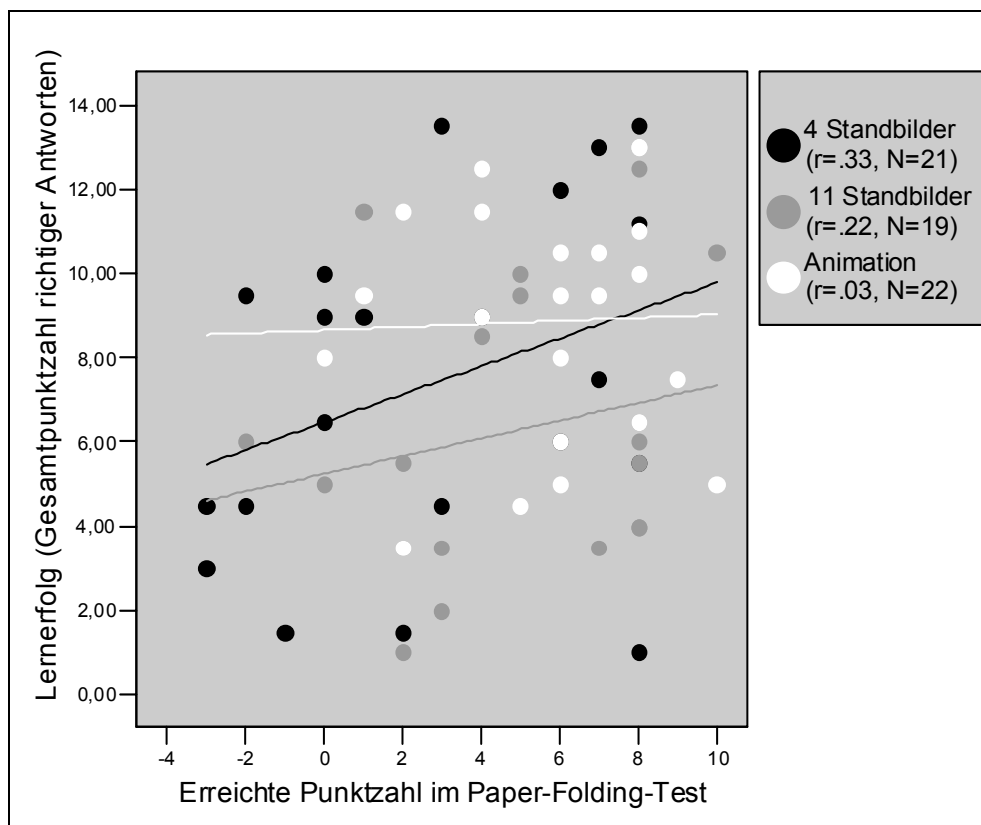
**Abbildung 5-6:** Streudiagramm mit Lernerfolg als Y-Achse und Punktzahl im Paper-Folding-Test als X-Achse in Abhängigkeit von der jeweiligen Visualisierungsart inklusive zugehöriger Regressionslinien (mit 75%-Vertrauensintervall) unter Ausschluss eines Outliers. Die 11-Standbilderversion ist nicht einbezogen.

**Tabelle 5-6:** Ergebnisse der Kovarianzanalyse zum Effekt der Spatial Ability (als Kovariate) beim Vergleich zwischen 4 Standbilder- und Animationsversion unter Ausschluss eines Outliers.

Abhängige Variable	Varianzquelle	SS*	df	MSE	F <sub>emp</sub>	p	η <sup>2</sup>
<b>Lernzuwachs gesamt</b>	Haupteffekt Geschlecht	3,50	1	3,50	0,41	0,528	0,01
	Haupteffekt Spatial Ability	58,40	1	58,40	6,78	0,013	0,16
	Haupteffekt Chemienote	38,59	1	38,59	4,48	0,041	0,11
	Haupteffekt Visualisierungsart	5,05	1	5,05	0,59	0,449	0,02
	Interaktionseffekt Spatial Ability - Visualisierungsart	44,74	1	44,74	5,19	0,029	0,13
	Fehler	310,25	36	8,62	---	---	---

\* Sequentielle Varianzzerlegung

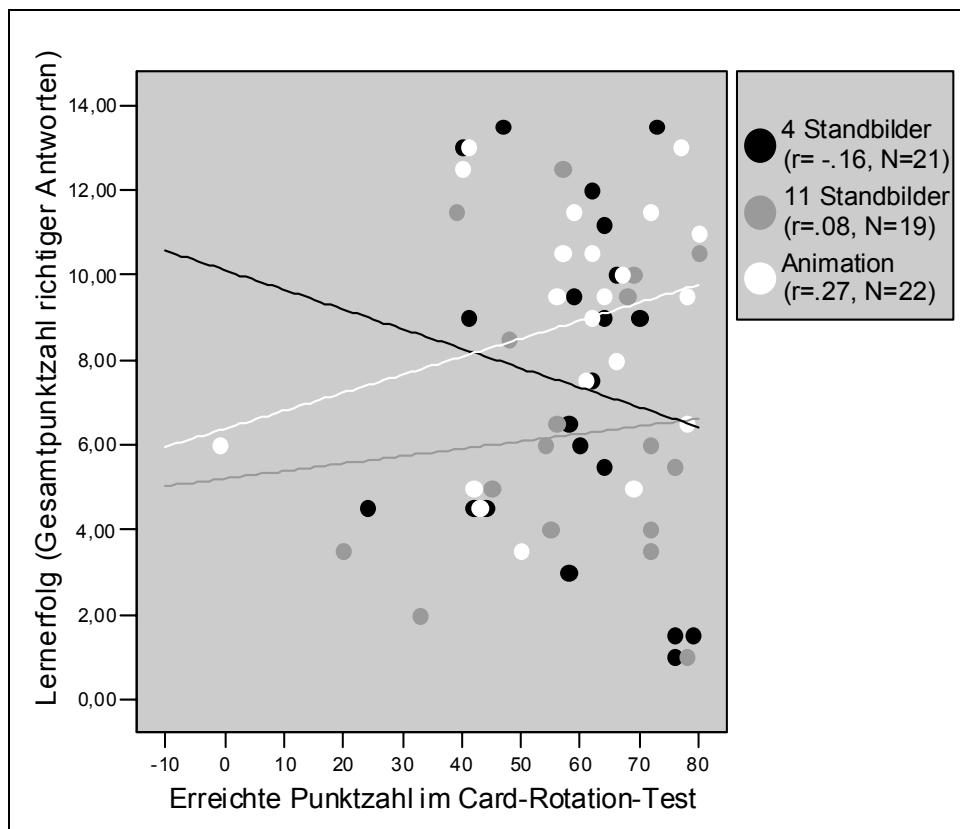
Beim Fokus auf die 11-Standbilderversion können keine derartigen Interaktionseffekte nachgewiesen werden. Weder gelingt dies hypothesenkonform in Relation zur 4-Standbilderversion (unter der Erwartung, dass die elf Standbilder ähnliche Auswirkungen haben wie die Animation), noch im umgekehrten Sinne, also im Vergleich zur Animation ( $F(1,36) < 1$ ,  $p = 0,71$ ,  $MSE = 13,14$ ,  $\eta^2 = 0,004$  bzw.  $F(1,37) < 1$ ,  $p = 0,60$ ,  $MSE = 9,59$ ,  $\eta^2 = 0,007$ ). Wie die Regressionslinien in Abbildung 5-7 visualisieren, verhielt sich die 11-Standbilderversion – rein deskriptiv – weder wie die Animations-, noch wie die 4-Standbilderversion hinsichtlich des Zusammenhangs zwischen Lernleistung und Ergebnis im Paper-Folding-Test, sondern liegt eher dazwischen, wenngleich insgesamt auf niedrigerem Leistungsniveau.



**Abbildung 5-7:** Streudiagramm mit Lernerfolg als Y-Achse und Punktzahl im Paper-Folding-Test als X-Achse in Abhängigkeit der jeweiligen Visualisierungsart inklusive zugehöriger Regressionslinien (auf die Darstellung der Vertrauensintervalle wurde zugunsten der Übersichtlichkeit verzichtet).

In dieser Untersuchung wurde die *Spatial Ability* nicht nur mit dem Paper-Folding-Test wie in der Pilotstudie erhoben, sondern zusätzlich mit dem Card-Rotation-Test.

Interessanterweise können bei der Messung mit diesem Instrument die oben dargestellten Effekte nicht nachgewiesen werden. In den Einzelvergleichen zwischen den Versionen manifestieren sich keinerlei ATI-Effekte der erwarteten Art (Tabelle 5-7), auch nicht unter Einbeziehung inhaltlich sinnvoller Kovariaten. Analog zu Abbildung 5-7 ist in Abbildung 5-8 das Streudiagramm der Probanden mit ihren jeweiligen Werten im Card-Rotation-Test in Relation zum Lernerfolg abgetragen. Wie deutlich wird, unterscheidet sich das Muster der Regressionslinien deskriptiv zwar recht deutlich – im Falle der vier Standbilder wird sogar ein negativer Effekt zwischen beiden Variablen suggeriert. Der ATI-Effekt ist jedoch nicht statistisch signifikant. Auch unter Ausschluss des auffälligen Outliers links (-1;6; vgl. Abb. 5-8) ist dies nicht der Fall.



*Abbildung 5-8: Streudiagramm mit Lernerfolg als Y-Achse und Punktzahl im Card-Rotation-Test als X-Achse in Abhängigkeit der jeweiligen Visualisierungsart inklusive zugehöriger Regressionslinien (auf die Darstellung der Vertrauensintervalle wurde zugunsten der Übersichtlichkeit verzichtet).*

*Tabelle 5-7: Interaktionseffekte von Visualisierungsart und Spatial Ability (bei Messung mit dem Card-Rotation-Test), getrennt für Einzelvergleiche zwischen den drei Visualisierungsarten.*

<b>Verglichene Visualisierungsarten</b>	<b>Varianzquelle</b>	<b>SS</b>	<b>df</b>	<b>MS</b>	<b><math>F_{emp}</math></b>	<b><math>p</math></b>	<b><math>\eta^2</math></b>
<b>Animation vs. 11 Standbilder vs. 4 Standbilder</b>	<i>Interaktionseffekt UV - Spatial Ability</i>	20,00	2	10,00	0,84	0,44	0,03
	<i>Fehler</i>	663,63	56	11,85	---	---	---
<b>11 Standbilder vs. 4 Standbilder</b>	<i>Interaktionseffekt UV - Spatial Ability</i>	9,08	1	9,08	0,64	0,43	0,02
	<i>Fehler</i>	507,41	36	14,10	---	---	---
<b>Animation vs. 4 Standbilder</b>	<i>Interaktionseffekt UV - Spatial Ability</i>	19,80	1	19,80	1,64	0,21	0,04
	<i>Fehler</i>	469,52	39	12,04	---	---	---
<b>Animation vs. 11 Standbilder</b>	<i>Interaktionseffekt UV - Spatial Ability</i>	1,81	1	1,81	0,19	0,66	0,01
	<i>Fehler</i>	350,34	37	9,47	---	---	---

#### 5.5.4 Weitere Berechnungen

Der *Cognitive Load* ist zwar nicht mehr Bestandteil einer separaten Hypothese, wird aber trotzdem noch weiterhin kontrolliert. Wie erwartet, zeigen sich wiederum keine gravierenden Unterschiede zwischen den drei Gruppen hinsichtlich des *Cognitive Load*. Wie Tabelle 5-8 veranschaulicht, werden alle Versuchsbedingungen als in etwa gleich belastend empfunden.

*Tabelle 5-8: Deskriptive Kennwerte für die Items der Cognitive Load Skala, abhängig von der Visualisierungsart.*

<b>Cognitive Load-Item</b>	<b>Visualisierungsart</b>	<b>N</b>	<b>M</b>	<b>SD</b>
Wie leicht oder schwer war das Lernmaterial zu verstehen?	4 Standbilder	21	3,10	0,25
	11 Standbilder	19	2,84	0,21
	Animation	22	2,86	0,23
Bei der Bearbeitung des Lernmaterials war meine mentale Anstrengung...	4 Standbilder	21	3,43	0,15
	11 Standbilder	19	3,63	0,23
	Animation	22	3,68	0,23
Gesamt	4 Standbilder	21	6,52	0,31
	11 Standbilder	19	6,47	0,37
	Animation	22	6,55	0,41

Zudem werden, analog zur ersten experimentellen Studie, die Korrelationen relevanter Variablen untereinander berechnet. Tabelle 5-9 zeigt diese in einer Korrelationsmatrix. Der *Cognitive Load* korreliert negativ mit dem Lernerfolg, was zu erwarten war, ebenso teilweise mit dem Vorwissen. Auffällig ist ferner die nur äußerst schwache Korrelation ( $r=0,16$ ) beider Tests zur *Spatial Ability* miteinander, was die Annahme zweier verschiedener Dimensionen (vgl. Kapitel 5.2) weiter stützt.

**Tabelle 5-9:** Mittelwerte, Standardabweichungen und Matrix der Produktmomentkorrelationen relevanter Variablen.

	M	SD		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lernerfolg gesamt (1)	7,57	3,52	r	1									
			N	62									
Lernerfolg Fakten- wissen (2)	5,76	2,59	r	0,94**	1								
			N	62	62								
Lernerfolg Verständnis (3)	1,81	1,40	r	0,78**	0,51**	1							
			N	62	62	62							
Chemienote (4)	2,84	0,98	r	-0,27*	-0,20	-0,31*	1						
			N	62	62	62	62						
Vorwissen (5)	1,33	0,94	r	0,40	0,39**	0,29*	-0,19	1					
			N	62	62	62	64	64					
Paper- Folding-Test (6)	4,44	3,49	r	0,25	0,28*	0,10	0,18	0,14	1				
			N	62	62	62	64	64	64				
Card- Rotation- Test (7)	58,58	16,02	r	0,05	0,66	0,01	0,07	0,08	0,16	1			
			N	62	62	62	64	64	64	64	64		
CL 1 (8)	2,94	1,04	r	-0,51**	-0,42**	-0,52**	0,18	-0,28*	-0,05	-0,19	1		
			N	62	62	62	64	62	62	62	62	62	
CL 2 (9)	3,58	0,93	r	-0,41**	-0,32**	-0,39**	-0,03	-0,21	-0,03	-0,12	0,40**	1	
			N	62	62	62	64	62	62	62	62	62	62
CL gesamt (10)	6,52	1,65	r	-0,56**	-0,46**	-0,55**	0,10	-0,30*	-0,05	-0,19	0,85**	0,82**	1
			N	62	62	62	64	62	62	62	62	62	62

\* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

\*\* Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Die Reliabilitäten der angewendeten Skalen fallen in dieser Studie nur zum Teil zufriedenstellend aus. Besonders die Vorwissensskala erbringt ein nicht ausreichendes Cronbachs Alpha von  $\alpha=0,43$ , das auch durch den Verzicht auf das bereits in der Pilotstudie problematische vierte Item lediglich auf  $\alpha=0,55$  gesteigert werden könnte. Möglicherweise ist dies aber durch das insgesamt sehr geringe Vorwissensniveau zu erklären. Auch die Reliabilität der Skala zum *Cognitive Load* zeigt dieses Mal ein nicht ausreichendes  $\alpha$  von 0,56.

Der Lernerfolg hingegen verzeichnet als Gesamtskala ein sehr gutes Alpha von 0,85 – wiederum wurden jedoch für die Berechnungen ausschließlich die Items verwendet, die sich explizit auf den Bereich der Visualisierung (Animation oder

Standbilder) beziehen und nicht auf die einbettende Lernumgebung. Diese erreichen ein ebenfalls befriedigendes Cronbachs Alpha von 0,79.

### 5.5.5 Follow-Up

In dieser Studie wurde zusätzlich ein Follow-Up-Test durchgeführt. Ca. dreieinhalb Monate nach der ersten Testung wurden den Probanden hierzu nochmals dieselben Fachfragen wie im Nachtest vorgelegt. Zwar zeigt sich für alle drei Versuchsbedingungen wie auch insgesamt ( $t(54)=6,02$ ,  $p<0,001$ ,  $d=0,70$ ) erwartungsgemäß ein signifikanter Abfall in den Lernleistungen, das Muster bleibt, deskriptiv betrachtet, aber gleich (Abbildung 5-10, Tabelle 5-10). Die Animation ist nach wie vor den 11 Standbildern signifikant überlegen (bei Berechnung des Kontrasttests:  $t(52)=2,14$ ,  $p<0,04$ ,  $d=0,75$ ), jedoch nicht (mehr) der 4-Standbilderversion.

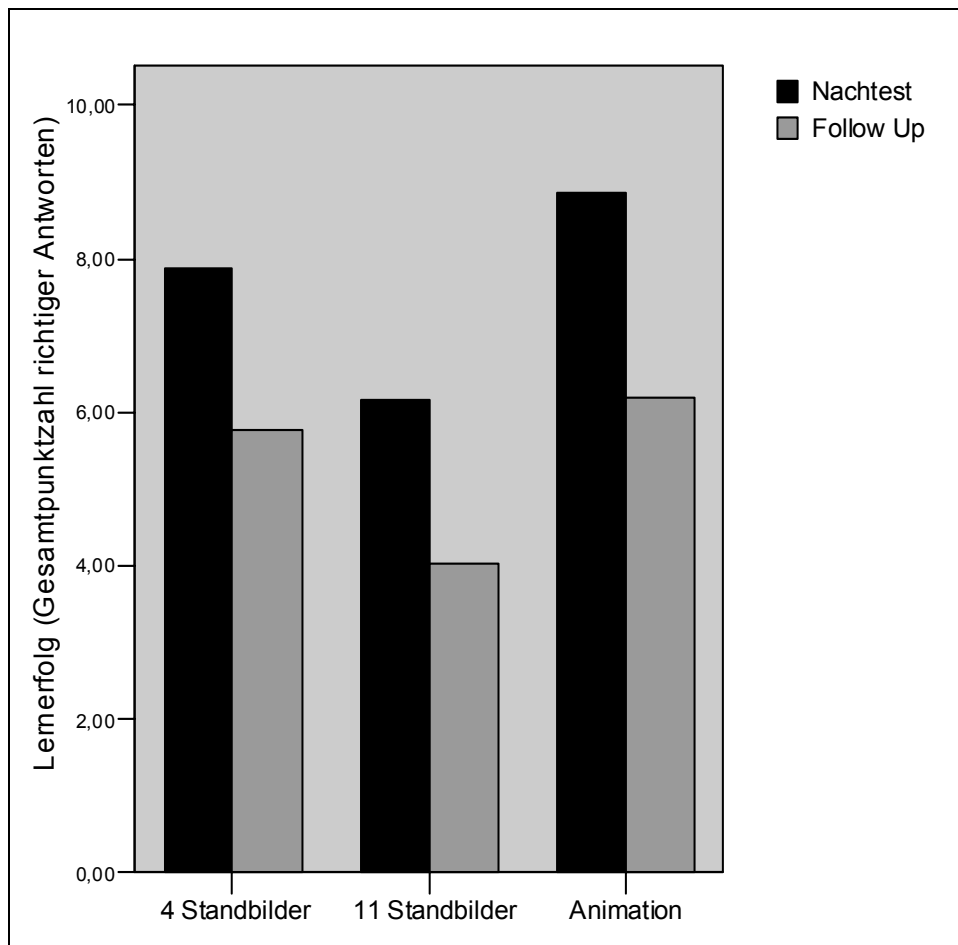
Eine *Repeated-Measures-Analyse* bestätigt diese Befunde. Die Haupteffekte der Zeit ( $F(1,52)=35,22$ ,  $p<0,001$ ,  $MSE=4,28$ ,  $\eta^2=0,40$ ) und der drei Visualisierungsarten ( $F(2,52)=3,52$ ,  $p<0,05$ ,  $MSE=16,53$ ,  $\eta^2=0,12$ ) werden statistisch signifikant, es zeigt sich aber keine Interaktion von Zeit und Visualisierungsart ( $F(2,52)<1$ ).

Interaktionseffekte mit dem räumlichen Vorstellungsvermögen lassen sich nicht mehr nachweisen.

*Tabelle 5-10: Deskriptive Kennwerte für die Visualisierungsarten im Follow-Up-Test.*

<b>Visualisierungsart</b>	<b>N</b>	<b>M</b>	<b>SD</b>
<i>4 Standbilder</i>	16	5,78	3,68
<i>11 Standbilder</i>	17	4,03	2,41
<i>Animation</i>	22	6,21	3,22
<i>Gesamt</i>	55	5,41	3,23





*Abbildung 5-10: Diagramm zur Darstellung der unterschiedlichen Lernerfolgsleistungen im Nachtest und Follow-Up-Test in Abhängigkeit zur Visualisierungsart.*

## 5.6 Diskussion

Vorliegende zweite Studie hatte zwei übergeordnete Ziele. Zum einen galt es, die in der Pilotstudie gefundenen Ergebnisse zu replizieren, zum anderen sollte anhand der neu eingeführten dritten Versuchsbedingung mit elf Standbildern die aus den Vorüberlegungen sowie den Ergebnissen der Metaanalyse abgeleitete Hypothese überprüft werden, ob Animationen prinzipiell durch Standbilder zu ersetzen seien – bei ausreichender Anzahl derselben. Es war also angestrebt, den in der Pilotstudie gefundenen Animationseffekt (die Überlegenheit der Animation hinsichtlich des Lernerfolgs) durch Erhöhung der Frequenz der Standbilder wieder „abzuschalten“.

Wie die Ergebnisse zeigten, wurde das erste Ziel erreicht, das zweite jedoch nicht.

Unter Einbeziehung des Vorwissens als Kovariate war die Animation wie erwartet der 4-Standbildversion statistisch signifikant überlegen. Die Pilotstudie kann in dieser Hinsicht als bestätigt angesehen werden.

Die Einführung der neuen Standbildversion hingegen hat ihr Ziel eindeutig verfehlt. Anstatt Lernergebnisse wie mit der Animation zu erreichen, waren die Ergebnisse dieser Version der der Animation signifikant unterlegen, sogar ohne Einbeziehung einer Kovariate. Offensichtlich erleichterten die zusätzlichen Standbilder die Bildung eines adäquaten mentalen Modells nicht, entgegen der Hypothese. Da bereits die vier Standbilder (nach Auskunft von Experten aus der Chemiedidaktik) vollkommen ausreichen können, um den Vorgang prinzipiell zu verstehen, könnte eine Argumentation zur Begründung dieses Befundes lauten, dass die zusätzlichen Standbilder *Extraneous Cognitive Load* beanspruchten, also unnötige, lediglich das Arbeitsgedächtnis belastende Elemente darstellten. Dies widerspräche jedoch dem Ergebnis, dass die Animation nach wie vor den Standbildversionen überlegen ist, da eine Begründung über die *Cognitive Load Theory* genau so auf die Animationsversion übertragen werden müsste und deshalb auch von dieser niedrigere Lernergebnisse zu erwarten wären. Zudem zeigten sich in der selbstberichteten kognitiven Belastung keine Unterschiede zwischen den Versionen.

Offensichtlich war also die 11-Standbildversion für das Lernen gänzlich ungeeignet, und zwar vermutlich aus anderen Gründen als die 4-Standbildversion. Eine wahrscheinlichere Begründung für diesen Befund liefert die Zeitkomponente: Wie im Kapitel der Versuchsdurchführung beschrieben wurde, sollte die potentielle Störvariable „Zeit der Betrachtung des Materials“ dadurch kontrolliert werden, dass den Lernenden die jeweilige Versuchsbedingung exakt gleich lang präsentiert wurde. Dies ist in vielen Untersuchungen zu Animationen und Standbildern nicht der Fall (z.B. Mayer et al., 2005) – eine Animation mit bestimmter Länge wird einigen, häufig papierbasierten, Standbildern entgegengestellt, die sich die Lernenden in ihrem eigenen Tempo anschauen dürfen. Zwar mag ein solches Vorgehen die (Schul-)Alltagsrelevanz erhöhen, auf diese Weise jedoch werden die Standbilder unzulässig „bevorteilt“, da den Lernenden mehr Zeit zum Verstehen zugestanden wird. Durch vorliegendes Design, in dem auch die Standbilder im Stile einer Diashow ohne Lernerkontrolle, dafür synchronisiert zum zugehörigen verbalen Kommentar weiter springen, ist diese Störvariable vollständig ausgeschaltet.

Nun könnte es aber sein, dass dieses Format gerade im Falle der elf Standbilder (mit einem also sehr häufigen Wechsel der Bilder) den Lernenden vor eine extrem ungewohnte Situation stellt, in der er es nicht schafft, in ausreichender Geschwindigkeit jeweils die entscheidenden Elemente der Bilder zu identifizieren und abzuspeichern. Lowe (1999) wies nach, dass Lernende tendenziell eher Elemente hoher Salienz statt hoher Wichtigkeit erinnern. Zwar mag der transiente Ablauf einer Animation einen solchen Vorgang ebenfalls erschweren, jedoch sind Lernende bei einer Animation möglicherweise eher auf deren speziellen Charakter eingestellt und können dies kompensieren. Bei vier Standbildern hingegen mag die Zeit für jedes einzelne Bild ausgereicht haben, die Bedeutung hinreichend zu erkennen (auch wenn die Bildung einer mentalen Animation offensichtlich ebenfalls nicht gelang, was durch die Unterlegenheit gegenüber der Animation belegt wird). Es könnte also Sinn machen, in einer weiteren Studie die Zeitkomponente auf eine andere Weise zu kontrollieren, diese Variable zu einem größeren Teil vom Lerner steuern zu lassen, um deren potentiell störenden Effekt zu eliminieren – dies könnte im Endeffekt zu den erwarteten Ergebnissen hinsichtlich der 11-Standbilderversion führen.

Weiterhin erscheint bei den Ergebnissen noch besonders die Rolle des räumlichen Vorstellungsvermögens interessant. Auch hier war in erster Linie angestrebt, die Befunde der Pilotstudie zu replizieren. Allerdings war zusätzlich mit dem Card-Rotation-Test eine weitere Messmöglichkeit der *Spatial Ability* eingeführt worden, um diese Eigenschaft auf eine mögliche Mehrdimensionalität abzutesten. Der moderierende Effekt, der in der Pilotstudie mit Hilfe des Paper-Folding-Tests gefunden wurde, konnte auch in dieser Studie nachgewiesen werden, wenngleich abermals zusätzliche Kovariaten, die letzte Schulnote in Chemie als Einschätzung des Vorwissens sowie das Geschlecht, in die Berechnung einbezogen werden mussten. Trotzdem scheinen sich die Hinweise auf diesen kompensierenden Effekt von *Spatial Ability* zu verdichten – hohes räumliches Vorstellungsvermögen kann Mängel im instruktionalen Design (Standbilder anstelle einer Animation) ausgleichen bzw., wichtiger noch, das Angebot einer dynamischen Visualisierung kann Mängel im räumlichen Vorstellungsvermögen ausgleichen, vermutlich mittels Bereitstellung eines Modells für eine mentale Animation.

Dass bei Nutzung der Ergebnisse des Card-Rotation-Tests dieser Effekt nicht nachgewiesen werden konnte, weist darauf hin, dass möglicherweise lediglich eine Unterkategorie des räumlichen Vorstellungsvermögens, die *Spatial Visualization*,

einen Einfluss beim Lernen mit Visualisierungen der hier untersuchten Art hat. Im Gegensatz zum Card-Rotation-Test, für den relativ einfache, einschrittige mentale Rotationen benötigt werden und der auf dem *Spatial Relations*-Faktor lädt (Carroll, 1993), erfordert *Spatial Visualization* mental komplexere sequentielle Leistungen, in die vermutlich auch höhere Exekutivfunktionen des Arbeitsgedächtnisses involviert sind. Die Ergebnisse der beiden ersten experimentellen Untersuchungen lassen vermuten, dass genau diese Fähigkeit bei der Bildung einer mentalen Animation aus Standbildern nötig ist.

Schließlich ist auf den Effekt des Vorwissens einzugehen: Hypothesenkonform zeigten die Ergebnisse dieser Studie neben einem erwartbaren signifikanten Haupteffekt des Vorwissens einen – bei Nichtberücksichtigung der offensichtlich völlig die angestrebten Ziele verfehlt habenden 11-Standbilderversion – statistisch signifikanten ATI-Effekt: Novizen profitieren in besonderem Maße von Animationen gegenüber (vier) Standbildern, während Lernende mit höherem (wenngleich absolut gesehen immer noch recht geringem) Vorwissen von beiden Versionen gleich gut abschneiden. Offensichtlich kann die Animation den Lernenden einen besseren Eindruck des zu vermittelnden Lernprozesses geben, was besonders Lernenden mit geringem Vorwissen entgegen kommt. Experten hingegen scheinen keine Probleme zu haben, auch mit Standbildern ausreichend zu verstehen und ein adäquates mentales Modell zu konstruieren. Novizen scheinen mit dieser Aufgabe überfordert und benötigen das durch die Animation bereitgestellte bewegte Modell als Vorlage für ihr mentales Modell.

Im Hinblick auf ATI verhält sich Vorwissen also vergleichbar wie *Spatial Ability*: Hohes Vorwissen kompensiert Mängel im instruktionalen Design bzw. gutes instruktionales Design (hier die Animation) kompensiert Defizite im Bereich des Vorwissens.

## 6 Dritte experimentelle Studie

### 6.1 Einführung

In der zweiten empirischen Studie wurden die aufgestellten Hypothesen nur teilweise bestätigt. Besonders die neu eingeführte Lernumgebung, die anstelle der Animation oder vier Standbildern elf Standbilder beinhaltet, konnte die Erwartungen nicht auch nur ansatzweise erfüllen. Anstelle einer Angleichung des Lernerfolgs an das Niveau der Animation führte diese Version zu ebenso niedrigen Leistungen wie die 4-Standbilderversion. Zurückgeführt wurde dieses unerwartete Ergebnis in vorangegangener Diskussion auf die Zeitkomponente – durch das systemgesteuerte schnelle Weiterspringen der einzelnen Bilder könnte es zu einer zusätzlichen Lernerschwerung aufgrund von Ungewohntheit und Schwierigkeiten bei der Erfassung der inhaltlich wesentlichen Komponenten gekommen sein.

Um diese Interpretation zu testen, wird nun eine weitere, dritte experimentelle Studie durchgeführt, die zwar weiterhin die drei etablierten Lernumgebungen beinhaltet, jedoch in abgewandelter Form. In jeder Version sind nun minimale Möglichkeiten der Interaktivität (Lawless & Brown, 1997, S. 122) dergestalt vorhanden, dass die Lernenden durch Klicken auf die entsprechenden Bedienelemente die Animation bzw. die Serie der Standbilder jederzeit stoppen, vor- oder zurückspulen können. Die Zeitkomponente wird nun nur noch dadurch kontrolliert, dass eine Maximalzeit zur Bearbeitung der gesamten Lernumgebung vorgegeben wird.

Mit Hilfe dieser modifizierten Lernmittel soll jetzt die ursprüngliche Annahme über die Rolle der Anzahl der Standbilder zum Lernen überprüft werden; gleichzeitig dient die Studie zur Replikation der weiteren Ergebnisse aus den vorangegangenen Studien.

### 6.2 Zusätzliche theoretische Erörterungen

Die oben angesprochene Modifikation der Lernumgebungen schneidet einen theoretischen Bereich an, auf den im Kapitel des theoretischen Hintergrunds noch nicht eingegangen worden ist und der nun kurz skizziert werden soll – die Unterscheidung zwischen Lernerkontrolle und Systemkontrolle des Programm- (und

somit implizit auch des Lern-)fortschritts, und damit die Rolle der Interaktivität in computerbasierten Lernumgebungen.

Die Neuerung in der dritten im Vergleich zur zweiten Studie, die Einführung minimaler interaktiver Möglichkeiten, stellt also zugleich den Wechsel von externer Kontrolle (*external pacing*; Leutner, 1992, S. 40; bzw. *program control*; Lawless & Brown, 1997, S. 118) zu Selbstkontrolle (*self-pacing*; Leutner, 1992, S. 38; bzw. *learner control*; Lawless & Brown, S. 118; bzw. *simple user interaction*; Mayer & Chandler, 2001, S. 390) dar. Der etablierteste dieser verschiedenen Begriffe für dasselbe Konstrukt ist sicherlich der des *self-pacing*, der bereits auf Skinner (1954) rückführbar ist und von diesem als Teil eines komplexen Lehrprogramms als lernvorteilhaft propagiert wird. In dieser Ansicht wird er nicht nur von verschiedenen Seiten in allgemeiner Hinsicht unterstützt (Merrill, 1975; Wittrock, 1978), auch speziell für den Bereich des computerbasierten Lernens konnten bereits Hinweise in dieser Richtung gefunden werden (Mayer & Chandler, 2001; Schwan & Riemp, 2004; Shyu & Brown, 1992): *Self-pacing* scheint dem Lernenden bezüglich der Generierung eines eigenen adäquaten mentalen Modells Vorteile zu erbringen im Gegensatz zur externen Vorgabe des Programmfortschritts, dem *external pacing*. Es gibt Hinweise darauf, dass sich dieser Vorteil besonders auf den Bereich des tieferen Verständnisses erstreckt und weniger auf die Wiedergabe puren Faktenwissens (Mayer & Chandler, 2001), allerdings haben die Autoren dieses Artikels lediglich eine einzige *Retention*-Aufgabe benutzt.

So vorteilhaft diese Form minimaler Interaktion auch zu sein scheint, so wichtig ist es darauf zu achten, den Lernenden durch ein Übermaß an Interaktivität kognitiv nicht zu überlasten (Chandler, 2004). Gerade Novizen dürfte es oftmals schwerfallen, in reinen interaktiven Simulationen den Überblick zu behalten und trotzdem angemessen zu lernen. Verschiedene Studien konnten keinen Vorteil für mehr als minimale Interaktivität zeigen (z.B. de Jong & van Joolingen, 1998; Nerdel, 2003; Höffler, 2003), trotz möglicher motivationaler Effekte (Lepper & Malone, 1987).

### 6.3 Fragestellung und inhaltliche Hypothesen

Trotz oder gerade wegen des im Vergleich zu den vorangegangenen Studien leicht veränderten Ausmaßes der Lernerkontrolle bleibt ein Großteil der bisherigen

Hypothesen gleich. Nur in manchen Fällen ist aufgrund einer eher *learner-paced*-zentrierten Ausrichtung eine Änderung der Ergebnisse zu erwarten.

Die ersten Hypothesen bleiben selbstverständlich unverändert, war die unerwartete Nichtbestätigung von Hypothese H1b in der zweiten Studie doch der Hauptgrund für die Veränderung des Designs der Studie:

***H1a: Mit der animierten Lernumgebung werden bessere Lernleistungen erzielt als mit der aus 4 statischen Bildern bestehenden Lernumgebung.***

***H1b: Es bestehen hinsichtlich der Lernleistung keine Unterschiede zwischen der Animation und der mit 11 Standbildern ausgestatteten Lernumgebung. Diese ist wiederum der 4-Standbilder-Version signifikant überlegen.***

Bezüglich einer besonderen Rolle der Animation für das Erlernen von tieferem Verständnis anstelle von basalem Faktenwissen konnten im Verlaufe der bisherigen Studien keine eindeutigen Ergebnisse gefunden werden. Dementsprechend lautet die zweite Hypothese weiterhin:

***H2: Die Überlegenheit der Animation zeigt sich sowohl bei Aufgaben, die Faktenwissen erfragen, als auch bei solchen, für die tieferes Verständnis erforderlich ist.***

Ebenfalls konnte bis dato kein Effekt für das themenspezifische Vorwissen nachgewiesen werden – ob die Zulassung von mehr Lernerkontrolle einen Moderatoreffekt des Vorwissens klarer zum Vorschein bringen kann, ist Spekulation. Sowohl eine Angleichung des Niveaus von Experten und Novizen als auch besondere Vorteile für die Experten, die sich evtl. die zusätzlichen Möglichkeiten der Interaktivität besonders zunutze machen könnten (Nerdel, 2003), sind möglich. Die dritte Hypothese wird deshalb zunächst unverändert beibehalten.

***H3: Die lernförderliche Wirkung von Animationen lässt sich besonders für Lernende mit niedrigem themenspezifischen Vorwissen nachweisen. Lerner mit hohem Vorwissen erzielen mit 4 stehenden Bildern mindestens genauso gute Ergebnisse wie mit Animationen oder 11 Standbildern.***

Beide vorangegangenen Studien fanden moderierende ATI-Effekte der *Spatial Ability*. Es erscheint denkbar, dass diese Annahme unter den Bedingungen des überarbeiteten Designs nicht länger aufrechtzuerhalten sein wird: Vorteile, die ein erhöhtes räumliches Vorstellungsvermögen Lernenden im Umgang mit Standbildern erbringen, könnten unerheblich werden, wenn sich der Lernende jedes Standbild in einer für ihn ausreichenden Zeit betrachten und abspeichern darf. Allerdings könnte die Fähigkeit, aus wenigen Standbildern eine angemessene mentale Animation zu generieren (als eine Facette räumlichen Vorstellungsvermögens), auch vollkommen zeitunabhängig sein und dementsprechend nach wie vor ein Vorteil beim Lernen mit Standbildern sein. Darum bleibt auch die vierte Hypothese unverändert:

***H4: Lernende mit niedrigem räumlichen Vorstellungsvermögen profitieren von Animationen im Vergleich zu 4 Standbildern in höherem Maße als Lernende mit hohem räumlichen Vorstellungsvermögen. Gleiches gilt für das Verhältnis von 11 Standbildern zu 4 Standbildern.***

## 6.4 Methode

### 6.4.1 Variablen

#### *Unabhängige Variable (UV)*

Als Unabhängige Variable in dieser Untersuchung dient weiterhin die *Art der Visualisierung*, also ob dem Lernenden eine Animation, eine Serie von vier oder elf Standbildern präsentiert wird.

#### *Abhängige Variablen (AV)*

Die Abhängige Variable stellt weiterhin der *Lernerfolg* dar, gemessen anhand der erzielten Punkte im Wissensnachtest. Dieser wird weiterhin differenziert in offene und Multiple-Choice-Fragen, ebenso in Fragen nach Faktenwissen und tiefer gehendem Verständnis.

Der *Cognitive Load* wird nach wie vor mit zwei adaptierten Items von Paas (1992) erhoben, auch wenn der angewandte Test weder durchgehend gute Reliabilitätswerte erzielte noch zwischen den Gruppen differenzieren konnte. Andere Messmethoden als die Selbstauskunft, wie etwa die sogenannte *Dual-Task-Methode* (Brünken, Steinbacher, Plass & Leutner, 2002), würden weiterhin einen zu hohen



Aufwand erfordern, der nicht durch die Rolle des *Cognitive Load* als lediglich am Rande interessierender Untersuchungsaspekt gerechtfertigt wird.

#### *Moderatorvariablen*

Mutmaßliche Moderatorvariablen bleiben weiterhin die *Höhe des themenspezifischen Vorwissens* (erhoben mit denselben vier Items aus der Pilotstudie) sowie die *Spatial Ability*, gemessen mit Hilfe sowohl des bereits eingesetzten Paper-Folding-Tests von Ekstrom et al. (1976) als auch mittels des ebenfalls bereits benutzten Card-Rotation-Tests derselben Autoren. Zwar sollen beide Tests das räumliche Vorstellungsvermögen erfassen, die zweite experimentelle Studie gab jedoch Hinweise darauf, dass womöglich unterschiedliche Facetten desselben Konstrukts getestet werden.

#### *Störvariablen*

Da das Design der Studie weitgehend gleich geblieben ist, ergeben sich im Grunde keine anderen potentiellen Störvariablen als bei der Pilotstudie (vgl. Kapitel 4.3.4). Lediglich die Zeitkomponente, wie bereits angesprochen, könnte einen verzerrenden Einfluss haben, da es nun denkbar erscheint, dass abhängig von individueller Bearbeitungsgeschwindigkeit manche Lernenden die dargebotenen Inhalte häufiger betrachten als andere. Die insgesamt zur Verfügung gestellte Zeit von 15 Minuten reichte aber (nach unsystematischer Befragung und Beobachtung zu urteilen) grundsätzlich aus, um sicherzustellen, dass jeder Proband zumindest einmal die komplette Lernumgebung bearbeitete.

### **6.4.2 Materialien**

#### *Computergestützte Lernumgebungen*

Die bereits in den vorangegangenen Studien verwendeten drei Lernumgebungen werden, wie bereits angedeutet, in modifizierter Form abermals eingesetzt. Die Änderungen beziehen sich auf den Teil, in dem entweder die Animation oder die vier bzw. elf Standbilder den Vorgang der Ablösung des Schmutzes durch Tenside darstellen. Im Gegensatz zu den bisherigen Versionen sind nun jeweils neben dem „Start“-Button weitere interaktive Steuerelemente hinzu gekommen, die den Benutzer in die Lage versetzen, nach Gutdünken die Animation bzw. den Ablauf der

Standbilder anzuhalten, auf den Anfang zu setzen oder (im Falle der Standbilder) ein Bild vor- oder zurück zu navigieren (vgl. Abbildung 6-1).

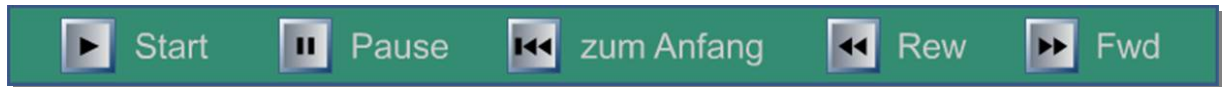


Abbildung 6-1: Die in dieser Studie modifizierte interaktive Steuerungsleiste.

Davon abgesehen sind die drei Versionen der Lernumgebungen exakt identisch mit denen aus der zweiten empirischen Studie und beinhalten denselben Audiokommentar (vgl. Anhang G).

### *Fragebögen*

Für diese Untersuchung werden die Fragebögen der Pilotstudie (wegen der Absicht, ebenfalls Studierende als Probanden zu rekrutieren) unverändert übernommen (Anhänge B und C), dazu kommen die bereits in der zweiten Studie eingesetzten Paper-Folding- (Anhang E) und Card-Rotation-Tests (Anhang F) zur Messung des räumlichen Vorstellungsvermögens.

### **6.4.3 Räumlichkeiten und Hardware**

Ein an der Universität Duisburg-Essen neu eingerichtetes Computerlaboratorium wird zur Untersuchung verwendet. In diesem können bis zu 15 Probanden gleichzeitig an Computern neuester Generation störungsfrei parallel arbeiten. Zudem ist jeder Rechner mit einem Kopfhörer ausgestattet.

### **6.4.4 Untersuchungsablauf**

Der Untersuchungsablauf unterschied sich nur geringfügig von dem der vorherigen Studien: Eine Untersuchung dauerte etwa 60 Minuten. Nach einer kurzen Begrüßung, Einführung und Aufklärung über die Anonymität der Untersuchung bekamen die Probanden den ersten Fragebogen (Anhang B) mit der Auflage zugeteilt, ihn zügig, aber trotzdem so genau wie möglich zu beantworten. Dies nahm in etwa 15 Minuten in Anspruch. Dabei wurden der Paper-Folding-Test (Anhang E) sowie der Card-Rotation-Test (Anhang F) als Bestandteil des ersten Fragebogens

aufgrund deren Speed-Test-Charakter gleichzeitig mit allen Teilnehmern durchgeführt.

Nach der Rückgabe der Fragebögen begann, gleichzeitig für alle Probanden, die Phase der individuellen Bearbeitung der jeweiligen Lernsoftware. Die Probanden wurden dabei angewiesen, still und für sich zu arbeiten und die Kopfhörer aufzusetzen. Für eventuell auftauchende technische Probleme stand der Versuchsleiter zur Verfügung, für fachliche bzw. den Programmablauf betreffende Fragen jedoch ausdrücklich nicht. Nach Ablauf von 15 Minuten – in dieser Zeit konnten die Probanden die Lernumgebung inklusive der Animation bzw. der Serien von Standbildern so ausführlich wie jeweils gewünscht bearbeiten – wurden die Rechner ausgeschaltet und der erste Teil des Posttests verteilt. Dieser sollte ebenfalls einzeln bearbeitet werden. Nach dessen Beendigung erhielt jeder Proband im Austausch den zweiten Teil des Nachtests, der teilweise Informationen in den Fragen enthielt, mit deren Hilfe Items des ersten Teils des Nachtests hätten gelöst werden können; deshalb diese Zweiteilung.

Sobald jeder diesen ebenfalls abgeschlossen hatte, bedankte sich der Versuchsleiter für die Teilnahme und bot noch eine Frage- und Diskussionsrunde zum Hintergrund der Untersuchung an.

Die Untersuchungen dieser Studie fanden an neun unterschiedlichen Terminen in der Zeit vom 13.11. bis zum 22.11.2006 in der Universität Duisburg-Essen statt. Die Teilnehmerzahlen lagen zwischen zwei und 15 Personen.

#### **6.4.5 Probanden**

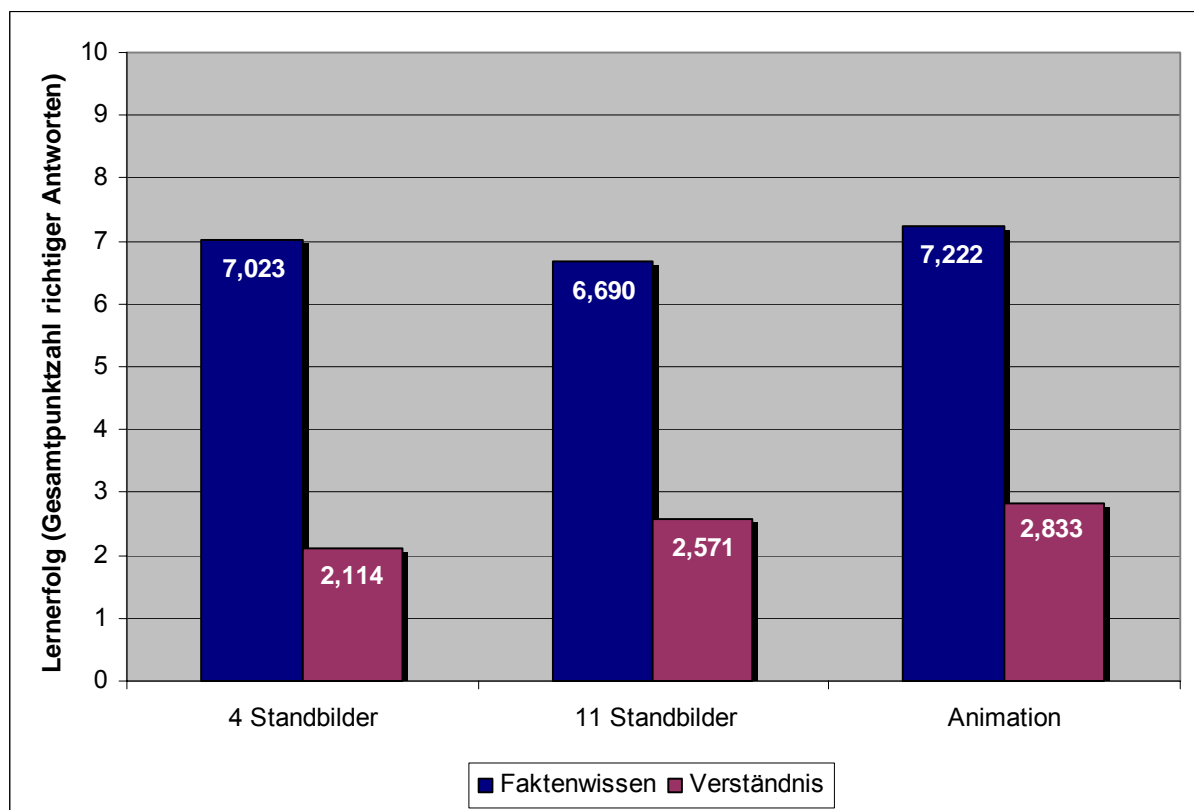
An der Studie nahmen insgesamt 61 Probanden teil. Fast alle waren Studierende des Lehramts Chemie (ein Teilnehmer strebte einen Diplom- zwei einen Magisterabschluss an). Das Durchschnittsalter betrug 20,9 Jahre, mit einem Minimum von 19 und einem Maximum von 37 Jahren. 86,9% der Probanden befanden sich im 1. Semester Chemie, waren also Novizen, nur sehr wenige im zweiten, dritten oder gar siebten Semester. 57,4% der Teilnehmer waren weiblich.

## 6.5 Ergebnisse

Die hier dargestellten statistischen Ergebnisse dieser Studie wurden wiederum mit den Statistikprogrammen SPSS 12.02 (2004) sowie GPOWER von Faul & Erdfelder (1992) mit Hilfe des Maßes der gepoolten Streuung berechnet. Die Darstellung der Ergebnisse erfolgt anhand der zuvor aufgestellten Hypothesen.

### 6.5.1 Hypothesen 1a, 1b und 2: Animationen versus Standbilder

Hypothese 1a drückt die Erwartung aus, die vorangegangenen Ergebnisse zu bestätigen, dass die Animationsversion insgesamt der 4-Standbilderversion überlegen sein sollte, während Hypothese 1b auf die 11-Standbilderversion fokussiert. Nach der Modifizierung der Lernumgebungen wurde erwartet, dass diese der Animation ebenbürtig sein, sich der 4-Standbilderversion aber überlegen zeigen sollte.



*Abbildung 6-1: Mittelwerte der erzielten Lernleistungen der Probanden mit Standbilder- oder Animationsversionen, aufgeteilt nach Faktenwissen und Verständnisaufgaben.*

*Tabelle 6-1: Deskriptive Kennwerte des Lernerfolgs, aufgeteilt nach Visualisierungsart.*

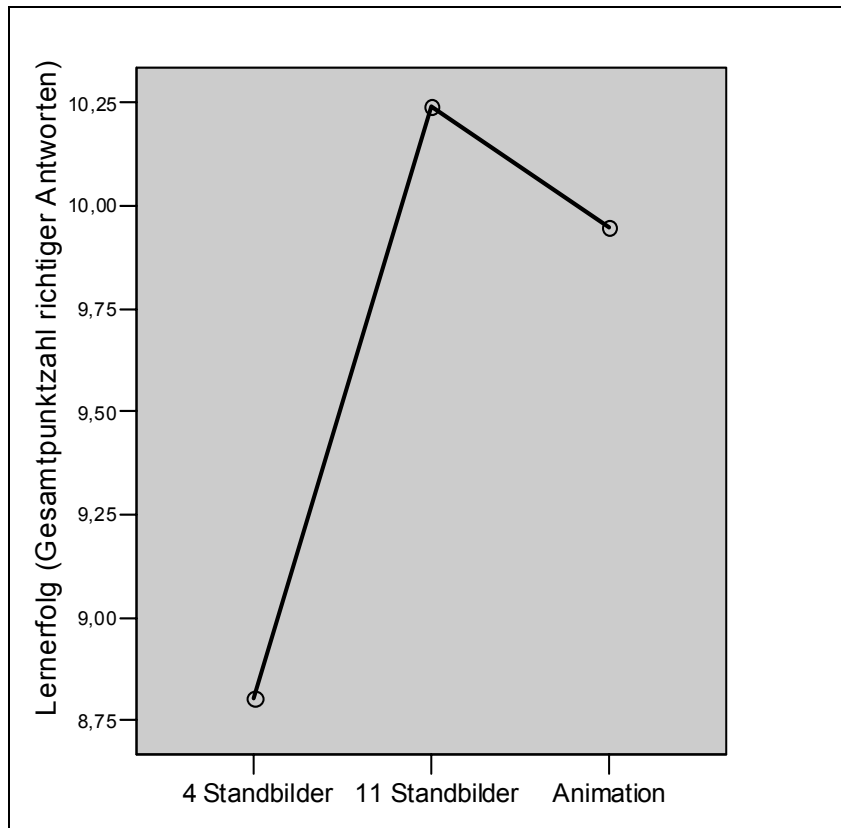
<b>Abhängige Variable</b>	<b>Visualisierungsart</b>	<b>N</b>	<b>M</b>	<b>SD</b>
<i>Lernerfolg gesamt</i>	<i>4 Standbilder</i>	22	9,14	3,31
	<i>11 Standbilder</i>	21	9,26	3,74
	<i>Animation</i>	18	10,06	2,39
<i>Lernerfolg Faktenwissen</i>	<i>4 Standbilder</i>	22	7,02	2,14
	<i>11 Standbilder</i>	21	6,69	2,63
	<i>Animation</i>	18	7,22	1,44
<i>Lernerfolg Verständnis</i>	<i>4 Standbilder</i>	22	2,11	1,53
	<i>11 Standbilder</i>	21	2,57	1,36
	<i>Animation</i>	18	2,83	1,45

Wie bereits die in Tabelle 6-1 und Abbildung 6-1 dargestellten deskriptiven Kennwerte andeuten, glückt die Bestätigung dieser Hypothesen zunächst nicht. Zwar fallen die Mittelwerte des Gesamtlernerfolgs in der erwarteten Reihenfolge aus (der Mittelwert der Animationsversion am höchsten, der der 4-Standbilderversion am niedrigsten), statistisch signifikant ist dieses Ergebnis jedoch nicht. Beim Vergleich der Gruppen via Varianzanalyse zeigen sich für den Gesamtlernerfolg keine signifikanten Unterschiede ( $F(2,58) < 1$ ,  $p = 0,64$ ,  $MSE = 10,45$ ,  $\eta^2 = 0,02$ ). Die hypothesenkonformen einseitigen Einzelvergleiche mittels a-priori-Kontrasttests kommen dem 5%-Signifikanzniveau zwar näher, verfehlen allesamt aber trotzdem diese Grenze (vgl. Tabelle 6-2).

*Tabelle 6-2: Ergebnisse der einseitigen a-priori-Kontrastvergleiche zwischen den drei Versionen der Lernumgebung. Ein negativer t-Wert kennzeichnet einen hypothesenkonträren höheren Mittelwert bei einer der Versionen. Dort gilt für p im Übrigen ein zweiseitiger Wert.*

<b>Abhängige Variable</b>	<b>Kontrast</b>	<b>df</b>	<b>t</b>	<b>p</b>	<b>d</b>
Lernerfolg gesamt	Animation – 4 Standbilder	58	0,894	0,19	0,31
	Animation – 11 Standbilder	58	0,764	0,22	0,25
	4 Standbilder – 11 Standbilder	58	0,127	0,45	0,03
Lernerfolg Faktenwissen	Animation – 4 Standbilder	58	0,291	0,39	0,11
	Animation – 11 Standbilder	58	0,768	0,22	0,25
	4 Standbilder – 11 Standbilder	58	-0,505	0,62	0,14
Lernerfolg Verständnis	Animation – 4 Standbilder	58	1,564	0,06	0,48
	Animation – 11 Standbilder	58	0,563	0,29	0,19
	4 Standbilder – 11 Standbilder	58	1,036	0,15	0,32

Eine explorative Datenanalyse lässt die vorhandenen hypothesenkonformen Tendenzen zwar deskriptiv deutlicher werden; unter Einbeziehung einer Reihe von Kovariaten kann für den Unterschied von Animations- und 4-Standbilderversion auch ein signifikanter Effekt wie in den vorangegangenen Studien ausgemacht werden. So muss das Vorwissen nicht nur durch den Vortest operationalisiert werden, sondern zusätzlich durch die letzte Chemie- sowie die Gesamtabiturnote. Wenn dann noch der *Cognitive Load* als vierte Kovariate in die Berechnungen einbezogen wird – was aufgrund eines möglichen Moderator- oder Mediatoreffektes dieser Variable recht problematisch ist und die Interpretierbarkeit des Ergebnisses in Frage stellt –, ergibt sich ein hypothesenkonformes Bild (Abbildung 6-2) sowie ein signifikanter einseitiger Kontrast beim Einzelvergleich von vier Standbildern und Animation, bezogen auf den Gesamtlernerfolg ( $F(1,50)=2,86$ ,  $p<0,05$ ,  $d=0,37$ ).



*Abbildung 6-2: Korrigierte Mittelwerte des Gesamtlernerfolgs der drei Versionen unter Einbezug zusätzlicher Kovariaten (Cognitive Load, Ergebnis im Vortest sowie letzte Chemie- und Abiturnote).*

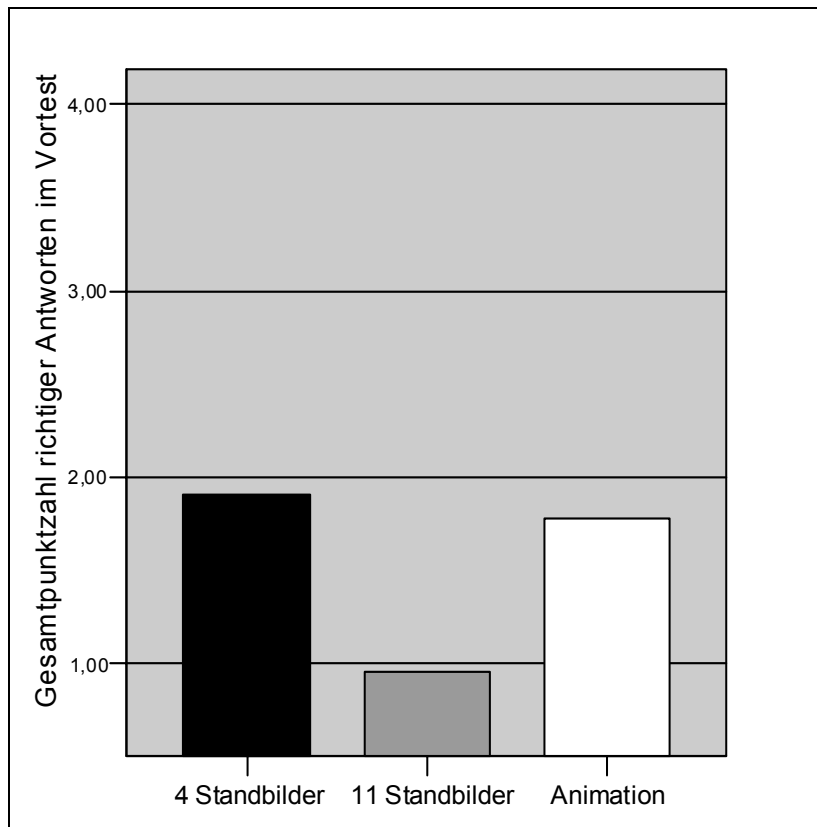
Zur zweiten Hypothese, die nach unterschiedlichen Effekten beim Erwerb von Faktenwissen und tieferem Verständnis fragt, wurden bereits in den vorangegangenen Studien uneinheitliche Ergebnisse gefunden. Auch in diesem Falle wird kein eindeutiger Vorteil der Animation beim Erwerb tieferen Verständnisses gefunden, allerdings (vgl. Tabelle 6-2) ergibt sich in dieser Untersuchung ein deutlich stärkerer Effekt als beim Faktenwissen. Deshalb ist auch nur der Einbezug einer einzigen Kovariate notwendig (die letzte Chemienote als Operationalisierung des allgemeinen Vorwissens), um einen statistisch signifikanten Kontrast zwischen der Animations- und der 4-Standbilderversion zu erhalten:  $F(1,56)=0,85$ ,  $p<0,04$ ,  $d=0,55$ .

### 6.5.2 Hypothese 3: Die Rolle des themenspezifischen Vorwissens

Ausgehend von den Ergebnissen der zweiten experimentellen Studie wird ein Moderatoreffekt des Vorwissens dahingehend erwartet, dass Lernende mit niedrigem

Vorwissen eher von Animationen profitieren, Experten hingegen auch von Standbildern.

Zunächst zeigt sich, dass auch in dieser Untersuchung das Vorwissen insgesamt wie erwartet recht niedrig ausgeprägt ist ( $M=1,54$ ,  $Md= 1,5$ ,  $SD=1,42$ ). Allerdings treten ungünstigerweise, trotz randomisierter Verteilung der Versuchspersonen auf die einzelnen Bedingungen, systematische Unterschiede im Vorwissen zwischen den Gruppen auf (vgl. Abbildung 6-3 sowie Tabelle 6-3): Das Ergebnis einer Varianzanalyse ( $F(2,58)=2,99$ ,  $p=0,058$ ,  $MSE=1,89$ ) liegt deutlich unter dem für das Nichtverwerfen der Nullhypothese üblichen 20%-Signifikanzniveau, im Einzelvergleich ist die 4-Standbilderversion der 11-Standbilderversion hinsichtlich des Vorwissens sogar auf dem 5%-Niveau signifikant überlegen ( $t(58)=2,28$ ,  $p<0,05$ ,  $d=0,74$ ).



*Abbildung 6-3: Darstellung der Mittelwerte im Vortest. Deutlich wird das zufällig niedrigere Vorwissensniveau bei der 11-Standbilderversion.*



*Tabelle 6-3: Deskriptive Kennwerte für die Ergebnisse des Vortests.*

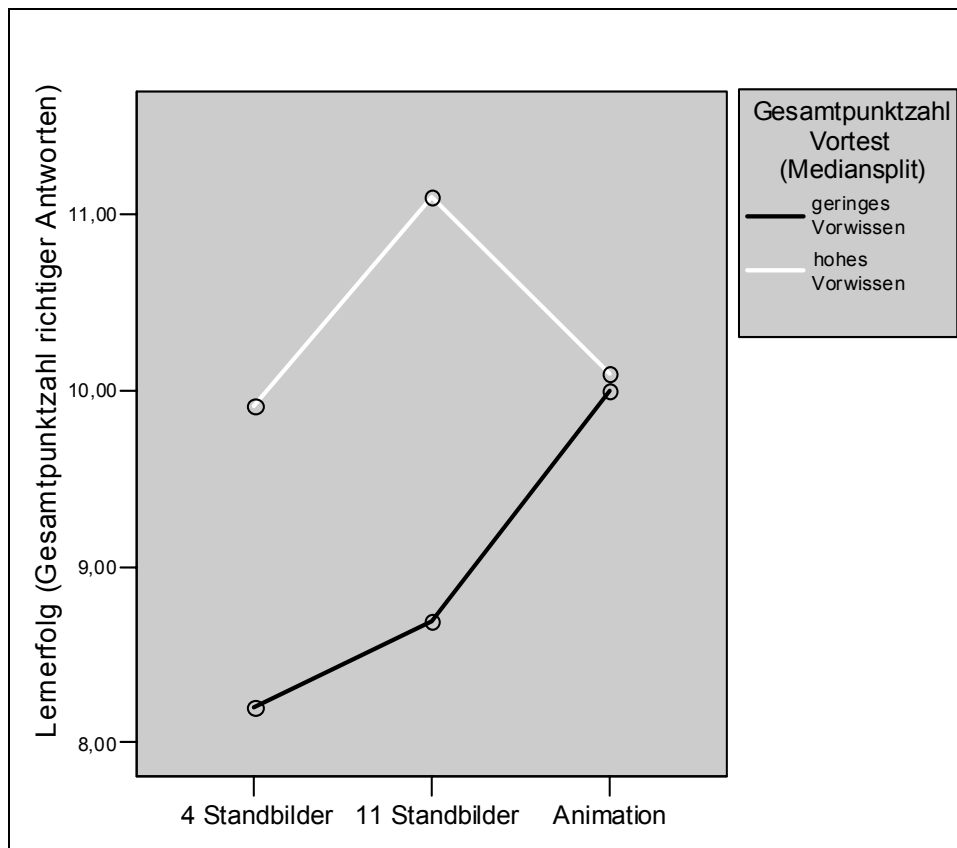
<b>Visualisierungsart</b>	<b>N</b>	<b>M</b>	<b>SD</b>
<i>4 Standbilder</i>	22	1,91	1,42
<i>11 Standbilder</i>	21	0,95	1,14
<i>Animation</i>	18	1,78	1,55
<i>Gesamt</i>	61	1,54	1,42

Eine solche Verteilung kann natürlich alle weiteren Berechnungen zur Rolle des themenspezifischen Vorwissens verzerren, weshalb das Vorwissen grundsätzlich als Kovariate sowie bei der im Anschluss folgenden Interpretation mit einbezogen werden muss.

Es wurde nun, parallel zu den Auswertungen der vorangegangenen Studien, ein Mediansplit zur Erzeugung zweier Gruppen, „Novizen“ (N=34) und „Experten“ (N=27) durchgeführt. Die deskriptiven Werte sind in Tabelle 6-4 aufgeführt.

*Tabelle 6-4: Deskriptive Kennwerte des Lernerfolgs für die Subgruppen themenspezifischen Vorwissens nach der Aufteilung per Mediansplit.*

<b>Visualisierungsart</b>	<b>Vorwissen</b>	<b>N</b>	<b>M</b>	<b>SD</b>
<i>4 Standbilder</i>	<i>niedrig</i>	10	8,20	4,05
	<i>hoch</i>	12	9,92	2,45
	<i>gesamt</i>	22	9,14	3,31
<i>11 Standbilder</i>	<i>niedrig</i>	16	8,69	4,02
	<i>hoch</i>	5	11,10	1,92
	<i>gesamt</i>	21	9,26	3,74
<i>Animation</i>	<i>niedrig</i>	8	10,00	2,28
	<i>hoch</i>	10	10,10	2,59
	<i>gesamt</i>	18	10,06	2,39
<i>Gesamt</i>	<i>niedrig</i>	34	8,85	3,66
	<i>hoch</i>	27	10,20	2,37
	<i>gesamt</i>	61	9,45	3,20



*Abbildung 6-4: Interaktionsdiagramm des Lernerfolgs der Gruppen mit geringem und hohem Vorwissen in Abhängigkeit von der Visualisierungsart.*

Eine zweifaktorielle ANOVA zum Nachweis eines möglichen Interaktionseffektes von Vorwissen und Visualisierungsart (Animation, vier oder elf Standbilder) findet kein statistisch signifikantes Ergebnis, auch wenn das Profildiagramm eine Interaktion suggeriert (Abbildung 6-4):  $F(2,55) < 1$ ,  $p = 0,56$ ,  $MSE = 10,33$ ,  $\eta^2 = 0,02$ ; vgl. Tabelle 6-5. Ein solcher Nachweis gelingt ebenfalls dann nicht, wenn die 11-Standbilderversion unberücksichtigt bleibt. Ebenfalls ist kein Haupteffekt des Vorwissens nachzuweisen.

*Tabelle 6-5: Ergebnisse der zweifaktoriellen Varianzanalyse für die Einflüsse der Visualisierungsart und des themenspezifischen Vorwissens (Mediansplit).*

<b>Abhängige Variable</b>	<b>Varianzquelle</b>	<b>SS*</b>	<b>df</b>	<b>MS</b>	<b>F<sub>emp</sub></b>	<b>p</b>	<b>η<sup>2</sup></b>
<b>Lernzuwachs gesamt</b>	<i>Haupteffekt Vorwissen**</i>	27,46	1	27,46	2,66	0,11	0,05
	<i>Haupteffekt Visualisierungsart</i>	8,38	2	4,19	0,41	0,67	0,02
	<i>Interaktionseffekt</i>	11,96	2	5,98	0,58	0,56	0,02
	<i>Fehler</i>	568,05	55	10,33	---	---	---

\* sequentielle Varianzzerlegung

\*\* Mediansplit

### 6.5.3 Hypothese 4: Die Rolle der Spatial Ability

In den beiden bisherigen Studien zeigen sich Belege für eine kompensierende Wirkung von räumlichem Vorstellungsvermögen, genauer der Facette *Spatial Visualization*, die mit dem Paper-Folding-Test erhoben wurde. Lernende mit niedrigem räumlichen Vorstellungsvermögen erzielten geringere Lernergebnisse beim Lernen mit Standbildern als Lernende mit hohem räumlichen Vorstellungsvermögen.

Wie bereits in der Erörterung der Hypothese (Kapitel 6.3) als Möglichkeit angesprochen, lässt sich dieser Effekt tatsächlich nicht ein weiteres Mal replizieren. Bei der Testung mit dem Paper-Folding-Test zeigen sich weder mit der Kovarianzanalyse über alle drei Versuchsbedingungen (vgl. Tabelle 6-6), noch in Einzelvergleichen signifikante ATI-Effekte. Tabelle 6-6 mit den Ergebnissen der Kovarianzanalyse macht deutlich, dass in dieser Untersuchung die *Spatial Visualization* lediglich in Form eines Haupteffektes ( $F(1,55)=7,60$ ,  $p<0,01$ ,  $MSE=9,32$ ,  $\eta^2=0,12$ ) eine Rolle spielt: Je höher das Ergebnis im Paper-Folding-Test, desto höher auch der Lernerfolg, und zwar unabhängig von der Visualisierungsart. Auch eine Kontrolle des Vorwissens ändert an diesen Ergebnissen nichts (hier nicht dargestellt).

*Tabelle 6-6: Ergebnisse der Kovarianzanalyse zum Effekt der Spatial Ability (als Kovariate, ohne Anwendung eines Mediansplits) mit dem Paper-Folding-Test.*

<b>Abhängige Variable</b>	<b>Varianzquelle</b>	<b>SS*</b>	<b>df</b>	<b>MS</b>	<b>F<sub>emp</sub></b>	<b>p</b>	<b>η<sup>2</sup></b>
<b>Lernzuwachs gesamt</b>	<i>Haupteffekt Spatial Visualization</i>	70,81	1	70,81	7,60	0,01	0,12
	<i>Haupteffekt Visualisierungsart</i>	14,98	2	7,49	0,80	0,45	0,03
	<i>Interaktionseffekt</i>	17,24	2	8,62	0,93	0,40	0,03
	<i>Fehler</i>	512,82	55	9,32	---	---	---

\* sequentielle Varianzzerlegung

Nach den Ergebnissen der zweiten Studie erwartungsgemäß zeigen sich bei der Erhebung der *Spatial Ability* mit dem Card-Rotation-Test (die *Spatial-Relations*-Komponente) ebenfalls keine moderierenden Effekte. Im Gegensatz zum Paper-Folding-Test kann bei der Verwendung des Card-Rotation-Tests darüber hinaus auch kein Haupteffekt des räumlichen Vorstellungsvermögens festgestellt werden (vgl. Tabelle 6-7).

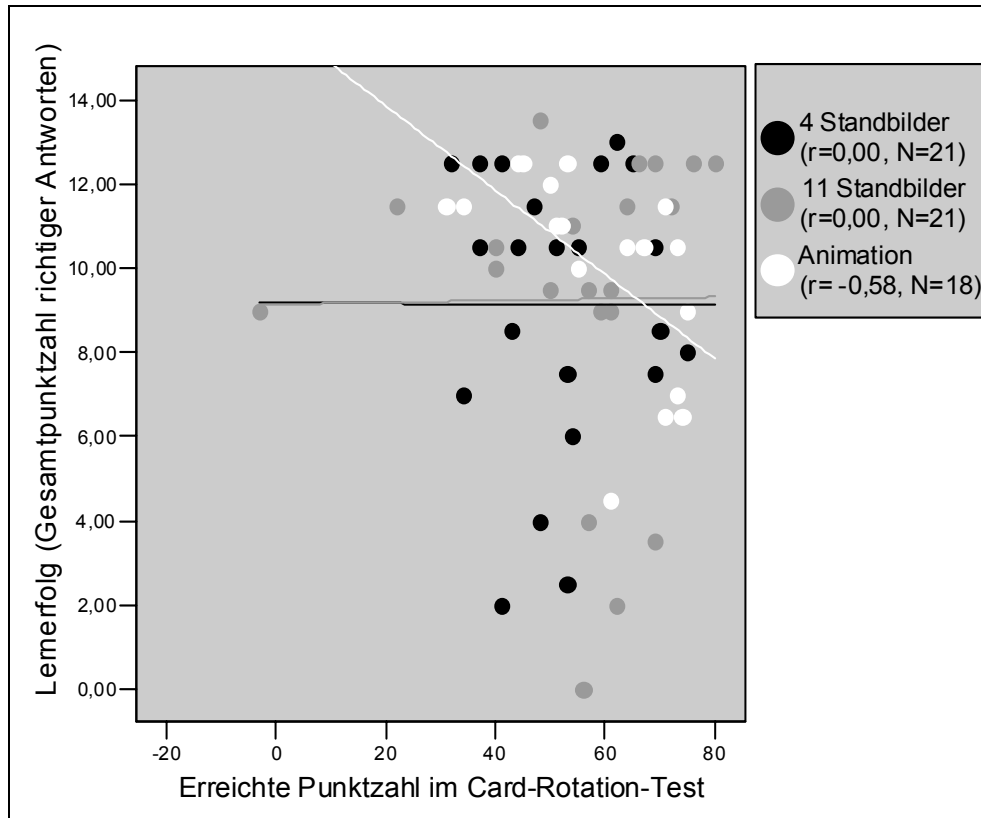
*Tabelle 6-7: Ergebnisse der Kovarianzanalyse zum Effekt der Spatial Ability (als Kovariate, ohne Anwendung eines Mediansplits) mit dem Card-Rotation-Test.*

<b>Abhängige Variable</b>	<b>Varianzquelle</b>	<b>SS*</b>	<b>df</b>	<b>MSE</b>	<b>F<sub>emp</sub></b>	<b>p</b>	<b>η<sup>2</sup></b>
<b>Lernzuwachs gesamt</b>	<i>Haupteffekt Spatial Relations</i>	4,95	1	4,95	0,48	0,49	0,01
	<i>Haupteffekt Visualisierungsart</i>	11,94	2	5,97	0,57	0,57	0,02
	<i>Interaktionseffekt</i>	25,73	2	12,87	1,23	0,30	0,04
	<i>Fehler</i>	573,23	55	10,42	---	---	---

\* sequentielle Varianzzerlegung

Weder die Einbeziehung von zusätzlichen Kovariaten noch das Testen im Einzelvergleich ändern etwas an dieser Nichtauffindung signifikanter Effekte. Abbildung 6-5 schließlich stellt das Streudiagramm der Verteilung des Lernerfolgs und der Punkte im Card-Rotation-Test in Abhängigkeit von der Visualisierungsart (ohne Berücksichtigung weiterer Kovariaten) dar. Es scheint bei den

Standbildversionen keinen Zusammenhang zwischen beiden Variablen zu geben, während für die Animationsversion sogar eine starke negative Beziehung suggeriert wird. Bei Ausschluss des auffälligen Outliers (-3;9; vgl. Abbildung 6-5 links) ändert sich nichts an diesem Bild.



*Abbildung 6-5: Streudiagramm nach Lernerfolg auf der Y-Achse und Punktzahl im Card Rotation-Test auf der X-Achse in Abhängigkeit von der jeweiligen Visualisierungsart inklusive zugehöriger Regressionslinien (auf die Darstellung der Vertrauensintervalle wurde aus Gründen der Übersichtlichkeit verzichtet).*

#### 6.5.4 Weitere Berechnungen

Durch die Neuimplementierung minimaler Kontrollfunktionen in die Lernumgebungen erscheint ein moderierender Einfluss des *Cognitive Load* nun noch unwahrscheinlicher; eine Hypothese dahingehend war von vornherein aufgrund der Erfahrungen der Pilotstudie nicht aufgestellt worden. Die deskriptiven Kennwerte (Tabelle 6-8) wie auch das Ergebnis der einfaktoriellen Varianzanalyse mit dem *Cognitive Load* als abhängiger Variable ( $F(2,58)=1,20$ ,  $p=0,31$ ,  $MSE=3,11$ ,  $\eta^2=0,04$ ),

zeigen, dass tatsächlich keine Unterschiede zwischen den Versionen hinsichtlich der kognitiven Belastung bestehen.

**Tabelle 6-8:** Deskriptive Kennwerte für die Items der Cognitive Load-Skala, abhängig von der Visualisierungsart.

<b>Cognitive Load-Item</b>	<b>Visualisierungsart</b>	<b>N</b>	<b>M</b>	<b>SD</b>
Wie leicht oder schwer war das Lernmaterial zu verstehen?	4 Standbilder	22	2,41	0,19
	11 Standbilder	21	2,90	0,19
	Animation	18	2,22	0,21
Bei der Bearbeitung des Lernmaterials war meine mentale Anstrengung...	4 Standbilder	22	3,55	0,19
	11 Standbilder	21	3,57	0,27
	Animation	18	3,39	0,26
Gesamt	4 Standbilder	22	5,96	0,35
	11 Standbilder	21	6,48	0,42
	Animation	18	5,61	0,41

Da jedoch zur Sichtbarmachung des Unterschieds zwischen der 4-Standbilderversion und der Animation in dieser Untersuchung die Einbeziehung des *Cognitive Load* als Kovariate von Nöten war (vgl. Kapitel 6.5.1), ist es dringend geraten, auch in dieser Untersuchung die Korrelationen zwischen den Variablen – und in diesem Falle besonders des *Cognitive Load* – zu beachten. Tabelle 6-9 gibt die Korrelationsmatrix relevanter Variablen wieder. Auffällig ist bei der Betrachtung der Matrix in diesem Zusammenhang die hohe negative Korrelation nur *eines* der *Cognitive-Load-Items* mit dem Vorwissen sowie dem Lernerfolg, nämlich die Frage nach der Schwierigkeit des Lernmaterials (*CL 1*). Die Frage nach der jeweiligen mentalen Anstrengung (*CL 2*) führt zu keinen höheren Korrelationen mit anderen Variablen. Darüber hinaus korrelieren das Vorwissen und die letzte Chemienote (in Punkten von Null bis 15) statistisch signifikant negativ mit der Abiturnote (negativ wegen der umgekehrten Polung von Schulnoten).

**Tabelle 6-9:** Mittelwerte, Standardabweichungen und Matrix der Produktmomentkorrelationen relevanter Variablen.

	M	SD		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Lernerfolg gesamt (1)	9,45	3,20	r	1										
			N	61										
Lernerfolg Fakten- wissen (2)	6,70	2,13	r	0,93**	1									
			N	61	61									
Lernerfolg Verständnis (3)	2,48	1,46	r	0,84**	0,58**	1								
			N	61	61	61								
Abiturnote (4)	2,51	0,54	r	-0,14	-0,14	-0,10	1							
			N	61	61	61	61							
Chemie- punkte (5)	10,03	2,53	r	0,01	-0,02	0,04	-0,37*	1						
			N	60	60	60	60	60						
Vorwissen (6)	1,54	1,42	r	0,27*	0,39**	0,02	-0,32*	-0,02	1					
			N	61	61	61	61	60	61					
Paper- Folding-Test (7)	4,05	3,61	r	0,34*	0,33	0,27*	-0,12	0,12	0,18	1				
			N	61	61	61	61	60	61	61				
Card- Rotation- Test (8)	54,80	15,35	r	-0,09	-0,03	-0,15	0,02	0,05	0,01	0,21	1			
			N	61	61	61	61	60	61	61	61	61		
CL 1 (9)	2,52	0,92	r	-0,31*	-0,39**	-0,11	0,22	0,11	-0,48**	-0,19	0,05	1		
			N	61	61	61	61	60	61	61	61	61	61	
CL 2 (10)	3,51	1,07	r	-0,11	-0,10	-0,09	-0,06	0,11	-0,10	-0,14	-0,06	0,57**	1	
			N	61	61	61	61	60	61	61	61	61	61	61
CL gesamt (11)	6,03	1,77	r	-0,23	-0,27*	-0,11	0,07	0,12	-0,31*	-0,19	-0,01	0,87**	0,90**	1
			N	61	61	61	61	60	61	61	61	61	61	61

\* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

\*\* Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Die Reliabilitäten der angewendeten Skalen fallen in dieser Studie weitgehend zufriedenstellend aus. Zwar liegt die Vorwissensskala wiederum im untersten Bereich eines akzeptablen Niveaus ( $\alpha=0,68$ ), alle weiteren Skalen sind aber deutlich besser: Die wichtige Gesamtskala des Lernerfolgs weist ein Cronbachs Alpha von 0,81 auf – einbezogen in die Berechnungen werden hier natürlich einmal mehr nur die sich explizit auf den Bereich der Visualisierung (Animation oder Standbilder) beziehenden Items und nicht diejenigen, deren Inhalt in der einbettenden Lernumgebung behandelt wurde. Trotzdem verbleibt ein akzeptabler Wert von  $\alpha=0,79$ .

Die Reliabilität der Skala zur kognitiven Belastung kann dieses Mal eine akzeptable Höhe von  $\alpha=0,72$  erreichen.

## 6.6 Diskussion

Diese dritte und abschließende Untersuchung hatte das vorrangige Ziel, die ursprüngliche Hypothese, dass eine Animation prinzipiell durch eine Serie von ausreichend vielen Standbildern zu ersetzen ist, mittels nunmehr angepasster Lernumgebungen schließlich doch noch zu bestätigen. Es muss konstatiert werden, dass dieses Ziel nicht in überzeugender Weise erreicht werden konnte. Die Einführung minimaler Interaktivität durch höhere Lernerkontrolle des Fortlaufs der Animation bzw. der Serie von Standbildern erbrachte zwar deutlich höhere Lernerfolge gegenüber den vorangegangenen Studien, wie vorausgesagt besonders im Falle der 11-Standbilderversion (die zumindest den anderen Versionen nicht länger signifikant unterlegen war), signifikante Überlegenheiten der Animation und der 11-Standbilderversion gegenüber der 4-Standbilderversion konnten jedoch nur unter Einbezug zahlreicher Kovariaten ausgemacht werden.

Da das Vorwissen per Zufall zwischen den Versuchsbedingungen ungleich verteilt war, musste dies natürlich in die Analysen miteinbezogen werden – dies stellt also noch keine interpretative Einschränkung des Effektes dar. Kritischer zu sehen und schwieriger zu interpretieren ist jedoch, dass die erwartete Überlegenheit der Animation gegenüber der 4-Standbilderversion erst unter zusätzlichen Einbezug des *Cognitive Load* als Kovariate in die ANCOVA mit sequentieller Varianzzerlegung statistisch signifikant wurde, wodurch auch das resultierende Bild (Abbildung 6-2) die erwarteten Unterschiede zwischen den drei Versionen deskriptiv wiedergab. Hätte der *Cognitive Load* tatsächlich die Rolle eines Mediators (der sich in den vorangegangenen Untersuchungen allerdings nicht bestätigen ließ), sollte jeglicher Effekt des Faktors auf die AV verschwinden, wenn der *Cognitive Load* kontrolliert wird. Da dies nicht der Fall ist, sondern im Gegenteil erst den erwarteten Effekt begünstigt, scheint eine andere Rolle vorzuliegen. Weil sich das Ausmaß des *Cognitive Load* nicht zwischen den Versuchsbedingungen unterschied, scheint es der Fall zu sein, dass diese Variable in diesem Fall lediglich Fehlervarianz (eventuell Motivationseffekte) bindet, was die Aussagekraft des Befundes, dass die Animation ein weiteres Mal den 4 Standbildern überlegen ist, weitgehend wieder herstellt. Zu



dieser Annahme passt, dass eines der beiden *Cognitive-Load-Items*, die Frage nach der mentalen Anstrengung der Lernenden während der Bearbeitung, nur eine sehr geringe Korrelation mit dem Lernerfolg (-0,11) aufweist. Hätte die kognitive Belastung tatsächlich eine mediierende Funktion, wäre hier eine hohe Korrelation zu erwarten gewesen.

Daneben lässt sich als Alternativerklärung auch nicht ganz ausschließen, dass die Ermöglichung des *self-pacing* nicht nur den Lernerfolg insgesamt verbessert hat, sondern in gewisser Weise auch als gleichmachendes Element zwischen den Versionen gedient haben mag – durch die erhöhte Lernerkontrolle war es in jeder Version möglich, den auf unterschiedliche Weise veranschaulichten Prozess in selbstbestimmtem Tempo und Intensität zu verstehen. Ein Indiz für die Richtigkeit dieser These wären auch die erwartungsgemäß nicht aufgefundenen Unterschiede in der kognitiven Belastung.

Ferner spricht für die Möglichkeit, dass *self-pacing* als gleichmachendes Element zwischen den Visualisierungsarten auftrat, dass in dieser Untersuchung erstmals kein Interaktionseffekt der *Spatial Ability* bzw. genauer der *Spatial Visualization* mehr nachgewiesen werden konnte. Wie bereits in der Erörterung der inhaltlichen Hypothesen (Kapitel 6.3) als Option in Betracht gezogen, scheint die zusätzliche Interaktivität den spezifischen Vorteil höheren räumlichen Vorstellungsvermögens beim Lernen mit wenigen Standbildern zu relativieren und auszugleichen. Zwar ist eine höhere Fähigkeit der *Spatial Visualization* weiterhin insgesamt ein Faktor, der das Lernen mit Visualisierungen der hier verwendeten Art erleichtert, was sich am statistisch signifikanten Haupteffekt zeigt, aber nicht mehr speziell für den Bereich der Bildung einer mentalen Animation aus Standbildern. Es muss jedoch betont werden, dass dies keine Widerlegung des in den vorangegangenen Studien gefundenen Effektes ist – die höhere Lernerkontrolle scheint lediglich ein weiterer möglicher Faktor zur Kompensation mangelnder Darstellungsgüte (wenige Standbilder statt einer Animation) neben hohem räumlichen Vorstellungsvermögen zu sein.

Trotzdem wären natürlich weitere Untersuchungen von Vorteil, die dieser Frage weiter auf den Grund gehen sowie in denen das Vorwissen der Probanden tatsächlich gleich verteilt ist und deshalb keinen verzerrenden Einfluss hat. Dies lässt sich im Rahmen vorliegender Dissertation allerdings leider nicht mehr realisieren.

## 7 Vergleich der drei experimentellen Studien

### 7.1 Einführung

Abschließend erscheint es nun sinnvoll, die Ergebnisse der drei empirischen Studien miteinander zu vergleichen und mögliche übergreifende Trends zu identifizieren. Zudem ermöglicht dieses Vorgehen einen Vergleich der drei Probandengruppen.

### 7.2 Ergebnisse

Führt man die Ergebnisse aller drei Studien in eine Datenmatrix zusammen, ist es möglich, Unterschiede zwischen den drei Stichproben zu ermitteln. Hierbei ist zunächst abzuklären, ob sich die Probandengruppen – in der ersten und dritten Studie Studierende zumeist in den ersten Semestern eines naturwissenschaftlichen Studiums mit Ziel Staatsexamen, in der zweiten Studie Schüler der 11. Jahrgangsstufe eines Gymnasiums und einer Gesamtschule – hinsichtlich ihres Vorwissens unterscheiden, da dies spätere Interpretationen von Unterschieden stark beeinflussen würde. Wie die deskriptiven Werte in Tabelle 7-1 bereits andeuten und das Ergebnis der einfaktoriellen Varianzanalyse bestätigt ( $F(2,147)=1,41$ ,  $p=0,25$ ,  $MSE=1,62$ ,  $\eta^2=0,02$ ), ist dies nicht der Fall. Die Probanden aller drei Gruppen haben im Mittel in etwa gleich wenig Vorwissen (Maximalpunktzahl: 8).

*Tabelle 7-1: Deskriptive Kennwerte des Vorwissens in den drei Untersuchungen.*

<b>Stichprobe</b>	<b>N</b>	<b>M</b>	<b>SD</b>
1. Studie	25	1,82	1,60
2. Studie	64	1,33	0,94
3. Studie	61	1,54	1,42
Gesamt	150	1,50	1,28

Sodann erscheint es interessant, mittels einer zweifaktoriellen Varianzanalyse sowohl mögliche Unterschiede im Lernergebnis zwischen den drei Stichproben (mit besonderem Augenmerk auf die dritte Untersuchung, in der wie bereits beschrieben die Lernumgebungen verändert wurden) als auch den Effekt der Visualisierungsart

(Animation, 4 oder 11 Standbilder) über die gesamte Stichprobe hinweg zu analysieren.

*Tabelle 7-2: Deskriptive Kennwerte des Gesamtlernerfolgs, aufgeteilt nach Visualisierungsart und Untersuchung.*

<b>Visualisierungsart</b>	<b>Untersuchung</b>	<b>N</b>	<b>M</b>	<b>SD</b>
4 Standbilder	1	12	6,50	3,48
	2	21	7,41	4,01
	3	22	9,14	3,31
	Gesamt	55	7,90	3,72
11 Standbilder	2	19	6,24	3,30
	3	21	9,26	3,74
	Gesamt	40	7,83	3,81
Animation	1	13	8,65	2,38
	2	22	8,86	2,83
	3	18	10,06	2,39
	Gesamt	53	9,22	2,61
Gesamt	1	25	7,62	3,10
	2	62	7,57	3,52
	3	61	9,45	3,20
	Gesamt	148	8,35	3,43

*Tabelle 7-3 Ergebnisse der zweifaktoriellen Varianzanalyse der Effekte von Visualisierungsart und Stichprobe unter zusätzlicher Einbeziehung des Vorwissens als Kovariate.*

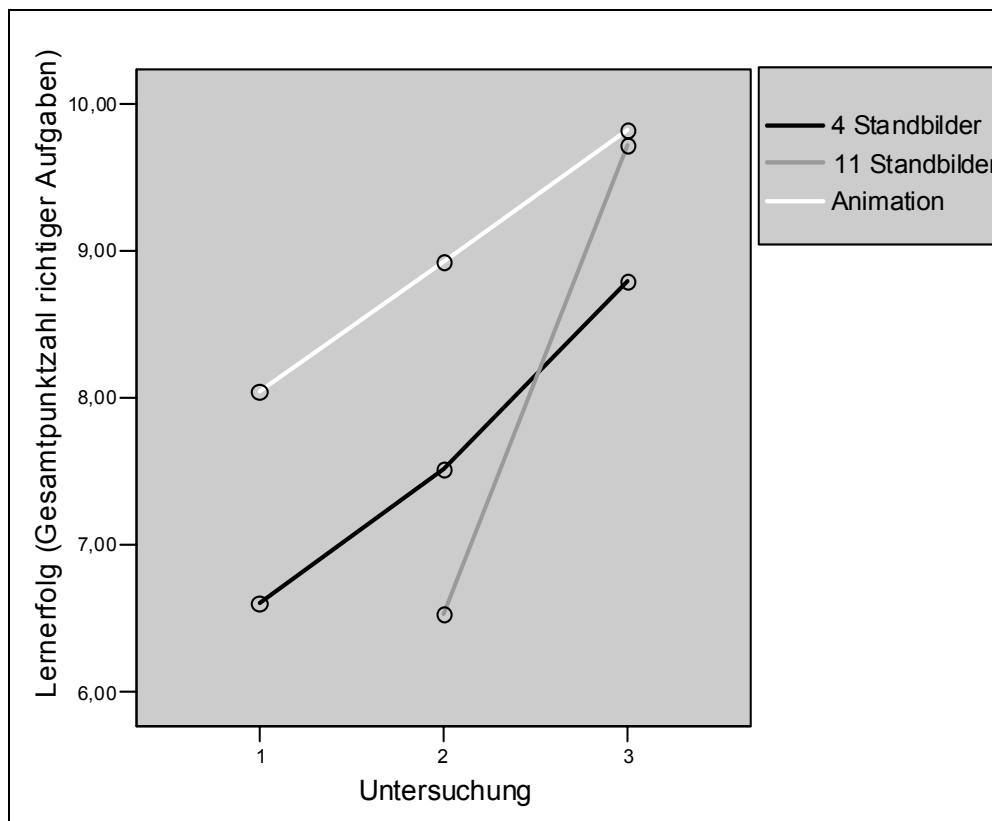
<b>Abhängige Variable</b>	<b>Varianzquelle</b>	<b>SS*</b>	<b>df</b>	<b>MS</b>	<b>F<sub>emp</sub></b>	<b>p</b>	<b>η<sup>2</sup></b>
<b>Lernzuwachs gesamt</b>	Haupteffekt Vorwissen	179,69	1	179,69	18,58	0,001	0,12
	Haupteffekt Stichprobe	119,10	2	59,55	6,16	0,003	0,08
	Haupteffekt Visualisierungsart	55,64	2	27,82	2,88	0,060	0,04
	Interaktionseffekt Visualisierungsart- Stichprobe	29,44	3	9,81	1,02	0,388	0,02
	Fehler	1344,45	139	9,67	---	---	---

\* sequentielle Varianzzerlegung

Die Ergebnisse dieser ANCOVA sind in Tabelle 7-3 dargestellt, die zugehörigen deskriptiven Kennwerte in Tabelle 7-2. Wie deutlich wird, verfehlt zwar zunächst der

Effekt für die Visualisierungsart knapp das statistische Signifikanzniveau ( $p=0,06$ ) – Einzelvergleiche zwischen den drei Versionen belegen aber die schon deskriptiv sichtbare (vgl. Tabelle 7-2) Überlegenheit der Animation, die auch in den einzelnen Untersuchungen unter Einbeziehung von Kovariaten gezeigt werden konnte. Während im Vergleich zur 4-Standbilderversion ein statistisch signifikanter Vorteil ( $p<0,05$ ) gefunden wird, gibt es – hypothesenkonform – keinen Unterschied im Vergleich zur 11-Standbilderversion ( $p=0,081$ ). Dies ist übrigens auch der Fall bei Nichteinbeziehung des Vorwissens als Kovariate.

Weiterhin findet sich ein hochsignifikanter Unterschied zwischen den drei Untersuchungen – der Lernerfolg der Probanden der dritten Studie ist mit einem Mittelwert von 9,45 (SD=3,20) deutlich höher als der der Probanden der ersten (M=7,62, SD=3,10) oder zweiten (M=7,57, SD=3,52) Studie, trotz keiner Unterschiede im Vorwissen. Dies wird auch durch Einzelvergleiche weiter belegt, die beide signifikant werden (mit  $p=0,014$  bzw.  $p=0,002$ ). Abbildung 7-1 illustriert besonders deutlich die Unterschiede des Lernerfolgs der Probanden zwischen den Untersuchungen wie auch zwischen den drei Visualisierungsarten.



**Abbildung 7-1:** Profildiagramm des Lernerfolgs der Probanden in den drei Untersuchungen in Abhängigkeit von der bearbeiteten Visualisierungsart. Das Vorwissen diente als Kovariate.

Natürlich ist bei diesem Berechnungsdesign problematisch, wie auch Abbildung 7-1 sehr deutlich macht, dass eine Zelle leer bleibt, da in der ersten Untersuchung ja keine 11-Standbilderversion existierte. Dies könnte eventuell zu größeren Verzerrungen der Ergebnisse führen. Darum werden dieselben Berechnungen ebenfalls unter Nichtberücksichtigung der ersten Untersuchung durchgeführt. Wie aus Tabelle 7-4 hervorgeht, ändern sich in diesem Fall die Ergebnisse aber nicht entscheidend: Der Haupteffekt für die Visualisierungsart verfehlt weiterhin das Signifikanzniveau ( $F(2,116)=2,10$ ,  $p=0,13$ ,  $MSE=10,04$ ,  $\eta^2=0,04$ ), in den Einzelvergleichen ist die Animation der 4-Standbilderversion nach wie vor einseitig statistisch signifikant überlegen ( $p<0,05$ ), der 11-Standbilderversion jedoch nicht ( $p=0,10$ ). Die deskriptiven Kennwerte sind in Tabelle 7-5 angegeben.

*Tabelle 7-4: Ergebnisse der zweifaktoriellen Varianzanalyse der Effekte von Visualisierungsart und Stichprobe unter zusätzlicher Einbeziehung des Vorwissens bei Nichtberücksichtigung der ersten Untersuchung.*

<b>Abhängige Variable</b>	<b>Varianzquelle</b>	<b>SS*</b>	<b>df</b>	<b>MS</b>	<b>F<sub>emp</sub></b>	<b>p</b>	<b>η<sup>2</sup></b>
<b>Lernzuwachs gesamt</b>	<i>Haupteffekt Vorwissen</i>	157,53	1	157,53	15,68	0,001	0,12
	<i>Haupteffekt Stichprobe</i>	87,10	1	87,10	8,67	0,004	0,07
	<i>Haupteffekt Visualisierungsart</i>	42,17	2	21,09	2,10	0,127	0,04
	<i>Interaktionseffekt Visualisierungsart- Stichprobe</i>	30,15	2	15,07	1,50	0,227	0,03
	<i>Fehler</i>	1165,10	116	10,04	---	---	---

\* sequentielle Varianzzerlegung

*Tabelle 7-5: Deskriptive Kennwerte des Gesamtlernerfolgs, aufgeteilt nach Visualisierungsart und Untersuchung unter Nichtberücksichtigung der ersten Studie.*

<b>Visualisierungsart</b>	<b>Untersuchung</b>	<b>N</b>	<b>M</b>	<b>SD</b>
4 Standbilder	2	21	7,41	4,01
	3	22	9,14	3,31
	<i>Gesamt</i>	43	8,30	3,73
11 Standbilder	2	19	6,24	3,30
	3	21	9,26	3,74
	<i>Gesamt</i>	40	7,83	3,81
Animation	2	22	8,86	2,83
	3	18	10,06	2,39
	<i>Gesamt</i>	40	9,40	2,68
<i>Gesamt</i>	2	62	7,57	3,52
	3	61	9,45	3,20
	<i>Gesamt</i>	123	8,50	3,49

Hinsichtlich des *Cognitive Load* zeigen sich keine signifikanten Unterschiede zwischen den Untersuchungen. Lediglich deskriptiv ist die kognitive Belastung in der dritten Studie niedriger (vgl. Tabelle 7-6), nicht jedoch im statistisch bedeutsamen Bereich ( $F(2,145)=1,32, p<.27, MSE=2,85, \eta^2=0,02$ ).

*Tabelle 7-6: Deskriptive Kennwerte des Cognitive Load in den drei Untersuchungen.*

<b>Stichprobe</b>	<b>N</b>	<b>M</b>	<b>SD</b>
1. Studie	25	6,40	1,58
2. Studie	62	6,52	1,65
3. Studie	61	6,03	1,77
<i>Gesamt</i>	25	6,30	1,69

Wie zu erwarten, erbringt die Zusammenfassung der Stichproben keine neuen Erkenntnisse hinsichtlich der *Spatial Ability* bzw. *Spatial Visualization*. Wie Tabelle 7-7 veranschaulicht, bleibt der Trend zwar gleich, auch der Haupteffekt für *Spatial Visualization* bleibt bestehen, nicht jedoch der Interaktionseffekt. Eine Nichtberücksichtigung der ersten experimentellen Studie wegen des Problems der leeren Zelle führt zu keinen qualitativ anderen Ergebnissen.

*Tabelle 7-7: Ergebnisse der Kovarianzanalyse zum Effekt der Spatial Ability (als Kovariate) mit dem Paper-Folding-Test bei der Zusammenfassung aller drei Studien.*

<b>Abhängige Variable</b>	<b>Varianzquelle</b>	<b>SS*</b>	<b>df</b>	<b>MS</b>	<b>F<sub>emp</sub></b>	<b>p</b>	<b>η<sup>2</sup></b>
<b>Lernzuwachs gesamt</b>	<i>Haupteffekt Spatial Visualization</i>	124,57	1	124,57	11,59	0,001	0,08
	<i>Haupteffekt Untersuchung</i>	140,78	2	70,39	7,20	0,001	0,09
	<i>Haupteffekt Visualisierungsart</i>	81,34	2	40,67	4,16	0,018	0,06
	<i>Interaktionseffekt</i>	12,94	2	6,47	0,66	0,518	0,01
	<i>Fehler</i>	1368,70	140	9,78	---	---	---

\* sequentielle Varianzzerlegung

### 7.3 Diskussion

Als auffälligstes Ergebnis dieses Vergleiches aller durchgeführten empirischen Studien ist sicherlich zu nennen, dass der Lernerfolg über alle Versuchsbedingungen hinweg in der dritten Studie deutlich höher war als in den beiden vorangegangenen Untersuchungen. Es läge nahe, dieses Ergebnis mit der Spezifität der Stichprobe zu begründen, dies erscheint jedoch unwahrscheinlich: Zum einen handelte es sich bei den Probanden nicht um eine besonders außergewöhnliche Gruppe im Gegensatz zu den anderen beiden Stichproben – genau wie die Versuchspersonen der Pilotstudie waren es Studierende der ersten Semester eines naturwissenschaftlichen Faches, zum Großteil mit der Zielrichtung Lehramt. Und dementsprechend zeigte sich auch kein Unterschied in der Vorwissensleistung zwischen den drei Gruppen.

Deshalb erscheint es plausibel anzunehmen, dass die Einführung von Möglichkeiten des *self-pacing*, der Selbstregulation der Lerngeschwindigkeit sowie der selbstgesteuerten Wiederholung, der entscheidende Faktor hinsichtlich der Erhöhung des Lernerfolgs war. Somit stellt dieses Ergebnis eine weitere Erhärtung der Erkenntnisse von z.B. Schwan & Riempp (2004) und Mayer & Chandler (2001) dar, die bereits auf die Vorteile minimaler Interaktivität für das Lernen mit dynamischen Visualisierungen hinwiesen. Neu ist, dass sich dieser Vorteil offenbar auch auf Serien statischer Bilder erstreckt. Beide Standbildversionen erzielten einen höheren Lernerfolg als in der zweiten Studie, die 11-Standbildversion konnte sich sogar überproportional steigern, wie Abbildung 7-1 veranschaulicht. Dies ist natürlich

besonders darauf zurückzuführen, dass diese Version offenbar für das Lernen ohne zusätzliche interaktive Möglichkeiten äußerst ungeeignet war. Wie bereits für andere Bereiche, etwa für den Vergleich von geschriebenem versus gesprochenem Text, scheint auch in dieser Untersuchung zu gelten: „In self-paced scenarios, apparently all learners gain benefit from minimal control options“ (Schmidt-Weigand, 2006, S. 101).

Im Gegenzug führten die zusätzlichen Interaktionsmöglichkeiten des Lerners offensichtlich nicht zu einer Erhöhung des *Cognitive Load*, wie die Ergebnisse zeigen. Im Gegensatz zu komplexen Simulationen, die in der Regel eine Vielzahl an interaktiven Optionen bieten (de Jong & van Joolingen, 1998) und denen eine hohe kognitive Belastung zugeschrieben wird (z.B. Bodemer & Plötzner, 2004; Kramer, 2005), scheint die in den hier vorliegenden Untersuchungen ermöglichte Form der Interaktivität sich nicht belastend auf das Arbeitsgedächtnis auszuwirken, zumal solcherlei Kontrolloptionen weitgehend selbsterklärend sein dürften. Er erscheint eher angebracht zu spekulieren, dass das Arbeitsgedächtnis entlastet und so das Lernen erleichtert wird.



## 8 Zusammenfassende Diskussion

Hauptziel dieser Dissertation war es, die Lerneffekte von dynamischen und nichtdynamischen Visualisierungen zu untersuchen und moderierende Variablen zu identifizieren. Hierzu wurde eine Metaanalyse berechnet, die zunächst den aktuellen Forschungsstand zu dieser Frage zusammenfassen und hypothesengenerierend mögliche Moderatoren aufzeigen sollte. Auf der Basis der Ergebnisse sowie darüber hinausgehender Überlegungen folgte sodann eine Serie von drei experimentellen Untersuchungen, deren Fokus auf der Hypothese lag, eine Animation im naturwissenschaftlichen Setting durch Standbilder ersetzen zu können. Darüber hinaus stand besonders der mutmaßlich moderierende Faktor des räumlichen Vorstellungsvermögens im Vordergrund des Interesses, zusätzlich wurden die Rollen des themenspezifischen Vorwissens sowie der kognitiven Belastung untersucht.

### 8.1 Zusammenfassung der Ergebnisse

Alle vier Studien zusammenfassend, die Metaanalyse und die drei eigenen experimentellen Untersuchungen, ist zunächst die durchgehend auffindbare Überlegenheit dynamischer gegenüber nichtdynamischen Visualisierungen auffällig. Entgegen den Thesen Tverskys et al. (2002) war sowohl in der publizierten Literatur eine deutliche Überlegenheit von Animationen zu finden, wie die Metaanalyse belegte, als auch in den eigenen experimentellen Studien, bei denen zumindest die 4-Standbilderversion praktisch durchgehend der Animation unterlegen war. Darüber hinaus konnten Hinweise dafür gefunden werden, dass eine erhöhte Anzahl von Standbildern diesen Vorteil teilweise ausgleichen kann. Gleiches gilt für erklärenden Text in den Standbildern, was die Metaanalyse klar belegte und in den eigenen Studien von vornherein eingesetzt wurde. Spezielle Hinweisreize hingegen scheinen Standbilder wenig substantiell zu verbessern.

Weiterhin ließen sich einige Richtlinien formulieren, wann der Einsatz von Animationen sinnvoll ist – und wann die aufwändige Entwicklung gespart werden kann: So förderte die Metaanalyse zutage, dass Animationen nicht für rein dekorative, sondern nur für so genannte repräsentationale Zwecke eingesetzt werden sollten, d.h. wenn die Animation tatsächlich den zu lernenden Gegenstand darstellt. Ferner scheint die Animation entgegen vorheriger Forschung (Mayer &

Moreno, 2002) durchaus auch dem Lernen basalen Faktenwissens zu dienen; allerdings sind die Ergebnisse der vier in dieser Dissertation dargestellten Untersuchungen in diesem Punkt uneinheitlich.

Daneben ist als ein wichtiger Punkt die moderierende Rolle des räumlichen Vorstellungsvermögens, genauer der Unterfunktion *Spatial Visualization* zu nennen, die mehrfach gezeigt werden konnte: Lernende mit hohem räumlichen Vorstellungsvermögen können Nachteile von Standbildern kompensieren oder umgekehrt kann eine Animation mangelnde Fähigkeit zur *Spatial Visualization* kompensieren.

Schließlich konnte ein interessanter, ursprünglich sogar explizit ausgeschlossener Punkt näher beleuchtet werden: die Rolle minimaler Interaktivität beim Lernen mit Animationen und Standbildern. Lediglich als Hilfskonstruktion in die dritte Studie eingeführt, um unerwartete Verzerrungen bei der 11-Standbilderversion auszugleichen, führte die erhöhte Lernerkontrolle über den Visualisierungsvorgang zu deutlichen Lernsteigerungen im Vergleich zu den anderen Studien – und das gleichermaßen für alle Versuchsbedingungen. Gleichwohl kann nicht ausgeschlossen werden, dass es sich hierbei lediglich um Stichprobenunterschiede handelt.

## 8.2 Theoretische Implikationen

Die Ergebnisse der in dieser Dissertation beschriebenen vier Studien haben vor dem Hintergrund des gegenwärtigen Forschungsstandes zum Lernen mit Animationen und Standbildern einige Implikationen theoretischer Natur. Zunächst muss der zurzeit vorherrschenden Meinung (Mayer et al., 2005; Tversky et al., 2002) widersprochen werden, nichtinteraktive Animationen seien in der Regel vergleichbaren Serien von Standbildern unterlegen. Nicht nur zeigte sich in den eigenen experimentellen Studien jeweils ein Vorteil der Animation gegenüber der Version mit wenigen Schlüsselbildern, sondern besonders legte die Metaanalyse offen, dass auch der überwiegende Teil der publizierten Literatur dem entgegensteht. Selbstverständlich gibt es auch Studien, die Vorteile von Standbildern fanden, doch das rechtfertigt sicher nicht eine konsequent ablehnende Haltung gegenüber dem Einsatz von Animationen (nachdem zunächst ein extrem hoher Enthusiasmus hinsichtlich der erwarteten Möglichkeiten von Animationen vorherrschte; z.B.

Hannafin & Rieber, 1989; Lepper & Malone, 1987; Proffitt & Kaiser, 1986). Wichtig ist in diesem Zusammenhang zum Beispiel, dass Animationen nicht ohne vorherige Reflexion über die Ziele der Visualisierung eingesetzt werden – aus rein dekorativen Gründen, die etwa der nicht leicht belegbaren Motivationssteigerung dienen sollen, sollten Animationen nicht verwendet werden; zumindest ist dann kein Lernvorteil gegenüber Standbildern nachweisbar.

Gerade auch bezüglich der Rolle der *Spatial Ability* beim Lernen mit Visualisierungen konnte in den hier vorgestellten experimentellen Studien gezeigt werden, dass Animationen Hilfestellungen leisten können: Lernende, die niedrige Fähigkeiten auf der *Spatial-Visualization*-Skala haben, hatten deutlich niedrigere Lernerfolge, wenn sie mit Standbildern anstelle von Animationen lernten. Lernende mit hohem räumlichen Vorstellungsvermögen hingegen lernten mit beiden Versionen gleich gut. Anstelle einer *ability-as-enhancer*-Funktion (Mayer, 2001; Mayer & Sims, 1994) scheint also in diesen Fällen eher eine kompensierende Funktion vorzuherrschen. Zumindest für bestimmte Lernergruppen, so die Folgerung aus diesen Ergebnissen, können Animationen sinnvoll sein und deren höheren Entwicklungsaufwand rechtfertigen. Der gefundene Unterschied zu den Ergebnissen von Mayer & Sims (1994), die *Spatial Ability* eher eine verstärkende statt einer kompensierenden Funktion zuschreiben, erklärt sich durch grundlegende Unterschiede im Design: Während in der hier vorliegenden Arbeit Animationen und Standbilder miteinander verglichen wurden, konzentrierten sich Mayer & Sims (1994) auf den bereits oben beschriebenen *Contiguity-Effect*. Zu einer Animation wurde eine verbale Erläuterung entweder gleichzeitig oder davor bzw. danach geliefert. Hätte räumliches Vorstellungsvermögen auch in diesem Fall einen kompensierenden Effekt, so hätten Lernende mit hoher *Spatial Ability* unter jeder Versuchsbedingung gleich viel lernen müssen. Es zeigte sich jedoch, dass Lernende mit den sequentiellen Versuchsbedingungen grundsätzlich Probleme hatten, unabhängig von ihrer räumlichen Vorstellungskraft. Erst bei günstigerer gleichzeitiger Darbietung von Animation und Erläuterung zeigten sich Unterschiede, indem Lernende mit hoher *Spatial Ability* mehr lernten als solche mit niedriger *Spatial Ability* – eben ein *ability-as-enhancer*-Effekt. Neben einem zu dieser Studie anzumerkenden methodischen Einwand, dass nämlich in zwei von zwei Experimenten *keine* statistisch signifikante Interaktion gefunden wurde (und in einem der beiden Experimente auch kein Haupteffekt für *Spatial Ability*), kann konstatiert werden, dass der Fokus dieser

Studie offensichtlich ein anderer ist als der der hier vorliegenden Arbeit und damit die jeweiligen Ergebnisse kaum im Widerspruch zueinander stehen. Bei Mayer & Sims (1994), so könnte argumentiert werden, war die Kontrollbedingung anscheinend deutlich zu schwer, um überhaupt akzeptable Lernergebnisse erzielen zu können (wie das zeitliche *Split-Attention-* bzw. *Contiguity-*Prinzip ja auch besagen). Mögliche kompensierende Effekte der *Spatial Ability* können unter diesen Bedingungen womöglich kaum zum Tragen kommen, weil eher Gedächtnisfertigkeiten von Nöten sind, die nicht mit der *Spatial Ability* identisch sind: Die verbale Erläuterung muss in der *phonological loop* (Baddeley, 1986; 1999) gehalten werden, bis die Animation erfolgt. In den hier vorliegenden experimentellen Untersuchungen hingegen wurden zwei verschiedene piktoriale Bedingungen miteinander verglichen – und mangelnde Eigenschaften der einen Version (die Nichtdynamik der Standbilder) konnten durch einen höheren Wert auf der *Spatial-Visualization*-Skala ausgeglichen werden.

Allerdings war ein Ziel dieser Dissertation auch der Versuch, den Nachweis zu erbringen, dass Animationen prinzipiell eben *doch* durch Standbilder ersetzt werden können, erhöht man bloß deren Anzahl auf ein ausreichendes Maß. Es muss allerdings festgestellt werden, dass diesbezüglich nur eher schwache Indizien gefunden werden konnten, kein klarer Beleg. Eine solche Tendenz war zwar deskriptiv erkennbar, eine statistische Signifikanz konnte jedoch nicht nachgewiesen werden. Somit hat die Animation weiterhin ihre Daseinsberechtigung im Kanon der Möglichkeiten zur Visualisierung von Sachverhalten.

Die Möglichkeiten, die eine umfassende interaktive Gestaltung von Animationen bietet (inklusive aller Gefahren, die damit durch eine kaum zu umgehende Erhöhung der kognitiven Belastung einhergehen), wurden in dieser Arbeit gezielt ausgeklammert. Jedoch zeigten sich selbst nach Einführung der Option minimaler Interaktivität (die im Übrigen auch in der Definition der Animation in der Metaanalyse eingeschlossen war) bereits deutliche zusätzliche Lerneffekte. Interessanterweise waren diese unter allen Versuchsbedingungen zu finden, auch Serien von Standbildern profitierten also von einem *self-paced*-Ansatz. Dieser wiederum überdeckte den oben beschriebenen Moderatoreffekt des räumlichen Vorstellungsvermögens; ein ausreichendes Maß an räumlichem Vorstellungsvermögen beim Arbeiten mit computerbasierten Visualisierungen scheint also möglicherweise nur dann nötig zu sein, wenn deren Design *Extraneous Cognitive Load* erzeugt. Entweder durch Animationen oder durch *self-pacing*

scheinen auch Lernende mit niedriger Fähigkeit zur *Spatial Visualization* erhöhte Lernerfolge erzielen zu können.

Schließlich konnte die Metaanalyse eine Erweiterung des gut belegten Multimediaeffekts (Mayer, 2001) aufzeigen: Während das Multimediaprinzip besagt, dass Texte in Verbindung mit Bildern effektiveres Lernen bewirken als Texte allein, deutete die Metaanalyse darauf hin, dass umgekehrt genauso gilt, dass Texte in Verbindung mit Bildern wirksamer sind als Bilder allein – zumindest jeweils in Relation zu Animationen mit oder ohne erläuternden Texten. Zwar ist dieses Ergebnis eher wenig überraschend, trotzdem aber eine aus theoretischer Sicht wichtige Erkenntnis.

### 8.3 Praktische Implikationen

Besonders für die Designer von virtuellen Lernumgebungen ergeben sich aus dieser Arbeit auch Implikationen eher praktischer Art. Zum einen ist in diesem Zusammenhang auf den bereits im letzten Abschnitt genannten Befund einzugehen, dass die Entwicklung einer Animation nicht den Aufwand wert ist, wenn sie lediglich zu rein dekorativen Zwecken eingesetzt werden soll. Gerade unter professionellen Entwicklern von Lernumgebungen scheint die Skepsis gegenüber dynamischen Lernumgebungen weit geringer ausgeprägt zu sein (d.h. praktisch nicht existent) als unter Instruktionspsychologen. Das belegt z.B. die große Menge an von Schulbuchverlagen veröffentlichter Lernsoftware auf dem Markt, die praktisch ausschließlich zumindest teilweise dynamisch gestaltet ist. Trotzdem sollte Entwicklern von Lernsoftware geraten werden, nicht reflexartig zu einer Animation zu greifen, wenn dies lediglich zu Ablenkung und erhöhter *Cognitive Load* führt. Nur wenn der zu lernende Sachverhalt auch tatsächlich dynamischer Natur ist, macht eine Animation Sinn. Ob es obendrein der Fall sein könnte, dass unter bestimmten Bedingungen eine Serie von Standbildern, wenn diese nur aus ausreichend vielen Bildern besteht, eine Animation vollständig ersetzen kann, konnte in vorliegender Arbeit allerdings nicht empirisch belegt werden.

Hinsichtlich eines ähnlichen Streits zwischen Instruktionspsychologen und Designern, der Frage der Realistik der Animation, konnten leider keine neuen Erkenntnisse gewonnen werden. Die Metaanalyse fand keine Vorteile für entweder eher schematische (wie sie Psychologen aufgrund geringerer *seductive details*

bevorzugen; vgl. Tversky et al., 2002) oder hochrealistische Animationen (wie sie Designer von Lernumgebungen und auch häufig Vertreter fachdidaktischer Forschung wegen höherer Wirklichkeitsnähe und der möglichen Verhinderung des Bildens von Fehlvorstellungen bevorzugen; vgl. Gonzalez, 1996). Zwar deutete die Metaanalyse zunächst auf einen Vorteil hochrealistischer gegenüber eher schematischen Animationen hin, die detailliertere Analyse ließ diesen scheinbaren Vorteil jedoch wieder zum Verschwinden bringen.

Der Befund, dass Lernende mit unterschiedlichem Grad räumlichen Vorstellungsvermögens unterschiedlich stark von Animationen oder Standbildern profitieren, könnte eventuell den Anstoß dazu geben, sich verstärkt mit der Entwicklung adaptiver Lernumgebungen (Leutner, 1992; 1993; Schwonke, Hauser, Nückles & Renkl, 2006; Triantafillou, Pomportsis, Demetriadis & Georgiadou, 2004) zu beschäftigen, die neben dem Vorwissen auch die *Spatial Ability* in die Variation des Angebots für den Lerner mit einbeziehen. Alternativ wäre es denkbar, dem Lernenden von vornherein Wahlmöglichkeiten zu bieten, ob er das Lernen mit dynamischen oder nichtdynamischen Visualisierungen bevorzugt, wenngleich bereits Hinweise dafür existieren (Leutner & Plass, 1998; Mayer & Massa, 2003), dass die Präferenz von Probanden nicht ihrer tatsächlichen Kompetenz entsprechen muss.

In diesem Zusammenhang spielt natürlich auch die Interaktivität eine Rolle, die zumindest in einer minimalen Ausprägung, wie sie in der dritten experimentellen Studie eingesetzt wurde, keinerlei negative Auswirkungen gehabt zu haben scheint, sondern im Gegenteil den Lernerfolg verbessert hat. Ein Einsatz von Interaktivität in dieser Form ist also eindeutig zu befürworten; sobald der Lernende jedoch zu viele Optionen hat, darauf deuten verschiedene Untersuchungen hin (z.B. de Jong & van Joolingen, 1998; Stark, Graf, Renkl, Gruber & Mandl, 1995; Höffler, 2003; Nerdel, 2003), kann es leicht zu negativen Folgen kommen, vermutlich ausgelöst durch eine kognitive Überlastung.

## 8.4 Fachdidaktische Implikationen

Was fachdidaktische Implikationen vorliegender Arbeit betrifft, so sollte zunächst die Frage des in die Metaanalyse zwar einbezogenen, jedoch nicht weiter vertieften Inhaltsbereichs von Animationen bzw. Standbildern angesprochen werden. Es wurde versucht, mittels einer Einteilung nach Fächern wie Chemie, Physik, Biologie etc. die

in die Metaanalyse einbezogenen Primärstudien zu kodieren und eventuelle Vorteile der einen oder anderen Visualisierungsart in bestimmten Bereichen zu identifizieren. Wie zu befürchten war, erwies sich diese Einteilung allerdings als deutlich zu grob und erbrachte keine weiteren statistisch verwertbaren Erkenntnisse. Eine detailliertere Analyse scheiterte jedoch an der häufig viel zu ungenauen Dokumentation der Lernumgebungen in den Primärstudien. Somit können aus dieser Arbeit leider keine Erkenntnisse abgeleitet werden, ob in bestimmten Domänen eher Animationen eingesetzt werden sollten als in anderen – sinnvoller wäre ohnehin eine feinkörnigere Einteilung, in der die spezifischen Charakteristika bestimmter Teilbereiche berücksichtigt werden.

Nichtsdestotrotz können auch für die Fachdidaktik einige Implikationen abgeleitet werden – dazu gehört zum Beispiel einmal mehr der dringende Rat, Animationen im Unterricht nicht blindlings um ihrer selbst und ihrer dekorativen Funktion willen einzusetzen, sondern nur gezielt zur Unterstützung der Verdeutlichung dynamischer Prozesse. Im Umkehrschluss soll hier auch nochmals darauf verwiesen werden, dass Schüler mit niedrigem räumlichen Vorstellungsvermögen deutliche Probleme haben, aus Standbildern eine adäquate mentale Animation zu bilden und den Sachverhalt zu begreifen (zumindest wenn sie die Präsentationszeit der Bilder nicht selbst steuern können). Da der klassische Unterricht vorrangig mit Texten und Bildern in Lehrbüchern bzw. Standbildern an der Tafel operiert, stellt dies möglicherweise einen Nachteil für diese Schüler dar. Diesem Problem könnte mit einem gezielten Einsatz von Animationen abgeholfen werden.

## 8.5 Abschließende Bemerkungen und Ausblick

Resümierend darf festgestellt werden, dass ein Gutteil der zu Beginn dieser Arbeit formulierten Ziele erreicht wurde – die Metaanalyse fasste erstmals zumindest die Mehrzahl der bisher zum Vergleich „Animation versus Standbilder“ publizierten Untersuchungen zusammen und fand eine überraschende Überlegenheit der Animationen, die dem zur Zeit gültigen Konsens unter Instruktionspsychologen (Tversky et al., 2002) widerspricht. Es konnten eine Reihe moderierender Variablen identifiziert werden, auch wenn es wünschenswert gewesen wäre, diese Liste noch deutlich zu verlängern. Dies war allerdings aufgrund häufig mangelnder Dokumentation der Primärstudien nicht möglich.

Als Erklärung für die Heterogenität früherer Forschungsergebnisse zum Unterschied von dynamischen und nichtdynamischen Visualisierungen könnten neben den durch die Metaanalyse identifizierten Variablen auch diejenigen dienen, die im Fokus der drei selbst durchgeführten experimentellen Untersuchungen standen. Dabei darf der schwache, aber mehrfach aufgefundene kompensierende Effekt des räumlichen Vorstellungsvermögens zumindest für das eingesetzte Material als relativ gesichert gelten – weitere Untersuchungen zumal mit anderen Inhaltsbereichen der Lernumgebung wären zur Absicherung des Effekts aber erforderlich. Fehlende Kontrolle der *Spatial Ability* in anderen Studien und so dort eventuell aufgetretene, im Hinblick auf *Spatial Ability* verzerrte Stichproben erscheinen aber als verzerrendes Moment der Ergebnisse durchaus plausibel.

Deutlich schwächer erscheint der Befund zur Variation der Anzahl an Standbildern. Zwar verweist der Trend auf das Potential der Angleichung der Wirksamkeit von Standbildern und Animationen durch Erhöhung der dargestellten Zwischenschritte, jedoch konnte dies nicht statistisch gesichert belegt werden. Eine erneute Überprüfung dieser Hypothese in weiteren Studien erscheint vielversprechend und geboten, konnte aber im Rahmen dieser Arbeit leider nicht mehr geleistet werden.

Schließlich erscheint die Rolle der minimalen Interaktivität für weitere Forschung interessant, steigerte diese doch den Lernerfolg insgesamt und machte erst den (ansatzweisen) Erfolg der Version mit zusätzlichen Standbildern sichtbar.



## 9 Literatur

Mit einem Sternchen (\*) markierte Literaturangaben sind Teil der Metaanalyse.

- Ainsworth, S. & Van Labeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction, 14*, 241-255.
- Alesandrini, K. L. (1982). Imagery-eliciting strategies and meaningful learning. *Journal of mental imagery, 6*, 125-140.
- Ayres, P. & Sweller, J. (2005). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 135-146). Cambridge: Cambridge University Press.
- Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. (1992). Working memory. *Science, 255*, 556-559.
- Baddeley, A. (1999). *Essentials of human memory*. Hove: Psychology Press.
- \*Baek, Y. K. & Layne, B. H. (1988). Color, graphics, and animation in a computer assisted learning tutorial lesson. *Journal of Computer Based Instruction, 15*, 131-135.
- Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research, 74*, 29-58.
- Bangert-Drowns, R. L. (1986). Review of developments in meta-analytic methods. *Psychological Bulletin, 99*, 388-399.
- Barnea, N. & Dori, Y. J. (1999). High-school chemistry students' performance and gender differences in a computerized molecular modeling learning environment. *Journal of Science Education and Technology, 8*, 257-271.
- Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.
- Bétrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 287-296). Cambridge: Cambridge University Press.
- Bétrancourt, M. & Tversky, B. (2000). Effect of computer animation on users' performance: A review. *Travail-Humain, 63*, 311-329.
- \*Blake, T. (1977). Motion in instructional media: some subject-display mode

- interactions. *Perceptual and Motor Skills*, 44, 975-985.
- Bodemer, D. & Plötzner, R. (2004). Encouraging the active processing of information during learning with multiple and interactive representations. In H. M. Niegemann, R. Brünken & D. Leutner (Eds.), *Instructional design for multimedia learning* (pp. 127-138). Münster: Waxmann Verlag.
- Brünken, R., Gall, T. & Leutner, D. (2001). Individuelle Unterschiede beim Lernen mit Neuen Medien: Zur Rolle der visuellen Aufmerksamkeit. Vortrag auf der 60. Tagung der Arbeitsgruppe Empirische Pädagogische Forschung. Bamberg: Universität Bamberg.
- Brünken, R., Plass, J.L., & Leutner, D. (2004) Assessment of cognitive load in multimedia learning with dual-task methodology: Auditory load and modality effects. *Instructional Science*, 32, 115-132.
- Brünken, R., Steinbacher, S. & Leutner, D. (2000). Räumliches Vorstellungsvermögen und Lernen mit Multimedia. In D. Leutner & R. Brünken (Hrsg.), *Neue Medien in Unterricht, Aus- und Weiterbildung: aktuelle Ergebnisse empirischer pädagogischer Forschung* (S. 37-46). Münster: Waxmann.
- Brünken, R., Steinbacher, S., Plass, J. L. & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology*, 49, 109-119.
- Carney, R. N. & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14, 5-26.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- \*Catrambone, R. & Seay, A. F. (2002). Using animation to help students learn computer algorithms. *Human Factors*, 44, 495-511.
- Chandler, P. (2004). The crucial role of cognitive processes in the design of dynamic visualizations. *Learning and Instruction*, 14, 353-357.
- Chandler, P. & Sweller, J. (1991). Cognitive Load Theory and the format of instruction. *Cognition and Instruction*, 8, 293-332.
- \*ChanLin, L.-J. (1998). Animation to teach students of different knowledge levels. *Journal of Instructional Psychology*, 25, 166-175.
- ChanLin, L.-J. (2000). Attributes of Animation for Learning Scientific Knowledge. *Journal of Instructional Psychology*, 27, 228-238.
- \*ChanLin, L.-J. (2001). Formats and prior knowledge on learning in a computer-

- based lesson. *Journal of Computer Assisted Learning*, 17, 409-419.
- Clark, J. M. & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3, 149-210.
- Cohen, J. (1966). Some statistical issues in psychological research. In B. B. Wolman (Ed.), *Handbook of clinical psychology* (pp. 95-121). New York: McGraw Hill.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences, 2nd edition*. Hillsdale: Lawrence Erlbaum Associates.
- Comenius, J. A. (1658). *Orbis sensualium pictus*. Nürnberg: Comenius.
- Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, 90, 1073-1091.
- \*Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, 94, 428-434.
- Dean, R. S. & Enemoh, P. A. C. (1983). Pictorial organization in prose learning. *Contemporary Educational Psychologist*, 8, 20-27.
- de Jong, T. (2005). The guided discovery principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 215-228). Cambridge: Cambridge University Press.
- de Jong, T. & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68, 179-201.
- Eliot, J. & Smith, I. M. (1983). *An international directory of spatial tests*. Windsor: NFER/Nelson.
- Ekstrom, R. B., French, J. W. & Harman, H. H. & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton: Educational Testing Service.
- Faul, F. & Erdfelder, E. (1992). *GPOWER: A priori, post-hoc, and compromise power analyses for MS-DOS* [Computerprogramm]. Bonn: Bonn University, Dept. of Psychology.
- Fletcher, J. D. & Tobias, S. (2005). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 117-133). Cambridge: Cambridge University Press.
- Fricke, R. & Treinies, G. (1985). *Einführung in die Metaanalyse*. Göttingen: Verlag Hans Huber.

- Gage, N. L. & Berliner, D. C. (1977). *Pädagogische Psychologie*. München: Urban & Schwarzenberg.
- Gerjets, P., Scheiter, K., & Catrambone, R. (2004). Designing instructional examples to reduce intrinsic cognitive Load: Molar versus modular presentation of solution procedures. *Instructional Science*, 32, 33-58.
- Ginns, P. (2005). Meta-analysis of the modality effect, *Learning and Instruction*, 15, 313-331.
- Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills: Sage.
- Gonzalez, C. (1996). Does animation in user interfaces improve decision making? In M. J. Tauber, V. Bellotti, R. Jeffries, J. D. Mackinlay & J. Nielsen (Eds.), *Proceedings of the ACM CHI 96 Human Factors in Computing Systems Conference* (pp. 27-34). Vancouver: ACM.
- Hannafin, M. & Rieber, L. (1989). Psychological foundations of instructional design for emerging computer-based instructional technologies. *Educational Technology Research & Development*, 37, 91-101.
- Harp, S. F. & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: on the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology*, 89, 92-102.
- Harp, S. F. & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, 90, 414-434.
- \*Hays, T. A. (1996). Spatial abilities and the effects of computer animation on short-term and long-term comprehension. *Journal of Educational Computing Research*, 14, 139-155.
- Hedges, L. V. & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.
- Hedges, L. V. (1994). Fixed effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 285-299). New York: Russell Sage Foundation.
- Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 1084-1102.
- Hegarty, M. (2004). Dynamic visualizations and learning: Getting to the difficult

- questions. *Learning and Instruction*, 14, 343-351.
- Hegarty, M. (2005). Multimedia learning about physical systems. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 447-465). Cambridge: Cambridge University Press.
- Hegarty, M. & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32, 175-191.
- 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.
- \*Höffler, T. N. (2003). *Animation, Simulation oder Standbilder? Zur Wirkung interaktiver Animationen auf das Verständnis komplexer biologischer Prozesse*. Unveröffentlichte Diplomarbeit, Kiel: Christian-Albrechts-Universität.
- Hunter, J. E. & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis methods: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, 8, 275-292.
- Isaak, M. I. & Just, M. A. (1995). Constraints on the processing of rolling motion: The curtate cycloid illusion. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1391-1408.
- \*Kaiser, M. K., Proffitt, D. R., & Anderson, K. (1985). Judgments of natural and abnormal trajectories in the presence and absence of motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 795-803.
- Kalyuga, S., Ayres, P., Chandler, P. & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38, 23-31.
- Kalyuga, S., Chandler, P. & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors*, 40, 1-17.
- Kalyuga, S., Chandler, P. & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, 92, 126-136.
- Kester, L., Paas, F. & van Merriënboer, J. J. G. (2006). Instructional control of cognitive load in the design of complex learning environments. Vortrag im Rahmen des Technology & Learning Symposiums, New York: New York University.
- Kramer, B. (2005). *Mentale Integration von Text und Bild beim Lernen mit Multimedia am Beispiel der olfaktorischen Signaltransduktion*. Elektronische Dissertation, Kiel: Universitätsbibliothek.

- Lafrance, M., Hecht, M. A., & Paluck, E. L. (2003). The contingent smile: a meta-analysis of sex differences in smiling. *Psychological Bulletin*, *129*, 305-334.
- \*Lai, S.-L. (2000). Increasing associative learning of abstract concepts through audiovisual redundancy. *Journal of Educational Computing Research*, *23*, 275-289.
- Large, A. (1996). Computer animation in an instructional environment. *Library & Information Science Research*, *18*, 3-23.
- Large, A., Beheshti, J., Breuleux, A., & Renaud, A. (1996). 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.
- Lawless, K. A. & Brown, S. W. (1997). Multimedia learning environments: Issues of learner control and navigation. *Instructional Science*, *25*, 117-131.
- Lepper, M. R. & Malone, T. W. (1987). Intrinsic motivation and instructional effectiveness in computer-based education. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, Learning, and Instruction, Volume 3: Conative and Affective Process Analyses* (pp. 255-286). Hillsdale: Lawrence Erlbaum Associates.
- Leutner, D. (1992). *Adaptive Lehrsysteme – Instruktionspsychologische Grundlagen und experimentelle Analysen*. Weinheim: Psychologie Verlags Union.
- Leutner, D. (1993). Guided discovery learning with computer-based simulation games: Effects of adaptive and non-adaptive instructional support. *Learning and Instruction*, *3*, 113-132.
- Leutner, D., & Plass, J. (1998). Measuring learning styles with questionnaires versus direct observation of preferential choice behavior in authentic learning situations: The visualizer/verbalizer behavior observation scale (VV-BOS). *Computers in Human Behavior*, *14*, 543-557.
- Leutner, D. & Rammsayer, T. (1995). Complex trait-treatment interaction analysis: A powerful approach for analyzing individual differences in experimental designs. *Personality and Individual Differences*, *19*, 493-511.
- Lewalter, D. (1997). *Lernen mit Bildern und Animationen: Studie zum Einfluss von Lernermerkmalen auf die Effektivität von Illustrationen*. Münster: Waxmann.
- \*Lewalter, D. (2003). Cognitive strategies for learning from static and dynamic visuals. *Learning and Instruction*, *13*, 177-189.
- Low, R. & Sweller, J. (2005). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 147-158).

- Cambridge: Cambridge University Press.
- Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education, 14*, 225-244.
- Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction, 13*, 157-176.
- Macromedia. (2004). *Macromedia Dreamweaver MX* [Computerprogramm]. San Francisco: Macromedia, Inc.
- Macromedia. (2004). *Macromedia Flash MX Professional* [Computerprogramm]. San Francisco: Macromedia, Inc.
- Marcus, N., Wong, A., Ayres, P., Paas, F. & Sweller, J. (2007). Instructional animations that foster motor skills. Vortrag auf der UNSW Cognitive Load Theory Conference, Sydney: University of New South Wales.
- Marshalek, B., Lohman, D. F. & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence, 7*, 107-127.
- Mautone, P. D. & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93*, 377-389.
- Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*, 1-19.
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge: Cambridge University Press.
- Mayer, R. E. (2005a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge: Cambridge University Press.
- Mayer, R. E. (2005b). Principles for reducing extraneous processing in multimedia learning: coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 183-200). Cambridge: Cambridge University Press.
- Mayer, R. E. (2005c). *The Cambridge handbook of multimedia learning*. Cambridge: Cambridge University Press.
- Mayer, R. E. & Chandler, P. (2001). When learning is just a click away: Does simple interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*, 390-397.
- Mayer, R. E. & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*, 715-726.
- Mayer, R. E., Hegarty, M., Mayer, S. & Campbell, J. E. (2005). When static media

- promote active learning: annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied*, 11, 256-265.
- Mayer, R. E. & Massa, L. J. (2003). Three facets of visual and verbal learners: cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology*, 95, 833-846.
- Mayer, R. E. & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review*, 14, 87-99.
- Mayer, R. E. & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, 82, 715-726.
- Metfessel, N. S., Michael, W. B. & Kirsner, D. A. (1969). Instrumentation of Bloom's and Krathwohl's taxonomies for writing educational objectives. *Psychology in the Schools*, 6, 227-231.
- Moreno, R. & Mayer, R. E. (1999). Multimedia-supported metaphors for meaning making in mathematics. *Cognition and Instruction*, 17, 215-248.
- Moreno, R. & Mayer, R. E. (2001). A coherence effect in multimedia learning: the case for minimizing irrelevant sounds in the design of multimedia messages. *Journal of Educational Psychology*, 92, 117-125.
- \*McCloskey, M. & Kohl, D. (1983). Naive physics: the curvilinear impetus principle and its role in interactions with moving objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 146-156.
- Merrill, M. D. (1975). Learner control: Beyond aptitude-treatment interactions. *AV Communications Review*, 23, 217-226.
- \*Michas, I. C. & Berry, D. C. (2000). Learning a procedural task: Effectiveness of multimedia presentations. *Applied Cognitive Psychology*, 14, 555-575.
- Milheim, W. D. (1993). How to use animation in computer-assisted-learning. *British Journal of Educational Technology*, 24, 171-178.
- Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P. & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology*, 130, 621-640.
- Narayanan, N. H. & Hegarty, M. (2002). Multimedia design for communication of dynamic information. *International Journal of Human-Computer Studies*, 57, 279-315.



- \*Nerdel, C. (2003). *Die Wirkung von Animation und Simulation auf das Verständnis von stoffwechselphysiologischen Prozessen*. Elektronische Dissertation, Kiel: Universitätsbibliothek.
- \*Nicholls, C. & Merkel, S. (1996). The effect of computer animation on students' understanding of microbiology. *Journal of Research on Computing in Education*, 28, 359-371.
- OECD (2006). *Are students ready for a technology-rich world? OECD Briefing Notes für Deutschland*. <http://www.oecd.org/dataoecd/49/7/36001057.pdf>. Gefunden am 31.01.2007.
- Orwin, R. G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics*, 8, 157-159.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics - A cognitive-load approach. *Journal of Educational Psychology*, 84, 429-434.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.
- Park, O.-C. (1998). Visual displays and contextual presentations in computer-based instruction. *Educational Technology Research and Development*, 46, 37-50.
- Park, O.-C. & Hopkins, R. (1993). Instructional conditions for using dynamic visual-displays - a review. *Instructional Science*, 21, 427-449.
- Plötzner, R. & Lowe, R. (2004). Editorial: Dynamic visualization and learning. *Learning and Instruction*, 14, 235-240.
- Proffitt, D. R. & Kaiser, M. K. (1986). Software Techniques: The use of computer graphics animation in motion perception research. *Behavior Research Methods, Instruments, & Computers*, 18, 487-492.
- Renkl, A. (2005). The worked-out examples principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 229-245). Cambridge: Cambridge University Press.
- Renninger, K. A., Hidi, S. & Krapp, A. (1992). *The role of interest in learning development*. Hillsdale: Erlbaum.
- \*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). Using computer animated graphics with science instruction with

- children. *Journal of Educational Psychology*, 82, 135-140.
- \*Rieber, L. P. (1991). Animation, incidental learning, and continuing motivation. *Journal of Educational Psychology*, 83, 318-328.
- Rieber, L. P. (1994). *Computers, graphics, and learning*. Madison: Brown & Benchmark.
- \*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.
- \*Rigney, J. W. & Lutz, K. A. (1976). Effect of graphic analogies of concepts in chemistry on learning and attitude. *Journal of Educational Psychology*, 68, 305-311.
- Rosenthal, R. (1979). The "file-drawer problem" and tolerance for null results. *Psychological Bulletin*, 86, 638-641.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park: Sage.
- Rustenbach, S. J. (2003). *Metaanalyse: Eine anwendungsorientierte Einführung*. Göttingen: Verlag Hans Huber.
- Salomon, G. (1979). *Interaction of media, cognition, and learning: An exploration of how symbolic forms cultivate mental skills and affect knowledge acquisition*. San Francisco: Jossey-Bass.
- Schmidt-Weigand, F. (2006). *Dynamic visualizations in multimedia learning: The influence of verbal explanations on visual attention, cognitive load and learning outcome*. Elektronische Dissertation, Gießen: Justus Liebig Universität.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49-69). Cambridge: Cambridge University Press.
- Schnotz, W. (2006). Lernen mit Animationen. In F. Lösel & D. Bender (Hrsg.), 45. *Kongress der Deutschen Gesellschaft für Psychologie* (S. 314-315). Lengerich: Pabst Science Publishers.
- Schnotz, W. & Bannert, M. (1999). Einflüsse der Visualisierungsform auf die Konstruktion mentaler Modelle beim Text- und Bildverstehen. *Zeitschrift für*

- Experimentelle Psychologie*, 46, 216-235.
- Schnotz, W. & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13, 141-156.
- Schnotz, W., Bockheler, J., & Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. *European Journal of Psychology of Education*, 14, 245-265.
- Schnotz, W., Picard, E. & Hron, A. (1993). How do successful and unsuccessful learners use texts and graphics? *Learning and Instruction*, 3, 181-199.
- Schnotz, W. & Lowe, R. (2003). External and internal representations in multimedia learning. *Learning and Instruction*, 13, 117-123.
- Schnotz, W., Seufert, T. & Bannert, M. (2001). Lernen mit Multimedia: Pädagogische Verheißungen aus kognitionspsychologischer Sicht. In R. K. Silbereisen (Hrsg.), *Bericht über den 42. Kongreß der Deutschen Gesellschaft für Psychologie* (S. 457-467). Göttingen: Hogrefe.
- Schuh, J., Gerjets, P., & Scheiter, K. (2005). Fostering the acquisition of transferable problem-solving knowledge with an interactive comparison tool and dynamic visualizations of solution procedures. In B. G. Bara, L. Barsalou & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 1973-1978). Mahwah: Erlbaum.
- Schulze, R. (2004). *Meta-analysis: a comparison of approaches*. Hogrefe & Huber: Göttingen.
- Schwan, S., Gassoffky, B. & Hesse, F. W. (2000). Do film cuts facilitate the perceptual and cognitive organization of activity sequences? *Memory and Cognition*, 28, 214-223.
- Schwan, S. & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction*, 14, 293-305.
- Schwartz, R. N. (2003). *The use of animation in education*. Unveröffentlichtes Manuskript, New York University.
- Schwonke, R., Hauser, S., Nückles, M., & Renkl, A. (2006). Enhancing computer-supported writing of learning protocols by adaptive prompts. *Computers in Human Behavior*, 22, 77-92.
- Shyu, H. S. & Brown, S. W. (1992). Learner control versus program control in interactive videodisc instruction: What are the effects in procedural learning? *International Journal of Instructional Media*, 19, 85-96.

- Skinner, B. F. (1954). The science of learning and the art of teaching. In (reprint) A. A. Lumsdaine & R. Glaser (Eds.) (1960), *Teaching machines and programmed learning* (pp. 99-113). Washington, D.C.: National Education Association of the United States.
- \*Spangenberg, R. W. (1973). The motion variable in procedural learning. *AV Communication Review*, 21, 419-436.
- \*Spotts, J. & Dwyer, F. M. (1996). The effect of computer-generated animation on student achievement of different types of educational objectives. *International Journal of Instructional Media*, 23, 365-375.
- SPSS. (2004). *SPSS 12.0.2 for Windows* [Computerprogramm]. Chicago: SPSS, Inc.
- Squire, L. R. & Kandel, E. R. (1999). *Gedächtnis: Die Natur des Erinnerns*. Heidelberg: Spektrum Akademischer Verlag.
- Stark, R., Graf, M., Renkl, A., Gruber, H. & Mandl, H. (1995). Förderung von Handlungskompetenz durch geleitetes Problemlösen und multiple Lernkontexte. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 27, 289-312.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 295-312.
- Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159-167). Cambridge: Cambridge University Press.
- \*Swezey, R. W. (1991). Effects of instructional strategy and motion presentation conditions on the acquisition and transfer of electromechanical troubleshooting skill. *Human Factors*, 33, 309-323.
- \*Szabo, M. & Poohkay, B. (1996). An experimental study of animation, mathematics achievement, and attitude toward computer-assisted instruction. *Journal of Research on Computing in Education*, 28, 390-402.
- Tabbers, H. K. (2002). *The modality of text in multimedia instructions: refining the design guidelines*. Elektronische Dissertation, Heerlen: Open University of the Netherlands.
- Tallmadge, G. A. (1977). *The joint dissemination review panel ideabook*. Washington, DC: National Institute of Education & U.S. Office of Education.
- Tarmizi, R. & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, 80, 424-436.

- Triantafillou, E., Pomportsis, A., Demetriadis, S. & Georgiadou, E. (2004). The value of adaptivity based on cognitive style: an empirical study. *British Journal of Educational Technology*, 35, 95–106.
- Tversky, B., Morrison, J.-B., & Bétrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human Computer Studies*, 57, 247-262.
- Urhahne, D. (2002). *Motivation und Verstehen*. Münster: Waxmann.
- Urhahne, D., Prenzel, M., von Davier, M., Senkbeil, M. & Bleschke, M. (2000). Computereinsatz im naturwissenschaftlichen Unterricht – Ein Überblick über die pädagogisch-psychologischen Grundlagen und ihre Anwendung. *Zeitschrift für Didaktik der Naturwissenschaften*, 6, 157-186.
- Weiss, R.-E., Knowlton, D.-S., & Morrison, G.-R. (2002). Principles for using animation in computer-based instruction: Theoretical heuristics for effective design. *Computers in Human Behavior*, 18, 465-477.
- Wirth, J. (2004). *Selbstregulation von Lernprozessen*. Münster: Waxmann
- Witrock, M. C. (1974). Learning as a generative process. *Educational Psychologist*, 11, 87-95.
- Witrock, M. C. (1978). Developmental processes in learning from instruction. *Journal of Genetic Psychology*, 132, 352-373.
- Witrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist*, 24, 345-376.
- \*Wright, P., Milroy, R., & Lickorish, A. (1999). Static and animated graphics in learning from interactive texts. *European Journal of Psychology of Education*, 14, 203-224.
- \*Yang, E. M., Andre, T., & Greenbowe, T. Y. (2003). Spatial ability and the impact of visualization/animation on learning electrochemistry. *International Journal of Science Education*, 25, 329-349.

## **10 Anhang**

Anhang A: Tabelle der in die Metaanalyse aufgenommenen Studien

Anhang B: Vortest der ersten und dritten empirischen Studie

Anhang C: Nachtest der ersten bis dritten empirischen Studie

Anhang D: Vortest der zweiten empirischen Studie

Anhang E: Paper-Folding-Test

Anhang F: Card-Rotation-Test

Anhang G: Abschrift des Audiokommentars der Lernumgebungen

## Anhang A: Tabelle der in die Metaanalyse aufgenommenen Studien

Anhang A

Studie	Gesamtstichprobengröße N	Stichprobengröße n des Einzelvergleichs	Adjustierte Stichprobengröße des Einzelvergleichs	Gewichtete Effektstärke d	Art der Stichprobe	Inhaltsbereich	Art der Animation	Grad der Realistik	Cues	Text	Rolle der Animation	Art der Lernleistung
Baek & Layne, 1988	119 <sup>a</sup>	46	46,0	1,17	Schüler	Mathematik: Durchschnittsgeschwindigkeit	Computerbasiert	2	n/a	Ja	Dekorativ	Prozeduralproblemlösend
		38	38,0	0,67	Schüler		Computerbasiert	2	n/a	Ja	Dekorativ	Prozeduralproblemlösend
Blake, 1977	84 <sup>b</sup>	28	28,0	1,34	Studierende	Bewegungsmodi von Schachfiguren	Videobasiert	4	Ja	Nein	Repräsentational	Deklarativ
		28	28,0	0,14	Studierende		Videobasiert	4	Ja	Nein	Repräsentational	Deklarativ
		28	28,0	1,27	Studierende		Videobasiert	4	Nein	Nein	Repräsentational	Deklarativ
		28	28,0	1,30	Studierende		Videobasiert	4	Nein	Nein	Repräsentational	Deklarativ
Catrambone & Seay, 2002, Exp. 2	188 <sup>c,d</sup>	188	94,0	0,82	Studierende	Computeralgorithmen	Computerbasiert	1	Nein	Ja	Dekorativ	Prozeduralproblemlösend
		188	94,0	-1,32	Studierende		Computerbasiert	1	Nein	Ja	Dekorativ	Prozeduralproblemlösend
ChanLin, 1998	135 <sup>a,c,e</sup>	50	25,0	0,92	Studierende	DNS-Rekombination	Computerbasiert	3	Ja	Ja	Repräsentational	Deklarativ
		50	25,0	-0,98	Studierende		Computerbasiert	3	Ja	Ja	Repräsentational	Prozeduralproblemlösend
		40	20,0	-0,01	Studierende		Computerbasiert	3	Ja	Ja	Repräsentational	Deklarativ
		40	20,0	0,31	Studierende		Computerbasiert	3	Ja	Ja	Repräsentational	Prozeduralproblemlösend
ChanLin, 2001	357 <sup>a,c,e</sup>	142	71,0	2,96	Schüler	Physik: Kräfte	Computerbasiert	2	Ja	Ja	Dekorativ	Prozeduralproblemlösend
		142	71,0	2,84	Schüler		Computerbasiert	2	Ja	Ja	Dekorativ	Deklarativ
		92	46,0	-4,61	Schüler		Computerbasiert	2	Ja	Ja	Dekorativ	Deklarativ

*Tabelle: Gesamtheit der in die Metaanalyse aufgenommenen Studien inklusive weiter analysierter Eigenschaften sowie adjustierter und nicht nichtadjustierter Stichprobengrößen und gewichteter Effektstärken.*



Studie	Gesamtstichprobengröße N	Stichprobengröße n des Einzelvergleichs	Adjustierte Stichprobengröße des Einzelvergleichs	Gewichtete Effektstärke d	Art der Stichprobe	Inhaltsbereich	Art der Animation	Grad der Realistik	Cues	Text	Rolle der Animation	Art der Lernleistung
		92	46,0	-2,28	Schüler		Computerbasiert	2	Ja	Ja	Dekorativ	Prozeduralproblemlösend
Craig, Gholson & Driscoll, 2002, Exp. 1	135 <sup>b,c,f</sup>	30	3,75	0,01	Studierende	Gewitterentstehung	Computerbasiert	3	Ja	Ja	Repräsentational	Deklarativ
		30	3,75	0,07	Studierende		Computerbasiert	3	Nein	Ja	Repräsentational	Deklarativ
		30	3,75	-0,01	Studierende		Computerbasiert	3	Ja	Ja	Repräsentational	Deklarativ
		30	3,75	0,04	Studierende		Computerbasiert	3	Nein	Ja	Repräsentational	Deklarativ
		30	3,75	0,04	Studierende		Computerbasiert	3	Ja	Ja	Repräsentational	Deklarativ
		30	3,75	0,12	Studierende		Computerbasiert	3	Nein	Ja	Repräsentational	Deklarativ
		30	3,75	0,03	Studierende		Computerbasiert	3	Ja	Ja	Repräsentational	Prozeduralproblemlösend
		30	3,75	-0,01	Studierende		Computerbasiert	3	Nein	Ja	Repräsentational	Prozeduralproblemlösend
Hays, 1996	116 <sup>a,c,g</sup>	77	77,0	0,20	Schüler	Diffusion	Computerbasiert	3	n/a	Ja	Repräsentational	Deklarativ
		67	67,0	1,58	Schüler		Computerbasiert	3	n/a	Ja	Repräsentational	Prozeduralproblemlösend
Hoeffler, 2003	115 <sup>c,f,h</sup>	59	29,5	-0,30	Studierende	Photosynthese	Computerbasiert	3	Ja	Ja	Repräsentational	Deklarativ
		59	29,5	-0,20	Studierende		Computerbasiert	3	Ja	Ja	Repräsentational	Prozeduralproblemlösend
Kaiser, Proffitt & Anderson, 1985, Exp. 1	105 <sup>i</sup>	51	51,0	1,88	Studierende	Bewegungsabläufe	Videobasiert	4	Ja	Nein	Repräsentational	Prozeduralproblemlösend
Lai, 2000	316 <sup>e,j,k</sup>	126	63,0	-1,82	Studierende	Grundlagen	Computerbasiert	2	Ja	Ja	Dekorativ	Deklarativ

Tabelle: (fortgesetzt)

Studie	Gesamtstichprobengröße N	Stichprobengröße n des Einzelvergleichs	Adjustierte Stichprobengröße des Einzelvergleichs	Gewichtete Effektstärke d	Art der Stichprobe	Inhaltsbereich	Art der Animation	Grad der Realistik	Cues	Text	Rolle der Animation	Art der Lernleistung
		126	63,0	-0,07	Studierende	des Programmierens	Computerbasiert	2	Ja	Ja	Dekorativ	Deklarativ
Lewalter, 2003	60 <sup>a,c,f,l</sup>	40	20,0	0,01	Studierende	Gravitation	Computerbasiert	3	Ja	Ja	Dekorativ	Deklarativ
		40	20,0	0,58	Studierende		Computerbasiert	3	Ja	Ja	Dekorativ	Prozeduralproblemlösend
McCloskey & Kohl, 1983, Exp. 2	72	72	72,0	-0,60	Studierende	Bewegungsabläufe	Videobasiert	4	Ja	Nein	Repräsentational	Prozeduralproblemlösend
Michas & Berry, 2000, Exp. 1	75 <sup>b,j</sup>	30	15,0	0,74	Studierende	Bandagieren einer Hand	Videobasiert	4	Nein	Nein	Repräsentational	Prozeduralmotorisch
		30	15,0	0,69	Studierende		Videobasiert	4	Ja	Nein	Repräsentational	Prozeduralmotorisch
Nerdel, 2003, Exp. 2 & 3	131 <sup>c,f,h</sup>	55	27,5	0,43	Schüler	Atmungskette	Computerbasiert	3	Ja	Ja	Repräsentational	Deklarativ
		55	27,5	0,49	Schüler		Computerbasiert	3	Ja	Ja	Repräsentational	Prozeduralproblemlösend
	150 <sup>c,f,h</sup>	60	30,0	0,06	Schüler	Photosynthese	Computerbasiert	3	Ja	Ja	Repräsentational	Deklarativ
		60	30,0	-0,30	Schüler		Computerbasiert	3	Ja	Ja	Repräsentational	Prozeduralproblemlösend
Nicholls & Merkel, 1996, Exp. 1	44 <sup>m</sup>	21	21,0	0,62	Studierende	Stickstoffzyklus	Computerbasiert	n/a	n/a	Ja	Repräsentational	Prozeduralproblemlösend
Rieber, 1989	192 <sup>a,c,n</sup>	32	8,0	1,49	Schüler	Newtons Gesetze der Bewegung	Computerbasiert	1	Ja	Nein	Repräsentational	Prozeduralproblemlösend
		32	8,0	0,54	Schüler		Computerbasiert	1	Ja	Nein	Repräsentational	Prozeduralproblemlösend
		32	8,0	1,31	Schüler		Computerbasiert	1	Ja	Nein	Repräsentational	Deklarativ

Tabelle: (fortgesetzt)

Studie	Gesamtstichprobengröße N	Stichprobengröße n des Einzelvergleichs	Adjustierte Stichprobengröße des Einzelvergleichs	Gewichtete Effektstärke d	Art der Stichprobe	Inhaltsbereich	Art der Animation	Grad der Realistik	Cues	Text	Rolle der Animation	Art der Lernleistung
		32	8,0	-0,12	Schüler		Computerbasiert	1	Ja	Nein	Repräsentational	Deklarativ
		32	8,0	0,27	Schüler		Computerbasiert	1	Ja	Nein	Repräsentational	Prozeduralproblemlösend
		32	8,0	0,24	Schüler		Computerbasiert	1	Ja	Nein	Repräsentational	Prozeduralproblemlösend
		32	8,0	0,04	Schüler		Computerbasiert	1	Ja	Nein	Repräsentational	Deklarativ
		32	8,0	0,40	Schüler		Computerbasiert	1	Ja	Nein	Repräsentational	Deklarativ
		32	8,0	0,13	Schüler		Computerbasiert	1	Ja	Ja	Repräsentational	Prozeduralproblemlösend
		32	8,0	0,34	Schüler		Computerbasiert	1	Ja	Ja	Repräsentational	Prozeduralproblemlösend
		32	8,0	0,24	Schüler		Computerbasiert	1	Ja	Ja	Repräsentational	Deklarativ
		32	8,0	0,34	Schüler		Computerbasiert	1	Ja	Ja	Repräsentational	Deklarativ
		32	8,0	0,39	Schüler		Computerbasiert	1	Ja	Ja	Repräsentational	Prozeduralproblemlösend
		32	8,0	0,22	Schüler		Computerbasiert	1	Ja	Ja	Repräsentational	Prozeduralproblemlösend
		32	8,0	0,32	Schüler		Computerbasiert	1	Ja	Ja	Repräsentational	Deklarativ
		32	8,0	-0,03	Schüler		Computerbasiert	1	Ja	Ja	Repräsentational	Deklarativ
		Rieber, 1990	119 <sup>a</sup>	79	79,0		-0,39	Schüler	Newtons Gesetze der Bewegung	Computerbasiert	1	Ja
Rieber, 1991	70 <sup>c,f</sup>	70	35,0	-0,37	Schüler	Newtons Gesetze der	Computerbasiert	1	Ja	Ja	Repräsentational	Prozeduralproblemlösend
		70	35,0	-0,03	Schüler		Computer-	1	Ja	Ja	Repräsentational	Deklarativ

Tabelle: (fortgesetzt)

Studie	Gesamtstichprobengröße N	Stichprobengröße n des Einzelvergleichs	Adjustierte Stichprobengröße des Einzelvergleichs	Gewichtete Effektstärke d	Art der Stichprobe	Inhaltsbereich	Art der Animation	Grad der Realistik	Cues	Text	Rolle der Animation	Art der Lernleistung
						Bewegung	basiert					
Rieber, Boyce & Assad, 1990	141 <sup>a,j</sup>	31	31,0	0,22	Studierende	Newtons Gesetze der Bewegung	Computerbasiert	1	Ja	Ja	Repräsentational	Prozeduralproblemlösend
		31	31,0	-1,13	Studierende		Computerbasiert	1	Ja	Ja	Repräsentational	Prozeduralproblemlösend
		31	31,0	1,21	Studierende		Computerbasiert	1	Ja	Ja	Repräsentational	Prozeduralproblemlösend
Rigney & Lutz, 1976	40 <sup>c,o</sup>	40	10,0	0,41	Studierende	Elektrochemische Vorgänge in der Batterie	Computerbasiert	n/a	n/a	Ja	Dekorativ	Deklarativ
		40	10,0	0,31	Studierende		Computerbasiert	n/a	n/a	Ja	Dekorativ	Deklarativ
		40	10,0	0,32	Studierende		Computerbasiert	n/a	n/a	Ja	Dekorativ	Deklarativ
		40	10,0	0,24	Studierende		Computerbasiert	n/a	n/a	Ja	Dekorativ	Deklarativ
Spangenberg, 1973, Exp. 1 & 2	120	40	40,0	1,45	Rekruten	Zusammenbau eines MG	Videobasiert	4	Nein	Nein	Repräsentational	Prozeduralmotorisch
		40	40,0	0,99	Rekruten		Videobasiert	4	Ja	Nein	Repräsentational	Prozeduralmotorisch
		40	40,0	0,99	Rekruten		Videobasiert	4	Nein	Nein	Repräsentational	Prozeduralmotorisch
Spotts & Dwyer, 1996	63 <sup>j</sup>	41	41,0	2,02	Studierende	Blutfluss im Herzen	Computerbasiert	2	Nein	Ja	Dekorativ	Deklarativ
Swezey, 1991	120 <sup>p</sup>	120	120,0	0,00	Studierende	Funktionen des Motors	Videobasiert	4	Nein	Nein	Dekorativ	Deklarativ
Szabo & Poohkay, 1996	174 <sup>a</sup>	117	117,0	0,82	Studierende	Dreiecks-konstruktion mit Kompass	Computerbasiert	2	Ja	Ja	Dekorativ	Deklarativ

Tabelle: (fortgesetzt)

Studie	Gesamtstichprobengröße N	Stichprobengröße n des Einzelvergleichs	Adjustierte Stichprobengröße des Einzelvergleichs	Gewichtete Effektstärke d	Art der Stichprobe	Inhaltsbereich	Art der Animation	Grad der Realistik	Cues	Text	Rolle der Animation	Art der Lernleistung
Wright, Milroy & Lickorish, 1999, Exp. 1	60 <sup>a,c,q</sup>	40	20,0	0,50	Adults	Britische	Computerbasiert	n/a	n/a	Ja	Dekorativ	Deklarativ
		40	20,0	-0,22	Adults	Geschichte	Computerbasiert	n/a	n/a	Ja	Dekorativ	Deklarativ
Yang, Andre & Greenbowe, 2003	263	263	263,0	7,84	Studierende	Elektrischer Stromkreis in einer Lampe	Computerbasiert	2	Ja	Ja	Repräsentational	Deklarativ <sup>rst</sup>

<sup>a</sup> Eine weitere, nicht berücksichtigte Versuchsbedingung war "Nur Text".

<sup>b</sup> Zwei verschiedene statische Versionen (mit/ohne cues) wurden mit einer Animationsversion verglichen. n wurde dementsprechend adjustiert.

<sup>c</sup> Die jeweiligen Stichprobengrößen n wurden entsprechend des Einbezugs multipler abhängiger Variablen adjustiert.

<sup>d</sup> Sogenannte "Near transfer"- und "far transfer"-Aufgaben wurden denselben Probanden dargeboten. Die Ergebnisse wurden zusammengefasst.

<sup>e</sup> Experten und Novizen wurden unterschieden und sowohl deklarative als auch prozedural-problemlösende Aufgaben vorgelegt.

<sup>f</sup> Denselben Probanden wurden deklarative wie auch prozedural-problemlösende Aufgaben vorgelegt; diese wurden separat ausgewertet.

<sup>g</sup> Prozedural-problemlösendes Wissen wurde eine Woche später als "long-term comprehension" erhoben.

<sup>h</sup> Weitere Versuchsbedingungen existierten in Form von interaktiven Animationen (Simulationen). Diese wurden nicht in die Metaanalyse einbezogen.

<sup>i</sup> Die Autoren geben nicht an, warum sie lediglich 51 ihrer 105 Probanden heranzogen, um Unterschiede zwischen dynamischen und nichtdynamischen Bedingungen zu analysieren.

<sup>j</sup> Weitere Versuchsbedingungen existierten, wurden aber nicht in die Metaanalyse einbezogen.

<sup>k</sup> Derselbe Test wurde als Posttest und Follow-Up-Test eine Woche später benutzt.

<sup>l</sup> Zwei verschiedene prozedural-problemlösende Testskalen wurden zusammengefasst.

<sup>m</sup> Die Autoren führen an, dass "wegen eines Missverständnisses" eine weitere Gruppe von Probanden nicht in weitere Analysen miteinbezogen werden sollte.

<sup>n</sup> Deklarative wie auch prozedural-problemlösende Aufgaben waren zusätzlich unterteilt in "near" und "far". Alle Bedingungen wurden separat analysiert.

<sup>o</sup> Vier verschiedene deklarative Tests wurden denselben Probanden vorgelegt und in der Metaanalyse separat analysiert.

<sup>p</sup> Verschiedene Wissenstests wurden benutzt, von denen aber lediglich einer ("conceptual knowledge") klar in die in der Metaanalyse benutzten Kategorien einzuteilen und dementsprechend weiter analysierbar war.

<sup>q</sup> Zwei verschiedene Wissenstests existierten; beide wurden separat analysiert.

## Anhang B: Vortest der ersten und dritten empirischen Studie

# 1. Fragebogen zur computergestützten Lernumgebung Tenside & Co

Liebe Studierende, lieber Studierender,

vielen Dank, dass Sie an dieser Untersuchung zur Wirksamkeit computergestützter Lernumgebungen zu der Wirkungsweise von Tensiden teilnehmen.

Ganz wichtig ist es für uns zu erfahren, inwieweit Sie mit diesem Inhaltsbereich bereits vertraut sind. Wir bitten Sie daher, die nachfolgenden Fragen so gut wie möglich zu beantworten. Natürlich gibt es dafür keine Noten o.ä. Deshalb ist es auch nicht schlimm, wenn Sie nur wenige Angaben zu den Fragen machen können. Beim anschließenden computergestützten Lernen werden Sie Gelegenheit haben, Ihre Wissensbasis aufzubauen, aufzufrischen oder zu erweitern.

Um die Vergleichbarkeit Ihrer Leistungen mit denen der anderen Teilnehmer zu gewährleisten, stehen zu Anfang einige persönliche Fragen zu Ihrer bisherigen Ausbildung. Wir möchten Sie bitten, auch diese Fragen so genau wie möglich zu beantworten. Alle von Ihnen zu diesem Zweck erhobenen Daten werden selbstverständlich absolut vertraulich behandelt und anonymisiert ausgewertet.

## *Ablauf der Untersuchung*

- Bitte füllen Sie zunächst diesen Fragebogen aus.
- Auf der vierten Seite werden Sie unten den Satz bemerken, erst umzublättern, wenn Sie dazu aufgefordert werden. Bitte legen Sie dann den Stift beiseite und warten Sie auf das Zeichen zum Fortfahren.
- Anschließend haben Sie die Möglichkeit, die Lerneinheit am Computer zu benutzen.
- Zum Abschluss folgen zwei weitere Fragebögen.

Haben Sie noch Fragen? Wenn nicht, dann beginnen Sie bitte jetzt mit der Bearbeitung dieses Fragebogens!

Damit die Angaben dieses und der späteren Fragebogens einander zugeordnet werden können, ist es notwendig, eine Kennzeichnung der Fragebögen vorzunehmen. Dazu benötigen wir Angaben, die einerseits individuell sind, andererseits keine Rückschlüsse auf Ihre Person erlauben, um die Vertraulichkeit zu wahren.

Deshalb erstellen Sie bitte Ihr Kennzeichen auf folgende Weise:

- Tragen Sie in die vorderen 2 Kästchen die **Initialen Ihres Vaters** ein.
- Tragen Sie in die hinteren 4 Kästchen den **Geburtstag (TTMM) Ihrer Mutter** ein.

**Kennzeichen:**

		-				
--	--	---	--	--	--	--

(Wenn Ihr Vater z.B. Max Mustermann hieße und Ihre Mutter am 7.12.1960 Geburtstag hätte, lautete Ihr Kennzeichen: **MM – 0712.**)

Bitte tragen Sie hier zur Sicherheit noch die Nummer ein, die an Ihrem Arbeitsplatz befestigt ist:

### Angaben zur Person:

Wie alt sind Sie? \_\_\_\_\_ Jahre.

Geschlecht:  weiblich  männlich

### Angaben zur Schulbildung:

*In einer Schule welcher Schulform haben Sie Ihr Abitur abgelegt?*

Gymnasium  Gesamtschule  Waldorfschule  andere: \_\_\_\_\_

Bitte geben Sie Ihre Gesamtabiturnote (1,0-4,0) an! \_\_\_\_\_

*Hatten Sie in der Oberstufe das Fach Chemie belegt?*

- ja, im **Grundkurs** bis einschließlich zur \_\_\_\_\_ Klasse.
- ja, im **Grundkurs mit anschließender Abiturprüfung** („Prüfungskurs“).
- ja, im **Leistungskurs**.
- nein, sondern bis einschließlich zur \_\_\_\_\_ Klasse.

Bitte geben Sie Ihre Abitur- bzw. letzte Zeugnisnote im Fach Chemie (in Punkten von 0-15) an! \_\_\_\_\_



**Angaben zum Studium:**

Was ist Ihr angestrebter Abschluss?  Diplom  Magister  Staatsexamen

Welches Hauptfach / welche Hauptfächer studieren Sie?

\_\_\_\_\_

Welche Nebenfächer studieren Sie? \_\_\_\_\_

In welchem Semester befinden Sie sich? \_\_\_\_\_

**Fragen zum Inhaltsbereich:**

Wie funktioniert die Schmutzablösung beim Wäschewaschen?

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Wie lösen Geschirrspülmittel Fette?

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Warum vermischen sich Essig und Öl z.B. in Salatsaucen nicht?

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Erläutern Sie den Begriff „Oberflächenspannung“!

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

## Anhang C: Nachtest der ersten bis dritten empirischen Studie

## 2. Fragebogen zur computergestützten Lernumgebung Tenside & Co

Zum Abschluss folgt nun ein weiterer Fragebogen, für dessen Bearbeitung Sie genügend Zeit haben sollten.

Es ist für uns sehr wichtig, dass Sie auch diesen so gut wie möglich zu beantworten versuchen.

Damit die Angaben dieses Fragebogens denen des ersten zugeordnet werden können, erstellen Sie bitte Ihr Kennzeichen auf die gewohnte Weise:

- Tragen Sie in die vorderen 2 Kästchen die **Initialen Ihres Vaters** ein.
- Tragen Sie in die hinteren 4 Kästchen den **Geburtstag (TTMM) Ihrer Mutter** ein.

**Kennzeichen:**   -

(Wenn Ihr Vater z.B. Max Mustermann hieße und Ihre Mutter am 7.12.1960 Geburtstag hätte, lautete Ihr Kennzeichen: **MM – 0712.**)

Bitte tragen Sie hier zur Sicherheit noch die Nummer ein, die an Ihrem Arbeitsplatz befestigt ist:

Zunächst zwei Fragen zu den Anforderungen des Lernmaterials: Bitte versuchen Sie, sich so genau wie möglich einzuschätzen:

### 1. Wie leicht oder schwer war das Lernmaterial zu verstehen?

sehr leicht	leicht	eher leicht	weder leicht noch schwer	eher schwer	schwer	sehr schwer
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 2. Bei der Bearbeitung des Lernmaterials war meine mentale Anstrengung...

sehr gering	gering	eher gering	weder gering noch hoch	eher hoch	hoch	sehr hoch
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Es folgen einige Fragen zum in der Lernumgebung veranschaulichten Inhaltsbereich. Bitte versuchen Sie, die Fragen so genau und umfassend wie möglich zu beantworten!

1.) *Beschreiben Sie bitte so genau wie möglich, wie der Waschvorgang funktioniert!*

---

---

---

---

---

---

---

---

---

---

2.) *Zeichnen Sie hier schematisch die Anordnung von Tensiden an Schmutz und Textilien in einer Waschlösung ein!*



3.) *Beim Waschen entsteht Schaum. Haben Sie eine Idee, wie eine Schaum-/Seifenblase entstehen könnte?*

---

---

---

---

---

---

---

---

*Sobald Sie fertig sind, melden Sie sich bitte. Es folgt dann noch ein letzter Fragebogen!*

**Hinweis: Bei den Multiple-Choice-Fragen ist jeweils nur eine einzige Antwort richtig!**

1. Die Abbildung stellt modellhaft ein Tensid-Anion dar. Beschriften Sie die Abbildung mit folgenden Begriffen: hydrophil, hydrophob, funktionelle Gruppe, Kohlenwasserstoffkette



2. Warum kann ein Wasserläufer auf dem Wasser laufen?

- A  Weil er leichter ist als Wasser
- B  Seine Beine sind hydrophil beschichtet
- C  Er ist zu schnell, als dass er untergehen könnte
- D  Wegen der Oberflächenspannung des Wassers

3. Was würde passieren, wenn man Geschirrspülmittel in das Wasser zugäbe?

- A  Der Wasserläufer ginge unter.
- B  Der Wasserläufer könnte schneller gleiten.
- C  Der Wasserläufer flöge davon.
- D  Nichts.

4. Welches Begriffspaar stellt einen Gegensatz dar?

- A  hydrophil-wasseranziehend
- B  hydrophob-wasserabweisend
- C  hydrophil-fettanziehend
- D  hydrophob-fettanziehend

5. Welche Ladung hat die Kohlenwasserstoffkette eines anionischen Tensids?

- A  positiv
- B  negativ
- C  nicht geladen
- D  unterschiedlich

6. Welche Ladung hat die funktionelle Gruppe eines anionischen Tensids?

- A  positiv
- B  negativ
- C  nicht geladen
- D  unterschiedlich

7. Wie lagern sich anionische Tenside an verschmutzten Textilien an?

- A  In Form einer Micelle
- B  Mit der funktionellen Gruppe an die Textilie
- C  Mit dem hydrophilen Teil an die Textilie
- D  Mit dem hydrophoben Teil an die Textilie

8. Wie kommt es zur Schmutzablösung? Welcher Vorgang ist dafür GENAU verantwortlich?

---

---

---

---

---

9. Warum reicht Wasser zur Schmutzablösung oft nicht aus?

---

---

---

10. Erläutern Sie den Begriff "Wasserstoffbrückenbindungen"!

---

---

---

11. Was sind Micellen?

---

---

---

---

12. Zeichnen Sie in die Abbildung ein, wie sich die Tenside während der Schmutzablösung anordnen!



13. Welche Eigenschaft der Tenside sorgt letzten Endes für die Ablösung des Schmutzes nach der Benetzung?

- A  Ihre Ladung.
- B  Ihre Form.
- C  Ihre Beweglichkeit.
- D  Ihre Hydrophilie/Hydrophobie.

14. Warum perlt Regen von imprägnierten Kleidungsstücken ab?

- A  Tenside im Imprägnierungsmittel machen den Stoff wasserabweisend.
- B  Der Stoff ist zu engmaschig, als dass das Wasser ihn benetzen könnte.
- C  Die Oberflächenspannung des Wassers verhindert eine Benetzung.
- D  Die Imprägnierung ist hydrophil.

15. Tenside werden oft auch in Cremes als "Stabilisatoren" verwendet. Warum?

- A  Sie stellen eine ausreichende Hautreinigung sicher.
- B  Sie sorgen dafür, dass sich Wasser- und Ölteile nicht trennen.
- C  Sie sorgen für die Einfettung der Haut.
- D  Sie stabilisieren durch Micellenbildung die Konsistenz der Creme.

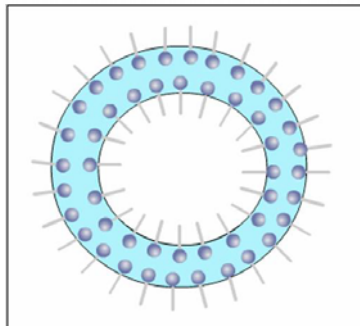
16. Nach dem Waschvorgang befindet sich die Wäsche in der schmutzigen Waschlauge. Warum verschmutzt die Wäsche dabei nicht wieder?

---

---

---

---



17. Links ist eine Seifenblase im Querschnitt dargestellt. Beschreiben und erklären Sie die Anordnung der Seifenanionen!

---

---

---

---

***Das war`s! Vielen Dank für Ihre Mitarbeit!***

## Anhang D: Vortest der zweiten empirischen Studie



---

# 1. Fragebogen zur computergestützten Lernumgebung Tenside & Co

Liebe Schülerin, lieber Schüler,

vielen Dank, dass Sie an dieser Untersuchung zur Wirksamkeit computergestützter Lernumgebungen zu der Wirkungsweise von Tensiden teilnehmen.

Ganz wichtig ist es für uns zu erfahren, inwieweit Sie mit diesem Inhaltsbereich bereits vertraut sind. Wir bitten Sie daher, die nachfolgenden Fragen so gut wie möglich zu beantworten. Natürlich gibt es dafür keine Noten o.ä. Deshalb ist es auch nicht schlimm, wenn Sie nur wenige Angaben zu den Fragen machen können. Beim anschließenden computergestützten Lernen werden Sie Gelegenheit haben, Ihre Wissensbasis aufzubauen, aufzufrischen oder zu erweitern.

Um die Vergleichbarkeit Ihrer Leistungen mit denen der anderen Teilnehmer zu gewährleisten, stehen zu Anfang einige persönliche Fragen zu Ihrer bisherigen Schullaufbahn. Wir möchten Sie bitten, auch diese Fragen so genau wie möglich zu beantworten. Alle von Ihnen zu diesem Zweck erhobenen Daten werden selbstverständlich absolut vertraulich behandelt und anonymisiert ausgewertet.

### *Ablauf der Untersuchung*

- Bitte füllen Sie zunächst diesen Fragebogen aus.
- Auf der vierten Seite werden Sie unten den Satz bemerken, erst umzublättern, wenn Sie dazu aufgefordert werden. Bitte legen Sie dann den Stift beiseite und warten Sie auf das Zeichen zum Fortfahren.
- Anschließend haben Sie die Möglichkeit, die Lerneinheit am Computer zu benutzen.
- Zum Abschluss folgen zwei weitere Fragebögen.

Haben Sie noch Fragen? Wenn nicht, dann beginnen Sie bitte jetzt mit der Bearbeitung dieses Fragebogens!

Damit die Angaben dieses und der späteren Fragebögen einander zugeordnet werden können, ist es notwendig, eine Kennzeichnung der Fragebögen vorzunehmen. Dazu benötigen wir Angaben, die einerseits individuell sind, andererseits keine Rückschlüsse auf Ihre Person erlauben, um die Vertraulichkeit zu wahren.

Deshalb erstellen Sie bitte Ihr Kennzeichen auf folgende Weise:

- Tragen Sie in die vorderen 2 Kästchen die **Initialen Ihres Vaters** ein.
- Tragen Sie in die hinteren 4 Kästchen den **Geburtstag (TTMM) Ihrer Mutter** ein.

**Kennzeichen:**

		-				
--	--	---	--	--	--	--

(Wenn Ihr Vater z.B. Max Mustermann hieße und Ihre Mutter am 7.12.1960 Geburtstag hätte, lautete Ihr Kennzeichen: **MM – 0712.**)

Bitte tragen Sie hier zur Sicherheit noch die Nummer ein, die an Ihrem Arbeitsplatz befestigt ist:

### Angaben zur Person:

Wie alt sind Sie? \_\_\_\_\_ Jahre.

Geschlecht:  weiblich  männlich

### Angaben zur Schulbildung:

*In einer Schule welcher Schulform gehen Sie im Moment zur Schule?*

Gymnasium  Gesamtschule  Realschule  andere: \_\_\_\_\_

*Bitte geben Sie Ihre letzte Zeugnisnote im Fach Chemie an!* \_\_\_\_\_

*Bitte geben Sie Ihre letzte Zeugnisnote im Fach Biologie an!* \_\_\_\_\_

**Einige Fragen zum Inhaltsbereich:**

*Wie funktioniert die Schmutzablösung beim Wäschewaschen?*

---

---

---

---

---

*Wie lösen Geschirrspülmittel Fette?*

---

---

---

---

---

*Warum vermischen sich Essig und Öl z.B. in Salatsaucen nicht?*

---

---

---

---

---

*Erläutern Sie den Begriff „Oberflächenspannung“!*

---

---

---

---

---

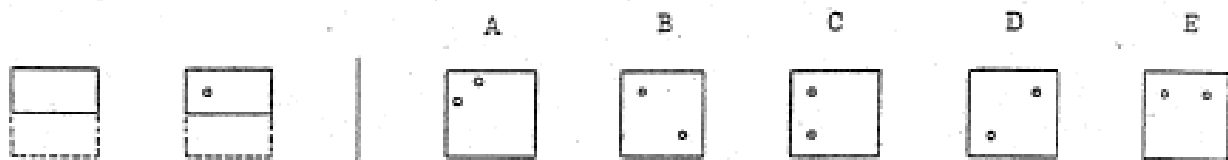
## Anhang E: Paper-Folding-Test

## PAPER FOLDING TEST

In diesem Test sollst du dir vorstellen, wie ein Stück Papier gefaltet wird. Jede Aufgabe besteht aus Abbildungen, die auf der linken Seite der senkrechten Linie und auf der rechten Seite der Linie stehen. Schau dir bitte das untere Beispiel an (*auf Beispiel zeigen*).

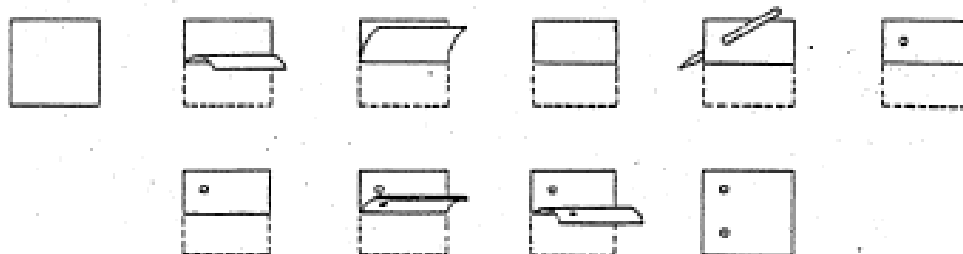
Die Abbildungen auf der linken Seite zeigen ein quadratisches Stück Papier, das Schritt für Schritt gefaltet wurde. Nachdem das Papier gefaltet wurde, wurden ein oder mehrere Löcher in das Papier gestanzt. Die Löcher wurden durch alle Lagen des gefalteten Papiers gestanzt. Die fünf Abbildungen auf der rechten Seite der senkrechten Linie zeigen, wie das Papier aussehen könnte, wenn man es wieder auseinanderfaltet. Das heißt, wo sich die Löcher dann befinden könnten. Nur eine dieser 5 Abbildungen ist richtig. Deine Aufgabe ist es, die richtige Abbildung herauszufinden. Kennzeichne diese Abbildung, indem du sie ankreuzt.

Beispiel



Die richtige Antwort in diesem Beispiel ist C und sollte deshalb angekreuzt werden.

Die untere Darstellung zeigt wie das Papier gefaltet wurde und warum C die richtige Antwort ist.

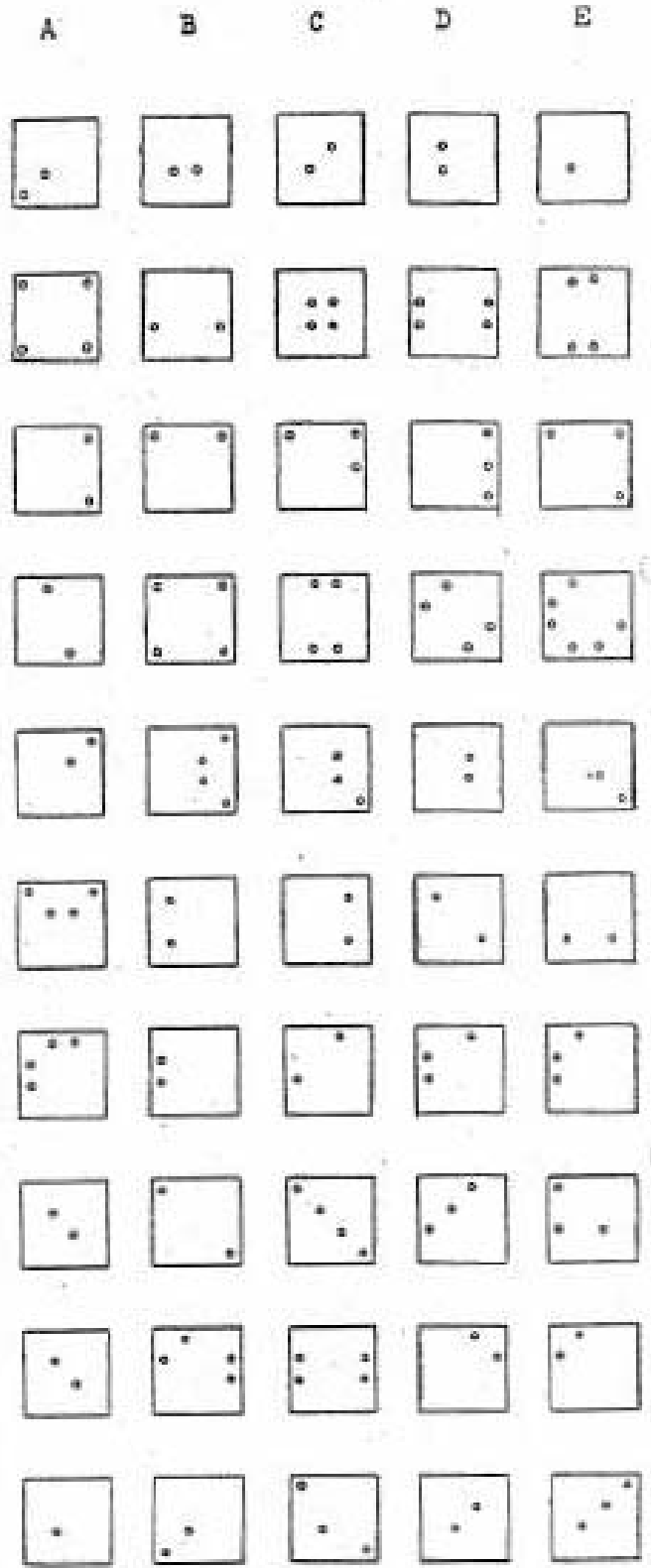
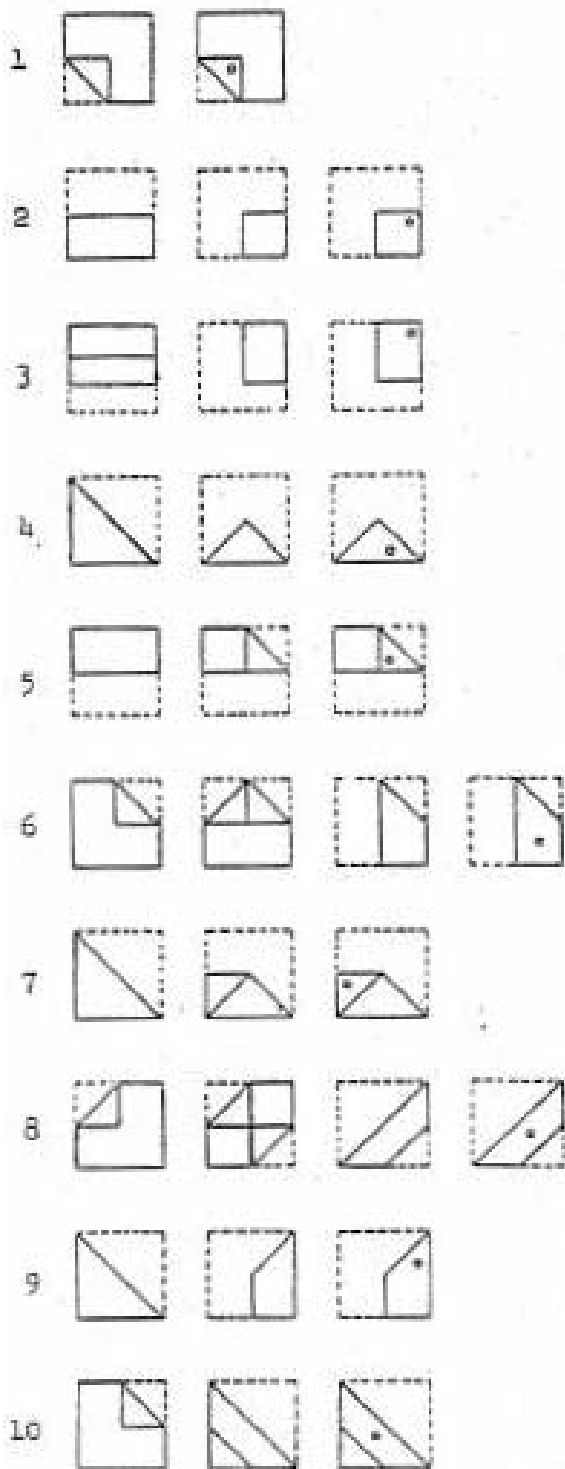


Alle Faltungen, die gemacht wurden, sind in den Abbildungen gezeigt. Das Papier wurde ansonsten nicht bewegt oder verändert.

Bitte erinnere dich daran, die richtige Antwort zeigt die richtige Position der Löcher wenn das Papier vollständig aufgefaltet wird.

Deine Punktzahl in diesem Test errechnet sich aus der Anzahl der richtig angekreuzten Antworten minus der Anzahl der falsch angekreuzten Antworten. Es ist also nicht vorteilhaft zu raten, es sei denn du kannst bestimmte Antworten ausschließen. Arbeite bitte so schnell du kannst, ohne dabei ungenau zu werden.

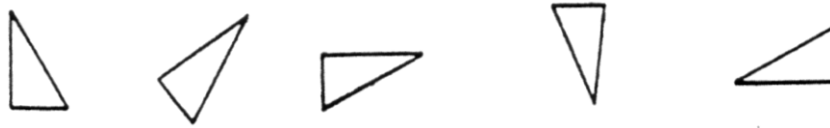
Du hast drei Minuten Zeit für diesen Test  
**Bitte erst umblättern, wenn du dazu aufgefordert wirst!**



## Anhang F: Card-Rotation-Test

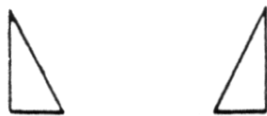
### Card Rotation Test

Dieser Test erfasst Ihre Fähigkeit, Unterschiede in Figuren zu erkennen. Schauen Sie sich bitte die 5 Dreiecke an, die hier aufgezeichnet sind:



Alle fünf Dreiecke zeigen die gleiche Figur, die so auf dem Papier gedreht wurde, dass sich verschiedene Positionen ergaben.

Schauen Sie sich nun die nächsten beiden Figuren an:



Diese beiden Figuren sind nicht gleich. Die erste kann nicht durch reines Drehen auf dem Papier dazu gebracht werden, wie die zweite auszusehen. Sie müsste dazu herumgeklappt oder verändert werden.

Jede Aufgabe besteht aus einer Figur, die auf der linken Seite der senkrechten Linie steht, und acht Figuren, die auf der rechten Seite der Linie stehen. Ihre Aufgabe ist, dir jede dieser acht rechten Figuren anzusehen und zu entscheiden, ob sie gleich oder nicht gleich zu der linken Figur ist. Kreuzen Sie das Kästchen neben dem „G“ an, wenn die rechte Figur gleich der linken Figur ist. Kreuzen Sie das Kästchen neben dem „N“ an, wenn die rechte Figur nicht gleich der linken Figur ist, d.h. nicht durch einfaches Drehen aus ihr hervorgehen kann.

Üben Sie in den folgenden Zeilen. Die erste Zeile ist bereits korrekt angekreuzt:

	 G <input checked="" type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input checked="" type="checkbox"/>	 G <input checked="" type="checkbox"/> N <input type="checkbox"/>	 G <input checked="" type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input checked="" type="checkbox"/>	 G <input type="checkbox"/> N <input checked="" type="checkbox"/>	 G <input type="checkbox"/> N <input checked="" type="checkbox"/>	 G <input type="checkbox"/> N <input checked="" type="checkbox"/>
	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>
	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>	 G <input type="checkbox"/> N <input type="checkbox"/>

Ihre Punktzahl in diesem Test errechnet sich aus der Anzahl der richtig angekreuzten Antworten minus der Anzahl der falsch angekreuzten Antworten. Es ist also nicht vorteilhaft zu raten, es sei denn, Sie haben eine Vermutung, ob die Figur gleich oder nicht gleich ist. Arbeiten Sie bitte so schnell wie möglich, ohne dabei ungenau zu werden.

Sie haben drei Minuten Zeit für diesen Test.

**Bitte erst umblättern, wenn Sie dazu aufgefordert werden!**



Kreuzen Sie an! G steht für die gedrehte, aber gleiche Figur. N steht für die nicht gleiche Figur!

1.										<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N
2.										<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N
3.										<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N
4.										<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N
5.										<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N
6.										<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N
7.										<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N
8.										<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N
9.										<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N
10.										<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N	<input type="checkbox"/> G <input type="checkbox"/> N

Vielen Dank!

## Anhang G: Abschrift des Audiokommentars der Lernumgebungen

**Abschrift des Audiokommentars der computerbasierten Lernumgebungen**

„In der ersten Phase der Schmutzablösung erfolgt die Benetzung von Schmutz und Faser durch die Tenside. Dabei lagern sie sich mit dem hydrophoben Teil an Schmutz und Faser an, während sich der hydrophile in Richtung des Wassers ausrichtet.

In der zweiten Phase stoßen sich die negativ geladenen Tenside gegenseitig ab und ordnen sich so an, dass sich an der Grenze zwischen Faser und Schmutz eine negativ geladene, hydrophile Schicht ausbildet. Dadurch wird nun der Schmutz langsam von der Faser abgehoben.

In der dritten Phase wird der Schmutz vollständig von der Faser getrennt. Faser und Schmutz werden sofort von weiteren Tensiden aus der Waschlauge benetzt, die sich teilweise zuvor in Micellen angeordnet hatten.

In der abschließenden vierten Phase wird der Schmutz in der wässrigen Lösung gehalten, da die Tenside eine hydrophile, negativ geladene Schicht um das Schmutzteilchen bilden, die von den ebenfalls negativ geladenen Tensiden auf der Faser abgestoßen werden. Abschließend werden die Schmutzteilchen mit der Waschlauge entfernt.“