

Nutzungsunterschiede bei verschiedenen Notizfunktionen  
in einem E-Learning-System

Diplomarbeit  
in der Fachrichtung Psychologie  
der Universität des Saarlandes  
vorgelegt von  
Björn Mittelsdorf  
(Betreuer: Prof. Dr. Werner Tack)  
(Betreuer: Dipl. psych. Joachim Wutke)  
Saarbrücken 2004

**Inhaltsverzeichnis**

1 Einleitung	6
2 Theorien und Modelle	9
2.1 Grundlegende Theorien	9
2.1.1 Felder der Ergonomie	9
2.1.2 Geistige Anforderungen der Arbeit	11
2.1.3 Eigenschaften eines eLearning-Systems	14
2.1.4 Medienforschung und Mediendidaktik	14
2.1.5 Methoden der Software-Evaluation	16
2.1.6 Mediendidaktik und allgemeine Didaktik	17
2.2 Grundlagen des Experimentes	18
2.2.1 Forschungsrichtungen der Mediendidaktik	18
2.2.2 Modelle der Textverarbeitung	19
2.2.3 Wissenserwerb und Metastrukturen	20
2.2.4 Nutzen der Notizerstellung nach Ciesinger, Muthig & Pikara	21
2.2.5 Nutzen der Notizerstellung nach Kiewra	23
2.2.6 Kritik der Modelle des Notiznutzens	23
2.2.7 Das Electronic Learner Satisfaction Instrument	24
2.2.8 Modell der Notizerstellung von Ciesinger & Piekara	25
2.2.9 Grenzen des Modells	29
2.2.10 Zusammenfassung der Forschungshypothesen	31

3 Methoden und Ergebnisse	31
3.1 Beschreibung der Entwicklung des Experimentes	31
3.2 Formulierung der statistischen Hypothesen	40
3.3 Durchführung des Experiments	40
3.4 Deskriptive Statistik	42
3.5 Prüfung der statistischen Hypothesen	44
3.6 Kritik der empirischen Vorgehensweise	45
4 Ausblick	48
5 Literatur	50
6 CD mit Anhängen	54

**Tabellenverzeichnis**

Tab. 1: Vier-Felder-Schema mediendidaktischer Fragestellungen	18
Tab. 2: Zugänge der Lehr/Lern-Forschung	19
Tab. 3: Balancierungstabelle	43

**Abbildungsverzeichnis**

Abb. 1: Bedeutungsbereiche der verschiedenen Medientypen nach Kerres	8
Abb. 2: Modell der Notizerstellung nach Ciesinger & Piekara	28
Abb. 3: Screenshot des Notizzustandes	34
Abb. 4: Visualisierung der Handlungsschleife	35
Abb. 5: Screenshot des Lernsystems	37

## **1 Einleitung**

Es wird allgemein davon ausgegangen, daß menschliche Informationsverarbeitung instrumentell und zielgerichtet ist. Desweiteren werden die zur Verfügung stehenden Ressourcen effizient genutzt. Doch wie soll eine solche Nutzung aussehen, und woran soll man sie erkennen?

Besonders in Zeiten, in denen neue Medien neuen Verwendungen zugeführt werden, wie zum Beispiel heutzutage der Computer, werden diese Fragen gestellt. Es gibt zum Beispiel Theorien die annehmen, daß der menschliche Geist eine Information besser enkodieren kann, wenn sie in verschiedenen Sinnesmodalitäten zur Verfügung gestellt wird (vgl. Hacker, 1994).

Das programmierte Lernen (Skinner, 1954) dagegen fordert ständige Reproduktionskontrollen, um die richtigen Antworten zu verstärken.

An dieser Stelle soll allerdings die Frage im Vordergrund stehen wie der autonome Wissenserwerb zu verbessern ist.

Was ist die Basis für einen erfolgreichen Informationserwerb?

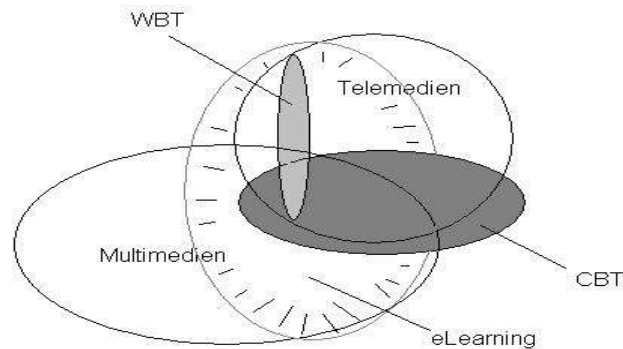
Gibt es hilfreiche Werkzeuge, und was passiert, wenn die Hilfsmittel keine Effizienzsteigerung bewirken?

Zwei in ihrem Verhalten, was die Positionierung selbstangefertigter Notizen angeht, unterschiedliche Werkzeuge sollen empirisch miteinander verglichen werden.

Als erstes wollen wir verschiedene Lernformen, die mit einem Computer realisiert werden können, aufzeigen (vgl. Kerres, 2001):

- Multimedia bezieht sich auf jegliches technische System, das verschiedene Datentypen (z.B. Text, Grafik, Ton und Bewegtbild) verarbeiten oder darstellen kann.
- Telemedien sind Techniken des Informationsaustausches, die genutzt werden können um räumliche und zeitliche Distanz zwischen Sender und Empfänger zu überbrücken. Man spricht in diesem Zusammenhang auch von asynchroner Kommunikation (Winterhoff-Spurk, 2004). Die Daten können, müssen aber nicht, multimedial sein.
- CBT (engl.: Computer Based Training) umfaßt alle Anwendungen des Computers in einer Lernsituation.
- WBT (engl.: Web Based Training) bezieht sich auf die Nutzung des WWW-Layers, sprich eines Teiles der Dienste, die für uns das Internet konstituieren, für die Kommunikation und Interaktion in einer Lernsituation.
- eLearning ist der am wenigsten trennscharfe Begriff, denn er beschreibt jegliches System, das WWW-Protokolle für seine Lehraufgabe einsetzt, sei dies E-Mail, Voice over IP oder auch ein Multi-User-Dungeon.

Abbildung 1: Bedeutungsbereiche der verschiedenen Medientypen nach Kerres (2001), S.14; die Definitionsfläche des eLearning wurde vom Autoren ergänzt



Diese Begriffe sind, wie schon gesagt, nicht als Gegensätze zu sehen, sondern als sich überlappende Eigenschaften, die der Systembeschreibung dienen.

Das System, das in dieser Arbeit benutzt wird, ist demnach ein eLearning-System, denn es verwendet das HyperTextTransferProtokoll und den HyperTextMarkupLanguage um den Lerntext für den Nutzer aufzubereiten.

Dies beinhaltet einerseits die Formatierung der Inhalte, andererseits die Implementierung von Lernwerkzeugen (siehe Kapitel 3).

Das System ist allerdings nicht multimedial, denn es enthält nur textuelle Information, es transportiert diese durch Zeit und Raum, folglich ist es telemedial. Die Frage nach CBT und WBT ist davon abhängig, ob die Implementation auf ein Inselsystem beschränkt ist, oder wir es mit einem tatsächlichen Lernnetzwerk zu tun haben. Beides wäre in diesem Fall einfach implementierbar.

Es ist auch nicht die Frage, wie wertvoll verschiedene Medien für die Lehre im Allgemeinen sind, sondern wie ein spezielles Bildungsanliegen wertvoll



und lebensfähig (im Engl.: viable) umgesetzt werden kann ( vgl. MORRIS et all., 1994).

Dieser zentralen Frage der Mediendidaktik soll nachgegangen werden, indem zwei unterschiedliche Implementierungen eines für die meisten CBT-Systeme relevanten Lernwerkzeugs miteinander verglichen werden. Das Hilfsmittel, das experimentell variiert wird, ist die Art der Repräsentation von Notizen bei einer Lernaufgabe. Der Unterschied zwischen den Prototypen, der auch didaktisch relevant ist, bezieht sich auf die Positionierung der Notizen

- entweder im zu lernenden Text oder
- in einer rechts vom Text befindlichen Seitenspalte.

## **2 Theorien und Modelle**

### **2.1 Grundlegende Theorien**

#### **2.1.1 Felder der Ergonomie**

Ziel der Software-Ergonomie ist nach Hacker (1994) die optimale Funktionsverteilung zwischen Mensch und Computer:

1. Hohe Effektivität
2. Optimierung der psychonervalen Beanspruchung
3. Vorbeugung gegen langfristige Beeinträchtigungen
4. Beiträge zur Persönlichkeitsentwicklung

(Dreifaches Gestaltungsziel der Arbeit: Nach Hacker, S.53, 3. und 4. im Original eine Kategorie)

Seiner Meinung nach ist die Gestaltung einer Software die Gestaltung eines Arbeitsablaufs, und folglich ist der Software-Gestalter ein Arbeitsgestalter.

Er wendet sich gegen eine Reduktion auf Schnittstellengestaltung, vielmehr müssen aus Sicht der Arbeitspsychologie auch Arbeitsinhalte und Arbeitsabläufe vom Entwickler berücksichtigt werden.

Hacker geht davon aus, daß, wirtschaftlich gesehen, die Nachbesserung einer Software meist teurer als die genaue Arbeitsanalyse ist. Leider nimmt er auch eine generelle Geringschätzung an, was die Rationalisierung geistiger Arbeiten angeht.

Die Optimierung körperlicher Arbeiten überwiegt, obwohl auch bei der geistigen Arbeit ein großes Algorithmierungspotential existiert. Im Regelfall werden Arbeitsprozesse so aufgeteilt, daß ein Dialog entsteht.

Hacker kritisiert die in diesem Dialog häufig passive Stellung des Nutzers, welcher nur gefragt wird, wenn der Computer eine ausführungsrelevante Information benötigt.

Die geistigen Anteile einer Arbeit können aufgeteilt werden in:

- Informationsübertragung: Wahrnehmung und kurzes Behalten der Information mit anschließender Wiedergabe in einem in der Regel anderen Format (Beispiel: Überführung von Daten aus einem Schema in eine andere Form)
- Informationsverarbeitung: Zuordnen und Auswählen von Informationen aus einer endlichen Menge nach bekannten Vorschriften. (Beispiel: Recherche einer Biographie)
- Informationsbearbeitung: Informationsverarbeitung mit zusätzlichen regelbasierten Verknüpfungen von Informationen und der Beurteilung nach Rahmenregeln (Beispiel: Erstellen einer Diplomarbeit)

- Informationserarbeitung: Zusätzlich zur Informationsbearbeitung müssen Informationen ohne vorgegebene Regeln generiert werden. (Beispiel: Schreiben eines Gedichts)

### **2.1.2 Geistige Anforderungen der Arbeit**

Hacker (1994) beschreibt die Anforderungen einer geistigen Tätigkeit als Denk- und Behaltensanforderungen.

Die Behaltensanforderungen teilen sich auf in:

- a) Aktivierung von Vorkenntnissen
- b) Komplexitätsgrad der neuen Information
- c) Abstraktionsgrad der neuen Information
- d) Operativitätsgrad

Die Reproduktion der Vorkenntnisse bewegt sich auf einem Kontinuum der mechanistischen Reproduktion einerseits und der kreativen Neuformulierung des Vorwissens andererseits, wodurch die Aktivierung bestehender Strukturen stufenlos in die Verarbeitung neuer Strukturen übergehen kann.

Der Komplexitätsgrad beschreibt das Ausmaß der Informationsverdichtung, allerdings scheint es keine optimale Dichte zu geben. Es ist vielmehr wichtig bei verschiedenen Tätigkeiten zwischen verschiedenen Auflösungsgraden wechseln zu können, und auch Informationen von einer Komplexitätsebene auf eine andere zu transportieren.

Der Abstraktionsgrad wird in der Arbeitspsychologie in Operationsklassen unterteilt, siehe hierzu auch Hacker, Rudolph & Schönfelder (1987). Hacker unterscheidet:

- materiell-gegenständliche Operationen

- Operationen im Abbildbereich ohne begriffliche Fassung
- Operationen mit konkreten Begriffen
- Operationen mit abstrakten Begriffen

Generell läßt sich sagen, daß eine abstraktere Information, als Schema gesehen, mehr Leerstellen im Sinne von möglicher Varianz aufweist. Diese Leerstellen erzeugen jedoch nicht nur hinderliche Mehrdeutigkeiten sondern bieten auch die Möglichkeit der schöpferischen Auseinandersetzung und Anwendung auf andere Themenbereiche.

Der Operativitätsgrad schließlich beschreibt die Passung zwischen Information und erforderlicher Handlung. Einerseits kann es hier zu Spannungen zwischen der Datenbasis und dem zu erzeugenden Produkt kommen, wenn zum Beispiel Daten neu eingegeben werden sollen und die Eingabemaske nicht dem Format der Originalinformationen entspricht.

Andererseits steigt die Beanspruchung des Nutzers, wenn das System Informationen anders repräsentiert als er es für seine Arbeit benötigt. Die Vermeidung solcher Umkodierzwänge wird unter dem Begriff der Kompatibilität diskutiert.

Die Kompatibilität ist laut Hacker wiederum ein wichtiger Baustein der Entlastung des Arbeitsgedächtnisses. Seiner Meinung nach ist diese Entlastung unabdingbar um eine Auseinandersetzung des Arbeiters mit seiner Aufgabe zu ermöglichen und Überlastung zu verhindern.

Um qualitative und quantitative Über- und Unterforderung zu vermeiden fordert Hacker, daß Arbeitsgänge nicht wie im Taylorismus maximal vereinfacht werden, sondern daß vielmehr vollständige und vielschichtige

Aufgaben mit sowohl fordernden als auch einfachen Anteilen gestaltet werden sollen.

Diese sollen dem Arbeiter eine holistische Repräsentation seiner Aufgabe ermöglichen, wodurch erst die kreative Auseinandersetzung mit dem Arbeitsprozeß selbst möglich wird.

Ein solcher Anspruch, der auch dem dreifachen Gestaltungsziel der Arbeit entspricht, mag im Sekundenbereich dem Taylorismus unterlegen sein, dafür fallen aber auch keine psychonervalen Erschöpfungszustände im Nutzer an, wodurch Verluste im Minutenbereich und -bei Betrachtung der Fehltag-Problematis- sogar im Tagesbereich vermieden werden können.

Es könnte schwierig sein diese Forderungen mit dem Konzept des programmierten Unterrichts im klassischen Sinne zu vereinbaren (siehe hierzu auch Blanckertz, 1986).

Der Dialog sollte niemals die einzige Aufgabe des Nutzers sein, sondern eingebettet in wahrnehmbare Arbeitsabläufe, so daß eine Aufteilung entsprechend der Stärken, Schwächen und Bedürfnisse von Mensch und Maschine gewährleistet ist.

Aus Sicht der medientechnischen Bildung ist der Lernaufwand für einen Arbeitsgang möglichst gering zu halten. Auch hier gilt, daß eine anschauliche, den für den Nutzer relevanten Prozeß begleitende Software, besser repräsentiert werden kann.

Dies sollte das weitverbreitete Auswendiglernen von Anweisungsschrittfolgen überflüssig machen.

### **2.1.3 Eigenschaften eines eLearningsystems**

Insgesamt lässt sich ein eLearningsystem, wie es in dieser Arbeit vorkommt, software-ergonomisch beschreiben als:

- Variierend komplex mit vielschichtigen Darstellungsmöglichkeiten
- Kreative Verarbeitung von Vorkenntnissen
- Informationsverarbeitung und Informationserarbeitung
- Repräsentation von Operationen mit abstrakten Begriffen
- Tätigkeit mit hohem Kompatibilitätsbedarf

### **2.1.4 Medienforschung und Mediendidaktik**

CLARK (1983) kritisiert das ein großer Teil der Medienforschung sich als Medienvergleichsforschung versteht. Medientypen, die verschiedene Eigenschaften aufweisen (u.a. Neuheit), werden miteinander verglichen, wodurch Unterschiede in den Ergebnissen nicht auf die Darstellung des gleichen Inhalts in verschiedenen Systemen, sondern auf elementare Eigenschaftsunterschiede in den Darbietungsmedien zurückzuführen sind.

Der unreflektierte Einsatz eines wie oben erläuterten Klassifikationsschemas der geistigen Arbeit kann dann zu Fehleinschätzungen der Homogenität der Aufgabenstellung in den Versuchsbedingungen führen.

Der Begriff des Bildungsmediums an sich unterliegt schon einer Unschärfe, beschreibt er doch zum Einen den sensorischen Informationsweg als auch technische Mittel der Wiedergabe, Übermittlung usw., zum Anderen aber die didaktisch aufbereitete Information.

Mit einer Einteilung von König & Riedel (1979) kann diese Doppeldeutigkeit umgangen werden:

- Lehrinhalte und -ziele die vermittelt werden sollen,
- (meist) kognitive Operationen des Lernenden
- Operationsobjekte (z.B. Lernaufgaben, Exkursionen)
- Hilfsmittel, d.h. Werkzeuge und Techniken

Dieser letzte Punkt darf jedoch nicht als Distributionsmedium verstanden werden, sondern bezieht sich auf Methoden des didaktischen Prozesses. Die Frage nach dem Lehrmedium bleibt von dieser Gliederung unberührt.

Der Wissenserwerb in einem eLearningsystem ist abhängig von den Lernzielen die pädagogisch und kulturell gesetzt werden. Wissensreproduktion benötigt mindestens Textbearbeitung wird aber in einem humanistischen System eher im Bereich der Texterarbeitung angesiedelt sein. Zum Beispiel können bei der Lektüre eines Buches über analytische Philosophie Lernregeln angewandt werden, um sich eine Aufzählung der geschilderten Modelle anzufertigen.

Für das Verständnis der verschiedenen Problemebenen helfen jedoch keine Regeln, vor allem nicht, wenn man bedenkt, daß der Lernende zu diesem Zeitpunkt eventuell nicht einmal weiß, wofür er dieses Wissen später einmal einsetzen möchte.

Der Interaktion von kognitiven Operationen und Werkzeugen soll nachgegangen werden, indem zwei unterschiedliche Implementierungen eines Lernwerkzeugs miteinander verglichen werden. Denn ein funktional beschriebenes Werkzeug kann immer wieder implementiert werden, das

Design eines Systems hingegen ist in der Regel einmalig und nicht replizierbar.

### **2.1.5 Methoden der Software-Evaluation**

Allgemein werden verschiedene Methoden der Evaluation genutzt (nach Oppermann & Reiterer, 1994):

- Subjektive Evaluationsmethoden (Interview, Fragebogen, lautes Denken)
- Objektive Evaluationsmethoden (Logfilerecording)
- Leitfadenorientierte Evaluationsmethoden (Expertenurteil anhand von Prüfkriterien)
- Experimentelle Evaluationsmethoden (Prüfung von Hypothesen, die aus Theorien abgeleitet werden, und Kausalbeziehungen aufdecken können)

In ihrer Arbeit schlagen sie auch vor, verschiedene Evaluationsmethoden je nach Fragestellung zu kombinieren. Hierbei sind die experimentellen Methoden ausgenommen, da diese nach Meinung der Autoren in der Praxis einen zu großen Aufwand darstellen.

Allerdings sind nur durch Experimente kreative Erweiterung eines Themas im Gegensatz zu einer Verwaltung von Mißständen möglich, denn die Produktion wird sich, solange sie nicht eines Besseren belehrt wird, immer mit dem Althergebrachten begnügen.

Außerdem ist bei Existenz eines theoretischen Modells, welches das Evaluationsobjekt beschreibt, das von Oppermann & Reiterer geschilderte Hauptproblem, nämlich die unendliche Zahl von möglichen unabhängigen und abhängigen Variablen, nicht gegeben.



Ein kognitives Modell beinhaltet eine endliche Zahl von relevanten Variablen. Diese Zahl kann sich zwar durch Revisionen des Modells verändern, wobei eine solche Revision genau wegen der Fundierung durch experimentelle Daten nötig und wissensmehrend sein kann.

Die Menge der für eine Theorie relevanten Größen ist aber zu einem bestimmten Evaluationszeitpunkt immer endlich.

Ein weiterer Punkt, der bedacht werden sollte, ist, daß die Kosten für Nachbesserungen in einem bereits ausgelieferten Programm im Regelfall höher sind als die einer sorgfältigen Entwicklung.

#### **2.1.6 Mediendidaktik und allgemeine Didaktik**

Die Mediendidaktik stellt die Frage nach dem Bildungspotential eines Mediums in Interaktion mit einem bildungsspezifischen Problem. Lernen, insbesondere die Autodidaktik jeglicher Art, wird zwar seit Jahrtausenden angewandt, aber nur selten didaktisch untersucht. Wenn überhaupt stehen sozio-kulturelle Faktoren des Lernens -meist im Zusammenhang mit der Lernsituation in einem Klassenverband- im Mittelpunkt (SEEL, 2000 oder auch Schank, 1995).

Die allgemeine Didaktik auf eine Didaktik des Schulunterrichts beschränkt zu sehen, wie es in der Mediendidaktik üblich ist, ist jedoch unangebracht. Man bedenke die erwähnte Asynchronität der Interaktion, die Hacker (1994) und Winterhoff-Spurk (2004) beschreiben. Die Stärkung von Konstruktion und Informationserarbeitung wie von AEBLI (1998), BLANCKERTZ (1986) und auch SEEL (2000) gefordert, ist dann auch Anregung für die Entwickler von

telemedialer Lehre und vielleicht sogar jeglicher Publikation die der Bildung dienen soll.

## 2.2 Grundlagen des Experiments

### 2.2.1 Forschungsrichtungen der Mediendidaktik

Folglich ist der Fokus dieser Arbeit das Versorgungssystem sowie die Bearbeitungswerkzeuge einer eLearning-Umgebung und nicht ihr medialer Inhalt.

Man sollte auch bedenken, daß Fehler in der Gestaltung in der Regel von der medientechnischen Bildung kompensiert werden müssen.

<i>Medien als</i>	<i>Produkt</i>	<i>Prozeß</i>
Medientechnische Systeme	(1) lernförderliche Infrastruktur	(3) Werkzeuge für Lehr- und Lernprozesse
Didaktische Medien	(2) Medienprodukte mit didaktischer Intention	(4) Konzeption, Entwicklung und Einsatz didaktischer Medien

*Tabelle 1: Vier-Felder-Schema mediendidaktischer Fragestellungen, nach Kerres (2001), S. 30*

Ziel dieser Arbeit ist dann der Vergleich zwischen der Lernförderlichkeit der Treatmentstufen, also eine Produktbewertung. Diese führen wir durch, indem wir aus der prozeduralen Beschreibung des Werkzeugs ein Kriterium für die Qualität des Produktes ableiten.

Die Zahl vielschichtig komplexer Medien nimmt zu, einige mögliche Ursachen sind die Stärkung des lebenslangen Lernens, höhere Klassenfrequenzen in allen Bildungssystemen aber auch die mediale Umorientierung der Lernenden (Nuria, 2001). Es muß also ein realistischer Lernprozeß evaluiert werden, der sich meist über längere Zeit erstreckt und vielfach unterbrochen ist.

Aus dem Instructional Design stammende Systeme, die stark an den Annahmen des programmierten Unterrichts orientiert sind, werden oft einmalig im direkten Wettbewerb verglichen, siehe Benchmarking, es kann aber auch Entwicklung und Etablierung des Systems prozeßhaft beschrieben werden. Auch hier gilt wieder die Unterscheidung in Inhalte und Werkzeuge.

	<i>Deskriptiv</i>	<i>Präskriptiv</i>	<i>Normativ</i>
Was?	Lehrziel, Lehrinhalte	Curriculum	Wertbegründungen für Lehrinhalte
Wie?	Lehrer-Lerner- Interaktion	Lehrmethoden	Ethische/kommunikative Standards des Lehrens
	Pädagogische Psychologie	Didaktisches Design	Erziehungs- philosophie

*Tabelle 2: Zugänge der Lehr-Lernforschung (aus Kerres, S.37)*

Forderungen sind unter anderem die Passung zwischen Lernstil und Lehrsystem (MOORE, 1993), da es interindividuell unterschiedliche Lernstile gibt, sollte die Funktionalität breit gefächert sein, um den verschieden ausgestatteten Nutzern gerecht zu werden. Aber auch die Beschreibung von Lerngewohnheiten ist notwendig, da diese zu Erfolgsunterschieden bei Anwendungen der gleichen Methode führen können, deren Ursachen wiederum in Neigungen oder Fähigkeiten liegen können.

### **2.2.2 Modelle der Textverarbeitung**

Kintsch & van Dijk (1978) stellten ein Textverarbeitungsmodell vor, das über eine bloße Regelgrammatik (vgl. Chomsky, 1957) hinausging. Ausgehend von

philosophischen Konzepten beschrieben sie die Erarbeitung der textuellen Information als Aufbau einer propositionalen Struktur.

Dieses Modell wurde später in eine Mikrostruktur, welche die aus dem Text aufgenommenen Informationen repräsentiert -analog zum ursprünglichen Modell- und eine Makrostruktur aufgeteilt. In dieser werden Thema und Interpretation repräsentiert.

Die nächste Erweiterung brachte einerseits eine Situationsstruktur, darin enthalten sind schemaartige Annahmen über den Kontext des Textes, und andererseits eine sogenannte Superstruktur. Schließlich wurde die Situationsstruktur als dynamisch angenommen, so daß sie durch heuristische Erkenntnisprozesse während der Textverarbeitung umgebaut werden kann.

Garrod & Sanford (1981) hingegen bauen auf einem Gedächtnismodell mit verschiedenen Speicherarten auf, und postulieren einen Focus, ähnlich der Scheinwerfer-Analogie, der die Informationsaufnahme koordiniert.

Für genauere Ausführung ist Hasebrook (1995) zu empfehlen, der auch die Basis dieser Ausführungen ist.

### **2.2.3 Wissenserwerb und Metastrukturen**

Laut Seel (2000) sollte Lernen im Sinne eines Wissenserwerbs, im Gegensatz zum Reiz-Reaktions-Lernen, als Repräsentation von Informationen und Bedeutungszusammenhängen mithilfe von Schemata und mentalen Modellen verstanden werden. Die Qualität eines Lernprozesses ist dann aber -im Gegensatz zu den Annahmen des Instructional Design- nicht nur von den Reizbedingungen abhängig, sondern ist das Resultat der aktiven und

konstruierenden Auseinandersetzung mit der Lernsituation, sowie natürlich des Vorwissens.

#### **2.2.4 Nutzen der Notizerstellung nach Ciesinger, Muthig & Piekara**

Es gibt zu diesem Zeitpunkt keine Theorie der Notizerstellung, die alle empirischen Befunde erklärt, einen Überblick geben zum Beispiel Cook & Mayer (1983), allerdings beherrschen zwei Arten von Ansätzen die Diskussion:

- Externe Speicherhypothese
- Enkodierungshypothese

Bei der externen Speicherhypothese wird angenommen, daß der Vorgang des Notierens an sich keinen Effekt hat, aber ein externer Speicher angelegt wird, der gezielt wieder abgerufen oder in späteren Lernphasen genutzt werden kann.

Ersteres wurde bisher kaum untersucht, über das wiederholte Studium von Notizen berichten zum Beispiel Einstein, Morris & Smith (1985).

Die Enkodierungshypothesen auf der anderen Seite besagen einen positiven Effekt des Notierens an sich. Die Theorien variieren in der Ursachenzuschreibung von Aufmerksamkeitsprozessen bis zu internem Rehearsal und der Transformation in eine subjektiv bedeutungsvollere Form. Sie stimmen aber darin überein, das die Notizanfertiung die interne Enkodierung verbessert.

Die Enkodierungshypothese wurde in der Arbeit von Piekara et al. noch einmal unterteilt:

- Allgemeine Enkodierungshypothese

- Eingeschränkte Enkodierungshypothese.

In der allgemeinen Formulierung ist der Akt des Notierens generell förderlich für die Enkodierungsleistung der gesamten Informationsmenge, während die eingeschränkte Variante eine Enkodierungsverbesserung nur für die notierten Informationen annimmt. Die eingeschränkte Enkodierungshypothese wurde empirisch bereits mehrfach bestätigt, die Befundlage der allgemeinen Hypothese ist uneindeutig.

Eine wichtige Frage ist in dieser Diskussion das Ausmaß der Interferenz zwischen Prozessen der Notizanfertigung und Prozessen der internen Wissensstrukturierung zum Beispiel durch die Beschränktheit beiderseits benötigter Ressourcen.

Piekara et al. sprechen sich dafür aus, daß das Notieren als Prozeß keine Auswirkungen hat, sondern daß die Speicherung einer Informationen von ihrer Relevanz für die Erreichung des Textverarbeitungszieles abhängt. Das Modell von Piekara, das im nächsten Abschnitt erläutert wird, sagt eine Abhängigkeit der Wahrscheinlichkeit für die Erstellung einer Notiz in Abhängigkeit von der Relevanz der Information voraus.

In diesem Fall wäre die gemeinsame Ursache, nämlich die Relevanz der Information sowohl für das Notieren als auch Behalten verantwortlich, ohne das die Prozesse miteinander interagieren müßten. Dies steht im Gegensatz zu Einstein et al. (1985), der von einer allgemeinen Verbesserung der Relevanzbewertungsfähigkeit durch den Prozeß des Notierens ausgeht. Allerdings stützen die Befunde von Ciesinger, Muthig & Piekara (1987) ihre Theorie.

### **2.2.5 Nutzen der Notizerstellung nach Kiewra**

Kiewra (1985) unterscheidet eine Prozeß- und eine Produktfunktion des Notierens, analog zu externer Speicherung und Elaboration.

Entgegen der bisherigen Forschung wendet er sich gegen einen monokausale Zuordnung der Behaltenssteigerung auf den Notizakt oder die Wiederholung der Notizinhalt, da nicht jeder Notizakt automatisch eine verbesserte Enkodierung nach sich zieht, und die Wiederholungshandlungen eines Lernenden über schlichtes Rehearsal, sprich die Aufrechterhaltung der Aktivierung im Kurzzeitspeicher hinausgehen.

Allerdings sieht er sehr wohl empirische Hinweise für den positiven Zusammenhang zwischen beiden Funktionen und der akademischen Leistung.

Hierbei steht nicht so sehr die Elaboration der Wissenstruktur im Vordergrund, wichtig sind in seiner Darstellung eher Notizhandlungen, die das Arbeitsgedächtnis entlasten. Dies deckt sich auch mit der Forderung von Hacker (1994) das Nadelöhr, welches das Arbeitsgedächtnis darstellt, zu entlasten.

### **2.2.6 Kritik der Modelle des Notiznutzens**

Beide Überblicke sind aus mediendidaktischer Sicht unvollständig, da die Vesorgungsmedien und auch die Speichermedien in den Vergleichen zwischen den empirischen Befunden nicht thematisiert werden.

Es wird auch selten ökologisch valide der Erwerb einer größeren Datenmenge über einen längeren Zeitraum hinweg analysiert. Genau dies ist jedoch für die Analyse des Lernens notwendig.

### **2.2.7 Das Electronic Learner Satisfaction Instrument**

Subjektive Methoden sind aus der Erhebung der ergonomischen Qualität nicht mehr wegzudenken, allerdings ist es methodisch nicht vertretbar, daß die Fragebogenitems von den Ergonomen meist ohne theoretische Grundlagen zusammengestellt werden. Insbesondere, da die Zufriedenheit als Einstellung zum Beispiel in der Konsumentenforschung sehr wohl formalisiert wird.

Wang (2003) hat aufbauend auf Torkzadeh (1991) durch eine Itemsammlung mit anschließender Expertenbefragung und Item-Streichung einen Fragebogen konstruiert, der in der Faktoranalyse vier Dimensionen ergab:

- Learner interface (Lernenden-System-Schnittstelle)
- Learning community (Lerngemeinschaft)
- Content (Inhalt)
- Personalization (Personalisierung)

Dieses Faktormodell erklärt 78% der Varianz in der Eichstichprobe. Ein solcher Fragebogen hat auch den Vorteil, daß die Items keine redundanten Informationen erzeugen, und die Versuchspersonen durch kürzere Fragebögen weniger ermüden.

Allerdings muß man nach Weiner (1986) einwenden, daß eine solche subjektive Bewertung unter anderem Wahrnehmungsverzerrungen unterliegt die aus Attributionsstilen der Bewerter entstehen.

Es liegt sozusagen im Ermessen des einzelnen, einen Erfolg auf eigene oder im System liegende Eigenschaften zurückzuführen.



### **2.2.8 Modell der Notizerstellung von Ciesinger & Piekara**

Im folgenden wird ein Modell von Ciesinger & Piekara (1986) vorgestellt, das die am Prozeß der Textverarbeitung im Hinblick auf die Erstellung von Notizen beteiligten Verrechnungsschritte beschreibt.

Menschliche Informationsbearbeitung ist instrumentell und zielgerichtet. Informationen werden durch kognitive Prozesse in einer Struktur organisiert und in bestehende Strukturen integriert. Eine klassische Alternative zum internen Memorieren ist die externe Repräsentation in Form einer Notiz. Die Erstellung von externen und internen Repräsentation benötigt interne in ihrer Kapazität beschränkte Ressourcen, wie zum Beispiel Aufmerksamkeit.

Die externe Speicherung unterliegt zusätzlich den Beschränkungen des Aufzeichnungsmediums, beispielsweise durch einen geringen Papiervorrat.

Das Verhalten ist ökonomisch in der Verteilung der mentalen Ressourcen auf die beiden verschiedenen Einspeicherungsarten und den später folgenden Wiederabruf. Folglich sollte auch nur zieldienliche Informationen gespeichert werden.

Ausgehend von der Ziel- und Wissensstruktur des Textverarbeiters, ergibt sich im Falle eines Lernenden eine Zielstruktur, die Leerstellen enthält. Externe Informationen können die bestehende Struktur ergänzen.

Die aufgenommene Information, repräsentiert in einer episodischen Struktur, wird auf ihre Zieldienlichkeit überprüft. Gibt es keine Ähnlichkeit zwischen einem Teil der Wissensstruktur und der neuen Information, wird die

Information als neu bewertet und auf ihre Relevanz für die Zielerreichung überprüft.

In einem semantischen Netz kann diese Feststellung durch die Zahl der Verbindungswege zwischen der neuen Information und der Zielstruktur dargestellt werden. Je besser die Passung zwischen den Leerstellen der Zielstruktur und der episodischen Struktur desto höher der Relevanzwert, welcher mit dem subjektiven Relevanzkriterium verglichen wird.

Überschreitet der Relevanzwert das Kriterium werden interne oder externe Speicherungsaktivitäten eingeleitet. Die kostengünstigere Speicherart wird präferiert. Diese Kosten setzen sich zusammen aus dem Einspeicheraufwand und den Zugriffskosten, wobei diese Bewertungen subjektiven Gewichtungen und Gewohnheiten beinhalten.

Das Erstellen einer externen Notiz wird als hochgradig geübte und automatisierte Tätigkeit angenommen, die kaum mentale Ressourcen benötigt. Die freibleibenden Anteile der Ressourcen können während der Notizanfertigung für andere Informationsverarbeitungsaufgaben genutzt werden, oder eine zur externen parallele interne Verarbeitung findet statt. Nur letzteres ist im Sinne der eingeschränkten Enkodierungshypothese behaltenswirksam.

Wenn eine hohe Zugriffssicherheit angestrebt wird, müßte die interne Speicherung in Anlehnung an Newell & Simon (1972), deutlich aufwendiger sein als die externe. Bei genauer Reproduktion verhält sich der Aufwand exponential zur Menge der Verknüpfungen. Der Zugriff auf die gespeicherte Information jedoch verhält sich umgekehrt. Der externe Speicher muß

eventuell seriell durchsucht werden. In einem der besten Fälle, existiert ein zusätzlich erstelltes Deskriptorenverzeichnis.

Im Falle der internen Verknüpfung der Information mit anderen Teilen der Wissenstruktur, fällt der assoziative Abruf immer leichter, je mehr Speicheraufwand betrieben wurde.

Eine Vorstellung der Entscheidungsprozesse kann man sich machen, indem man, wie die Autoren empfehlen, einige Marker generiert und den Fluß der Information entlang der Kanten durch die Abbildung des Modells verfolgt, statt zu versuchen sich einen statischen Überblick zu verschaffen.



- Einspeicherungskosten

ergibt sich die Valenz der entsprechenden Speicherhandlung.

Wenn wir also annehmen, daß die Menge an relevanten Informationen in einem Text interindividuell vergleichbar ist, ergibt sich bei Manipulation der Größen Zugriffsaufwand und Einspeicherungskosten eine Veränderung in der Zahl der erstellten Notizen. Wird das Notieren teurer, werden mehr Informationen intern gespeichert, wird es billiger, werden mehr externe Notizen erstellt. Diese Steigerung ist allerdings beschränkt, da, egal wie günstig die Speicherung ist, nur Relevantes Objekt der Bearbeitung sein kann.

### **2.2.9 Grenzen des Modells**

Das Modell ist kein Laufzeitmodell. Es beschreibt zwar die Verrechnungsschritte, macht aber keine Annahmen, über deren Ressourcenbedarf, die Informationssynchronisierung oder den Zeitverbrauch. Dies ist allerdings von den Autoren nicht beabsichtigt. Da sie keine Hypothesen über die Bearbeitungs- und Entscheidungszeiten aufstellen wollen, ist dieses Modell in seiner funktionalen Form ausreichend.

Es ist für die Vorhersage von Reproduktionswahrscheinlichkeiten einzelner Informationen aus einem Text völlig ausreichend.

Allerdings spricht auch nichts dagegen unter Zuhilfenahme der entsprechenden Daten, das Modell in ein Produktionssystem, wie zum Beispiel ACT-R (Anderson, 1990) zu überführen. Dieses erweiterte Modell könnte dann sogar einzelne Handlungszeiten vorhersagen.

Allerdings müßte eine solche Modellierung auch die Art der Repräsentationsstruktur und die beteiligten Konstrukte der Textrepräsentation

bestimmen, Arbeitsgedächtnis und Aufmerksamkeit sind zwar ankerkannte Konzepte, sie sind jedoch auch nicht spezifiziert genug, um Hypothesen über die Handlungszeiten aufzustellen. Die Zeitkosten von Aufmerksamkeitswechslern beschreiben Anderson, Lebiere & Matessa (1997), Monsell & Rogers (1995) die Kosten von erwarteten Aufgabenwechslern, wie es in diesem Fall beim Schreiben einer Notiz während des Lesens eines Textes gegeben ist.

Modelle des Gedächtnisses gibt es viele, für einen prozedural orientierten Beitrag siehe Barlas, Goldstein & Weber (1995).

Klärungsbedarf besteht auch bei der Auflösung von Handlungskonflikten. Einerseits, wenn es um die Auswahl einer Option aus einer multidimensional beschriebenen Menge geht oder aus anderen Gründen die Ressourcen nicht für eine vollständige Bewertung der Optionen ausreichen. Hier können adaptive Entscheidungsmodelle Anregungen liefern, siehe hierzu Bettman, Johnson & Payne (1993).

Der direkte Konflikt zwischen zwei Handlungen benötigt die Formulierung einer Handlungsregulation, siehe beispielsweise Skell (1999).

Die Existenz von Lernstilen und der Einfluß von Gewohnheiten auf die Handlungskontrolle wird kurz angesprochen. Im Gegensatz zu anderen Kritikpunkten benötigt die Integration von Handlungsursachen, die nicht auf Valenz- und Erwartungs-Mal-Wert-Berechnungen basieren, keine prozedurale Umformung des ursprünglichen Modells.

Eng verwandt mit dem Thema der Gewohnheiten sind die Heuristiken (Gigerenzer, 2000), es ist zum Beispiel vorstellbar, daß bei gewissen

Konfigurationen des Reizmaterials Neuheit und Relevanz nicht berechnet sondern heuristisch geschätzt werden.

### **2.2.10 Zusammenfassung der Forschungshypothesen**

Wenn wir also zwei verschiedene Notizfunktionen überprüfen, so sollten die Nutzer der günstigeren Funktion bei der Bearbeitung des gleichen Textes mehr externe Speicherhandlungen ausführen, als die „benachteiligte“ Gruppe.

Desweiteren sollte das bessere System von seinen Nutzern auch subjektiv als das zufriedenstellendere beschrieben werden.

Es sollte aber auch ein Zusammenhang zwischen dem Attributionsstil der Versuchspersonen und ihren Bewertungen des Lernsystems geben.

## **3 Methoden und Ergebnisse**

### **3.1 Beschreibung der Entwicklung des Experiments**

Es wurde ein eLearning-System erstellt. Als Stimulusmaterial dienten Auszüge aus einem Reader zum Thema Narzissmus, der von Joachim Wutke gepflegt wird. Die Menge des Materials war so bemessen, daß eine vollständige Bearbeitung unmöglich sein sollte. Um die Texte zu verwalten, wurde eine MySQL-Datenbank erstellt, in dieser wurden auch alle anderen Daten gespeichert, die während der Nutzung des Systems erzeugt wurden:

- Inhalt von Notizen
- Genaue Positionen der Notizen
- Zeitindizes der Seitenaufrufe (die Differenzen lassen Rückschlüsse auf die Verweilzeiten zu)
- Ziel und Art jeder User-Action, in der episodischen Reihenfolge
- Identität der Seiten, auf denen ein Logout stattfand

Um auf die Daten zuzugreifen wurde mysqld und der Apache-Webserver benutzt. Beide gehören zur Basisausstattung von Knoppix 3.4.

Im Experimentalsystem wurden die Schriften Arial, für feste Fonts, und Times, für proportionale Fonts, verwandt, die Schriftgröße des Inhaltsindizes war 8, alle anderen Schriften hatten die Größe 12. Die Navigationsleiste, die sich am Ende des Textkörpers identisch wiederholt, ist tatsächlich hellblau, der Inhaltsindex mittelgrau und der Textkörper -inklusive der Notizspalte im Falle von  $T_{\text{right}}$  hellgrau. Eingefügte Notizen waren in beiden Versuchsbedingungen lila, im Falle von Inside, diente die farbliche Hervorhebung der Lesbarkeit, bei Right wurde sie nur der Vergleichbarkeit wegen beibehalten. Die genauen RGB-Werte können im Skript basis.php nachgelesen werden.

Mithilfe einer unsichtbaren Tabelle wurde der Bildschirm aufgeteilt, so daß der Inhaltsindex eine Breite von 20% des Bildschirms hatte, die Notizspalte falls vorhanden desgleichen und der Textkörper dementsprechend 60% beziehungsweise bei Wegfall der Notizspalte 80% der Breite einnahm.

Da im Falle von  $T_{\text{inside}}$  die Notizen in den Text eingefügt werden, sollte durch diesen Zuschlag verhindert werden, daß der modifizierte Text sich nach unten aus der Bildschirmgebrenzung herausschiebt. Es muß jedoch festgehalten werden, daß insofern die beiden Treatmentstufen abgesehen von der eigentlichen Modifikation nicht vollständig gleich sind. Die Annahme scheint jedoch berechtigt, daß beide Stufen trotz dieses Unterschiedes in der Spaltenbreite vergleichbar sind.



Durch die Nutzung des Vollbildmodus konnte gewährleistet werden, daß der Textkörper vollständig auf dem 19 Zoll großen Monitor dargestellt werden konnte. Dadurch wurde das Bewegen des Textes über den Bildschirm unnötig. Der Grundgedanke war, daß die Handlung des Scrollens vom basalen User-Logging nicht erfaßt werden konnte. Im Laufe der Durchführung beschwerten sich allerdings mehrere Versuchspersonen über den flimmernden Bildschirm, obwohl die Wiederholfrequenz von 65 Hertz während der Entwicklung und entwicklungsbegleitenden Evaluation ausreichend erschien. Durch den Vollbildmodus des Konqueror-Browsers der zur Darstellung der Lernumgebung diente waren keine Elemente des Unix-Betriebssystems, das auf Knoppix Version 3.4 beruhte und erst recht keine der KDE 2.0-Oberfläche, für die Vpn sichtbar. Um eine zeitliche Orientierung zu ermöglichen stand eine Uhr auf dem Tisch an dem die Versuchsteilnehmer saßen. Um Erschöpfungszuständen entgegenzuwirken, bekamen die Teilnehmer außerdem zusätzlich zu der Bildschirmpause Mineralwasser und Kekse soviel sie wollten.

Es folgt eine Aufzählung der möglichen Handlungen, in Klammern wird die Position des entsprechenden auslösenden Links genannt, man vergleiche hierzu die Abbildungen 3 und 5, die ein Screenshots des Alphasystems ist. Es gibt allerdings einige Abweichungen in den Farben und Schrifttypen zwischen Darstellung und Experiment.

- Eine Seite vorwärts blättern (Navigationsleiste)
- Eine Seite rückwärts blättern (Navigationsleiste)
- Logout (Navigationsleiste)

- Eine Seite über den Inhaltsindex direkt ansteuern (Linke Seitenspalte)
- Eine neue Notiz auf der aktuellen Seite erstellen (Textfenster)
- Eine bestehende Notiz auf der aktuellen Seite verändern (Text der Notiz)

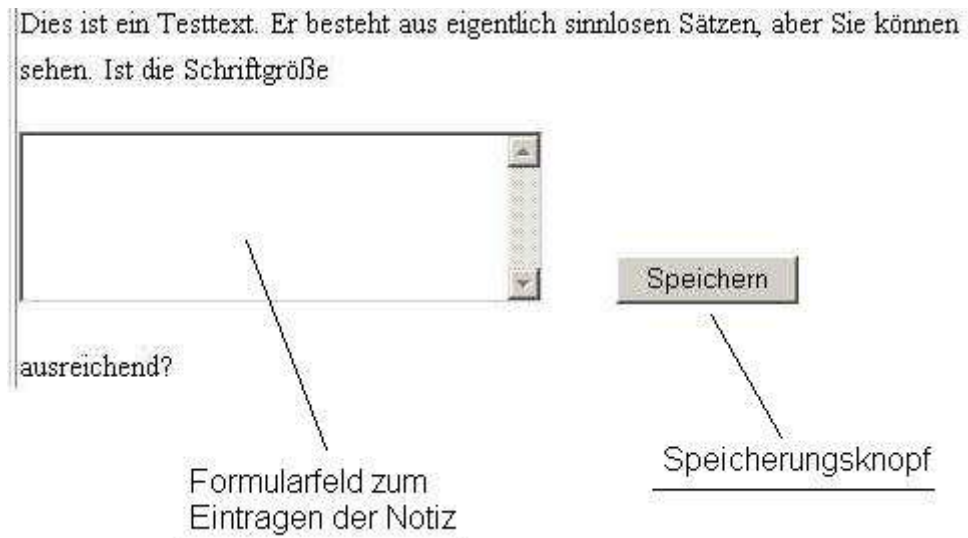


Abbildung 3: Screenshot des Systems im Notizzustand, Erläuterungen sind unterstrichen

Falls eine Notiz geändert oder erstmalig erstellt werden sollte, wurde die entsprechende Textseite erneut aufgebaut mit einem Formularfeld hinter dem Wort das als Notizanker angeklickt wurde. In Abbildung 3 wäre dieses Wort Schriftgröße gewesen. Nach der Übergabe der Notiz an das Skript `basis.php` wurde die entsprechende Textseite erneut aufgebaut diesmal mit der Notiz. Je nach Treatmentbedingung entweder hinter dem Textanker in den fortlaufenden Text integriert oder aber in der rechten Seitenspalte.

Grundsätzlich bestand die Interaktion zwischen Nutzer und System aus einer Schleife von Informationsdarbietung und Berechnung der neuen Darbietung, wobei der Auslöser für den nächsten Durchlauf immer eine Aktion des Nutzers war.

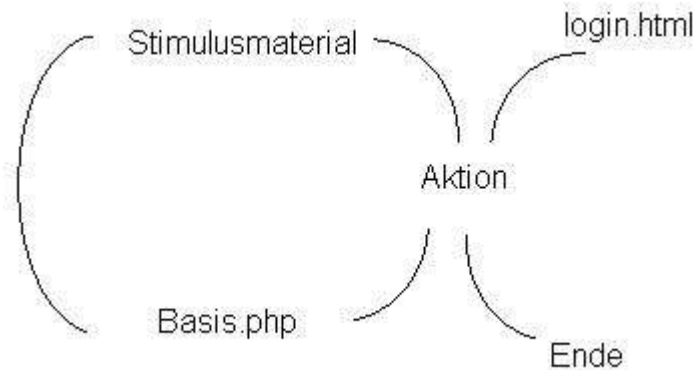


Abbildung 4: Visualisierung der Handlungsschleife des Systems

Durch eine GOMs-Analyse lassen sich die Möglichkeiten der Vpn im System folgendermaßen beschreiben:

?Besitze ich alle zielrelevanten Informationen?

Ja: Logout

Nein: Weiterlesen

?Ist das gelesene relevant?

Nein: zurück zu Weiterlesen

Ja: Speicherung der Information

?Welche Speicherung ist subjektiv und momentan günstiger?

Intern: Interne Speicherung

Danach: Zurück zu Weiterlesen

Extern: Externe Speicherung

Aufruf des Notizfensters

?Wie soll die Notiz formuliert werden?

Gar nicht: Zurück zu Weiterlesen

?Welche Abrufhilfen kann ich in die Notiz integrieren?

Schlagwörter

Bezugspunkt

Schreiben der Notiz

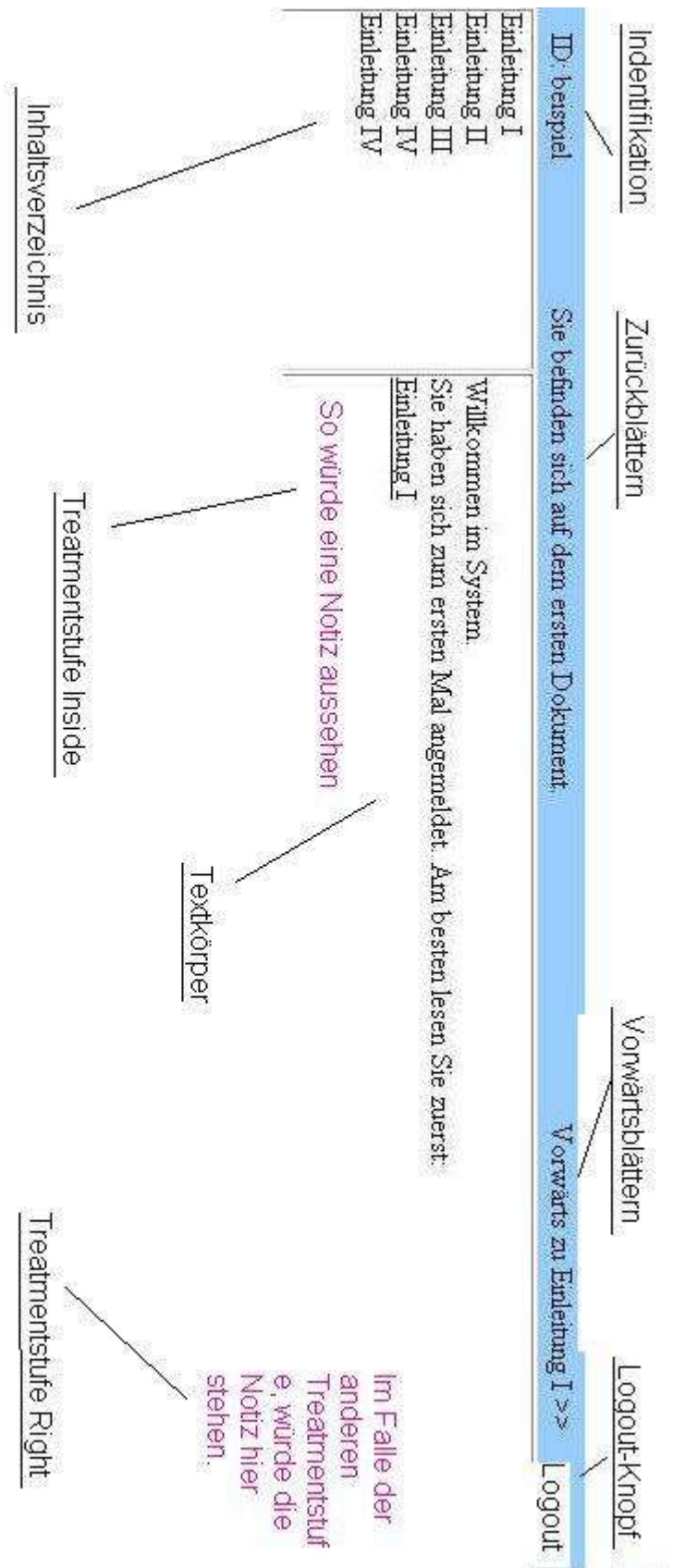


Abbildung 5: Screenshot der Lernumgebung, unterstrichene Begriffe sind eingefügte Erläuterungen

Der Fragebogen wurde mithilfe eines PHP-Skriptes erzeugt, danach allerdings in statischer Form gespeichert als `fragebogen_statisch.html` und nachbearbeitet. Eine automatische Produktion reduziert die Wahrscheinlichkeit von Fehlern in der Festlegung der Werte, die durch die Fragebogen-Elemente zurückgegeben werden. Die Item-Scores der Versuchspersonen wurden in einer separaten Datei gespeichert.

Zuerst werden deskriptive Faktoren erhoben:

- Alter
- Geschlecht
- Nationalität und Muttersprache
- Tätigkeit und Tätigkeitsdauer

Desweiteren wurde das Vorwissen der Vpn hinsichtlich der Psychodynamik im Allgemeinen und des Narzissmus im Speziellen erfragt, zusätzlich dazu eine Einschätzung, wie groß das Wissen über Narzissmus nach Absolvierung des Experimentes war.

Auch wenn aus Sicht der empirischen Hypothesen die Frage nach der Behaltensleistung irrelevant ist, sollte den Vpn trotzdem das Gefühl einer authentischen Lernerfahrung vermittelt werden, abgesehen davon, können diese Fragen helfen Vp zu identifizieren, die mit offenen Augen geschlafen haben.

Danach folgen die ins Deutsche übertragenen Items von Wang (2003), da jedoch das Experimentalsystem keine vollständige Lernumgebung ist, wurden die unangebrachten Items gestrichen. Es blieben die Items Q2-Q10, Q16, Q18, Q25 und Q26. Zusätzlich zu den beiden Globalbewertungen Q25 und Q26

wurde als Ergänzung die Frage gestellt, ob die Vpn das System auch in Zukunft benutzen würden (Item: QAA). Die Antwortmöglichkeiten bei all diesen Fragen waren siebenfach gestuft und hatten die Ausprägungen:

- völlige Zustimmung,
- Zustimmung,
- geringe Zustimmung,
- Neutral,
- geringe Ablehnung,
- Ablehnung und
- völlige Ablehnung.

Danach wurde mit drei in Anlehnung an Weiner (1986) gestalteten Fragen, QBB, QCC und QDD, die Attributionsneigung der Vpn erhoben, auf die Nutzung eines vollständigen Questionnaires wurde aufgrund der Fragebogenermüdung verzichtet. Als letztes folgte ein freies Textfeld um den Vpn freie Kritik zu ermöglichen (Qfrei). Zusätzlich zu diesem Freifeld existierte zu jeder Frage ein freies Kommentarfeld. Es ist aus Sicht der Auswertung sinnvoller, zusätzliche Kommentare direkt in die Datenerhebung zu integrieren, selbst wenn man sie nicht beachtet, als daß man vom VI erwartet, daß er solche Äußerung schnell und adäquat genug mitschreibt.

Das Skript, das die Antworten abspeicherte, überprüfte zusätzlich die Vollständigkeit der Antworten, um die Wahrscheinlichkeit von Missing Values zu minimieren. Html-Fragebögen haben genau wie alle Hypertext-Dokumente den Nachteil, daß sie, sobald sie länger sind als ein Bildschirm, sehr unübersichtlich werden.

### 3.2 Formulierung der statistischen Hypothesen

Wie oben schon ausgeführt sind die Vorhersagemöglichkeiten des Modells von Ciesinger & Piekara (1986) beschränkt, daraus folgt, daß nicht die produzierten Inhalte sondern nur die Produktionsakte an sich bewertet werden.

Die Hypothesen sind zweiseitig formuliert, da aus den theoretischen Überlegungen die Überlegenheit der einen Treatmentstufe über die andere nicht zwingend hergeleitet werden kann.

Die  $H_0$  lautet dementsprechend:

$$\mu_{\text{right}} = \mu_{\text{inside}}$$

Die  $H_1$  lautet:

$$\mu_{\text{right}} \neq \mu_{\text{inside}}$$

Desweiteren muß, wenn die Häufigkeit der Notizen ein valides Konstrukt der Systemqualität ist, die Summe der Item-Scores von Wang mit der Zahl der Notizen korrelieren:

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

Außerdem soll der Zusammenhang zwischen Ursachenattribution und Systembewertung überprüft werden:

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

Für alle drei Hypothesen gilt als Signifikanzniveau  $\alpha = 0,05$ .

### 3.3 Durchführung des Experimentes

22 Versuchspersonen nahmen an dem Experiment teil, die Versuchsteilnahme wurde mit einem Speiseeis und zusätzlich im Falle von Studierenden der



Psychologie im Grundstudium mit zwei Versuchspersonenstunden entgolten. Die Werbung bestand aus einem Anhang am schwarzen Brett der Fachrichtung, der in den digitalen Anhängen enthalten ist.

Die verschiedenen Abschnitte des Experimentes lassen sich folgendermaßen gliedern:

1. Begrüßung durch den VI
2. Verlesen von Einleitungstext I
3. Anmelden der Vp im System durch den VI (Ab diesem Moment läuft die Zeit der ersten Lerneinheit.)
4. Vp liest den Einleitungstext II
5. Vp beginnt mit der Textbearbeitung
6. Ende der ersten Lerneinheit nach 45 Minuten
7. Fünf Minuten Bildschirmpause
8. Zweite Lerneinheit (Dauer: 45 Minuten)
9. Bearbeitung des Fragebogens
10. Verabschiedung der Versuchsperson

Die Gesamtdauer des Experimentes lag knapp unter zwei Stunden.

Auf eine Aufteilung der Lerneinheiten auf verschiedene Erhebungstage wurde aus Gründen der Versuchspersonenmortalität verzichtet.

Die Vpn wurden begrüßt, und es wurde ihnen Einleitungstext I vorgelesen, der einen groben Überblick über den Ablauf des Experimentes gibt. Wie in Einleitungstext I erbeten meldeten sich die Vpn dann im System an, und navigierten entweder durch einen zentral positionierten Link oder den Inhaltsindex auf den Einleitungstext II. Nach der Instruktion begannen sie mit

der Textbearbeitung. Nach Ablauf der ersten Lerneinheit wurden die Vpn vom VI aufgefordert sich auszuloggen. Während der anschließenden fünfminütigen Pause wurde den Vpn geraten sich zu bewegen, es stand ihnen frei die Toilette aufzusuchen. Nach dem erneuten Anmelden im System konnten die Teilnehmer über einen Link auf der Begrüßungsseite direkt auf die zuletzt bearbeitete Seite zurückkehren. Nach dem Ablauf der zweiten Lerneinheit wurde den Vpn der Fragebogen vorgelegt. Im Zuge der Verabschiedung wurden die Teilnehmer über die Fragestellung des Experimentes aufgeklärt und es wurde ihnen angeboten, den vollständigen Reader per E-Mail zugestellt zu bekommen.

### **3.4 Deskriptive Statistik**

Von den 22 Versuchspersonen diente die erste Vp der abschließenden Testung für die Lauffähigkeit des Experimentes und geht nicht in die Ergebnisse ein (ID: w18m).

Aufgrund der Balancierung zwischen den Geschlechtern mußte eine weitere Vp mit der Faktorkombination weiblich und  $T_{\text{inside}}$  gestrichen werden (ID: w29c). Die zu streichende Vp wurde zufällig bestimmt (Heller & Rosemann, 1981).

Alle folgenden Aussagen beziehen sich auf diese Stichprobe mit einem N von 20.

Die Versuchspersonen verteilten sich auf zwei zufällige Stufen des Treatmentfaktors „Art der Notizdarstellung“, es gab keine Meßwiederholung, entsprechend ist die Haupthypothese sowohl mit einem t-Test für unabhängige Stichproben, als auch mit einer Varianzanalyse für einen zufälligen

zweigestuften Faktor möglich. Für die Gleichwertigkeit dieser beiden Verfahren siehe Bortz (1999).

Die Notizhäufigkeit als Variable mit absolutem Skalenniveau, erfüllt die Voraussetzung der Intervallskalierung, die für beide Verfahren gilt.

Das Alter der Teilnehmer hatte einen Mittelwert von 25 Jahren und eine Varianz von 23,9 bei einer Spannweite von 20, einem Minimum von 20 und einem Maximum von 40.

Aufgrund dieser hohen Varianz des Alters sind die Tätigkeitszeiten ähnlich weit gestreut: Tätigkeit in Fachsemestern: Mittelwert 7 und Varianz 77,3 bei einer Spannweite von 39.

Es waren 10 männliche und 10 weibliche Versuchspersonen, die balanciert auf die Faktorstufen verteilt sind.

		<i>TREAT</i>		<i>Gesamt</i>
		<i>inside</i>	<i>right</i>	
<i>SEX</i>	<i>weiblich</i>	5	5	10
	<i>maennlich</i>	5	5	10
<i>Gesamt</i>		10	10	20

*Tabelle 3: Verteilung der Vpn auf die Treatmentstufen und den Balancierungsfaktor*

Das Vorwissen in der Psychodynamik wurde im Mittel mit 0,4 bei einer Varianz von 0,5 angegeben, hinsichtlich des Wissens über Narzissmus werteten sich die Vp im Mittel mit 0,65 bei einer Varianz von 0,59.

Achtzehn Versuchspersonen waren deutscher Nationalität, genauso viele gaben Deutsch als ihre Muttersprache an, und keine Vp hatte auf beiden Variablen einen Wert, der ungleich Deutsch war.

Es kann also angenommen werden, daß die Texte für alle Teilnehmer verständlich waren.

Fünfzehn Versuchspersonen gaben als Tätigkeitsfeld die Psychologie an, drei sind Studierende der Informatik, eine ist Lehrling im Elektronik-Bereich und eine Versuchsperson gab arbeitssuchend an. Die Nicht-Psychologen verteilen sich gleichmäßig auf beide Faktorstufen.

Die Angaben über den subjektiv geschätzten Lernzuwachs hatten einen Mittelwert von 0,75 und eine Standardabweichung von 0,55, allerdings gaben fünf Teilnehmer an, daß sie keinen Wissenszuwachs hatten.

Die Häufigkeit von Notizanfertigungen hatte einen Mittelwert von 12,35 und eine Standardabweichung von 20,16, das Minimum war Null und das Maximum 77.

Der Mittelwert der summierten Itemscores von Wang belief sich auf 24,5 bei einer Standardabweichung von 6,72.

Die Summe der Items, welche die Attributionsneigung erheben, haben einen Mittelwert von 4,25 bei einer Standardabweichung von 1,37.

Die theoretische Spannweite dieser Summe läuft von 0 bis 18, woraus man schließen kann, daß alle Versuchspersonen sich als grundsätzlich eigenverantwortlich für ihre Leistungen sehen.

### **3.5 Prüfung der statistischen Hypothesen**

Zur Prüfung der Haupthypothese wird ein t-Test für unabhängige Stichproben ausgeführt. Die Varianzen der Gruppe  $T_{\text{inside}}$  und  $T_{\text{right}}$  sind homogen (F-Wert: 9,295,  $p = 0,007$ ), dementsprechend gilt die Ergebniszeile für homogene Varianzen. Der empirische t-Wert ist 0,779 bei 18 Freiheitsgraden ist dieser

Wert nicht signifikant. Der zweiseitige p-Wert ist 0,446. Damit muß die  $H_1$  verworfen werden. Wir nehmen die Nullhypothese an, was in diesem Fall bedeutet, daß die Notizfrequenz keine Unterscheidung zwischen den Treatmentstufen erbracht hat.

Die zweite Hypothese betreffend die Korrelation zwischen der Zahl der angefertigten Notizen und den summierten Itemscores nach Wang muß auch verworfen werden. Der Zusammenhang, bestimmt mit der Pearson-Korrelation, ist minimal negativ  $r = -0.006$  und nicht signifikant ( $p = 0,98$ ). Also muß auch in diesem Fall die  $H_1$  abgelehnt werden, was bedeutet, daß es keinen Zusammenhang zwischen den beiden Variablen gibt.

Die dritte Hypothese, die der Absicherung der bereits verworfenen zweiten Haupthypothese dienen sollte, muß auch verworfen werden. Der Pearson-Korrelationskoeffizient ist zwar größer,  $r$  ist gleich 0,414, aber nicht signifikant, denn der p-Wert ist 0.07.

Somit treffen in allen drei Fällen die Nullhypothesen zu, alle Zusammenhänge die angenommen wurden, konnten empirisch nicht nachgewiesen werden.

### **3.6 Kritik der empirischen Vorgehensweise**

Es lassen sich mehrere mögliche Ursachen für das Ausbleiben signifikanter Ergebnisse finden:

Es können Fehler in der theoretischen Entwicklung der Vorhersagen vorliegen. Die meisten der in Kapitel 2 aufgeführten Theorien sind zwar empirisch untermauert worden, eine genaue Beschreibung der Wirkmechanismen und vor allem der Geltungsbereiche der Theorien fehlen

jedoch. So beziehen sich die Textverarbeitungsmodelle, die beschrieben wurden auf Erzählungen und nicht auf Sachtexte.

Andererseits können auch bei korrekten Schlußfolgerungen die experimentellen Rahmenbedingungen anders sein als von der Theorie verlangt. Ciesinger und Piekara (1986) betonen die Rolle von Motivation und Zielerwartung für die Textbearbeitung. In experimentellen Situationen ist die Homogenität der Motivationen über die Versuchspersonen hinweg grundsätzlich problematisch, in diesem speziellen Fall aber wurde es versäumt durch eine detaillierte und plausible Versuchsanweisung die Zielstrukturen der Versuchsteilnehmer anzugleichen.

Desweiteren mag die Evaluation von Computer-Software grundsätzlich zu den allgemeinen Fragestellungen gehören, die Nähe der Fragestellung zur differentiellen Psychologie ist jedoch augenfällig. So waren, wie schon gesagt, die Zielstrukturen der Vpn unterschiedlich, doch auch Variationen in so grundlegenden Größen wie der Lesegeschwindigkeit wurden unterschätzt. So gab es beispielsweise eine Versuchsperson die in 90 Minuten das komplette Experiment absolvierte, inklusive der Bearbeitung aller Texte.

Auch die Nutzung der Notizfunktion wurde nicht einheitlich genutzt. Über die Hälfte der Teilnehmer machte überhaupt keine Notizen, eine Versuchsperson benutzt die Notizen um Textstellen zu markieren und mindestens eine andere Vp schreibt in die Notizen Nachrichten an den Software-Entwickler. Allerdings sind die Nicht-Notierer gleichmäßig auf beide Treatmentstufen verteilt.

Insgesamt muß dieser Untersuchung jegliche Konstruktvalidität abgesprochen werden.

Bei einem größeren Stichprobenumfang wäre es vielleicht noch möglich gewesen, nachträglich verschiedene Gruppen von Systemnutzern zu identifizieren, wobei man jedoch bedenken sollte, daß auch große Stichproben eine möglichst umfangreiche Beschreibung der intervenierenden Variablen nicht ersetzen können.

Doch auch in der Planung der Inferenzstatistik gab es Fehler. Zum einen wurde nicht bedacht, daß die Versuchspersonenvektoren der Verweilzeiten und der Notizlängen nicht gleich lang sein würden, wodurch die Auswertung mit einer Varianzanalyse nicht mehr möglich war (vgl. Bortz, 1999).

Durch die fehlende Kontrolle der Handlungsabläufe, zum Beispiel das Anbieten einer Textmarkerfunktion oder auch die Beschränkung des Notizinhalts auf echte Wörter, erwies sich diese Vorgehensweise sowieso als unangebracht.

Daß die Notizhäufigkeiten jedoch nicht einmal als Varianzanalyse mit Meßwiederholung bearbeitet werden konnte, wobei jede Lerneinheit ein Meßzeitpunkt gewesen wäre, lag daran, daß die Datenbankmodellierung zwar die Besitzer jeder Notiz zuordnete, aber nicht die Lerneinheit in der sie angefertigt wurde.

Das Verweilzeitenkonzept war zu grob, es ist nicht unterscheidbar für das System, ob eine Versuchsperson auf Dokument Eins, mental eine Notiz erstellt, dann auf Dokument Eins in den Notizzustand wechselt um die sie einzuspeichern, und danach zu Dokument Zwei weiterblättert, oder ob das

System zuerst in den Notizzustand versetzt wird, und die bedienende Person erst dann mit der Externalisierung der Information beginnt.

Auch eine hochauflösende Beobachtung der Versuchsperson mit Hilfe von Eye-Tracking und exaktem User-Logging jedes Tastenanschlags, kann diese Fragen nicht endgültig lösen -ansonsten wäre die kognitive Wende als Ende des Behaviorismus auch unnötig gewesen- zusammen mit einem empirisch ermittelten Produktionensystem könnte jedoch auf viele der in dieser Arbeit angesprochenen Prozesse zurückgeschlossen werden.

Eine Voruntersuchung, um die Relevanz der einzelnen Informationen zu erheben, hätte die differentielle Varianz vielleicht kompensieren können.

#### **4 Ausblick**

Die objektive Evaluation hat Vorteile, die nicht abgestritten werden können, sie kann parallel zu Handlungsabläufen aus den Daten, welche die Nutzer im Zuge ihrer Arbeit generieren, entwickelt werden. Bei einer Vertrauensbasis zwischen Nutzern und Evaluatoren können so Beobachtungseffekte minimiert werden.

Außerdem ist es eine Tatsache, daß subjektive Meinungen nicht zwingend die tatsächlichen Umstände wiedergeben.

Bei einer erneuten Untersuchung der Fragestellung dieser Arbeit sollte die Gestaltung des Datenmodells noch enger mit der Entwicklung der Operationalisierung verknüpft werden, und aus Gründen der Sparsamkeit alle Daten in einem Speichermedium repräsentiert werden. Desweiteren ist es für die Entwickler der Software-Anteile eines solchen Experimentes angenehm eine Programmiersprache zu benutzen die einen echten Debugger zur



Vefügung stellt, und nicht wie im Falle von PhP, die Entwicklung des Source-Codes aus vielen Trial-and-Error-Schleifen besteht.

Wie in der Kritik schon erwähnt sollte das User-Logging, wenn es verwertbare Daten liefern soll, eine viel größere Detailtiefe aufweisen, durch eine Verbindung von Eye-Tracking und Aufzeichnung der Tastenanschläge sowie der Mausbewegungen könnten Wechsel des Aufmerksamkeitsfokus und andere interne Prozesse wenigstens ansatzweise isoliert werden.

Um die Zielstrukturen und differentiellen Unterschiede zu kontrollieren muß eine konsequente deskriptive Analyse der Versuchspersonen vorgenommen werden und die Aufgabenstellungen müssen ökologisch valide sein.

Dies bedeutet auch, das die Teilnehmer intrinsisch motiviert sein, echtes Interesse am Lernstoff vorweisen und vergleichbare Ziele haben müssen.

Eine logische Konsequenz wäre die begleitende Analyse einer echten Lernumgebung, wobei die angesprochenen Probleme dadurch nur reduziert nicht nivelliert werden.

Desweiteren reicht es dann nicht mehr intervenierende Variablen zu erheben, sondern geschichtete Stichproben müssen eingesetzt werden. Unter dem Aspekt der Mortalität und der Vielzahl an Schichtungsvariablen wäre dies ein sehr ehrgeiziges Unternehmen.

Trotz all dieser Bedenken sollte man jedoch nicht aus den Augen verlieren, daß jede andere Methode der Evaluation mit Problemen dieser Größenordnung ringen muß.

**Literatur**

- Aebli, H. (1998). Zwölf Grundformen des Lernens: Eine allgemeine Didaktik auf psychologischer Grundlage. Stuttgart: Klett-Cotta.
- Anderson, J., Lebiere, C. & Matessa, M. (1997). ACT-R: A Theory of higher level cognition and its relation to visual Attention. Human-Computer-Interaction, 12, 439-462.
- Anderson, J. (1990). The adaptive character of thought. Hillsdale, NJ: Erlbaum.
- Barlas, S., Goldstein, W. & Weber, E. (1995). And let us not forget memory: the role of memory processes and techniques in the study of judgement and choice. Psychology of Learning and motivation, 32, 33-136.
- Bettman, J., Johnson, E. & Payne, J. (1993). The adaptive decision maker. Cambridge: Cambridge University Press.
- Blanckertz, H. (1986). Theorien und Modelle der Didaktik. München: Juventa.
- Bortz, J. (1999). Statistik für Sozialwissenschaftler. Berlin: Springer.
- Chomsky, N. (1957). Syntactic Structures. The Hague: Mouton.
- Ciesinger, K. & Piekara, F. (1986). Eine Theorie über das interne und externe Speichern von Information (Bericht Nr. 4). Münster: Universität Münster, Psychologisches Institut III.
- Ciesinger, K., Muthig, K. & Piekara, F. (1987). Notizanfertigen und Behalten. Zeitschrift für Pädagogische Psychologie, 1(4), 267-280.
- Clark, R. (1983). Reconsidering research on learning from media. Review of Educational Research, 53, 445-459.

- Cook, L. & Mayer, R. (1983). Reading strategies for meaningful learning from prose. In M. Presley & J. Leven (Eds.). Cognitive strategy research: Educational Applications(87-131). New York: Springer.
- Doll, W. & Torkzadeh, G. (1988). The measurement of end-user computing satisfaction. MIS Quarterly, 12(2), 259-274.
- Einstein, G., Morris, J. & Smith, S. (1985). Note-Taking, individual differences and memory for lecture information. Journal of Educational Psychology, 77, 522-532.
- Garrod, S. & Sanford, A. (1981). Understanding written language. New York: John Wiley.
- Gigerenzer, G. (2000). Adaptive thinking. Oxford: Oxford University Press.
- Hacker, W. (1994). Arbeits- und organisationspsychologische Grundlagen der Software-Ergonomie. In E. Eberleh (Hrsg.), Einführung in die Software-Ergonomie (53-93). Berlin: deGruyter.
- Hacker, W., Rudolph, E. & Schönfelder, E. (1987). Tätigkeitsbewertungssystem für geistige Arbeit mit/ohne Rechnerunterstützung TBS-GA. Berlin: Göttingen.
- Hasebrook, J. (1995). Multimedia-Psychologie: Eine neue Perspektive menschlicher Kommunikation. Heidelberg: Spektrum Akademischer Verlag.
- Heller, K. & Rosemann, B. (1981). Planung und Auswertung empirischer Untersuchungen. Stuttgart: Klett-Cotta.
- Kerres, M. (2001). Multimediale und telemediale Lernumgebungen. Wien: Oldenburg.

- Kiewra, K. (1985). Investigating Notetaking and Review: A Depth of Processing Alternative. Educational Psychologist, 20(1), 23-32.
- Kintsch, W. & vanDijk, T. (1978). Towards a model of discourse comprehension and production. Psychological Review, 85, 363-394.
- König, E. & Riedel, H. (1979). Unterrichtsplanung I. Konstruktionsgrundlagen und Kriterien (2. Aufl.). Weinheim: Beltz.
- Monsell, S. & Rogers, R. (1995). Costs of predictable task switch between simple cognitive tasks. Journal of Experimental Psychology, 124(2), 207-231.
- Moore, M. (1993). Theory of Transactional distance. In D.Keegan (Hrsg.), Theoretical Principles of distance education (22-38). London: Routledge.
- Morris, P., Ehrmann, S., Goldsmith, R., Howard, K. & Kumar, V. (1994). Valuable viable software in education: Case studies and analysis. New York: Primis Division of MacGraw Hill.
- Newell, A. & Simon, H. (1972). Human Problem Solving. Englewood Cliffs, NJ: Prentice Hall.
- Nuria, V. (2001). Multimediaprodukte für das Internet: Psychologische Grundlagen. München: Oldenburg.
- Oppermann, R. & Reiterer, H. (1994). Software-ergonomische Evaluation. In E. Eberleh (Hrsg.), Einführung in die Software-Ergonomie (335-371). Berlin: deGruyter.

- Schank, R. (1995). What We Learn When We Learn by Doing. (Technical Report No. 60). Northwestern University: Institute for Learning Sciences.
- Seel, N. (2000). Psychologie des Lernens: Lehrbuch für Pädagogen und Psychologen. Basel: E. Reinhardt.
- Skell, W. (1999). Anfragen und Antworten zu Lernen und Handlungsregulation. Forschungsbericht 28. Dresden: Technische Universität Dresden. Institut für Arbeits-, Organisations- und Sozialpsychologie.
- Skinner, B. (1954). The science of Learning and the Art of teaching. In Smith & Moore (Eds.), Programmed Learning. New York.
- Wandmacher, J. (1993). Software-Ergonomie. Berlin: deGruyter.
- Wang , Y. (2003). Assessment of Learner Satisfaction with Asynchronous Electronic Learning Systems. Information & Management, 41, 75-86.
- Weiner, B. (1986). An Attributional Theory of Motivation and Emotion. New York: Springer.
- Winterhoff-Spurk, P. (2004). Medienpsychologie. Stuttgart: Kohlhammer.

## Anhänge

Aus Gründen der Übersichtlichkeit finden sie alle Anhänge, also:

- Php-Skripte
  - basis.php (Lernsystem)
  - questionnaire.php (Verarbeitung des Fragebogens)
- Datenbankstruktur inklusive der Texte
- Rohdaten
  - VI-Logs
  - Fragebögen
  - aus der Datenbank extrahierte User-Logging-Daten
- Instruktionstexte
  - I und (einleitung.html)
  - II (liegt in der Datenbank)
- Fragebogen

auf der beigefügten CD.

Die meisten der Dateien entsprechen dem Unix-Plain-Text-Format.

Sie lassen sich in anderen Betriebssystemen öffnen, allerdings benutzen die verschiedenen Systeme unterschiedliche Marker für das Zeilenende.

Im Falle der Rohdaten, die als HTML gespeichert sind, ist das irrelevant, die Datei gesamt enthält jedoch einen Fall pro Zeile. Um die Daten sinnvoll betrachten zu können müssen die Zeilenenden konvertiert werden.

Bitte fragen Sie Ihren Systemadministrator oder benutzen Sie OpenOffice.

Zusätzlich liegen die Dateien noch einmal in gepackter Form vor. Das Komprimieren gewährt eine höhere Immunität gegenüber Datenfehlern.

## EIDESSTATTLICHE ERKLÄRUNG

Ich versichere hiermit an Eides Statt, dass ich die von mir eingereichte Diplomarbeit bzw. die von mir namentlich gekennzeichneten Teile selbständig verfasst und ausschließlich die angegebenen Hilfsmittel benutzt habe.

Saarbrücken, den 04.08.2004

Björn Mittelsdorf

## EINVERSTÄNDNISERKLÄRUNG

Ich erkläre mich einverstanden/nicht einverstanden, dass meine Diplomarbeit an Personen, die nicht mittelbar oder unmittelbar an meiner Prüfung beteiligt sind, ausgeliehen wird.

Saarbrücken, den 04.08.2004

Björn Mittelsdorf