

Technische Universität Braunschweig

Institut für Sozialwissenschaften
Medienwissenschaften

Vorstellung der Untersuchung “Räumliches
Vorstellungsvermögen und Lernen mit Multimedia”

Seminararbeit

Karsten Dambekalns (2544375)
Betreuer: Dr. Thomas Huk
Braunschweig, 18. Juli 2002

Inhaltsverzeichnis

| | | |
|----------|---|------------|
| 0 | Einleitung | 2 |
| 1 | Hintergrund der Untersuchung | 3 |
| 1.1 | Multimediale Lehr- und Lernsysteme | 3 |
| 1.2 | Informationskodalität | 3 |
| 1.3 | Lernereigenschaften und Interaktionseffekte | 4 |
| 1.4 | Mentale Modelle | 4 |
| 1.5 | Ziele der untersuchten Studie | 5 |
| 2 | Das Versuchsdesign | 6 |
| 2.1 | Experimente und Versuchspersonen | 6 |
| 2.2 | Vortest zum räumlichen Vorstellungsvermögen | 6 |
| 2.3 | Test des Lehrsystemes | 7 |
| 2.4 | Kriteriumsorientierte Tests | 8 |
| 3 | Ergebnisse und Auswertung der Experimente | 10 |
| 3.1 | Vortest räumliches Vorstellungsvermögen | 10 |
| 3.2 | Haupttest Lehrsystem | 11 |
| 4 | Diskussion der Vorgehensweise | 16 |
| 5 | Fazit | 18 |
| | Literaturverzeichnis | I |
| | Index | III |

0 Einleitung

Der vorliegende Text stellt die Studie *Räumliches Vorstellungsvermögen und Lernen mit Multimedia* [Brü⁺00] von Roland Brünken, Susan Steinbacher und Detlev Leutner vor.

Ziel ist es nicht, sich eingehend mit den inhaltlichen Aspekten der Studie zu befassen¹. Vielmehr soll es darum gehen, die formellen Aspekte der verwendeten statistischen Methoden zu beleuchten, und festzustellen ob die Autoren die gewählten Verfahren sinnvoll und korrekt eingesetzt haben.

Zum besseren Verständnis wird trotz dieser Zielsetzung zu Beginn ein Überblick über die von den Autoren der Studie untersuchte Thematik gegeben. Im Anschluss daran wird der Aufbau des Experimentes beschrieben, um eine Grundlage für die folgende Beschreibung der Ergebnisse sowie der verwendeten Analyseverfahren zu schaffen. Der nächste Abschnitt versucht zu klären, ob in der Studie Voraussetzungen und Besonderheiten der verwendeten Analyseverfahren berücksichtigt wurden, und alle relevanten Informationen im Text enthalten sind.

¹Hierfür fehlt dem Autor auch das nötige Fachwissen

1 Hintergrund der Untersuchung

Dieses Kapitel vermittelt einen Überblick über den theoretischen Hintergrund der untersuchten Studie. Erläuterungen zu Fragen der Informationskodalität, zum Begriff der mentalen Modelle sowie zu den untersuchten Merkmalen sollen beim Verständnis des Textes helfen.

1.1 Multimediale Lehr- und Lernsysteme

Multimediale Lehr- und Lernsysteme finden immer weitere Verbreitung. Der Begriff Multimedia sorgte lange für Diskussionen, einige definieren Multimedia über den technischen Aspekt; dieser Ansatz war lange weitestgehend akzeptiert, und wird auch heute noch benutzt.

Andere wiederum verstehen unter Multimedia Computerprogramme, die mindestens drei der Elemente Text, Ton, Grafik, Animation, Foto, Film und Daten² kombinieren.

In der aktuellen Diskussion steht jedoch die Interaktivität immer stärker im Mittelpunkt des Interesses. Der Anwender kann also den Ablauf des Programms, die Präsentation des multimedialen Angebotes nach seinen Wünschen beeinflussen. Hier wird Multimedia also weniger über die Hardware definiert, sondern über die Interaktivität mit dem Benutzer.

Diese Definitionsfragen spielen hier eine untergeordnete Rolle, im Rahmen der Untersuchung geht es jedoch um Fragen, die ein System im groben Sinn der zweiten Definition voraussetzen. Insbesondere die Rolle von Text und Bild in der Informationsvermittlung ist Gegenstand der Studie.

Die immer stärkere Verbreitung solcher Lehr- und Lernsysteme wirft Fragen nach der Wirksamkeit der genutzten Methodiken auf. Denn sowohl die Produktion als auch die Einführung und Nutzung solcher Systeme verursachen Kosten. Diese wollen durch einen im Vergleich zu klassischen Lehrmethoden wie auch zu anderen multimedialen Systemen höheren Lehr- beziehungsweise Lernerfolg gerechtfertigt werden.

1.2 Informationskodalität

In Rahmen zahlreicher Untersuchungen wurde in der Vergangenheit bereits die Informationsverarbeitung beim Lernen mit multimedialen CBT-Systemen untersucht. Hierbei beschäftigte man sich besonders mit der Frage, ob und wie die Informationsverarbeitung von der Art und Weise (*Kodalität*) abhängt, in der Informationen präsentiert werden (Text, Bild, Sprache, Video, etc.).

Die vorliegenden Untersuchungen finden hierbei durchgängig Vorteile für Systeme, die Informationen in mehreren unterschiedlichen Kodalitäten darbieten. Mayer nennt dies *multimedia dual coding effect* [Maye97]. Schwierig ist die Generalisierbarkeit des dual coding effect vor

²Auch die übrigen Elemente stellen für den Computer Daten dar, insofern ist diese Unterscheidung fragwürdig.

allem deshalb, weil bei diesbezüglichen Untersuchungen Lehrsysteme verglichen wurden, deren inhaltliche Äquivalenz nicht kontrolliert wurde (reine Textsysteme und Text-/Bildsysteme).

Die feststellbaren Effekte der Auswirkungen auf den Lernerfolg, die von der Gestaltung des Lehrsystems abhängen, werden als *Treatment-Effekte* bezeichnet.

1.3 Lernereigenschaften und Interaktionseffekte

Neben den Treatment-Effekten wird zunehmend untersucht, ob auch Eigenschaften des Lernenden den Lernerfolg mit multimedialen Systemen beeinflussen (*Aptitude-Effekt*), und ob es Wechselwirkungen zwischen Lernereigenschaften und Instruktionsformen (*Aptitude-Treatment-Interaktion*) gibt.

Eine der hierbei untersuchten Lernereigenschaften³ ist das räumliche Vorstellungsvermögen (*spatial ability*). Es anzunehmen, dass die individuelle Ausprägung des räumlichen Vorstellungsvermögens den Erfolg multimedialer Lehrsysteme beeinflusst.

Zur Definition des Begriffes “Räumliches Vorstellungsvermögen” gibt es verschiedene Vorschläge. Brünken, Steinbacher & Leutner nehmen in Übereinstimmung mit [MaSi94] an, dass die Fähigkeit zur räumlichen Visualisierung von Informationen – insbesondere wenn es um die Verarbeitung multikodal präsentierter Informationen geht – einen zentralen Aspekt des räumlichen Vorstellungsvermögens darstellt. Darunter versteht man die Fähigkeit, Objekte mental in zwei oder drei Dimensionen rotieren oder falten, und sich die Lage- und Formveränderungen dieser Objekte vorstellen zu können.

Wie das räumliche Vorstellungsvermögen die Informationsverarbeitung beeinflusst, dazu gibt es zwei konkurrierende Theorien [CrSn77]. Die *ability-as-enhancer*-Hypothese geht davon aus, dass ein hohes räumliches Vorstellungsvermögen die Verarbeitung bildlich dargestellter Informationen fördert [MaSi94]. Die *ability-as-compensator*-Hypothese geht von einer kompensatorischen Wirkung aus; ein stark ausgeprägtes räumliches Vorstellungsvermögen kann fehlende piktoriale Information ersetzen.

Mayer & Sims fanden in [MaSi94] und [Mayer97] empirische Belege für einen ATI-Effekt zwischen einem instruktionalen Gestaltungsmerkmal und räumlichem Vorstellungsvermögen und interpretierten die Befunde als “Unterstützungseffekt” hohen räumlichen Vorstellungsvermögens.

1.4 Mentale Modelle

Momentan geht man davon aus, dass unterschiedliche Kodalitäten (piktorial/analog vs. verbal/propositional) in spezifischen Verarbeitungsprozessen beim Lernen mit Multimedia resultieren.

³Weitere Eigenschaften sind spezifisches Vorwissen, Erfahrung in der Nutzung von CBT-Programmen, Alter und Geschlecht, etc.

Aufbauend auf [Witt94] und [Paiv86] vertritt Mayer in [Maye97] die These eines mehrstufigen (selection, organisation, integration), kodalitätsspezifischen Verarbeitungsprozess multimedial präsentierter Informationen. Textuelle und piktoriale Informationen werden hierbei in getrennten Systemen verarbeitet und erzeugen modalitätsspezifische *mentale Modelle*. Die Verknüpfung zwischen diesen Modellen (*referential connections*) im Verarbeitungsprozess macht demnach den Vorteil multimedialer Informationsvermittlung aus.

Mentale Modelle sind interne Abbildungen der aufgenommenen Informationen, welche gegenständliche kognitive Operationen am repräsentierten Objekt ermöglichen. Auch von anderen Autoren werden diese mentalen Modelle für den effizienten Wissenserwerb als sehr wichtig angesehen.

1.5 Ziele der untersuchten Studie

Die Studie von Brünken, Steinbacher & Leutner verfolgt vor dem dargestellten Hintergrund zwei Ziele.

Es soll die Replizierbarkeit der von Mayer & Sims in [MaSi94] beobachteten ATI-Effekte untersucht werden, wenn Systeme verglichen werden, die sich ausschließlich in der Kodalität unterscheiden, inhaltlich jedoch äquivalent sind.

Darüber hinaus wird der Einfluss nicht nur der Präsentationskodalität, sondern auch der Kodalität des Informationsabrufes auf diese ATI-Effekte betrachtet.

2 Das Versuchsdesign

Dieses Kapitel vermittelt einen Überblick über die eingesetzten Versuchspersonen und das Design der durchgeführten Experimente.

2.1 Experimente und Versuchspersonen

Es wurden zwei inhaltlich identische Experimente mit unterschiedlichen Versuchspersonen durchgeführt. Die Versuchspersonen wurden für das Experiment zufällig in zwei Gruppen aufgeteilt. Zusammensetzung und wichtigste Daten der Versuchspersonen:

Experiment 1

- 40 Schüler einer gymnasialen Mittelstufe
- 18 Jungen und 22 Mädchen
- mittleres Alter 14,2 Jahre, $\sigma_n=1,38$ Jahre
- Gruppengröße $n_1 = n_2 = 20$

Experiment 2

- 42 Studenten erziehungswissenschaftlicher Studiengänge der PH Erfurt
- 8 Männer und 32 Frauen⁴
- mittleres Alter 20,98 Jahre, $\sigma_n=1,89$ Jahre
- Gruppengröße $n_1 = n_2 = 21$

Beide Experimente bestanden aus einem Vortest zur Feststellung der Ausprägung des räumlichen Vorstellungsvermögens und einem Test im $2*2*2$ -Design mit den zwei *between-subjects*-Faktoren Präsentationskodalität und räumliches Vorstellungsvermögen, sowie dem *within-subjects*-Faktor Abrufkodalität.

2.2 Vortest zum räumlichen Vorstellungsvermögen

Die üblichen Tests zur Feststellung des räumlichen Vorstellungsvermögens konzentrieren sich entsprechend der Definition in Abschnitt 1.3, Seite 4 auf die Manipulation von Objekten. In der

⁴Hier fehlen offensichtlich 2 Personen ($8 + 32 \neq 42$), dieser Fehler ist jedoch bereits in [Brü⁺00] enthalten.

Regel werden Ausgangs- und Zielkonfiguration einiger Objekte gezeigt, und die Versuchsperson muss feststellen, ob die Zielkonfiguration durch rotieren, falten oder ähnliche Operationen auf der Ausgangskonfiguration erreicht werden kann. Die im Test gezeigte Leistung ist ein Indikator für die Verfügbarkeit des mentalen Modells des manipulierten Objektes.

Brünken, Steinbacher & Leutner legten den Versuchspersonen in Übereinstimmung mit [MaSi94] zwei Subtests des *Manual for Kit of Factor-Referenced Cognitive Tests* [Eks⁺76] vor, einen *Paper-Folding-Test* und einen *Cards-Rotation-Test*.

Beide Tests waren als *paper-pencil-Tests*⁵ ausgelegt, und mit einer selbst erstellten deutschsprachigen Instruktion versehen. Die Tests bestanden aus je zwei Teilen mit insgesamt 20 Items. Im Paper-Folding-Test waren 20 korrekte Antworten möglich, im Cards-Rotation-Test 160. Die Durchführung der Tests erfolgte mit einer Zeitbeschränkung auf 2 mal 3 Minuten pro Test.

Paper-Folding-Test

In jedem Item werden auf der linken Seite des Aufgabenblattes die Ausgangslage eines Blattes Papier sowie mehrere Faltungsoperationen gezeigt. Letzte Operation ist das Stanzen des Blattes durch alle gefalteten Lagen.

Auf der rechten Seite sind fünf Zielkonfigurationen dargestellt, die die Stanzungen im Papier im aufgefalteten Zustand zeigen. Die Versuchsperson muss die richtige Zielkonfiguration identifizieren.

Cards-Rotation-Test

Auf der linken Seite jedes Items wird ein zweidimensional grafisches Objekt gezeigt, auf der rechten Seite acht Zielkonfigurationen des selben Objektes.

Für jede Zielkonfiguration hat der Proband festzustellen, ob sie durch Rotation aus der Ausgangslage entstehen kann, oder nicht.

2.3 Test des Lehrsystemes

Gegenstand des Experiments war ein multimedialer Stadtführer durch die historische Altstadt von Florenz. In getrennten Bildschirmfenstern wurden kunsthistorische Informationen zu einzelnen Sehenswürdigkeiten sowie deren topographische Lage dargestellt. Die beschriebenen Sehenswürdigkeiten waren so angeordnet, dass ihre Aneinanderreihung eine lineare Abfolge von Richtungswechseln auf dem Stadtplan bildete.

Das Lehrsystem lag in zwei nur in der Präsentationskodalität (*between-subjects*-Faktor) unterschiedlichen Versionen vor. In der Präsentationsbedingung *Grafik* (piktoriale Aufgabenko-

⁵Man sieht hier, dass die Nutzung von Anglizismen triviale Ausdrücke enorm aufwertet. Auch auf Deutsch kann man sowohl Papier falten als auch Karten rotieren, auch mit Papier und Bleistift...

dalität) wurde die topographische Information durch schematische Stadtplanausschnitte, in der Präsentationsbedingung *Text* (textuelle Aufgabenkodalität) durch verbale Beschreibungen. Für eine ausführliche Programmbeschreibung verwiesen Brünken, Steinbacher & Leutner auf den inzwischen veröffentlichten Artikel [Brü⁺01].

Mittels eines aus zwei Subtest bestehenden Tests wurde der Erwerb prozeduralen Wissens – welches den Einsatz der in Abschnitt 1.4, Seite 4 vorgestellten mentalen Modelle erfordert – überprüft.

Subtest A – verbale Aufgabenkodalität

Der erste Teilttest bestand aus 10 vierstufigen Multiple-Choice-Items. Die Aufgaben hatten die Form *Stellen Sie sich bitte vor, Sie befinden sich auf dem im CBT beschriebenen Weg am Punkt X und sehen in Richtung Y. Welches Gebäude sehen Sie als nächstes/übernächstes vor/hinter sich?* (vgl. [Brü⁺00], S. 40).

Es ist anzunehmen, dass mit dieser Formulierung ausgedrückt werden soll, dass jede Frage die Identifikation *eines* Gebäudes fordert, die Fragen jedoch im Detail variieren.

Subtest B – piktoriale Aufgabenkodalität

Der zweite Teilttest bestand aus 11 Multiple-Choice-Items. Jedes Item zeigte einen (im CBT nicht gezeigten) Stadtplanausschnitt mit drei aufeinander folgenden Sehenswürdigkeiten. Dieser Stadtplanausschnitt wurde räumlich rotiert (mit einer Schrittweite von 90°) präsentiert. Die Aufgaben des Teilttests hatten folgende Form: *Sehen Sie sich bitte den Detailplan genau an. In welcher Richtung befindet sich Norden?* Die Antworten können über vier mit der Himmelsrichtung Norden bezeichnete Schaltflächen⁶ angegeben werden.

2.4 Kriteriumsorientierte Tests

In beiden Teiltests kamen *kriteriumsorientiert konstruierte* (nach [Klau87]) Multiple-Choice-Items zum Einsatz. Der Begriff der kriteriumsorientierten Tests wurde im Jahr 1963 von Robert Glaser begründet:

Was ich kriteriumsbezogene Messung nennen werde, hängt von einem absoluten Qualitätsstandard ab, während das, was ich normbezogene Messung nenne, von einem relativen Standard abhängt. (vgl. [Glas63] und [Stri73, Seite 10],)

Er wurde von Glaser also verwendet, um den Gegensatz von kriteriumsorientierter und *normorientierter* Messung zu verdeutlichen. Der Begriff der Kriteriumsorientierung wird allerdings

⁶Nach dieser Bezeichnung handelt es sich zumindest bei Subtest B scheinbar nicht um einen *paper-pencil*-Test. [Brü⁺00] enthält hierzu keine näheren Angaben.

als nicht optimal angesehen, da der Begriff *Kriterium* verschiedene Sachverhalte bezeichnen kann:

- das Lernkriterium im Sinne eines lernpsychologischen Versuches. Dies ist ein vorgegebener Leistungsstandard, mit welchem Probanden in Können und Nichtkönnen klassifiziert werden können.
- einen Leistungsbereich der mehr oder weniger gut beherrscht wird. Der Beherrschungsgrad ist dabei ein kontinuierlich variabler Wert zwischen “überhaupt nicht” und “perfekt”.
- eine abhängige Variable im Zusammenhang mit der Validierung eines Tests, beispielsweise die schulische Leistung als Vorhersagekriterium eines Intelligenztests. In diesem Sinn spricht man von der Kriteriumsvalidität eines Tests.

Besonders wichtig ist hier, dass ein kriteriumsorientierter Test gerade nicht durch die dritte Bedeutung definiert ist, sondern sich die meisten Autoren zum Thema auf die beiden ersten Bedeutungen beziehen. Ein kriteriumsorientierter Test kann, muss aber nicht kriteriumsvalid sein.

Meist wird ein kriteriumsorientierter Test dadurch definiert, dass er nicht normorientiert ist. Gemeint ist damit, dass er sich nicht an einer Durchschnitts- oder Realnorm orientiert. Allerdings liegt auch kriteriumsorientierten Tests eine Norm zugrunde, wenn auch die *Idealnorm* des gegebenen Lehrzieles. Diese *curriculare Norm* ist eine Idealnorm, weil sie von empirischen Werten unabhängig gültig ist, und als Norm einer absoluten Messung die Positionierung auf einem festgesetzten Standard erlaubt. Dies zeigt, dass auch der Begriff *normorientiert* nur begrenzt zur (negativen) Definition dieser Art von Test herangezogen werden kann.

Entscheidend für einen kriteriumsorientierten Test ist, dass er *kontentvalide* ist, das heißt die Gesamtheit einer Menge von Aufgaben enthält oder repräsentiert. Zur Konstruktion kontentvalider Tests ist also vorher präzise eine Aufgabenmenge zu definieren.

Auf diesen Erkenntnissen aufbauend, gibt [Klau87, Seite 11] folgende Definition:

Kriteriumsorientiert ist ein Test, der die Gesamtheit einer wohldefinierten Menge von Aufgaben enthält oder repräsentiert und der zu dem Zweck konstruiert ist,

– *die Fähigkeit des Probanden zur Lösung der Aufgaben der definierten Menge zu schätzen oder/und*

– *ihn gemäß dieser Fähigkeit einer Klasse von Probanden zuzuordnen.*

Der entscheidende Punkt, der die Autoren der untersuchten Studie zur Auswahl kriteriumsorientierter Tests bewegt hat, dürfte sein, dass die gewonnenen absoluten Werte direkt Intervallkalenqualität besitzen. Dies ist für die spätere Auswertung mit Hilfe der Messwiederholungs-Varianzanalyse von Bedeutung.

3 Ergebnisse und Auswertung der Experimente

Dieses Kapitel stellt die Ergebnisse der im vorhergehenden Kapitel beschriebenen Experimente dar. Sowohl die deskriptiven, als auch die Ergebnisse der durchgeführten Messwiederholungs-Varianzanalyse werden vorgestellt, und die daraus möglichen Schlüsse erläutert.

3.1 Vortest räumliches Vorstellungsvermögen

Die Ergebnisse des Vortests dienten der Zuordnung der Versuchspersonen zur Variablen *Räumliches Vorstellungsvermögen*. Hierzu wurden die Rohwerte eines jeden Probanden in Prozentwerte transformiert, und der Mittelwert über beide Teiltests bestimmt. Mittels Mediansplit wurden die Probanden dann der Variablenausprägung “hohes räumliches Vorstellungsvermögen” ($X_i \geq Med$) oder “niedriges räumliches Vorstellungsvermögen” ($X_i < Med$) zugeordnet.

Diese Zuordnung erfolgte *ex-post* ([Brü⁺00, Seite 40]). Vermutlich ist hiermit gemeint, dass die Zuordnung der Versuchspersonen zu den Gruppen erst nach Durchführung des Haupttests, bei der Auswertung der Ergebnisse vorgenommen wurde. Über die Gründe hierfür schweigt sich die Studie aus.

Tabelle 1 stellt die deskriptiven Ergebnisse des zur Feststellung des räumlichen Vorstellungsvermögens durchgeführten Vortests in beiden Experimenten in Relation zur Variablen *Aufgabekodalität* im Test des Lehrsystems dar.

| Experiment | Aufgabekodalität | Mittlere % korrekt | σ_n |
|------------|------------------|--------------------|------------|
| 1 | textuell | 58,21% | 19,71 |
| | grafisch | 61,24% | 18,16 |
| 2 | textuell | 60,03% | 8,95 |
| | grafisch | 62,93% | 14,16 |

Tabelle 1: Die Ergebnisse des Vortests zum räumlichen Vorstellungsvermögen.

Die Unterschiede zwischen den Gruppen sind in beiden Experimenten statistisch nicht signifikant, wie die Ergebnisse der Überprüfung mittels *T-Test* zeigen (siehe Tabelle 2). Dies bedeutet, dass die im Hauptexperiment beobachteten Unterschiede zwischen den Gruppen mit unterschiedlicher Aufgabekodalität nicht allein auf Unterschieden in der Ausprägung des räumlichen Vorstellungsvermögens beruhen.

| Experiment | <i>t</i> -Wert | <i>p</i> |
|------------|-----------------|----------|
| 1 | $t(38) = 0,523$ | 0,60 |
| 2 | $t(40) = 0,794$ | 0,43 |

Tabelle 2: Signifikanzprüfung der Ergebnisse des Vortests mittels T-Test.

Der T-Test ist parametrisches Verfahren zur Signifikanzprüfung der Unterschiede zwischen

zwei (in diesem Fall unverbundenen) Stichproben. Aus den Mittelwerten der Stichproben, dem Stichprobenumfang, der Standardabweichung und der Anzahl der Freiheitsgrade wird der *t*-Wert berechnet. Zu gegebenen Untersuchungsparametern (Freiheitsgrade und Irrtumswahrscheinlichkeit α) gibt es einen dazugehörigen *kritischen Wert* t^* . Liegt t zwischen $-t^*$ und t^* , so ist der Unterschied nicht signifikant, die Nullhypothese wird nicht verworfen. Die Varianzanalyse (s.u.) ist eine Verallgemeinerung des T-Tests auf mehr als zwei Stichproben.

3.2 Haupttest Lehrsystem

Die deskriptiven Ergebnisse des Haupttests sind in Abbildung 1 dargestellt. Die Tabelle zeigt die mittlere Prozentzahl korrekt gelöster Aufgaben in den Tests (Mean), die jeweilige Standardabweichung (SD, *standard deviation*) sowie die Anzahl der Items (N) für die acht Faktorstufenkombinationen aus *Präsentationskodalität*, *Aufgabenkodalität* und *räumlichem Vorstellungsvermögen* (SA).

| Experiment 1 | | | | | |
|------------------|------|--------------|---------------------|--------------|---------------------|
| n = 40 Schüler | | CBT Text | | CBT Grafik | |
| | | Textaufgaben | Piktoriale Aufgaben | Textaufgaben | Piktoriale Aufgaben |
| SA (+) | N | 10 | 10 | 11 | 11 |
| | Mean | 39 | 31.82 | 37.27 | 67.77 |
| | SD | 21.32 | 23.57 | 20.05 | 17.89 |
| SA (-) | N | 10 | 10 | 9 | 9 |
| | Mean | 30 | 34.55 | 31.11 | 33.33 |
| | SD | 14.91 | 28.04 | 16.16 | 17.01 |
| Experiment 2 | | | | | |
| n = 42 Studenten | | CBT Text | | CBT Grafik | |
| | | Textaufgaben | Piktoriale Aufgaben | Textaufgaben | Piktoriale Aufgaben |
| SA (+) | N | 10 | 10 | 11 | 10 |
| | Mean | 63 | 66.36 | 79.09 | 86.78 |
| | SD | 20.58 | 21.45 | 22.01 | 16.94 |
| SA (-) | N | 11 | 11 | 10 | 11 |
| | Mean | 51.82 | 46.28 | 58 | 81.82 |
| | SD | 23.59 | 22.8 | 18.68 | 19.17 |

Abbildung 1: Deskriptive Ergebnisse der Tests zum prozeduralen Wissen.

Es fällt auf, dass die Gruppe der studentischen Versuchspersonen mit im Mittel 66,67% korrekt gelösten Aufgaben die Gruppe der Schüler mit im Mittel 38,81% weit hinter sich lässt. Dieser Unterschied wird in der Studie nicht weiter betrachtet⁷.

⁷Leider, zeigt er doch, dass Studieren doch einen positiven Effekt hat...

In Abschnitt 2.1 war bereits von *within-subjects*- und *between-subjects*-Variablen die Rede. Eine *within-subjects*-Variable ist eine unabhängige Variable, die an jeder Versuchsperson mehrfach, jeweils für unterschiedliche Ausprägungen getestet wird. Wird beispielsweise ein Test zur Lernleistung durchgeführt, und wird je Person ein Test mit einer und zwei Minuten Testdauer durchgeführt, so stehen nach Abschluss für jede Versuchsperson zwei Ergebnisse zur Verfügung. Die Testdauer ist damit eine *within-subjects*-Variable, da jede Versuchsperson für jede Faktorenstufe getestet wird. *Within-subjects*-Variablen werden oft als *repeated-measures*-Variablen bezeichnet.

Between-subjects-Variablen sind unabhängige Variablen, die je Faktorenstufe mit einer eigenen Gruppe von Versuchspersonen getestet werden. Werden z.B. vier Lehrmethoden verglichen, und jede Lehrmethode wird an einer eigenen Gruppe von Versuchspersonen getestet, so ist die Lehrmethode eine *between-subjects*-Variable. Sind alle Variablen eines Experimentes *between-subjects*-Variablen, so spricht man von einem *between-subjects-Design*.

In der vorliegenden Studie ist die Aufgabenkodalität eine *within-subjects*-Variable. Jede Versuchsperson führte zwei (inhaltlich unterschiedliche) Tests zum prozeduralen Wissen durch, einmal mit Textaufgaben, ein zweites Mal mit grafischen Aufgaben. Dadurch wird sowohl eine bessere Fehlerstreuung erreicht, als auch eine höhere Anzahl an Ergebnissen gewonnen⁸.

Messwiederholungs-Varianzanalyse

Die statistische Signifikanz der Unterschiede zwischen den einzelnen Gruppen erfolgte mit der Messwiederholungs-Varianzanalyse (RM-ANOVA). Varianzanalysen sind eine Kategorie von parametrischen statistischen Analyseverfahren zum Mittelwertsvergleich von zwei oder mehr Stichproben. Festgestellt werden soll die Signifikanz der Wirkung einer oder mehrerer unabhängiger Variablen (Faktoren) auf eine oder mehrere abhängige Variablen. Kommen in einer Untersuchung mehrere unabhängige Variablen zum Einsatz, so liefert die Varianzanalyse nicht nur für die einzelnen Variablen einen Signifikanztest, sondern auch für die Wechselwirkung *interaction* der Variablen untereinander.

Die unterschiedlichen Ausprägungen der Faktoren werden als Faktorstufen bezeichnet, diese Bezeichnung wurde im vorliegenden Text auch bereits verwendet. Die Faktoren sind häufig nominal skaliert. Während das Skalenniveau der Faktoren jedoch beliebig sein kann, sind die abhängigen Variablen immer intervallskaliert. In der vorliegenden Studie handelt es sich um eine mehrfaktorielle univariate Varianzanalyse, da mehrere Faktoren, jedoch nur eine abhängige Variable vorliegen (vgl. [Glas78] und [Sieb02]).

Bei der Messwiederholungs-Varianzanalyse werden die zu untersuchenden Faktorstufen (sofern es sich um eine *within-subjects*-Variable handelt) wiederholt an jeder Versuchsperson er-

⁸Ohne noch mehr der ohnehin knappen Probanden zu verschleifen.

hoben (siehe auch Seite 12). Neben vielen Problemen, die man sich mit der Messwiederholung einhandelt (Sequenzeffekte, Lerneffekte, ...), gibt es einige Vorteile:

- Da die Faktorstufen innerhalb der Person variieren, spart man Versuchspersonen.
- Man kann die Variation zwischen Personen von der Fehlervariation trennen, und erhält so unter Umständen eine erhebliche Reduktion der Fehlervarianz.

Zur Analyse werden die Ergebnisse der Messwiederholungen in der Regel zu einem Mittelwert zusammengefasst, sofern nicht die Effekte der Wiederholung selbst von Interesse sind. Nach einer solchen Behandlung der Ergebnisse entsteht für die so behandelten abhängigen Variablen ein Design mit lediglich unabhängigen Faktoren, auf welches nun eine Varianzanalyse ohne Berücksichtigung der Messwiederholung angewendet werden kann.

Die Voraussetzungen für die Anwendung einer Messwiederholungs-Varianzanalyse entsprechen weitgehend denen der Varianzanalyse *ohne* Messwiederholung:

- Intervallskalen-Qualität der abhängigen Variablen
- wenige Extremwerte und keine Ausreisser in den Messwerten
- Normalverteilung der Ergebnisse in den Gruppen
- Varianzhomogenität

Es kommt als weitere Voraussetzung die sogenannte Zirkularitätsannahme (auch Sphäritätsannahme) hinzu. Diese besagt, daß die Varianz der Differenzen von je zwei Messzeitpunkten (in der Population) identisch sein müssen.

Gegen eine Verletzung der Zirkularitätsannahme ist die Messwiederholungs-Varianzanalyse nicht robust. Verletzungen führen zu artifiziell niedrigen p -Werten (Überschreitungswahrscheinlichkeiten). Die erforderliche Korrektur der Freiheitsgrade (z.B. Greenhouse-Geisser oder Huynh-Feldt) wird in publizierten Arbeiten nicht immer durchgeführt.

Ergebnisse der Messwiederholungs-Varianzanalyse

Die Analyse der Ergebnisse hinsichtlich der Signifikanz mit der Messwiederholungs-Varianzanalyse ergab die in Tabelle 3 dargestellten Ergebnisse. Für einen schnellen Überblick und zum Erkennen der Interaktionseffekte ist eine Darstellung wie in Abbildung 2 und Abbildung 3 besser geeignet. Man erkennt hier deutlich, wenn sich die Mittelwertskurven überschneiden. In einem solchen Fall spricht man (Signifikanz vorausgesetzt) von einer *disordinalen Interaktion*. Überschneiden sich die Kurven *nicht* so spricht man (wiederum Signifikanz vorausgesetzt) von einer *ordinalen Interaktion*.

| Faktor bzw. Faktorenkombination | Experiment 1 | | Experiment 2 | |
|--|--------------|-------|--------------|-------|
| | F(1, 36) | p | F(1, 38) | p |
| Aufgabenkodalität | 4,26 | 0,046 | 7,522 | 0,009 |
| Aufgabenkodalität * Präsentationskodalität | 5,884 | 0,02 | 9,916 | 0,003 |
| räumliches Vorstellungsvermögen * Aufgabenkodalität * Präsentationskodalität | 7,532 | 0,009 | 5,479 | 0,025 |

Tabelle 3: Ergebnisse der Varianzanalyse.

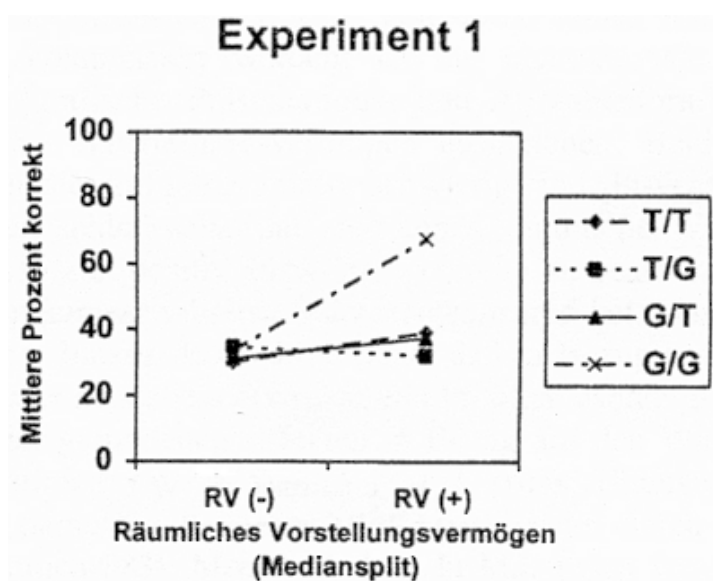


Abbildung 2: Grafische Darstellung der Ergebnisse aus Experiment 1.

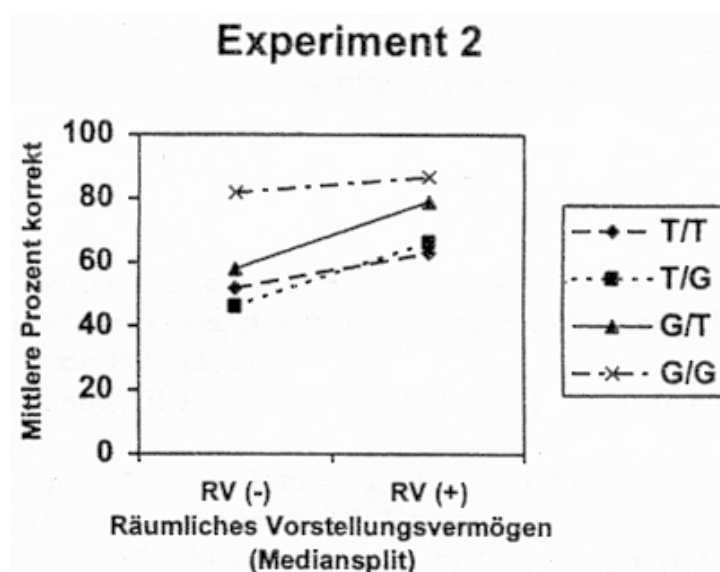


Abbildung 3: Grafische Darstellung der Ergebnisse aus Experiment 2.

Es ist deutlich erkennbar, dass die piktorial kodierten Aufgaben besser gelöst wurden, als die textuell kodierten Aufgaben. Dieser Effekt wird durch die piktoriale Präsentation aufgabenrelevanter Informationen signifikant verstärkt. In diesem Fall (sowohl Informationen als auch Aufgaben piktorial kodiert) zeigt ein stark ausgeprägtes räumliches Vorstellungsvermögen ebenfalls den Effekt einer signifikant besseren Lösung der Aufgaben. Dieser Effekt ist in Experiment 1 deutlich stärker als in Experiment 2.

Insgesamt legen die Ergebnisse also den Schluß nahe, daß sich hohes räumliches Vorstellungsvermögen insbesondere dann positiv auswirkt, wenn die zu bearbeitende Aufgabe für den Lerner einen subjektiv hohen Schwierigkeitsgrad hat. [Brü⁺00, Seite 45]

4 Diskussion der Vorgehensweise

Dieser Abschnitt soll darlegen, ob die Vorgehensweise der Autoren der untersuchten Studie den Anforderungen an eine solche Arbeit genügt. Neben eher nebensächlichen Dingen wie der Tatsache, dass eine Erklärung zur Anzahl der Items in den Subtests (10 bzw. 11 Items) fehlt, oder die Mischung von deutschen Bezeichnungen und teils englischen Abkürzungen im Text und den Tabellen bzw. Grafiken unglücklich ist, gibt es einige weitaus wichtigere Dinge in dieser Hinsicht zu betrachten.

Die Versuchspersonen

Die Gruppengrößen von 40 bzw. 42 Personen werden durch die Anwendung der Messwiederholungs-Varianzanalyse auf 80 beziehungsweise 84 Versuchspersonen “vergrößert”. Dies ist wohl ein ausreichender Stichprobenumfang, insbesondere wenn man die üblichen Stichprobengrößen im universitären Umfeld in Betracht zieht. Die Zusammensetzung der Gruppe der studentischen Versuchspersonen ist in Bezug auf die Verteilung der Geschlechter nicht ideal. Auf eventuelle Effekte dieser Ungleichverteilung wird nicht eingegangen.

Der Testaufbau

Die Nutzung kriteriumsorientiert konstruierter Items ist sinnvoll, da es darum geht, die Einflüsse der unterschiedlichen Modalität der Informationsvermittlung zu messen. Hier ist es eher von untergeordneter Bedeutung, eine Leistung im Vergleich zu den anderen Versuchspersonen zu beurteilen, die Einordnung auf einer idealen Skala (basierend auf den maximal möglichen Punkten) scheint sinnvoll.

Die Studie hatte das Ziel, die in vorliegenden Untersuchungen vermisste inhaltliche Äquivalenz des untersuchten Lehrsystems sicherzustellen. Dies ist nach der Beschreibung des CBT-Programmes geglückt. Für eine genaue Beschreibung des eingesetzten CBT-Programmes wird jedoch auf ein (zum Zeitpunkt des Erscheinens) noch nicht veröffentlichtes Manuskript verwiesen⁹.

Die Fragen zur Überprüfung des gewonnenen prozeduralen Wissens sind inhaltlich jedoch nicht äquivalent. Die Frage nach der Himmelsrichtung und die Frage nach weiteren Sehenswürdigkeiten scheinen nicht den gleichen Schwierigkeitsgrad zu haben, bzw. kontrollieren nicht den gleichen Lernprozess.

⁹Der in [Brü⁺00] genannte Artikel erschien als [Brü⁺01] erst 2001.

Voraussetzungen zur Anwendung der Varianzanalyse

Zwei der wichtigsten Voraussetzungen zur Anwendung der Messwiederholungs-Varianzanalyse sind erfüllt:

- Intervallskalen-Qualität der abhängigen Variablen: die Ergebnisse liegen als Prozentwerte vor.
- Varianzhomogenität: die Varianzen in den Gruppen sind recht ähnlich.

Ob Extremwerte oder Ausreisser in den Messwerten vorliegen, kann nicht beurteilt werden, da keine Angaben hierzu (beispielsweise ein Boxplot) vorliegen. Auch über die Normalverteilung der Ergebnisse in den Gruppen werden keine Angaben gemacht. Ebenso fehlen Angaben zur Zirkularitätsannahme, beziehungsweise zu eventuell durchgeführten Korrekturen in diesem Bereich.

5 Fazit

Brünken, Steinbacher & Leutner ziehen in ihrem Text eigene Schlüsse aus den gewonnenen Ergebnissen, die jedoch eher fachspezifische Bedeutung haben. Für die vorliegende Arbeit war die Zielsetzung eine andere; es galt formelle Fehler und andere Ungereimtheiten aufzuspüren.

Nach den Ausführungen im letzten Kapitel bin ich der Meinung, die Autoren haben eine recht solide Arbeit vorgelegt. Teilweise fehlen Angaben, besonders im Zusammenhang mit den durchgeführten statistischen Analysen (beispielsweise zur Qualität der Daten aus dem Vortest oder zu den Voraussetzungen der Varianzanalyse). Auch die Zusammensetzung der Gruppe der studentischen Versuchspersonen ist in der Geschlechterverteilung nicht optimal, auf eventuell daraus resultierende Effekte gehen die Autoren nicht ein.

Die Zielsetzung der inhaltlichen Äquivalenz der untersuchten Systeme scheint geglückt, auch wenn für eine tiefergehende Beurteilung auf eine zum Veröffentlichungszeitpunkt der Studie noch nicht erschienene Publikation verwiesen wurde. Die Anzahl der Versuchspersonen ist im Vergleich mit anderen Studien aus dem universitären Umfeld recht hoch, darüber hinaus verdoppelt die Untersuchung mit Hilfe der Messwiederholungs-Varianzanalyse den Umfang der Messdaten.

Immerhin zwei der wichtigsten Voraussetzungen der Varianzanalyse (Intervallskalen-Qualität der abhängigen Variablen, Varianzhomogenität) sind erfüllt, zu weiteren Voraussetzungen schweigt die Studie sich leider aus.

Kleinere Ärgernisse, wie die blinde Übernahme von Abkürzungen aus der verwendeten Statistik-Software und die daraus resultierenden Unklarheiten in der Beschriftung der Tabellen und Abbildungen, trüben zwar den Gesamteindruck, ein stärker an den fachlichen Aspekten der Arbeit interessierter Leser mag dies jedoch weniger stark wahrnehmen.

Bleibt mir zu hoffen, dass ich bei späteren eigenen Arbeiten im Feld der Statistik diese Dinge beachte und eine bessere Arbeit abliefere.

Literaturverzeichnis

- [Brü⁺00] BRÜNKEN, R.; STEINBACHER, S.; LEUTNER, D.: *Räumliches Vorstellungsvermögen und Lernen mit Multimedia*. In: LEUTNER, D.; BRÜNKEN, R. (Hrsg.): *Neue Medien in Unterricht, Aus- und Weiterbildung. Aktuelle Ergebnisse empirischer pädagogischer Forschung*. Münster: Waxmann, 2000, S. 37–46
- [Brü⁺01] BRÜNKEN, R.; STEINBACHER, S.; SCHNOTZ, W.; LEUTNER, D.: Mentale Modelle und Effekte der Präsentations- und Abrufkodalität beim Lernen mit Multimedia. In: *Zeitschrift für Pädagogische Psychologie*, (2001) 15, S. 16–27
- [CrSn77] CRONBACH, L. J.; SNOW, R. E.: *Aptitudes and instructional methods*. New York: Irvington, 1977
- [Damb98] DAMBEKALNS, K.: *Evaluation von Multimedia*. 1998. – Referat im Rahmen des Seminars "Ausgewählte Kapitel der Medienpädagogik", WS 1997/98
- [Eks⁺76] EKSTROM, R. B.; FRENCH, J. W.; HARMAN, H. H.: *Manual for Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service, 1976
- [Glas63] GLASER, R.: Instructional technology and the measurement of learning outcomes: Some questions. In: *American Psychologist*, (1963) 18, S. 519–521
- [Glas78] GLASER, W. R.: *Varianzanalyse*. Stuttgart, New York: Gustav Fischer Verlag, 1978
- [Gros02] GROSS, R.: *U-Test und T-Test*. 2002. – Referat im Rahmen des Seminars "Medienwissenschaften goes Statistics", SS 2002
- [Klau87] KLAUER, K. J.: *Kriteriumsorientierte Tests*. Göttingen: Verlag für Psychologie Dr. C. J. Hogrefe, 1987
- [Lane] LANE, D. M.: *Hyperstat online: An Introductory Statistics Book and Online Tutorial for Help in Statistics Courses*. – <http://davidmlane.com/hyperstat/>
- [MaSi94] MAYER, R. E.; SIMS, V. K.: For whom is a picture worth ten thousand words? Extensions of a dual-coding theory of multimedia learning. In: *Journal of Educational Psychology*, (1994) 86, S. 389–401
- [Maye97] MAYER, R. E.: Multimedia learning: Are we asking the right questions? In: *Educational Psychologist*, (1997) 32, S. 1–19
- [Paiv86] PAIVIO, A.: *Mental representations: a dual-coding approach*. New York: Oxford University Press, 1986

- [Sieb02] SIEBERT, K.: *ANOVA - Analysis of Variance*. 2002. – Referat im Rahmen des Seminars "Medienwissenschaften goes Statistics", SS 2002
- [Stri73] **STRITTMATTER, P. (Hrsg.)** : *Lernzielorientierte Leistungsmessung*. Weinheim: Beltz, 1973
- [Witt94] WITTROCK, M. C.: Learning as a generative process. In: *Educational Psychologist*, (1994) 11, S. 87–95

Index

ability-as-compensator, 4
ability-as-enhancer, 4
Aptitude-Effekt, 4
Aptitude-Treatment-Interaktion, 4

between-subjects, 6, 11, 12

curriculare Norm, 9

deskriptiv, 10
disordinalen Interaktion, 13
dual coding effect, 3

Kodalität, 3
kontentvalide, 9
kriteriumsorientierte Tests, 8
kritischer Wert, 11

mentale Modelle, 4, 8
Messwiederholungs-Varianzanalyse, 10, 12
Multimedia, 3
multimedia dual coding effect, 3
Multiple-Choice, 8

normorientierte Tests, 8

ordinalen Interaktion, 13

parametrisches Verfahren, 10

Räumliches Vorstellungsvermögen, 4
RM-ANOVA, 12

spatial ability, 4
Sphärizitätsannahme, 13

T-Test, 10
Treatment-Effekt, 4

Varianzanalyse, 11

Versuchspersonen, 6

within-subjects, 6, 11

Zirkularitätsannahme, 13