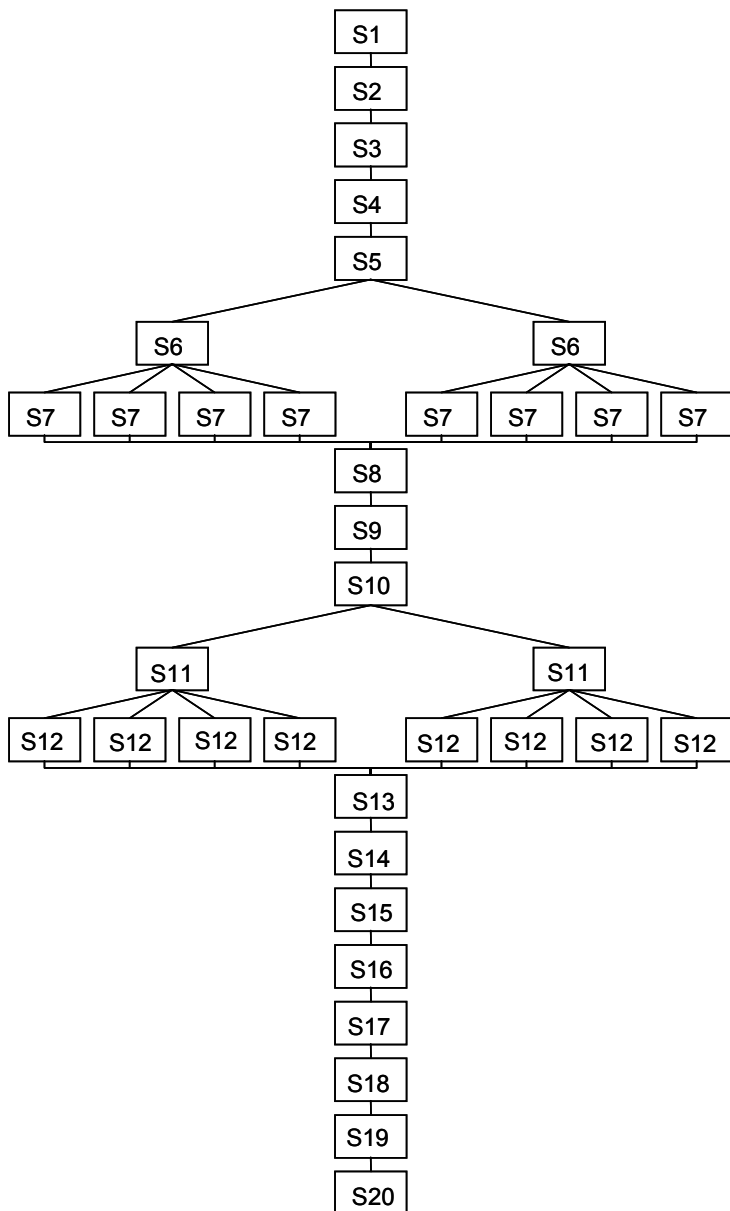


Lars Kaczmirek

## Information und Selektion. Gebrauchstauglichkeit der Ergebnisseiten von Suchmaschinen

---



Veröffentlichung der  
gleichnamigen Diplomarbeit

## **Schriftenreihe: Mannheimer sozialwissenschaftliche Abschlussarbeiten**

### **Vorwort**

Die Fakultät für Sozialwissenschaften der Universität Mannheim bildet in den Fächern Politikwissenschaft, Soziologie, Psychologie und Erziehungswissenschaft qualifizierten Forschungs- und Führungsnachwuchs aus. Viele Abschlussarbeiten der Studierenden zeugen von dem hohen wissenschaftlichen Niveau der Mannheimer Ausbildungsinhalte, die insbesondere gekennzeichnet sind von der empirisch-analytischen Ausrichtung unter Betonung quantitativer Methoden.

Die Ergebnisse und Inhalte vieler dieser Arbeiten sind publikationswürdig. Aus diesem Grund bietet die Fakultät für Sozialwissenschaften ihren besten Absolventen die Möglichkeit, ihre Arbeiten einem breiteren Publikum zu präsentieren und hat hierfür diese Schriftenreihe ins Leben gerufen. Diese Schriftenreihe soll dazu beitragen, die wissenschaftlichen Ergebnisse der besten Abschlussarbeiten dem Fachpublikum zugänglich zu machen. Damit sind sie für weitere Untersuchungen verfügbar und können eventuell eine Grundlage für weitere Forschungen bieten.

In dieser Reihe werden nur Abschlussarbeiten veröffentlicht, die von beiden Gutachtern mit „sehr gut“ bewertet und für veröffentlichungswürdig befunden wurden.

Prof. Dr. Josef Brüderl

Dekan der Fakultät für Sozialwissenschaften

## **Vorwort des Erstgutachters**

Mit dem Wachstum des Internet entstehen täglich neue Informationsquellen. Anwender sehen sich mit der wachsenden Schwierigkeit konfrontiert, die für sie relevanten Informationen zu identifizieren. Suchmaschinen helfen bei der Vermittlung und Auswahl der vielfältigen Angebote, indem sie Listen mit Internetangeboten anhand zuvor angegebener Suchphrasen zusammenstellen. Die letztendliche Entscheidung über die persönliche Relevanz der einzelnen Treffer verbleibt jedoch beim Anwender.

Lars Kaczmirek widmet sich in der vorliegenden Arbeit der Frage, welche Aspekte auf Ergebnisseiten von Suchmaschinen zur Entscheidungsfindung der Anwender beitragen und wie die Auswahl persönlich relevanter Treffer am besten unterstützt werden kann. Er verknüpft dazu psychologische Entscheidungstheorie und Aspekte der Ergonomie mit informationswissenschaftlichen Konzepten zum Suchprozess und Informationsbedürfnissen. In einem komplexen Webexperiment gelangen sowohl die aus den Informationswissenschaften gängigen Kriterien wie Präzision und Vollständigkeit als auch psychologische Variablen wie Entscheidungssicherheit, Zufriedenheit und Entscheidungszeiten zur Anwendung.

Die Diplomarbeit wurde mit dem Otto-Selz-Preis 2004 von den Professorinnen und Professoren der Fachgruppe Psychologie an der Universität Mannheim für herausragende Leistungen verliehen. In der Begründung heißt es dazu: Ein Thema von hoher Aktualität und praktischer Relevanz wurde in innovativer, einfallsreicher Weise bearbeitet, wobei Ergebnisse aus mehreren psychologischen Teildisziplinen und Nachbardisziplinen eingeflossen sind. Die Arbeit zeichnet sich durch hohes methodisches Problembewusstsein, konstruktiv-kritischen Umgang mit der umfassend aufgearbeiteten Primärliteratur, kreative Verknüpfung von Grundlagen- und Anwendungswissenschaft und einen hohen Grad an Interdisziplinarität aus. Auch in empirisch-handwerklicher Hinsicht (Herleitung und Begründung des Untersuchungsdesigns, statistische Analyse und Interpretation der Ergebnisse) liegt die Arbeit weit über dem, was man von einer Diplomarbeit im Fach Psychologie erwarten kann.

Die vorliegende Arbeit leistet damit einen wichtigen Beitrag zur gezielten Erforschung derjenigen Aspekte, die sich für die Anwenderfreundlichkeit bei Ergebnisseiten von Suchmaschinen verantwortlich zeigen. Gleichzeitig kann das entwickelte Rahmenmodell als Grundlage für weitere Forschung zu den verschiedenen Aspekten der Triade Mensch-Maschine-Aufgabe dienen.

Prof. Dr. Walter Bungard

### **Bitte zitieren als**

Kaczmirek, L. (2006). *Information und Selektion. Gebrauchstauglichkeit der Ergebnisseiten von Suchmaschinen*. Mannheim: Universitätsbibliothek Mannheim (MADOC). Verfügbar unter: <http://bibserv7.bib.uni-mannheim.de/madoc/volltexte/2006/1314/>.

URN: [urn:nbn:de:bsz:180-madoc-13140](http://nbn-resolving.org/urn:nbn:de:bsz:180-madoc-13140)

### **Teile der Arbeit sind erschienen als**

Kaczmirek, L. (2003). *Gebrauchstauglichkeit der Ergebnisseiten von Suchmaschinen*. In G. Szwillus, & J. Ziegler (ed.), *Mensch & Computer 2003: Interaktion in Bewegung* (S. 337-347). Stuttgart: B. G. Teubner.

### **Anpassungen**

Dieses Dokument entspricht der gleichnamigen Diplomarbeit. Die Gestaltung wurde jedoch überarbeitet, Fehler entfernt, der Anhang um einige Erläuterungen erweitert, sowie ein Autorenverzeichnis hinzugefügt.

## **Zusammenfassung**

Bei Ergebnisseiten von Suchmaschinen nutzen Anwender die Informationen in den Linkbeschreibungen, um zu entscheiden, ob ein Link für die Suche relevant ist und weiterverfolgt werden soll. Die Linkbeschreibung ist damit ein Kernelement im Selektionsprozess der Suche. Bisher ungelöst ist die Frage, welcher Text aus den verlinkten Dokumenten zur Beschreibung am besten geeignet ist.

Ein neues Modell für eine systematische und differenzierte Forschung über Ergebnisseiten von Suchmaschinen wird vorgestellt. Dazu werden die drei Faktoren Mensch, Maschine, Aufgabe und ihre Implikationen erläutert. Berücksichtigt werden neben den im Informationretrieval gängigen Kriterien Präzision und Vollständigkeit auch die Sicherheit, mit der Entscheidungen getroffen werden, die Zufriedenheit, sowie Entscheidungszeiten.

In einem Web-Experiment wurden die Linkbeschreibungen zu den verlinkten Internetseiten und die Spezifität der Suchaufgabe variiert. Die Linkbeschreibungen entsprachen den aktuell konkurrierenden Ansätzen: (a) die Suchwörter im Kontext des Vorkommens im Dokument darzustellen, (b) Text vom Anfang des Dokuments darzustellen und (c) Schlüsselwörter zur Beschreibung zu verwenden.

Teilnehmer trafen Selektionsentscheidungen auf Ergebnisseiten anhand verschiedener Suchaufgaben mit anschließenden Bewertungen ihrer Entscheidungen und der Ergebnisseiten. Die Daten von 191 Personen erlaubten einen Vergleich der Gebrauchstauglichkeit verschiedener Linkbeschreibungen, sowie Aussagen zu den Effekten durch unterschiedliche Aufgabenschwierigkeiten.

Die Ergebnisse zeigen, dass eine Linkbeschreibung durch 'Anfangstext mit 400 Zeichen redaktionellem Inhalt' den anderen Bedingungen überlegen ist und die Entscheidung am besten unterstützt. Es wird daher vorgeschlagen, dass Suchmaschinen zur Beschreibung der Links redaktionellen, zusammenhängenden Text verwenden. Dabei sollte die Menge auf die Bedürfnisse des Anwenders variabel einstellbar sein.

## **Abstract**

When searching the web with search engines, users apply information provided by the link descriptions of result pages to decide whether the link is relevant for their search and should be selected. The link description is therefore a key element in the selection process of the search. However, the question of how the link description should be structured to allow an efficient selection process presents a challenge.

This thesis puts forth a new framework for a systematic analysis of result pages which integrates the three main factors relevant for a differentiated research: user, machine and task. Besides the criteria precision and recall, which are well known in information retrieval, it also considers confidence in decisions, satisfaction and time needed for selection.

In a web experiment based on the proposed framework, effects of different link descriptions and task specificity were examined. The link descriptions were composed according to the following three currently competing strategies: (a) keywords in context, (b) the beginning text of a document and (c) essential keywords derived from a document.

Participants decided which links to select on the basis of different information needs and then evaluated their decisions and the viewed result pages. The data provided by 191 participants permitted an evaluation of the different kinds of link descriptions and of the effects caused by varying the difficulty of the search task.

The results show that a link description composed of the first 400 characters of the content of a document is superior to the other representations, and therefore optimally supports the user's decision. Based on these findings, the thesis proposes a link description that uses continuous text as a method to summarise a document. The user should be able to control the amount of displayed text according to his or her needs.

Mannheimer sozialwissenschaftliche Abschlussarbeiten, Nr. 003/2006

Universität Mannheim

# Information und Selektion

## Gebrauchstauglichkeit der Ergebnisseiten von Suchmaschinen

Diplomarbeit

im Fachbereich Psychologie  
am Lehrstuhl für Wirtschafts- und Organisationspsychologie

vorgelegt bei Prof. Dr. Walter Bungard

betreut von Dipl.-Psych. Christian Liebig und  
Dipl.-Psych. Hans-Joachim Schütze

vorgelegt von Lars Kaczmirek  
Augartenstr. 83  
68165 Mannheim  
lars@kaczmirek.de

Mannheim, Februar 2003

*Here do I choose, and thrive I as I may.*

*William Shakespeare,  
The Merchant of Venice*

## **Danksagung**

Mein herzlicher Dank gebührt meinen Eltern, die mir das Studium ermöglichten. Ihnen ist diese Arbeit gewidmet.

Besonderer Dank gilt Sarah Heinz für die richtigen Worte zur richtigen Zeit.

Ebenso möchte ich mich bei meinen Betreuern Christian Liebig und Hans-Joachim Schütze für die stets konstruktiven Gespräche bedanken und dafür, dass sie mir die Konzeption und Durchführung dieses Webexperiments ohne externen Auftraggeber gestatteten.

Bedanken möchte ich mich weiterhin bei allen Personen, die mit einem Link von ihrer Webseite auf meine Untersuchung aufmerksam gemacht haben. Auch den vielen Teilnehmern gilt mein Dank. Ohne sie wäre diese Untersuchung nicht möglich gewesen.



## **Erklärungen für den Prüfungsausschuss und die Institutsbibliothek**

Ich versichere, dass ich die beiliegende Diplomarbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Ich stimme ausdrücklich zu, dass meine durch Christian Liebig und Hans-Jochachim Schütze betreute Diplomarbeit mit dem Titel „Information und Selektion. Gebrauchstauglichkeit der Ergebnisseiten von Suchmaschinen“ nach Beendigung der Diplom-Prüfung zu wissenschaftlichen Zwecken im Bereich der Fakultätsbibliothek PE aufgestellt und zugänglich gemacht wird (Veröffentlichungen nach §§6 Abs. 1 UrhG), und hieraus im Rahmen des § 51 UrhG zitiert werden kann. Sämtliche Verwertungsrechte nach § 15 UrhG verbleiben beim Verfasser der Diplomarbeit.

Lars Kaczmirek

Mannheim im Februar 2003

# Inhalt

<b>1</b>	<b>EINLEITUNG</b>	<b>1</b>
1.1	Hintergrund der Arbeit	1
1.2	Zielsetzung und Forschungsfragen	2
1.3	Aufbau der Arbeit	3
<b>2</b>	<b>RAHMENMODELL MENSCH MASCHINE AUFGABE</b>	<b>4</b>
2.1	Überblick und Komponenten	4
2.2	<b>Maschine: Informationspräsentation bei Suchmaschinen</b>	<b>7</b>
2.2.1	Information Retrieval	7
2.2.2	Suchmaschinen im World Wide Web	9
2.2.3	Präsentation von Suchergebnissen	10
2.2.3.1	Signifikante Sätze	12
2.2.3.2	Signifikante Wörter	12
2.2.3.3	Informationen aus HTML-Tags	13
2.2.3.4	Zusätzliche Informationen	13
2.2.3.5	Klassifikation von Dokumenten	13
2.2.3.6	Zitierende Dokumente	14
2.2.3.7	Informationen aus Datenbanken	14
2.2.3.8	Fazit	15
2.2.4	Geordnete Listen zur Repräsentation von Informationen (Ergebnisseiten)	15
2.2.5	Zusammenfassung und Ableitungen	18
2.3	<b>Aufgabe: Klassifikation und Verhaltensanalyse</b>	<b>18</b>
2.3.1	<i>Klassifikation von Informationsbedürfnissen</i>	19
2.3.1.1	Drei Dimensionen bei Marchionini	19
2.3.1.2	Zwei Dimensionen bei Navarro-Prieto, Scaife und Rogers	20
2.3.1.3	Eindimensionale Ansätze	21
2.3.2	Prozess der Suche	22
2.3.3	Hierarchische Aufgabenanalyse und Task Description	24
2.3.4	Anwendung von Task Description auf Ergebnisseiten	25
2.3.5	Zusammenfassung und Ableitungen	26
2.4	<b>Mensch: Entscheidungspsychologie und Ergonomie</b>	<b>27</b>
2.4.1	Die Selektionsentscheidung	27
2.4.2	Zum Zusammenhang von Verhalten und Kognitionen	30
2.4.3	Die Bedeutung von Wissen	31
2.4.4	Ergonomie	34
2.4.4.1	Aufgabenangemessenheit	34
2.4.4.2	Selbstbeschreibungsfähigkeit	35
2.4.4.3	Steuerbarkeit	36
2.4.4.4	Erwartungskonformität	36
2.4.4.5	Fehlertoleranz	36

---

2.4.4.6	Individualisierbarkeit	37
2.4.4.7	Lernförderlichkeit	37
2.4.4.8	Kritische Würdigung	37
2.4.5	Zusammenfassung	38
<b>2.5</b>	<b>Aktuelle Forschungsparadigmen</b>	<b>39</b>
<b>3</b>	<b>METHODEN DER EMPIRISCHEN UNTERSUCHUNG</b>	<b>42</b>
<b>3.1</b>	<b>Untersuchungsdesign</b>	<b>42</b>
3.1.1	Untersuchungsparadigma	42
3.1.2	Operationalisierung der Faktoren Beschreibung und Spezifität	44
3.1.3	Aufbau und Ablauf der Studie	46
3.1.3.1	Allgemeine Designprinzipien	46
3.1.3.2	Eingangsseite und Eingangsfragebogen	48
3.1.3.3	Übungsdurchgänge	48
3.1.3.4	Ergebnisseiten	50
3.1.3.5	Einschätzungen der Treffer	51
3.1.3.6	Bewertungen der Ergebnisseiten	51
3.1.3.7	Direkter Vergleich der Beschreibungsvarianten	51
3.1.3.8	Abschluss der Untersuchung	52
3.1.4	Exkurs Technische Umsetzung	52
3.1.5	Vortests, Durchführung und Ableitungen	54
3.1.5.1	Beobachtete Teilnahme mit lautem Denken	54
3.1.5.2	Gruppendiskussion	54
3.1.5.3	Feldtest	55
3.1.5.4	Designänderungen	55
3.1.6	Anforderungen an die Gebrauchstauglichkeit	56
3.1.6.1	Effektivität	56
3.1.6.2	Effizienz	57
3.1.6.3	Zufriedenheit	58
3.1.7	Fazit	59
<b>3.2</b>	<b>Forschungsleitende Fragen</b>	<b>60</b>
3.2.1	Maschine	60
3.2.2	Aufgabe	60
3.2.3	Mensch	60
<b>3.3</b>	<b>Auswertungsverfahren</b>	<b>60</b>
<b>4</b>	<b>ERGEBNISSE</b>	<b>62</b>
<b>4.1</b>	<b>Überprüfung der Daten und der Implementation</b>	<b>62</b>
4.1.1	Relevanz, Prüfen der Treffertauglichkeit	62
4.1.2	Unvollständige Angaben (Missings)	64
4.1.3	Dropouts, Prüfen der Randomisierung	65
4.1.4	Manipulationsüberprüfung (Treatmentcheck)	68
4.1.5	Reihenfolgeeffekte	68

---

4.1.6	Probleme in den Übungsdurchgängen	68
<b>4.2</b>	<b>Beschreibung der Stichprobe</b>	<b>69</b>
<b>4.3</b>	<b>Effekte in der Ebene Maschine und Aufgabe</b>	<b>72</b>
4.3.1	Genauigkeit	72
4.3.2	Vollständigkeit	73
4.3.3	Erfolgsrate	74
4.3.4	Entscheidungszeit	74
4.3.5	Zufriedenheit	75
4.3.6	Entscheidungssicherheit	76
4.3.7	Direkter Vergleich	77
<b>4.4</b>	<b>Moderierender Einfluss der Ebene Mensch</b>	<b>78</b>
<b>4.5</b>	<b>Skalenbildung Zufriedenheit</b>	<b>79</b>
<b>4.6</b>	<b>Zusammenfassende Interpretation</b>	<b>80</b>
<b>5</b>	<b>DISKUSSION</b>	<b>82</b>
<b>5.1</b>	<b>Gesamtfazit</b>	<b>83</b>
<b>5.2</b>	<b>Anmerkungen zum Design</b>	<b>83</b>
<b>5.3</b>	<b>Teststärkenanalyse</b>	<b>84</b>
<b>5.4</b>	<b>Validität der Untersuchung</b>	<b>86</b>
<b>5.5</b>	<b>Implikationen für die Praxis</b>	<b>87</b>
<b>5.6</b>	<b>Ausblick</b>	<b>88</b>
<b>6</b>	<b>LITERATUR</b>	<b>90</b>
<b>Anhang A:</b>	<b>Mittelwerte der Varianzanalysen</b>	<b>99</b>
<b>Anhang B:</b>	<b>Korrelationen zur Ebene Mensch</b>	<b>103</b>
<b>Anhang C:</b>	<b>t-Tests zu systematischen Verzerrungen bei Dropouts</b>	<b>106</b>
<b>Anhang D:</b>	<b>Zitierte Suchmaschinen und Webangebote</b>	<b>107</b>
<b>Anhang E:</b>	<b>Texte zur Teilnehmerrekrutierung</b>	<b>108</b>
<b>Anhang F:</b>	<b>Webangebote, die auf die Untersuchung verwiesen</b>	<b>109</b>
<b>Anhang G:</b>	<b>Teststärkeberechnungen</b>	<b>111</b>
<b>Anhang H:</b>	<b>Webseiten der Voruntersuchungen</b>	<b>112</b>
<b>Anhang I:</b>	<b>Webseiten der Hauptuntersuchung</b>	<b>126</b>
<b>GLOSSAR</b>		<b>148</b>
<b>AUTORENVERZEICHNIS</b>		<b>149</b>

## Tabellen

<b>Tabelle 1:</b> Vergleich von Untersuchungen über Ergebnisseiten mit einer Einordnung in die Ebenen Mensch, Maschine und Aufgabe.....	40
<b>Tabelle 2:</b> Entscheidungsmatrix.....	56
<b>Tabelle 3:</b> Variablen im Untersuchungsdesign. ....	59
<b>Tabelle 4:</b> T-Test für eine Stichprobe zur Prüfung der Treffertauglichkeit.....	64
<b>Tabelle 5:</b> Little's MCAR Test zur Prüfung auf zufällige Ausfallmuster und prozentuale Anzahl fehlender Werte. ....	65
<b>Tabelle 6:</b> Häufigkeiten in den einzelnen Faktorstufen und zwischen Teilnehmern und Abbrechern..	68
<b>Tabelle 7:</b> Korrektheit der Angaben in den Übungsdurchgängen zur Selektion.....	69
<b>Tabelle 8:</b> Beschreibung der Stichprobe.....	70
<b>Tabelle 9:</b> Verwendete Bildschirmauflösung und Browser. ....	70
<b>Tabelle 10:</b> Genauigkeit. ....	72
<b>Tabelle 11:</b> Vollständigkeit. ....	73
<b>Tabelle 12:</b> Erfolgsrate.....	74
<b>Tabelle 13:</b> Entscheidungszeiten.....	75
<b>Tabelle 14:</b> Zufriedenheit. Multivariater Test mit Wilks $\lambda$ . ....	76
<b>Tabelle 15:</b> Entscheidungssicherheit. ....	77
<b>Tabelle 16:</b> Direkter Vergleich aller vier Beschreibungsvarianten. Multivariater Test mit Wilks $\lambda$ . ....	78
<b>Tabelle 17:</b> Differenzierungsfähigkeit der Items zur Zufriedenheit. ....	80
<b>Tabelle 18:</b> Teststärken bei $\alpha = 0,05$ ; $N=191$ ; Effektstärke $f^2 = 0,01$ . ....	85
<b>Tabelle 19:</b> Beobachtete Mittelwerte für Genauigkeit.....	99
<b>Tabelle 20:</b> Beobachtete Mittelwerte für Vollständigkeit.....	99
<b>Tabelle 21:</b> Beobachtete Mittelwerte für Entscheidungssicherheit.....	100
<b>Tabelle 22:</b> Beobachtete Mittelwerte der Entscheidungszeiten.....	100
<b>Tabelle 23:</b> Beobachtete Mittelwerte Erfolgsrate.....	101
<b>Tabelle 24:</b> Beobachtete Mittelwerte Zufriedenheit. ....	101
<b>Tabelle 25:</b> Geschätzte Mittelwerte des direkten Vergleichs aller vier Bedingungen im Faktor Beschreibung.....	101
<b>Tabelle 26:</b> Geschätzte Mittelwerte der Zufriedenheit aller Items. ....	102

## Abbildungen

<b>Abbildung 1:</b> Rahmenmodell zur Erforschung der Gebrauchstauglichkeit der Ergebnisseiten von Suchmaschinen. ....	5
<b>Abbildung 2:</b> Modell einer Suchmaschine im Zusammenspiel mit dem Internet und Anwendern. ....	9
<b>Abbildung 3:</b> Klassifikationsschema zu den Möglichkeiten der Präsentation von Ergebnismengen nach Drori (2000a). ....	11
<b>Abbildung 4:</b> Aufbau einer Ergebnisseite am Beispiel <i>Altavista</i> mit den Suchwörtern „Mannheim Psychologie“ (gekürzt). ....	16
<b>Abbildung 5:</b> Eigenschaften einzelner Treffer einer Ergebnisseite am Beispiel <i>Excite</i> mit den Suchwörtern „Skinner Box“. ....	18
<b>Abbildung 6:</b> Flussdiagramm zu linearem und hierarchischem Suchverhalten auf Ergebnisseiten....	26
<b>Abbildung 7:</b> Integratives Modell zum Selektionsverhalten auf Ergebnisseiten. ....	31
<b>Abbildung 8:</b> Ablauf der Untersuchung. ....	47
<b>Abbildung 9:</b> Auswirkungen mehrdeutiger Treffer auf die Kriterien. ....	63
<b>Abbildung 10:</b> Dropoutanalyse. ....	66
<b>Abbildung 11:</b> Anwerbung. ....	71
<b>Abbildung 12:</b> Genauigkeit hinsichtlich Beschreibung und Spezifität. ....	73
<b>Abbildung 13:</b> Vollständigkeit hinsichtlich Beschreibung und Spezifität. ....	73
<b>Abbildung 14:</b> Erfolgsrate hinsichtlich Beschreibung und Spezifität. ....	74
<b>Abbildung 15:</b> Entscheidungszeiten hinsichtlich Beschreibung und Spezifität in Sekunden pro Webseite. ....	75
<b>Abbildung 16:</b> Zufriedenheit hinsichtlich Beschreibung und Spezifität. ....	76
<b>Abbildung 17:</b> Entscheidungssicherheit hinsichtlich Beschreibung und Spezifität. ....	77
<b>Abbildung 18:</b> Direkter Vergleich aller vier Beschreibungsvarianten getrennt nach Trefferrelevanz. .	78

# 1 Einleitung

Das Internet bietet eine Fülle von Informationen. Sein stetiges Wachstum führt zu täglich neuen Angeboten, die online abgerufen werden können. Die Qualität der Angebote ist allerdings sehr verschieden, da es mit der Hilfe von Webpublishing-Programmen heutzutage relativ einfach ist, beliebige Inhalte online zu publizieren. So lassen sich dort sowohl wissenschaftliche Abhandlungen zu einer Menge von Themen als auch viele persönliche Meinungen und Erfahrungsberichte finden (Tillmann, 2000). Auch Unternehmen nutzen das Internet als Werbepattform, zur Selbstdarstellung und zum Verkauf ihrer Produkte. Generell kann jede Institution oder Privatperson Inhalte publizieren, die sie für die Allgemeinheit oder einen ausgewählten Kreis weltweit zur Verfügung stellen möchte. Das Internet stellt damit eine ungeheuer große Ansammlung von Daten zur Verfügung, aus der sich jeder bedienen kann. Dabei sind allerdings einige Hürden zu überwinden: Wie lässt sich dasjenige Angebot zielsicher aus der unüberschaubaren Menge von Webseiten herausfinden, das ein Anwender benötigt? Welche Techniken kann er zum Finden von Informationen anwenden? Wie setzt er diese optimal ein?

Suchmaschinen bieten solch eine Möglichkeit, interessierende Webangebote zu finden. Anwender können ihre Suchanfrage formulieren und erhalten eine Menge von Ergebnisseiten, auf denen Webangebote aufgeführt sind. Diese Auswahl basiert dabei immer auf den in der Suchanfrage verwendeten Wörtern.

## 1.1 Hintergrund der Arbeit

Ein aktueller Bericht des Pew Internet Project zeigt eindrucksvoll die große Bedeutung von Suchmaschinen für den Beruf und Alltag von Internetnutzern auf. Demnach haben acht von zehn Amerikanern bereits Suchmaschinen zum Finden von Informationen verwendet und mehr als jeder Vierte, d.h. 33 Millionen Menschen, verwenden täglich Suchmaschinen (Fox, 2002). Nach einer repräsentativen Umfrage der Forschungsgruppe Wahlen (2002) verfügt in Deutschland jeder zweite Erwachsene über einen Internetanschluss und hat damit Zugang zu einer enormen Datenmenge.

Mit Hilfe von Suchmaschinen lässt sich allerdings nur ein Teil der Dienste des Internets durchsuchen. Von diesen Diensten ist wiederum nur ein Teil von Suchmaschinen erfasst (Lawrence & Giles, 1999). Das World Wide Web (WWW) ist wohl der bekannteste Dienst, in dem Webseiten, die beliebige Inhalte darstellen können, durch Hyperlinks miteinander verknüpft sind (Nielsen, 1990). Die Webseiten werden mit Hilfe von Browsern (z.B. MS Explorer, Netscape, Opera) betrachtet. Aber auch Daten in Newsgroups (ein Nachrichtensystem) und Dateien auf FTP-Servern (FTP bezeichnet ein Protokoll zur Datenübertragung) können mittlerweile durch Browser dargestellt und von Suchmaschinen gefunden werden. Trotz dieser Einschränkungen beträgt allein die Anzahl an WWW-Dokumenten, die mit Suchmaschinen erreicht werden können, Größenordnungen von annähernd einer Milliarde Webseiten (Karzauninkat, 1998a).

Bei der Suche nach Informationen können allerdings auch alternative Wege beschritten werden. Kataloge bzw. Indizes bieten nach Kategorien vorsortierte Linklisten zu einer enormen Menge an Themen. Diese eignen sich besonders, wenn Anwender sich einen Überblick über eine Fragestellung verschaffen möchten. Sie können aber ebenfalls als Ausgangspunkt zum freien Browsen verwendet werden. Bei einem konkreten Informationsbedarf (z.B. zu einem aktuellen politischen Ereignis) bietet sich jedoch der Einsatz von Suchmaschinen an (RRZN, 2000).

Die Forschung in diesem Bereich zeichnet sich durch eine hohe Interdisziplinarität aus, wobei die Schwerpunkte kurz angerissen werden sollen: Einerseits kann die Funktionsweise von Suchmaschinen selber Gegenstand von Untersuchungen sein, die im Bereich des Information Retrieval mit wesentlichen Überschneidungspunkten zur Informatik behandelt werden (vgl. Baeza-Yates & Ribeiro-Neto, 1999). Rankingalgorithmen der Links auf Ergebnisseiten von Suchmaschinen weisen Überschneidungen mit der Mathematik auf (z.B. Kleinberg, 1998). Dem Such- und Entscheidungsverhalten von Anwendern wird ebenfalls ein großes Interesse gewidmet und von den Bereichen der Medien- und Kommunikationswissenschaften und der Bibliotheks- und Informationswissenschaften behandelt (vgl. Choo, Detlor & Turnbull, 1999; Cooke, 1999; Jansen & Pooch, 2000; Wirth & Schweiger, 1999b).

Die Sortierung bzw. das Ranking der Trefferliste durch Suchmaschinen entsprechen nicht immer derjenigen Reihenfolge, die auch durch Anwender gewünscht wird. Auf einer Ergebnisseite sehen sich Anwender demnach der Herausforderung gegenüber herauszufinden, welche Links sie auswählen sollen. Die Texte in den Beschreibungen der Links helfen Anwendern dabei zu entscheiden, ob ein Link für die Suche relevant ist oder ignoriert werden sollte. Dabei bestehen zahlreiche Möglichkeiten, solche Beschreibungen auf verschiedene Weise zu gestalten (Drori, 2000b). Auf diesem Gebiet fanden bisher nur wenige Untersuchungen mit sehr unterschiedlichem Design statt (z.B. Drori, 2001; Lergier & Resnick, 2001; Unz, Capstick & Erbach, 1999). Dies erstaunt umso mehr, als täglich Millionen von Anwendern mindestens ebenso viele Ergebnisseiten betrachten und evaluieren (Karzauninkat, 1998a). Die vorliegende Arbeit möchte daher aufgrund der Aktualität und dürftigen Forschungslage auf diesem Gebiet einen psychologischen Beitrag zur Forschung über Ergebnisseiten von Suchmaschinen leisten.

## **1.2 Zielsetzung und Forschungsfragen**

Die Beschreibungen der Links bzw. Treffer der Ergebnisseiten bieten eine Grundlage für die Entscheidungen der Anwender, welche Treffer als relevant oder irrelevant einzustufen sind. Unklar ist bisher, welche Texte solch eine Entscheidung am besten unterstützen. Diese sollen einerseits kurz sein, um schnelle Entscheidungen bezüglich der Relevanz zu ermöglichen, und andererseits lang genug, um Zusammenhänge mit dem Gesuchten und dem verlinkten Dokument erkennen zu lassen. Die Befundlage auf diesem Gebiet ist uneinheitlich, spezielle methodische Standards zur Untersuchung existieren nicht.



Daraus ergeben sich mehrere Forschungsfragen und -aufgaben. Um eine systematische und differenzierte Forschung über die Gebrauchstauglichkeit der Ergebnisseiten von Suchmaschinen zu ermöglichen, wird ein Rahmenmodell benötigt. Dieses ermöglicht die Einordnung vorhandener Untersuchungen und die Ableitung weiterer Forschungsvorhaben. Inhaltlich stellt sich die Frage nach der Bedeutung verschiedener Arten von Beschreibungen für den Entscheidungsprozess. Da die Suchanfrage direkt mit dem Suchprozess verbunden ist, kann sinnvollerweise auch nach dem Einfluss von verschiedenen Suchaufgaben auf den Entscheidungsprozess gefragt werden.

Die vorliegende Arbeit untersucht die folgenden Leitfragen, die sich in diesem Zusammenhang ergeben:

1. Welche Konstrukte bzw. Variablen beeinflussen das Entscheidungsverhalten auf Ergebnisseiten?
2. Welche Untersuchungsdesigns eignen sich für den Vergleich verschiedener Textbeschreibungen?
3. Wie sollte eine Beschreibung von Treffern gestaltet sein, damit diese die Entscheidungen der Anwender optimal unterstützt?
4. Welchen Einfluss haben Unterschiede in den Suchaufgaben auf die Güte der Entscheidungen?
5. Welchen Einfluss haben Unterschiede zwischen den Anwendern auf die Güte der Entscheidungen?

Die vorliegende Arbeit möchte einerseits ein theoretisches Rahmenmodell für die weitere Forschung liefern und andererseits ein experimentelles Untersuchungsdesign entwickeln und erproben, um den inhaltlichen Fragen nachzugehen.

### **1.3 Aufbau der Arbeit**

Die vorliegende Arbeit gliedert sich den Forschungsfragen entsprechend in zwei Teile. Im theoretischen Teil der Arbeit (Kap. 2) wird ein Rahmenmodell für die Erforschung zur Gebrauchstauglichkeit (usability) der Ergebnisseiten von Suchmaschinen herausgearbeitet. Das Ergebnis ist ein Rahmenmodell, für das sich die drei Hauptkomponenten Mensch, Maschine und Aufgabe identifizieren lassen. Zunächst wird ein Überblick über den Aufbau und die Komponenten des Rahmenmodells gegeben (Kap. 2.1). Danach werden die einzelnen Komponenten beschrieben, die im Folgenden auch als Ebenen bezeichnet werden.

Das Kapitel zur Ebene Maschine (Kap. 2.2) befasst sich mit der Funktionsweise von Suchmaschinen, und ein Klassifikationsschema zu verschiedenen Varianten von Trefferbeschreibungen und der Aufbau von Ergebnisseiten werden dargestellt. Das anschließend folgende Kapitel zur Ebene Aufgabe (Kap. 2.3) beschreibt, worin sich Suchaufgaben unterscheiden und wie ein typischer Suchprozess vorzustellen ist. Weiterhin wird darauf eingegangen, welches Suchverhalten sich auf Ergebnisseiten beobachten lässt. Das Kapitel zur Ebene Mensch (Kap. 2.4) stellt verschiedene theoretische Ansätze zum Entscheidungsverhalten dar und zieht die Verbindung zum beobachtbaren Verhalten des

vorherigen Kapitels. In einem eigenen Abschnitt wird auf die Bedeutung von Wissen für eine erfolgreiche Suche eingegangen. Anschließend folgt eine Interpretation und kurze Darstellung der EN ISO 9241-10 auf Ergebnisseiten hin. Abgeschlossen wird der theoretische Teil durch einen Überblick über die bisherigen Untersuchungen zu Beschreibungsvarianten von Treffern (Kap. 2.5).

Im zweiten Teil der Arbeit (Kap. 3) wird das Rahmenmodell in einem Webexperiment umgesetzt, um die inhaltlichen Fragestellungen zu verfolgen. Einen großen Raum nimmt hierbei die Beschreibung des Untersuchungsdesigns mit seinen unabhängigen und abhängigen Variablen ein (Kap. 3.1). Nach der Entwicklung von operationalisierbaren Fragen für die Untersuchung (Kap. 3.2) folgt ein kurzer Abschnitt über die eingesetzten Auswertungsverfahren (Kap. 3.3).

Im Ergebnisteil der Arbeit (Kap. 4) stellt die Überprüfung der korrekten Operationalisierung einen wichtigen Aspekt dar (Kap. 4.1), bevor dann die Ergebnisse zusammen mit einer Stichprobenbeschreibung berichtet werden (Kap. 4.2, 4.3 und 4.4). Abgeschlossen wird der Ergebnisteil durch Vorschläge zu einer möglichen Skalenbildung zur Zufriedenheit (Kap. 4.5).

Der letzte Teil der Arbeit wird durch die Diskussion gebildet (Kap. 4.6), in der zunächst ein Gesamtfazit gezogen wird. Danach folgen Ausführungen über die inhaltliche Bedeutsamkeit der Ergebnisse und der Generalisierbarkeit. Implikationen für die Praxis werden betrachtet und in einem Ausblick zukünftige Forschungsvorhaben skizziert.

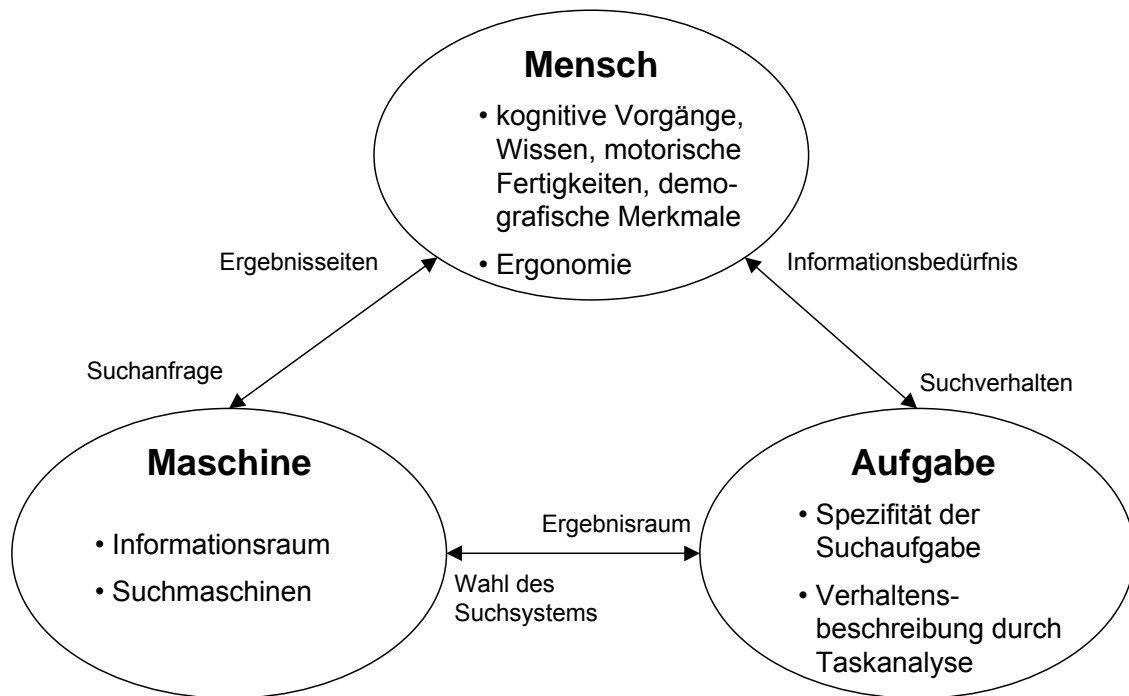
## 2 Rahmenmodell Mensch Maschine Aufgabe

### 2.1 Überblick und Komponenten

Um ein systematisches Vorgehen bei einer empirischen Untersuchung zu gewährleisten, ist es notwendig, dass ein theoretischer Rahmen existiert. Er hilft, die Facetten einer Untersuchung zu begründen, zeigt auf, welche Schwerpunkte sich setzen lassen und wie die Struktur des zu untersuchenden Objekts beschaffen ist. Im Theorieteil der vorliegenden Arbeit wird solch ein Rahmenmodell zur Gebrauchstauglichkeit der Ergebnisseiten bei Suchmaschinen entwickelt (Abbildung 1). Um die späteren Ausführungen leichter einordnen zu können, werden die Schlussfolgerungen aus der Literaturrecherche, wie sie sich zu einem Rahmenmodell zusammenfügen, an den Anfang der Arbeit gestellt.

Eine Analyse der Literatur zur Gebrauchstauglichkeit (usability) von Suchmaschinen und den Standardwerken zur Mensch-Maschine-Interaktion lässt drei zentrale Ebenen erkennen, die als die größten Einflussfaktoren in der Forschung zu Suchmaschinen gelten können. Diese werden im Folgenden mit *Mensch* (user), *Maschine* (system/machine) und *Aufgabe* (task) bezeichnet (siehe Abbildung 1). Die Metaphorik von Ebenen bietet sich an, da sich diese einerseits durch zwei Dimensionen auszeichnen, die ihre Entsprechung im weiteren Modell finden, und andererseits eine Anordnung in einem mehrdimensionalen Raum

ermöglichen, der Überschneidungen zwischen ihnen zulässt. Auch ist es möglich, die zu Ergebnisseiten gehörenden Objekte und Prozesse durch die verschiedenen Ebenen zu betrachten. Dadurch werden im Folgenden unterschiedliche Schwerpunkte gesetzt und verschiedene Ableitungen ermöglicht.



**Abbildung 1:** Rahmenmodell zur Erforschung der Gebrauchstauglichkeit der Ergebnisseiten von Suchmaschinen.

Die Ebene Mensch beinhaltet eine Struktur von Objekten, die Beschreibungen der Anwender selbst ermöglicht. Eine Dimension enthält dabei diejenigen Variablen, durch die sich *Unterschiede zwischen Anwendern* beschreiben lassen. Dies sind insbesondere kognitive Strategien und Prozesse beim Umgang mit Ergebnisseiten. Anwender unterscheiden sich außerdem hinsichtlich ihres Wissens über Suchmaschinen, das Internet und den Suchgegenstand. Darüber hinaus finden sich Unterschiede in den motorischen Fertigkeiten, zum Beispiel die Sicherheit, mit der die Maus bedient wird, sowie in allgemeinen demografischen Merkmalen wie Alter, Geschlecht usw. Die zweite Dimension enthält *allgemeinpsychologische und psychophysiologische Gesetzmäßigkeiten*, wie sie für alle Menschen gelten. Diese münden in Ergonomierichtlinien, aus denen wiederum allgemeine Strukturen für Ergebnisseiten abgeleitet werden können.

Die zweite Ebene, Maschine, umfasst allgemein betrachtet das gesamte System, auf dem und mit dem die Suche durchgeführt wird. Erstens ist dies der *Informationsraum*, in dem die Suche erfolgt. Dies entspricht der Struktur des Internets mit all seinen Webdokumenten und Angeboten. Zweitens enthält sie eine Beschreibung des Werkzeugs, mit dem gesucht wird, und befasst sich daher mit *Funktion und Aufbau von Suchmaschinen*. Dies schließt auch die formalen Eigenschaften zur Definition von Ergebnisseiten sowie die Dialoge, mit deren Hilfe Anwender mit Suchmaschinen interagieren, mit ein.

Die Ebene Aufgabe schließlich beinhaltet alle mit der Suchaufgabe im direkten Zusammenhang stehenden Aspekte von Ergebnisseiten. Dies ist zunächst die Klassifizierung von Suchaufgaben selbst, die allgemeiner auch als Informationsbedürfnisse aufgefasst werden können. Hierbei spielt insbesondere eine Rolle, wie breit bzw. wie spezifisch eine Suchaufgabe ausfällt (*Spezifität*). Der zweite Aspekt betrifft die *Aufgabenanalyse* (task analysis), womit das gezeigte Suchverhalten in elementare Bestandteile zerlegt und so eine genauere Betrachtung der Anforderungen an Ergebnisseiten ermöglicht wird.

Wie oben angedeutet stehen die verschiedenen Ebenen aber nicht isoliert neben-, sondern auch in Beziehung zueinander (Abbildung 1). Zu Beginn einer Suche stellen Anwender zunächst einen Mangel an Informationen fest. Dieses Informationsbedürfnis ergibt die Suchaufgabe und führt zu Suchverhalten. Ein Suchsystem muss gewählt werden. Der von einer Suchmaschine erfasste Teil des Internets bildet den Suchraum, in dem mit dieser Suchmaschine gesucht werden kann. Die Suchmaschine weist einen bestimmten Deckungsgrad auf. Der Ergebnisraum ist wiederum eine Teilmenge davon und entspricht den für die Suche relevanten Webangeboten. Anwender werden nach der Wahl des Suchsystems eine Suchanfrage stellen und Ergebnisseiten zurück erhalten. Durch die Interaktion mit den Ergebnisseiten zeigen sie wiederum ein Suchverhalten.

Bevor mit der Vertiefung des Rahmenmodells begonnen wird, sollen noch andere Rahmenmodelle mit zunächst scheinbar hoher Ähnlichkeit von dem hier dargestellten Modell abgegrenzt werden. Im Bereich der Informatik findet sich ein häufig zitierter Ansatz, der die Bezeichnungen Mensch, Aufgabe, Technik als Komponenten von Informationssystemen nennt und als Grundlage für deren Entwicklung herangezogen werden kann (Heinrich, 1993, S. 14). Die Komponenten sind hierbei in einem systemischen Zusammenhang zu verstehen, werden auf abstraktem Niveau definiert und sind auf die Konstruktion von kompletten Informationssystemen ausgerichtet. So bezeichnet die Komponente Mensch Individuen und Gruppen und deren Rolle im Entwicklungsprozess. Als Aufgabe wird das im Informationssystem umzusetzende Problem bzw. die benötigte Informationsinfrastruktur und Technik, beispielsweise Hardware, Software und Programmiersprachen, bezeichnet. Eingebettet werden die Komponenten in die Methoden und Werkzeuge der Systemplanung, wie sie zum Programmieren eines Informationssystems benötigt werden. Die drei Begriffe Mensch, Aufgabe, Technik werden ebenfalls im Kontext der Software-Ergonomie bei Maaß (1993) verwendet. Dort wird das Umfeld der Systemgestaltung durch die vier Faktoren Mensch, Aufgabe, Technik und Organisation eingefasst. „Benutzer- und aufgabenorientierte [System-]Gestaltung muss danach individuelle, technische, aufgabenspezifische sowie gesamtorganisatorische Gegebenheiten beachten.“ (Maaß, 1993, S. 195). Dieser Ansatz enthält durch die Nennung der Begriffe lediglich die Grobstruktur des hier dargestellten Rahmenmodells, wobei Anforderungen von Organisationen für Suchmaschinen im Internet eine untergeordnete Rolle spielen, da diese mit dem Anspruch konstruiert werden, für Anwender weltweit gebrauchsfähig zu sein.

Im Kontext der Informationssuche im Internet identifizieren Navarro-Prieto, Scaife und Rogers (1999) drei Faktoren und entwickeln ihr Rahmenmodell, „the interactivity

framework". Das Ziel der Autoren ist dabei die Identifizierung derjenigen Aspekte, die bei der Forschung über Suchverhalten beachtet werden müssen. Ihr Schwerpunkt entspricht damit einem Teilaspekt des Rahmenmodells der vorliegenden Arbeit. Ihre drei Faktoren, user, task und external representation genannt, beschreiben sie als (1) Erfahrung der Anwender und kognitive Strategien, (2) die Art der Suchaufgabe und (3) die Art der Präsentation von Informationen und wie die Information mit dem Anwender interagiert. Deutlich zu erkennen ist, dass diese Faktoren jeweils nur eine Dimension auf jeder Ebene im hier beschriebenen Rahmenmodell ansprechen. Zudem konzentrieren sich die Autoren in ihrem Konferenzbeitrag auf die knappe Darstellung einer Untersuchung zu Suchverhalten unter dem Aspekt der Art der Aufgabe, auf die im späteren Teil dieser Arbeit noch eingegangen wird (Kap. 2.3.1.2). Festzustellen ist, dass die eben diskutierten Rahmenmodelle zwar in ihren Bezeichnungen der Faktoren Ähnlichkeit mit dem hier vorgeschlagenen Rahmenmodell haben, jedoch inhaltlich entweder andere Ziele verfolgen oder nur unzureichend ausgeführt sind, sodass sie sich nicht ohne größere Ausarbeitung und weitere Systematisierung für die hier angestrebte Forschung verwenden lassen. Aus diesem Grund wird das oben beschriebene neue Rahmenmodell zur Forschung über die Gebrauchstauglichkeit der Ergebnisseiten von Suchmaschinen vorgeschlagen.

Die folgenden Kapitel widmen sich der Aufgabe, die einzelnen Komponenten des Rahmenmodells genauer zu explizieren. Dabei sollen die Bestimmungsstücke identifiziert werden, die eine erfolgreiche Forschung konstituieren. Forschungsschwerpunkte aus verschiedenen Disziplinen werden zusammengeführt, die in der Synthese ein hohes Potenzial an Weiterentwicklungsmöglichkeiten offenbaren.

## **2.2 Maschine: Informationspräsentation bei Suchmaschinen**

In diesem Kapitel werden die Grundlagen zum Aufbau und der Funktionsweise von Suchmaschinen dargelegt und die Teilkomponenten von Ergebnisseiten analysiert. Dies schafft die Voraussetzungen für spätere systematische Variationen von Ergebnisseiten unter Forschungsaspekten, und mögliche Forschungsfragen können leichter abgeleitet werden. Da Suchmaschinen Werkzeuge darstellen, um an Informationen zu gelangen, erscheint es angebracht, diese zunächst in den allgemeineren Zweig des Information Retrieval einzuordnen. Anschließend werden der Aufbau und die Funktionsweise von Suchmaschinen selbst erläutert. Die Thematik wird weiter vertieft, indem die verschiedenen Möglichkeiten, Suchergebnismengen zu präsentieren, klassifiziert und beschrieben, sowie die typische Ergebnisseite von Suchmaschinen – geordnete, textuelle Listen – in Aufbau und Eigenschaften analysiert werden.

### **2.2.1 Information Retrieval**

Möchten Anwender Informationen in einem Datenbestand bzw. dem Internet finden, so können sie für diesen Zweck ein Information Retrieval System bzw. eine Suchmaschine verwenden. Informationen sind in dieser Arbeit allgemein betrachtet alle Dinge, die in der Lage sind, den Wissensstand einer Person zu verändern. Information wird damit im Sinne

Marchioninis (1995) verstanden. Entsprechend ist ein Informationsbedürfnis ein motivationaler Zustand, der dadurch entsteht, dass eine Person Bereiche ihres Wissens (vgl. dazu Kap. 2.4.3) umstrukturieren möchte.

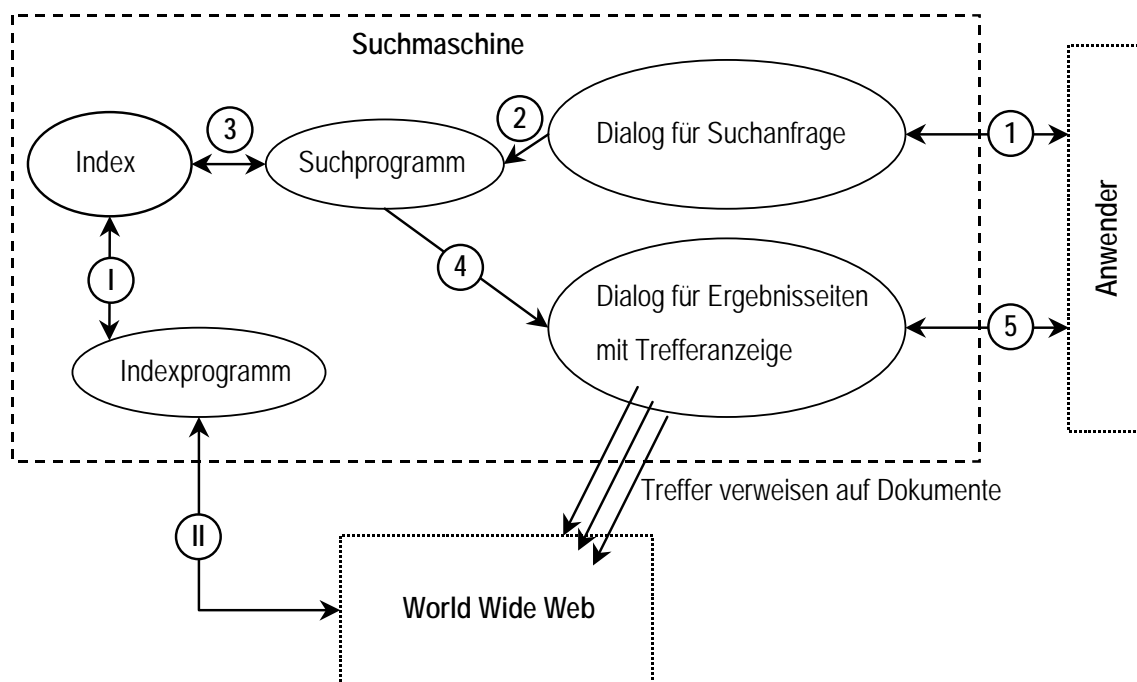
Das einfache Grundprinzip von Information Retrieval (IR) erscheint unmittelbar verständlich. Zunächst teilt der Anwender dem IR-System sein Informationsbedürfnis mit. Daraufhin durchsucht die Software ihren Datenbestand und präsentiert dem Anwender das gefundene Ergebnis. Dieses enthält im Idealfall verschiedene für den Anwender relevante Treffer, die die gefundenen Dokumente repräsentieren. Die Ergebnisseite liefert somit indirekt die relevanten Dokumente, indem sie auf diese verweist.

Selbstverständlich ist auch die Suche nach anderen Medieninhalten wie Tonaufnahmen oder Bilder möglich (Jansen, Goodrum & Spink, 2000). Dies sind jedoch Spezialfälle und werden üblicherweise von eigens dafür angefertigten Programmen (beispielsweise Filesharing-Tools (vgl. dazu Pursche, 2002)) und speziellen Suchmethoden (Garber & Grunes, 1992) unterstützt, weshalb sie in dieser Arbeit nicht weiter behandelt werden. Stattdessen konzentriert sich die vorliegende Arbeit auf den größeren und geschichtlich älteren Bereich der textorientierten Suche. Entsprechend konzentriert sich die vorliegende Arbeit auf das Internet als eine Ansammlung von Dokumenten.

Das Internet als unüberschaubarer und rasant wachsender Datenbestand bringt neue Herausforderungen durch Dokumente, die im Volltext statt als Zusammenfassungen verfügbar sind, wie dies in früheren IR-Systemen üblich war und auch noch in heutigen Bibliothekssystemen wie dem OPAC (Online Public Access Catalogue) den Normalfall darstellt. Heutige Suchmaschinen unterscheiden sich daher in einigen Punkten von ihren Vorgängern (Bekavac, 1999). Moderne Systeme verwenden statistische Sortieralgorithmen (Ranking), die bei großen Treffermengen eine bessere Sortierung ermöglichen als frühere, einfachere Sortierungen nach Titel, Datum oder Autor. Hinzu kommt eine an die Dynamik des Internets angepasste Systemstruktur, wie sie weiter unten erläutert wird. Aber auch das Verhalten der Anwender hat sich deutlich verändert und entspricht nicht mehr dem Bild der Anwender aus der klassischen Information Retrieval Literatur (Silverstein & Henzinger, 1999). Gab es in Zeiten kommandozeilenorientierter Eingabesysteme noch professionelle Vermittler zwischen System und Informationssuchendem, die umfangreiche boolesche Ausdrücke (logische Verknüpfungen zwischen den Suchwörtern) verwendeten, um relevante und wohldefinierte Treffer zu erzielen (Baeza-Yates & Ribeiro-Neto, 1999), so führen heutige Anwender ihre Suche meist direkt selber mit relativ ungenauen Suchanfragen durch, was zu hohen Treffermengen führt. Die Tatsache, dass sowohl erfahrene wie auch unerfahrene Anwender (vgl. Kap. 2.4.3) Suchmaschinen einsetzen, setzt hohe Anforderungen an die Konstruktion gebrauchstauglicher Suchmaschinen. Nach dieser thematischen Eingrenzung von Information-Retrieval-Systemen auf textorientierte Suchmaschinen für die vorliegende Arbeit und der Gegenüberstellung der wichtigsten Unterschiede zwischen früheren und neueren IR-Systemen behandelt das nächste Kapitel die Kerneigenschaften von Suchmaschinen.

### 2.2.2 Suchmaschinen im World Wide Web

Der Faktor Maschine beinhaltet neben der technischen Realisierung von Suchmaschinen auch die Schnittstelle (Interface), über die der Anwender mit dem Suchsystem kommunizieren kann. Die Schnittstelle ermöglicht es dem Anwender, im Dialog mit dem System Informationen auszutauschen, wohingegen die technische Realisierung aus Anwenderperspektive als eine „black box“ angesehen werden kann. Um den Zusammenhang der Dialoge mit ihren Repräsentationen von Informationen zum Gesamtsystem der Suchmaschinen als Werkzeuge der Suche zu veranschaulichen, wird daher nun das Zusammenspiel der verschiedenen Komponenten einer typischen Suchmaschine erläutert (Abbildung 2).



**Abbildung 2:** Modell einer Suchmaschine im Zusammenspiel mit dem Internet und Anwendern.

Beim Einsatz von Suchmaschinen sieht sich der Anwender zunächst einem Eingabedialog gegenüber (1). Mit ihm kann er verschiedene Einstellungen sowohl für seine Suche als auch für die Ergebnisausgabe vornehmen (Notess, 2001). Beispielsweise lässt sich der Suchraum auf eine bestimmte Sprache einschränken und auch die Anzahl an dargestellten Treffern pro Seite festlegen. Das Kernstück ist jedoch ein Eingabefeld, mit deren Hilfe der Anwender sein Informationsbedürfnis dem System mitteilt, indem er Wörter zu einem Suchausdruck verbindet, der seiner Ansicht nach mit der gesuchten Information in Zusammenhang steht. Dabei lassen sich mehrere Wörter mit Hilfe der booleschen Algebra miteinander verknüpfen und spezielle Ausdrücke für Phrasen- oder Feldsuche einsetzen, um die Suchanfrage genauer zu spezifizieren (vgl. dazu Karzauninkat, 1998b). Um Fehlermeldungen zu vermeiden, machen Suchmaschinen verschiedene Annahmen, falls die Eingabe nicht eindeutig ist. So gehen die meisten Systeme von einer logischen Und-Verknüpfung aus, wenn keine explizit boolesche Verknüpfung angegeben wird, was immer dann der Fall ist, wenn mehrere Wörter einfach hintereinander eingegeben werden. Dies führt zu einer

sinnvollen Reduzierung der relevanten Dokumente, kann jedoch gelegentlich auch Verwirrung auslösen, sollten die Annahmen der Anwender darüber, wie die Suchmaschine funktioniert, fehlerhaft sein (Greene, Devlin, Cannata & Gomez, 1990). Eine ähnlich versteckte Annahme seitens der Suchmaschine wäre, dass bei einem großgeschriebenen Wort auch nur solche großgeschriebenen Wörter gefunden werden, wohingegen ein kleingeschriebenes Wort keine Unterscheidung zwischen Groß- und Kleinschreibung für die Suche zur Folge hat. Komplizierter wird die Bedienung von Suchmaschinen dadurch, dass diese zum Teil sehr unterschiedliche Eingabesyntax verwenden (vgl. Reibold, 2000).

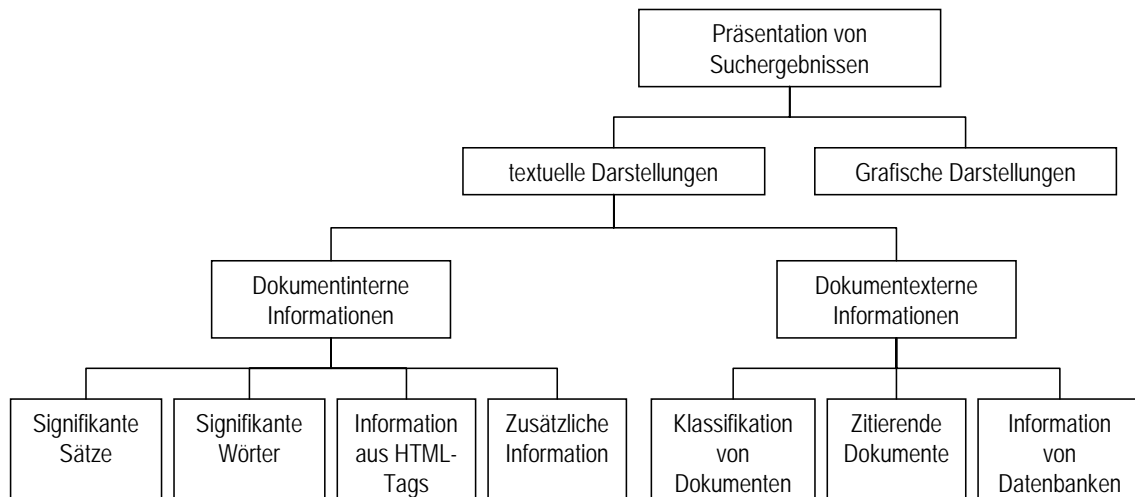
Nach der Eingabe wird die Suchanfrage an ein Suchprogramm weitergegeben (2), das wiederum in einem Index die relevanten Dokumente nachschlägt (3). Der Index enthält ein komprimiertes Verzeichnis (Baeza-Yates & Ribeiro-Neto, 1999; Karzauninkat, 1998b) von Internetdokumenten und ihren Inhalten. Dieses wird von einem Indexprogramm verwaltet (I), das durch Robots, sogenannte Spider und Crawler (vgl. Koster, 2002) mit dem Internet in Verbindung steht (II). Aufgabe der Robotprogramme ist es, das Internet nach Inhalten zu durchforsten und dadurch den Index ständig zu erweitern und zu aktualisieren. Die vom Suchprogramm ermittelten Treffer werden einem komplexen statistischen Rankingverfahren (vgl. Grossan, 1997) unterzogen, an deren Ende eine nach Relevanz geordnete Liste von Treffern steht (4). Diese erhält der Anwender in einem Ergebnisdialog zurückgemeldet (5). Jeder Treffer repräsentiert ein gefundenes Dokument und wird mit Informationen über das Dokument dargestellt (s. Kap. 2.2.4). Auf diesen sogenannten Ergebnisseiten ist es wiederum Sache des Anwenders, mit Hilfe der Informationen zu jedem Treffer die für sein Informationsbedürfnis passenden Dokumente auszuwählen. Auf den für den Anwender und die vorliegende Arbeit besonders bedeutsamen Ergebnisseiten mit den verschiedenen Möglichkeiten der Ergebnispräsentation wird im Folgenden näher eingegangen.

### **2.2.3 Präsentation von Suchergebnissen**

Werden zum Finden von Informationen Suchmaschinen verwendet, so liefern diese aufgrund einer Suchanfrage eine Menge an Treffern aus ihrem Index an den Anwender zurück. Diese umfassen oft hunderte oder gar tausende Angaben, die ein Anwender unmöglich alle evaluieren kann. Jeder Treffer repräsentiert ein gefundenes Dokument. Eine Schwierigkeit besteht nun darin, die richtigen Informationen auszuwählen, die bei den Treffern angezeigt werden sollen, sodass der Anwender einerseits den Überblick über die Treffermenge behält und andererseits jedoch in der Lage ist, die für ihn relevanten Dokumente zu erkennen: „Determining how much information to show the user of the system is a major design choice in information access interfaces.“ (Baeza-Yates & Ribeiro-Neto, 1999, S. 259). Auf dem Weg zur optimalen Darstellung hat die Forschung zum Information Retrieval zahlreiche Werkzeuge entwickelt (Baeza-Yates & Ribeiro-Neto, 1999; Bekavac, 1999; Drori, 2001). Um eine Einordnung zu ermöglichen, beschreibt dieses Kapitel ein Klassifikationssystem von Drori (2000a) mit verschiedenen Techniken der Ergebnisdarstellungen und bettet verschiedene Forschungsrichtungen zur Textsummarisation in diesem ein. Dadurch wird deutlich, welche differenzierte Möglichkeiten existieren, die von Treffern repräsentierten Dokumente näher zu beschreiben.



Nach dem Klassifikationsschema von Drori (2000a) lassen sich zwei grundsätzliche Ansätze zur Präsentation der Ergebnismenge unterscheiden (Abbildung 3). Grafische Darstellungen (vgl. Baeza-Yates & Ribeiro-Neto, 1999; Bekavac, 1999; Drori, 2001; Stenmark, 1997) verwenden grafische Objekte, um die gefundenen Treffer zu beschreiben und die Beziehungen zwischen ihnen darzustellen. Die Suchmaschine Kartoo<sup>1</sup> bildet zum Beispiel die Verbindungen zwischen verschiedenen gefundenen Treffern ab und zeigt verbindende zentrale, alternative Suchwörter, um die Verfeinerung der Suche zu unterstützen. Solche grafischen Visualisierungen eignen sich vor allem für explorative Suche im Sinne von Rosenfeld und Morville (1998), wobei vorwiegend der Suchraum erforscht werden soll (vgl. Kap. 2.3.1.3). Eine Sammlung weiterer Visualisierungsmöglichkeiten findet sich bei Dodge und Kitchin (2001). Es lässt sich feststellen, dass zwar eine hohe Anzahl verschiedener Visualisierungsansätze existiert, sich jedoch bisher keiner auf breiter Basis durchsetzen konnte, was wohl auch darin seine Begründung findet, dass viele Darstellungen zwar durchaus ästhetischen Wert besitzen, jedoch keinesfalls intuitiv für die praktische Anwendung erschließbar sind.



**Abbildung 3:** Klassifikationsschema zu den Möglichkeiten der Präsentation von Ergebnismengen nach Drori (2000a).

So entspricht der heutige Standard bei Suchmaschinen auch aufgrund der einfacheren Umsetzbarkeit der zweiten Möglichkeit, der textuellen Darstellung, auf die auch die vorliegende Arbeit ihren Schwerpunkt legt. Hierbei können sowohl dokumentinterne also auch dokumentexterne Informationen für die Beschreibung eines Treffers verwendet werden. Interne Informationen lassen sich aus den gefundenen Dokumenten selbst extrahieren, wohingegen externe Informationen aus anderen Dokumenten, dem Kontext, in dem das gefundene Dokument sich befindet oder separat gespeicherten Informationen über das gefundene Dokument erstellt werden. Sieben verschiedene Kategorien am Fuß der Hierarchie stellen Designer von Ergebnisseiten vor die Herausforderung zu entscheiden, welche Methoden sie für ihre Suchmaschine implementieren sollen. Signifikante Sätze und

<sup>1</sup> <http://www.kartoo.com/> Die Internetadressen der in dieser Arbeit zitierten Suchmaschinen finden sich in Anhang D.

signifikante Wörter enthalten die beiden Hauptmethoden, um Zusammenfassungen der relevanten Dokumente zu erzeugen, damit diese bei den Treffern angezeigt werden können.

### **2.2.3.1 Signifikante Sätze**

Zusammenfassungen mit Hilfe von signifikanten Sätzen lassen sich auf zwei grundsätzliche Weisen generieren. Zum einen können diejenigen Sätze eines Dokumentes extrahiert werden, in denen die gesuchten Wörter vorkommen. Komplexere Methoden jedoch analysieren Textabsätze und extrahieren signifikante Sätze, die die Kernaussagen der gefundenen Dokumente wiedergeben (Endres-Niggemeyer, 1998). Amitay und Paris (2000) beschreiben drei Forschungsrichtungen zur Textsummarisation, die sich in den Bereich signifikanter Sätze einbetten lassen. Absatzorientierte Summarisation extrahiert ganze Textabschnitte, die jeweils ein eigenes Thema betreffen. Diese Technik wird daher angewendet, wenn verschiedene Themen in einem Dokument behandelt werden und der Bedarf nach einer Zusammenfassung besteht. Die zweite Forschungsrichtung wird als diskursmodellorientierte Extraktion und Summarisation bezeichnet. Sie nutzt Eigenschaften der natürlichen Sprache wie Reihenfolge der Wörter, Wortwahl und vordefinierte Hinweissätze, um eine Sammlung von Fakten und Hinweisen zu generieren, die darstellen, was im Dokument erwartet werden kann. Aufgrund der Komplexität der Sprache sind diese Verfahren bisher an spezifische Themenbereiche und Schreibstile, wie zum Beispiel wissenschaftliche Artikel in der Medizin gebunden und eignen sich daher nur geringfügig für die große Vielfalt an Textsorten im Internet. Satzorientierte Abstraktionstechniken wiederum suchen nach Wiederholungen von Begriffen und Sätzen und möchten damit die salientesten Aussagen zusammenstellen. Dabei erfahren Hervorhebungen durch Titel oder Fettschrift besondere Beachtung. Dieser Ansatz scheint für Webdokumente besonders geeignet, da damit auch kleine Informationseinheiten behandelt werden können und Inhaltselemente im weit verbreiteten Dokumenttyp HTML (Hypertext Markup Language) in ihren Funktionen zum Beispiel als Überschrift bereits extra ausgezeichnet werden (vgl. dazu Münz, 2001). Weitere Evidenz erfährt dies dadurch, dass auch die Rankingalgorithmen von Suchmaschinen insofern Ähnlichkeiten mit diesem dritten Verfahren aufweisen, da Suchmaschinen ein Dokument für umso relevanter erachten, je höher die Ähnlichkeit ist zwischen den Ergebnissen satzorientierter Abstraktionstechniken und den Suchwörtern der Suchanfrage. Allen diesen Methoden ist gemeinsam, dass sie die Treffer auf Ergebnisseiten mit Zusammenfassungen der repräsentierten Dokumente zu beschreiben versuchen.

### **2.2.3.2 Signifikante Wörter**

Signifikante Wörter stellen die zweite Kategorie möglicher Informationsquellen dar. Hierbei werden im Gegensatz zu ganzen Sätzen einzelne Schlüsselwörter identifiziert. Dies kann zum Beispiel durch statistische Vergleiche der Worthäufigkeiten im betreffenden Dokument zu Worthäufigkeiten in anderen Texten geschehen. Findet sich ein Wort signifikant häufiger in einem Dokument als dies statistisch zu erwarten wäre, so wird es der Zusammenfassung hinzugefügt, da angenommen wird, dass es eine zentrale Rolle im Dokument spielt. Schlüsselwörter können jedoch auch durch den Autor des gefundenen Dokumentes direkt

vorgegeben und als solche gekennzeichnet sein, wie dies in Schlagwortkatalogen von Bibliotheken der Fall ist.

### **2.2.3.3 Informationen aus HTML-Tags**

Neben Zusammenfassungen von Textinhalten bietet es sich an, auch andere Informationen über Dokumente darzustellen. Mittels des Standardformates HTML lassen sich Strukturmerkmale von Textabschnitten spezifizieren. Dies geschieht durch sogenannte HTML-Tags (Münz, 2001). Überschriften werden zum Beispiel durch spezielle Überschriftentags gekennzeichnet. Aber auch andere Informationen über den normalen Text hinaus lassen sich mit Hilfe von Meta-Tags kennzeichnen. Durch Meta-Tags ist es beispielsweise möglich, die Sprache eines Dokumentes im Dokument zu kodieren und damit auch leicht zu identifizieren. Überschriften und Absätze sind leicht erkennbar und mittels weiterer Meta-Tags können beliebige Zusatzinformationen wie Autor, Zusammenfassungen und auch die oben genannten Schlüsselwörter dem Dokument als Metainformationen mitgegeben werden. HTML-Tags und hierbei insbesondere die für den durchschnittlichen Anwender nicht sichtbaren Meta-Tags sind jedoch aufgrund des leichten Missbrauchs durch Anbieter nicht ganz unproblematisch (Lennartz, 1999). Einige Anbieter versuchen, durch extensive Verwendung von oft gesuchten Schlüsselwörtern ohne erkennbaren Zusammenhang zu ihrer beworbenen Seite auf die oberen Rankingplätze bei Suchmaschinen zu gelangen und laufen damit dem eigentlichen Zweck von Html-Tags, sinnvolle Zusatzinformationen zu liefern, zuwider. Glücklicherweise existieren Filtermethoden, die dieser Problematik größtenteils Herr werden (vgl. Google, 2002a).

### **2.2.3.4 Zusätzliche Informationen**

Die vierte Kategorie, zusätzliche Informationen, bildet eine Restkategorie für alle sonstigen Informationen im Dokument. Denkbar wäre eine Liste der Titel aller zitierten Dokumente, um das Thema eines Dokuments besser darzustellen unter der Annahme, dass diese thematisch eng mit dem Dokument zusammenhängen.

Zusammenfassend lässt sich also feststellen, dass Methoden zur Darstellung dokumentinterner Informationen ausschließlich bestehende Textteile extrahieren, wobei diesen jeweils eine besondere Bedeutung für ein Dokument unterstellt werden. Ein anderer Weg wird mit dokumentexternen Informationen eingeschlagen. Diesen Methoden ist gemeinsam, dass sie die Zugehörigkeit von Dokumenten zu Kategorien nutzen und so helfen, den Kontext von Dokumenten herauszustellen.

### **2.2.3.5 Klassifikation von Dokumenten**

Wie oben bereits angeklungen, stellen heutige Anwender von Suchmaschinen typischerweise relativ ungenaue Suchanfragen, indem sie nur wenige Worte verwenden. Obwohl der Anwender eine bestimmte Bedeutung im Sinn hat, die seinem Informationsbedürfnis entspricht, fehlt ihm häufig das spezielle Wissen, um diese spezielle Bedeutung durch einen umfangreichen Suchausdruck deutlich zu machen (vgl. Kap. 2.4.3). Aufgrund der Vieldeutigkeit der Sprache kann eine Suchmaschine nur schwer feststellen,

welche Bedeutung eines Wortes der Anwender bei seiner Eingabe im Sinn hatte. Als Ergebnis liefern Suchmaschinen daher zahlreiche Dokumente aus den unterschiedlichsten Kontexten, die vom Rankingalgorithmus aufgrund der fehlenden Information der Bedeutung alle gleich behandelt werden. Klassifikationen von Dokumenten bieten einen Ausweg aus dieser Problematik. Ergebnisseiten können die verschiedenen gefundenen Kontexte anzeigen und die Anwender wählen lassen, in welchem dieser Kontexte sie relevante Dokumente suchen. Dies wird beispielsweise von den Suchmaschinen Vivísimo<sup>2</sup> und Northern Light<sup>3</sup> realisiert. Zu einer gewählten Kategorie werden jeweils die darin relevanten Treffer angezeigt, wobei der Anwender in immer tieferen Subkategorien seine Suche solange verfeinern kann, bis er auf relevante Dokumente stößt.

Um eine Klassifikation von Dokumenten zu erreichen, können Suchmaschinen, die mit Katalogen zusammenarbeiten, angeben, aus welcher Kategorie des Katalogs die gefundenen Dokumente stammen. Des Weiteren können Algorithmen die Zugehörigkeit eines Dokuments bestimmen, indem sie Gemeinsamkeiten zwischen verschiedenen Dokumenten aufspüren und die Dokumente entsprechend anordnen oder bereits bekannten Kategorien zuordnen (Weiss et al., 1996). Mit Hilfe lernfähiger Modelle können Algorithmen auf das Erkennen bekannter Kategorien trainiert werden. Dafür wird das Modell auf eine bereits zu einer Kategorie zugeordnete Sammlung von Dokumenten angewendet. Treffer werden dann von dem Modell einer Kategorie zugeordnet und in diesen thematischen Clustern dargestellt (Dumais, Cutrell & Chen, 2001; Grundman & Michalek, 2000). Aber auch speziellere Kontextinformationen können verwendet werden.

#### **2.2.3.6 Zitierende Dokumente**

Zitierende Dokumente, solche also, die auf gefundene Dokumente verweisen, liefern weitere Informationen über das Dokument unter der Annahme, dass diese Dokumente inhaltliche Gemeinsamkeiten aufweisen. Von zitierenden Dokumenten können beispielsweise deren Titel oder die zitierenden Absätze verwendet werden. Solche Informationen werden von Googles<sup>4</sup> PageRank-Technologie<sup>TM</sup> verwendet, die im Ranking die Verweise anderer Dokumente berücksichtigt (Google, 2002b). Theoretische Grundlagen zur Bestimmung zuverlässiger, relevanter Quellen anhand solcher Verweise finden sich bei Kleinberg (1998). Ein weiteres Beispiel zur praktischen Anwendung bietet Amitay (2001) mit ihrem InCommonSense System, das Textstellen, die andere Dokumente beschreiben und zitieren, extrahiert und zur genaueren Beschreibung von Dokumenten heranzieht.

#### **2.2.3.7 Informationen aus Datenbanken**

Die letzte Kategorie bezeichnet die Möglichkeit, Informationen aus einer Datenbank zu verwenden, um ein Dokument näher zu beschreiben. Dies ist selbstverständlich nur möglich, wenn das zu findende Dokument bereits Bestandteil einer Datenbank ist. In diesem Fall

---

<sup>2</sup> <http://vivisimo.com/>

<sup>3</sup> <http://www.northernlight.com/>

kann beispielsweise das Thema der Datenbank angegeben werden. Auch hier können die verschiedensten Angaben wie Autor, Datum der Erstellung, Schlüsselwörter und viele andere Daten extrahiert werden, sofern sie in der Datenbank vorhanden sind.

#### **2.2.3.8 Fazit**

Deutlich wird, dass es eine ungeheure Vielzahl an Möglichkeiten gibt, einen Treffer näher zu beschreiben, um ein gefundenes Dokument angemessen zu repräsentieren. Diese lassen sich jedoch klassifizieren. Die Taxonomie von Drori (2000a) klassifiziert Angaben danach, aus welcher Datenquelle die Informationen stammen. Die Schwachpunkte eines solchen Vorgehens sollten die gewählten Beispiele verdeutlichen. Daran zeigt sich, dass Schlüsselwörter im Dokument nicht eindeutig einer Kategorie zuzuordnen sind, da diese aus unterschiedlichen Datenquellen stammen können und dann entsprechend vielseitig eingeordnet werden. Identische Inhalte können demnach in unterschiedlichen Datenquellen vorkommen und eine exakte Einordnung ist daher nicht immer möglich. Trotz dieser Schwäche bietet der Ansatz von Drori eine Entscheidungshilfe bei der Frage, welche Datenquellen zur Darstellung herangezogen werden können.

Nach diesem theoretischen Überblick widmet sich das nächste Kapitel den Ergebnisseiten selbst und stellt damit den Zusammenhang der oben erläuterten Möglichkeiten, Informationen über Dokumente darzustellen, zu den Ergebnisdialogen her.

#### **2.2.4 Geordnete Listen zur Repräsentation von Informationen (Ergebnisseiten)**

Der folgende Abschnitt beschreibt den Ergebnisdialog und bietet damit den praxisrelevanten Zusammenhang zu den vorhergehenden Kapiteln. Dazu wird zunächst die grobe Struktur der Ergebnisseite geschildert, anschließend Hinweise zum Ergebnisdialog gegeben, um schließlich im Detail die Eigenschaften der Trefferdarstellung zu analysieren.

Auf einer Webseite lassen sich drei funktional unterschiedliche Komponenten identifizieren: Inhalt, Navigation und Werbung (Abbildung 4).

Eine Ergebnisseite beginnt üblicherweise mit einem Bereich zur Navigation, in der die Suche weiter verfeinert werden kann. Darunter kann ein Werbeblock stehen, häufig ein Banner und Links mit Werbeinhalten. Es folgen die Inhalte mit Informationen zur Suche und den Treffern bzw. gefundenen Dokumenten. Hierbei kann die Frage gestellt werden, welches Layout der Trefferliste für den Anwender am günstigsten ist. Ergebnisse lassen vermuten, dass geringe Abstände zwischen einzelnen Treffern diese besser voneinander absetzen und von Anwendern favorisiert werden im Gegensatz zu keinen oder großen Abständen (Bernard, Chaparro & Thomasson, 2000).

---

<sup>4</sup> <http://www.google.de/>

The image shows a screenshot of an AltaVista search results page for the query "Mannheim Psychologie". The page is annotated with several boxes and arrows:

- Navigation:** A box at the top right points to the search bar and navigation links.
- Werbung:** A box at the top left points to a banner advertisement for "PARIS AB €98" by opodo.
- Inhalt:** A box in the middle points to the main search results.
- Werbung:** A box on the right points to an advertisement for "Aktuelle Angebote! CDs & DVDs bis €4,-!" by ebay.
- Navigation:** A box at the bottom right points to the footer navigation area.
- Navigation:** A box at the bottom left points to a banner advertisement for "LONDON AB €58" by opodo.

The search results include links to various university and research pages, such as "Universität Mannheim - Psychologie", "Abteilung Klinische Psychologie", and "13. Kongress der Deutschen Gesellschaft für Psychologie".

Abbildung 4: Aufbau einer Ergebnisseite am Beispiel *Altavista* mit den Suchwörtern „Mannheim Psychologie“ (gekürzt).

Auch bei der Entscheidung, ob 10, 50 oder 100 Treffer auf einer Seite dargestellt werden sollen, bevorzugen Anwender die mittlere Bedingung (Bernard, Baker & Fernandez, 2002). Letzteres Ergebnis wird von Suchmaschinen standardmäßig nicht umgesetzt, da der Seitenaufbau bei nur 10 Treffern deutlich schneller stattfindet und die damit verbundene Geschwindigkeit ebenfalls ein wichtiger Faktor für den Anwender ist (vgl. Selvidge, 2003). Bei einer Untersuchung von Khan und Locatis (1998b) zeigt sich ein anderes Bild. Hier sind drei Links der Bedingung mit sechs dargestellten Links überlegen. Die hohe Diskrepanz zwischen den beiden Untersuchungen ist auf den hohen Einfluss zahlreicher Randbedingungen zurückzuführen (z.B. Suchaufgabenstellung, Bewertungskriterien). Darin zeigt sich aber auch deutlich die Notwendigkeit eines Rahmenmodells zur Forschung.

Der Fuß der Seite wird wiederum von einem Navigationsbereich eingenommen. Dieser kann genutzt werden, um weitere Ergebnisseiten aufzurufen, die Suchanfrage zu modifizieren oder Informationen zur Suchmaschine abzurufen. Weitere Werbung kann rechts von den Suchergebnissen eingeblendet sein. Im Informationsbereich über die Suche erscheint es sinnvoll, verschiedene Metainformationen über die durchgeführte Suche anzugeben. Ausführliche Angaben hierzu finden sich bei Schulz (2001). Die wesentlichsten Angaben sind im Folgenden aufgeführt:

- Gesamtanzahl der gefundenen Dokumente
- Anzahl der Ergebnisseiten insgesamt
- Nummerierung der aktuell angezeigten Ergebnisseite
- Verwendete Suchkriterien: Die verwendeten Suchwörter, automatisch verwendete Synonyme, rechtschreibkorrigierte Wörter und nicht verwendete Stoppwörter (dies sind Wörter, die so oft vorkommen, dass nach ihnen nicht gesucht werden kann; Bspw. der, die, das, a, an, the werden häufig nicht indexiert)
- In welchem Bereich gesucht wurde (WorldWideWeb, Website, Sub-Site, vorherige Ergebnismenge)
- Fehlermeldungen mit Hilfestellungen, wenn keine Dokumente gefunden wurden
- Hilfestellungen, um die Suche einzuschränken, wenn sehr viele Dokumente gefunden wurden

Diese Angaben sollten üblicherweise im Kopfbereich jeder Ergebnisseite stehen. Bei den einzelnen Treffern können ebenfalls verschiedene Angaben unterschieden werden. Jeder Treffer ist ein Ordnungspunkt in einer Liste und lässt sich durch mehrere Eigenschaften beschreiben (Kleinfeldt & Baphna, 2000); Abbildung 5):

- Die Position in der Liste. Je weiter oben ein Dokument platziert ist, desto höher war die vom Suchmaschinenalgorithmus geschätzte Relevanz des Links.
- Der Titel des Dokuments, der auch einen Verweis (Link) auf das Zieldokument enthält.
- Eine Beschreibung des Dokuments. Üblicherweise eine Zusammenfassung, in der der Anfang des Dokuments oder die gesuchten Wörter im Kontext der umliegenden Wörter dargestellt werden. Diese Beschreibungen können nach den in Kapitel 2.2.3 beschriebenen Verfahren generiert werden.
- Zusätzliche Informationen über das Dokument wie URL (uniform resource locator), Dokumenttyp, Dokumentgröße, Autor, Datum der letzten Änderung usw.
- Zusätzliche Funktionen. Diese Eigenschaft wird zwar von Kleinfeldt und Baphna nicht beschrieben, darf hier jedoch nicht fehlen, da sie mittlerweile von vielen Suchmaschinen angeboten wird. Am Ende des Treffers finden sich verschiedene Links, die beispielsweise eine Dokumentvorschau ermöglichen (z.B. bei Fireball<sup>5</sup>). Aber auch die direkte Möglichkeit, eine erneute Suche über ähnliche Seiten durchzuführen (z.B. bei Google<sup>6</sup>, sog. relevance Feedback, s. Kap. 2.4.4.1), sowie eine Seite zu übersetzen (z.B. bei Abacho<sup>7</sup>) gehören in diese Kategorie. In Abbildung 5 wird eine Zusatzfunktion bei Excite<sup>8</sup> dargestellt, mit der die Möglichkeit besteht, dieselbe Suche auf die angegebene Internetdomäne, d.h. auf alle Seiten, die die angegebene URL als Anfangsadresse haben, zu beschränken.

Allerdings findet die Reihenfolge, in der eine Suchmaschine die Ergebnisse aufführt, nicht immer eine Entsprechung im Informationsbedürfnis des Anwenders. Die Ursachen liegen teilweise am Sortieralgorithmus der Suchmaschine, aber ebenso an den mehrdeutigen Eingaben der Anwender (s. Kap. 2.4.3). Sowohl an Verbesserungen der Algorithmen als auch an der Unterstützung zur besseren Formulierung der Suchanfrage, damit diese das

---

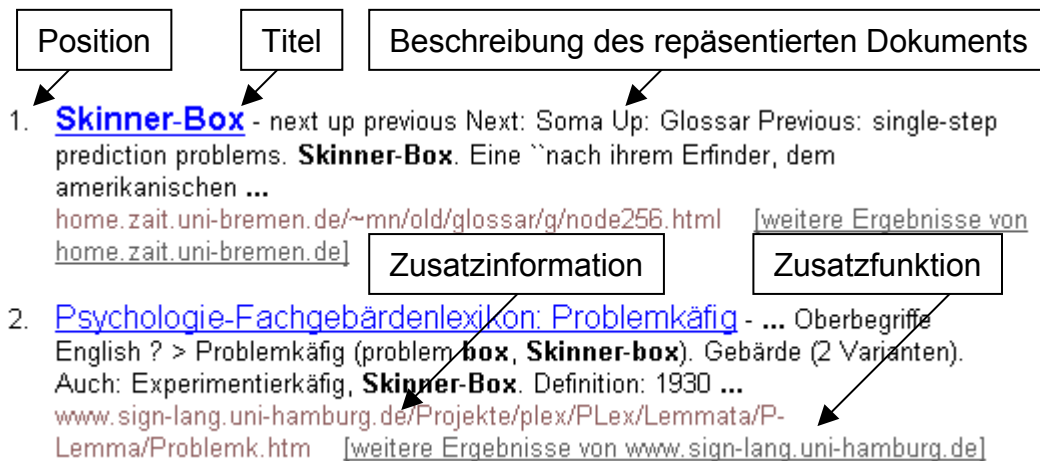
<sup>5</sup> <http://www.fireball.de/>

<sup>6</sup> <http://www.google.de/>

<sup>7</sup> <http://www.abacho.de/>

<sup>8</sup> <http://www.excite.de/>

Informationsbedürfnis besser beschreibt, wird geforscht (Baeza-Yates & Ribeiro-Neto, 1999; Karzauninkat, 1998b; Kleinfeldt & Baphna, 2000; McArthur & Bruza, 2000).



**Abbildung 5:** Eigenschaften einzelner Treffer einer Ergebnisliste am Beispiel *Excite* mit den Suchwörtern „Skinner Box“.

## 2.2.5 Zusammenfassung und Ableitungen

Die obigen Abschnitte über das System stellten Suchmaschinen als Teilbereich des Information Retrieval (IR) dar. Wesentliche Unterschiede herkömmlicher IR-Systeme zu den neueren Suchmaschinen und ihre Funktionsweise wurden deutlich. Geklärt wurde die für die weitere Arbeit wichtige Struktur von Ergebnislisten mit den Beschreibungsmöglichkeiten gefundener Dokumente sowie Eigenschaften einzelner Treffer. Dies ist insofern bedeutsam, als Anwender anhand der Informationen auf der Ergebnisliste darüber entscheiden, hinter welchen Treffern sie Antworten auf ihre Fragen erwarten. Aufgrund begrenzter Zeit- und Personalressourcen und begrenztem Wissen können keineswegs alle Möglichkeiten zur Beschreibung von gefundenen Dokumenten auf den Ergebnislisten von Suchmaschine realisiert werden. Es stellt sich somit für jeden Suchmaschinenbetreiber die Designfrage, welche Informationen er darstellen soll. Die vorliegende Arbeit sucht eine Antwort, indem sie der Frage nach der optimalen Information auf Ergebnislisten nachgeht. Verschiedene Beschreibungsmöglichkeiten von gefundenen Dokumenten werden daher im Empirieteil untersucht und gegeneinander abgewogen.

Vom Rahmenmodell her betrachtet stellt die Maschine die Realisierung einer Schnittstelle mit den passenden Dialogen, d.h. ein System dar, mit deren Hilfe Anwender ihre Suchaufgaben überhaupt erst durchführen können. Wie sich diese Aufgaben genauer fassen lassen und welchen Einfluss sie auf die Suche haben wird im nächsten Kapitel geklärt.

## 2.3 Aufgabe: Klassifikation und Verhaltensanalyse

Nachdem im vorangehenden Kapitel die Ebene Maschine des Rahmenmodells dargestellt wurde, steht nun die zweite Ebene, die Aufgabe, im Mittelpunkt. Bei der Literaturrecherche wird in diesem Zusammenhang klar, dass je nach Forschergruppe zwei verschiedene



Aspekte unter diesem Begriff gefasst werden. Diese sollen allerdings in der vorliegenden Arbeit differenzierter auseinander gehalten werden.

Ein erster Ansatz legt den Schwerpunkt auf das Informationsbedürfnis der Anwender, das zu einer Aufgabe im Sinne einer Suchaufgabe führt. Ein zweiter Ansatz untersucht das aus dem Informationsbedürfnis folgende Suchverhalten, also die Aktivität des Informationssuchenden im Bezug zur Aufgabe. Im Folgenden werden diese beiden Aspekte unter der Ebene der Aufgabe differenziert und als zwei sich ergänzende Teilaspekte verstanden.

Zunächst wird das die Suche antreibende Informationsbedürfnis näher beleuchtet, indem drei mögliche Klassifikationsschemata dargestellt werden, die zwischen ein, zwei und drei Dimensionen dieses Bedürfnisses postulieren. Die wichtigsten Aussagen aller drei Schemata werden daraufhin in einer für die vorliegende Arbeit verdichteten Synthese zusammengeführt. Anschließend wird das Modell eines Suchprozesses dargestellt und die im vorigen Kapitel behandelten Ergebnisse in den Kontext von Suchprozessen eingegliedert. Der letzte Abschnitt befasst sich erneut schwerpunktmäßig mit Ergebnisseiten, dieses Mal jedoch aus dem Blickwinkel der Anwender. Dazu wird das Verhalten auf Ergebnisseiten einer Aufgabenanalyse unterzogen und der Bogen zum ersten Aspekt, den Informationsbedürfnissen, gespannt, beziehungsweise den daraus resultierenden Suchanfragen.

### **2.3.1 Klassifikation von Informationsbedürfnissen**

Zu Beginn jeder Suche steht ein Bedürfnis nach Information. Die Erfüllung der Informationsnachfrage ist der Motivator der Suche und bestimmt das Ziel der Suche (Marchionini, 1995). Aus dem Blickwinkel einer Suchmaschine ist das Ziel die Bereitstellung von Dokumenten, wohingegen das Ziel aus der Sicht eines Suchenden die Extraktion von Informationen zum Beheben seines Informationsdefizits ist. Die für eine Suche bedeutenden Gesichtspunkte solch eines Informationsbedürfnisses werden in diesem Abschnitt dargestellt. Solch eine Klassifizierung der Suchaufgaben ist sinnvoll, da die Art von Suchaufgaben das Suchverhalten maßgeblich beeinflusst (Lergier & Resnick, 2001). Wird dies nicht berücksichtigt, kann es in Untersuchungen zu uneinheitlichen Verhaltensmustern und Ergebnissen kommen, wie noch im Kapitel zum aktuellen Forschungsparadigma (Kap. 2.5) deutlich werden wird. Informationsbedürfnisse, Suchanfragen und die zugehörigen Antworten können auf einer bis drei Dimensionen betrachtet werden, wie die folgenden Ausführungen zeigen.

#### **2.3.1.1 Drei Dimensionen bei Marchionini**

Marchionini (1995) ordnet die Antworten auf eine Suchanfrage auf den drei Kontinuen Spezifität (specificity), Quantität (quantity) und erwarteter Zeitumfang (timeliness, wörtlich Rechtzeitigkeit) an. Die Spezifität eines Zieles reicht von einfachen Fakten bis hin zu Meinungen und Interpretationen. Einfache Fakten sind klar strukturierte und abgrenzbare Informationseinheiten wie zum Beispiel ein Geburtsdatum. Dem Anwender ist klar, wonach er sucht und sobald er die Information gefunden hat, ist die Suche beendet, wobei er großes

Vertrauen in die Richtigkeit und Vollständigkeit der gefundenen Antwort hat. Am anderen Ende des Kontinuums, bei Meinungen und Interpretationen, ist die Bestimmung einer ausreichenden Antwort komplizierter und damit auch ein Ende der Suche nicht klar bestimmbar. Der Anwender wird oft mehrere, zum Teil widersprüchliche Antworten aus verschiedenen Quellen erhalten und muss daher die Qualität der Antworten gegeneinander abwägen. Solch eine Syntheseleistung zu einer eigenen Meinung erfordert deutlich mehr kognitiven Aufwand und stellt eine höhere Leistung dar als das Aneignen einfachen Faktenwissens (Bloom, 1972). Beispiele für unspezifische Aufgaben sind die Fragen nach den gesundheitlichen Folgen von Ozon oder ob Verhaltenstherapie oder tiefenpsychologische Therapie wirkungsvoller sei.

Die Quantität steht in direkter Beziehung zur Spezifität. Sie beschreibt die Größe der Antwort, gemessen in Bits (die kleinste binäre Informationseinheit). Spezifische Antworten haben normalerweise eine geringere Quantität als unspezifische Antworten. Die Suche nach dem Börsenwert einer Aktie hat in diesem Sinne eine geringere Quantität als die Aufgabe, sich eine Meinung zur Börsensituation insgesamt zu bilden.

Die dritte Dimension umfasst den erwarteten Zeitaufwand bis eine Antwort erreicht wurde. Schnelle Antworten können nur für hochspezifische, quantitativ kleine Ziele erwartet werden und tragen nur wenig zur Erfüllung einer damit verbundenen übergeordneten Aufgabe bei, beispielsweise das Nachschlagen einer exakten Literaturangabe beim Schreiben einer Diplomarbeit. Informationsbedürfnisse mit sehr geringem Zeitumfang stellen daher oft nur einen Zwischenschritt dar, bei der Aufgabe ein umfassenderes Informationsbedürfnis abzudecken. Es lässt sich erkennen, dass Marchioninis Darstellung (1995) die Spezifität als Hauptdimension nahe legt, die die anderen beiden Dimensionen maßgeblich beeinflusst.

### **2.3.1.2 Zwei Dimensionen bei Navarro-Prieto, Scaife und Rogers**

Im Gegensatz zu Marchionini erwähnen Navarro-Prieto, Scaife und Rogers (1999) nur zwei Dimensionen. Zusätzlich zum Kontinuum der Spezifität klassifizieren sie die Antwort danach, ob sie in verteilter Struktur oder kategorischer Struktur vorliegt. Dies übe einen Einfluss auf die Suchstrategie aus. Als kategorische Struktur wird das Vorhandensein einer Kategorie in Suchmaschinen oder Web-Katalogen zum Thema oder das Vorliegen der Antwort auf einer Webseite bezeichnet. Von verteilter Struktur kann dann gesprochen werden, wenn keine Kategorie oder Überblicksseite zur Antwort existiert und die Antwort daher auf verschiedenen Seiten verteilt ist.

Kritisch anzumerken ist jedoch, dass die Autoren offen lassen, anhand welcher Kriterien Suchende diese Differenzierung zu Beginn der Suche treffen können, um dann eine passende Suchstrategie zu wählen. Denn inwieweit Informationen kategorisch oder verteilt vorliegen wird auch dadurch bestimmt, ob vor dem Anwender bereits andere Personen dasselbe Informationsbedürfnis hatten und die Antwort daher auf *einer* Webseite zur Verfügung steht oder eine eigene Kategorie in einem Katalog erstellt wurde. Denn trotz der theoretisch sehr hohen Anzahl an möglichen Suchanfragen finden sich viele Wiederholungen, worauf sich

zum Beispiel die Suchmaschine Ask Jeeves<sup>9</sup>, deren Besonderheit darin besteht, dass sie natürliche Fragen versteht, spezialisiert hat. Zudem befindet sich die Struktur des Internet in ständigem Wandel und daher lässt sich für eine zu findende Antwort eine Unterscheidung nach der Verteilungsstruktur nicht einmalig und dauerhaft vor der Suche feststellen, sondern diese zeigt sich dem Anwender erst während derselben. Da sich diese postulierte Dimension nur bei Experten auswirkte (Navarro-Prieto et al., 1999), ist zu vermuten, dass das Vorwissen um die Kategorisierbarkeit der Antwort in einem Katalog den ausschlaggebenden Faktor darstellte und nicht, ob die Information tatsächlich kategorisch vorlag. Bei der Darstellung ihrer Annahmen nehmen die Autoren ausdrücklich auf das Kontinuum der Spezifität nach Shneiderman (1997) und damit auf eindimensionale Ansätze als Ausgangspunkt ihrer Überlegungen Bezug.

### **2.3.1.3 Eindimensionale Ansätze**

Nur eine Dimension wird von Shneiderman (1997) beschrieben. Er klassifiziert die Suchaufgabe auf einem Kontinuum von spezifischem Faktenfinden bis hin zu unstrukturiertem freien Browsen bekannter Datenbanken und Überprüfung der Verfügbarkeit von Informationen. Dies entspricht im Wesentlichen der Spezifität von Marchionini (1995), wobei Shneiderman die Spezifität in Worten des Suchverhaltens definiert. Auch Rosenfeld und Morville (1998) legen den Schwerpunkt auf eine Dimension und beschreiben die Spezifität als Unterschiede zwischen Informationsbedürfnissen im Sinne unterschiedlicher Erwartungen an die Suche. Auf diesem Kontinuum unterscheiden sie (1) das Suchen von Informationen, deren Vorhandensein bekannt ist und die eindeutig eingegrenzt werden können (known-item searching), (2) die Existenzprüfung von Informationen bei bekanntem Informationsbedürfnis, aber unklarer Zielbeschreibung (existence searching), (3) die explorative Suche für Fälle, in denen mehr über ein Thema gelernt werden soll, aber unbekannt ist, nach welchen Informationen genau gesucht wird (exploratory searching) und (4) die umfassende Suche, wenn ein Themengebiet komplett abgedeckt werden muss (comprehensive searching, research). Ratgeber für erfolgreiches Suchen betonen ebenfalls die Bedeutung der Spezifität für das Suchverhalten und leiten daraus empfehlenswerte Suchstrategien ab (DeMicheli, 1998; Lawrence & Giles, 1998c). So kann zum Beispiel für eine unspezifische, möglichst umfassende Literatursuche für eine wissenschaftliche Arbeit das Schneeballsystem verwendet werden, das den Kreis relevanter Quellen kontinuierlich vergrößert. Danach werden Literaturlisten von Büchern nach weiteren relevanten Büchern durchgesehen und die Literaturlisten dieser Bücher ebenfalls, solange bis kaum neue Angaben gefunden werden.

Trotz der Herangehensweisen der verschiedenen Ansätze, mehr als eine Dimension zur Klassifikation von Informationsbedürfnissen bzw. gesuchten Antworten heranzuziehen, kann abschließend festgestellt werden, dass Spezifität die bedeutendste Dimension darstellt und daher im Untersuchungsdesign dieser Arbeit besondere Berücksichtigung erfährt. Bei der Wortwahl verschiedener Autoren für ihre Klassifikation von Suchaufgaben zwischen

---

<sup>9</sup> <http://www.ask.com/>

Informationsbedürfnis, Suchanfrage und Antworten zeigen sich zwar die von ihnen gesetzten Schwerpunkte im Prozess der Suche, jedoch keineswegs völlig unterschiedliche Konzepte. Nach der Klassifikation von Suchaufgaben ist daher nun auf das Suchverhalten von Anwendern einzugehen.

### 2.3.2 Prozess der Suche

Lässt sich ein Schema im Verhalten der Suchenden während des Suchprozesses identifizieren, so können daraus hilfreiche Zusatzfunktionen (siehe Kap. 2.2.4) für Ergebnisseiten abgeleitet und Methoden zur Verhaltensmessung entwickelt werden. In diesem Zusammenhang wurden bisher sehr unterschiedliche Ansätze verfolgt, um Suchverhalten im Internet zu beschreiben. Neben klassischen, quantitativen Methoden wie der GOMS-Analyse (Peck & John, 1992) stehen Pfadmodelle (Hoelscher & Strube, 1999), Kategorisierungen verschiedener Suchmodi (Choo, Detlor & Turnbull, 1998; Choo et al., 1999), dichotome Unterscheidungen zwischen Experten- und Laienvorgehen (Navarro-Prieto et al., 1999; Weber & Groner, 1999) sowie umfassendere Ansätze, die metaphorische Bezüge zur Nahrungssuche herstellen. Hierbei sind besonders die „Berry picking Technique“ von Bates (1998) sowie das „Foraging-Modell“ von Pirelli und Card (1999) zu erwähnen. In der vorliegenden Arbeit sind solche Ansätze von Interesse, die sich speziell mit dem Suchverhalten bei Suchmaschinen befassen. In ihrem grundlegenden Artikel „Clarifying Search. A User-Interface Framework for Text Searches“ schlagen Shneiderman, Byrn und Croft (1997) vier Phasen vor: *formulation* (Suchanfrage formulieren und Suchmaschine wählen), *action* (Senden der Suchanfrage), *review of results* (Ergebnisse evaluieren) und *refinement* (reformulieren der Suche oder Abbruch). Detaillierter lässt sich der Vorgang in acht Schritte zerlegen (in Anlehnung an Baeza-Yates & Ribeiro-Neto, 1999, S. 263):

1. Informationsbedarf feststellen
2. Auswahl einer Suchmaschine
3. Formulierung der Suchanfrage (Suchsyntax bestimmen)
4. Senden der Suchsyntax an die Suchmaschine
5. Empfangen der Ergebnisseite
6. Evaluation der Ergebnisseite
7. Entweder Beenden der Suche oder
8. Reformulierung der Suche und weiter mit Punkt vier

Solch ein lineares Modell bildet das Suchverhalten in einer vereinfachenden Form ab. Anwender springen allerdings auch auf vorherige Stufen zurück und wechseln nach Bedarf flexibel zwischen einzelnen Suchsystemen als Alternative zu Suchmaschinen, wie die Suche in Katalogen und dem Browsen im Internet (Hoelscher & Strube, 1999). Des Weiteren formulieren neuere Modelle ein dynamisches Informationsbedürfnis anstelle des statischen, in dem die Suchanfrage solange reformuliert wird, bis die Ergebnisseite das Informationsbedürfnis erfüllt. Tatsächlich lernen Anwender während der Suchanfrage aus der Ergebnisliste und dem Besuchen der vorgeschlagenen Seiten, entwickeln möglicherweise weiterführende Fragen und passen dadurch ihr Informationsbedürfnis an (Bates, 1998). Diese hohe Flexibilität und Dynamik im Suchprozess kann dazu führen, dass sich das

ursprüngliche Suchziel verändert und der Anwender andere Ziele verfolgt, als er ursprünglich geplant hatte.

Zuweilen geschieht es, daß man auf der Suche nach einer bestimmten Information von einer anderen Information so ‚beschlagnahmt‘ wird, daß über deren aktueller Dominanz das ursprüngliche Ziel irrelevant oder vergessen wird. Dies bezeichnet man als ‚Serendipity‘-Effekt. (Kuhlen, 1992, S. 312)

Zusätzliche Anstrengung und Konzentration ist notwendig, um mehrere Suchvorgänge, mehrere Suchanfragen in verschiedenen Fenstern bzw. parallel mehrere Treffer zur selben Zeit zu verfolgen. Dieses Mehr an kognitivem Aufwand wird als „cognitive overload“ bezeichnet (Thimbleby, Jones & Theng, 1997). Damit in Zusammenhang steht das Lost-in-Hyperspace-Phänomen, die Tendenz, die Orientierung in nicht-linearen Informationen zu verlieren (Thimbleby et al., 1997). Letzteres tritt besonders dann auf, wenn Anwender sich assoziativ von einem Link zum nächsten bewegen. Am Ende der Assoziationskette lässt sich dann nur schwer rekonstruieren, wo sich der Ausgangspunkt befand (Kuhlen, 1992). All diese negativen Effekte zeigen deutlich, dass Ergebnisseiten die gesuchten Informationen auf eine einfache und brauchbare Weise darstellen sollten.

Die hier gewählte schematische Darstellung des Suchprozesses nach Baeza-Yates und Ribeiro-Neto (1999) ist besonders geeignet, um die verschiedenen Stadien zu verdeutlichen, auf die sich diese Arbeit konzentriert. Hier ist insbesondere festzustellen, dass jedem Schritt eigene Forschungsvorhaben gewidmet werden könnten. Die Bedeutung des Informationsbedürfnisses (1) ist bereits im vorigen Kapitel deutlich geworden. Die ausgewählte Suchmaschine (2) sollte zum Informationsbedürfnis passen, optimalerweise findet sich eine Spezialsuchmaschine, die das Thema adressiert. Unter Umständen sind sogar völlig andere Suchmethoden effizienter. Für Adressauskünfte können die „gelben Seiten“ sinnvoller sein, und auch Experten liefern schnell Antworten in ihrem Fachgebiet. Die Formulierung der Suchanfrage (3) kann durch die Angabe von Synonymen zur Suchanfrage unterstützt werden (vgl. McArthur & Bruza, 2000). Die besondere Leistung einer gelungenen Suchanfrage besteht darin, einerseits spezifisch genug zu formulieren, um irrelevante Treffer zu vermeiden, und andererseits offen genug zu formulieren, um alle wirklich relevanten Dokumente auch in den Ergebnisseiten zu erhalten. Zwischen dem Senden der Anfrage (4) und dem Erhalten der Ergebnisseite (5) steht die suchmaschinenspezifische Technik mit ihren von den Betreibern streng gehüteten Rankingalgorithmen. Die Kernkompetenzen zu diesem Gebiet finden sich in der Mathematik und Informatik. Bei der Evaluation der Ergebnisseiten (6) werden vom Anwender erneut Entscheidungen (vgl. Kap. 2.4.1) und Eingaben erwartet. Gefundene Dokumente werden aufgerufen und die Suche bei erfolgreicher Deckung des Dokuments mit dem Informationsbedürfnis beendet (7). Viel häufiger jedoch wird es vorkommen, dass einige oder alle vorhergehenden Schritte wiederholt werden (8). Das Verhalten der Anwender während der Evaluation (6) wird gemäß dem Schwerpunkt der vorliegenden Arbeit im Folgenden einer genaueren Aufgabenanalyse unterzogen.

### 2.3.3 Hierarchische Aufgabenanalyse und Task Description

Für eine Aufgabenanalyse (Task Analysis) bieten sich zahlreiche unterschiedliche Methoden an (einen Überblick bietet Jonassen, Hannum & Tessmer, 1989). Eine weit verbreitete Technik in der Mensch-Computer-Interaktion ist die Aufgaben-Zerlegung (Dix, Finley, Abowd & Beale, 1995), wie sie auch in der hierarchischen Task-Analyse (Diaper, 1989) angewendet wird. Dabei wird eine Aufgabe in mehrere Ziele aufgeteilt und diese dann wiederum in mehrere Unteraufgaben zerlegt. Die übergeordnete Aufgabe ‚Evaluation der Ergebnisseite‘ könnte für jeden Treffer in die Einzelziele ‚Evaluation eines Treffers‘ zerlegt werden und diese wiederum in die Evaluation jeder einzelnen Eigenschaft eines Treffers (vgl. dazu Kap. 2.2.4). Am unteren Ende der Hierarchie würde dann eine Entscheidungsfrage stehen. Für die Evaluierung der Eigenschaft Rangposition in der Ergebnisliste würde die Frage beispielsweise lauten: ‚Kann ich den Treffer aufgrund der Rangposition als für mein Informationsbedürfnis wichtig oder unwichtig klassifizieren, d.h. soll ich das repräsentierte Dokument lesen oder nicht?‘ Auf welche Weise solche Einzelentscheidungen verbunden werden und zur Gesamtentscheidung führen, das repräsentierte Dokument zu lesen oder nicht, wird in einem eigenen Kapitel zum Entscheidungsverhalten genauer untersucht (s. Kap. 2.4.1). Im Moment genügt es festzustellen, dass es sich widersprechende Annahmen über die zugrundeliegenden kognitiven Vorgänge gibt.

Für die vorliegende Arbeit ist die Analysetechnik „Task Description“ ergiebig, wie sie von Miller (1966) beschrieben wurde. Ihre Wurzeln liegen in Rahmen der Systemtheorie und der beschreibenden Verhaltenspsychologie und sie befasst sich insbesondere mit Mensch-Maschine-Interaktionen. Sie helfen bei der Entwicklung von Interfaces und dadurch auch bei der benutzerfreundlichen Gestaltung von Ergebnisseiten, die wie in Kapitel 2.2.2 verdeutlicht eine Form von Interface sind und sich daher besonders für die vorliegende Arbeit eignen. Auch in der Task-Description können bei Bedarf verschiedene Abstraktionsniveaus wie in der hierarchischen Task-Analyse unterschieden werden. Der Unterschied und gleichzeitige Vorteil der Task-Description liegt darin, dass Verhalten modelliert werden kann, ohne unsichere Annahmen über kognitive Vorgänge treffen zu müssen. Entsprechend können daraus objektive Messmethoden zum Verhalten entwickelt werden.

Das Verhalten auf Ergebnisseiten lässt sich besonders anschaulich in Flussdiagrammen darstellen. Ursprünglich wurden die verschiedenen Elemente einer Task-Description zwar in Tabellen dargestellt, jedoch lassen sich alle wichtigen Elemente in Flussdiagrammen integrieren und ermöglichen dadurch den Prozess, insbesondere Verhaltenswiederholungen, also Schleifen genauer zu modellieren. Auslösende Reize und Handlungen ergeben sich leicht aus dem dargestellten Ablauf (viereckige Formen). Indikatoren, die Entscheidungen veranlassen, werden in rautenförmigen Formen mit abgehenden Verbindungen für die Entscheidungen beschrieben. Start- und Endpunkte des Verhaltens werden in Kreisen gezeichnet. Eine Voraussetzung zur Anwendung der Task-Description ist, dass sich das Verhalten auf Ergebnisseiten beobachten und in einzelne Schritte zerlegen lässt sowie, dass

die Zusammensetzung dieser Schritte ein erfolgreiches Suchverhalten ergibt. Da sich die Forschung zum Suchverhalten hauptsächlich mit der Zerlegung in Verhaltensabschnitte und deren Verbindung zu erfolgreichem Suchverhalten beschäftigt, soll diese Annahme auch hier akzeptiert werden.

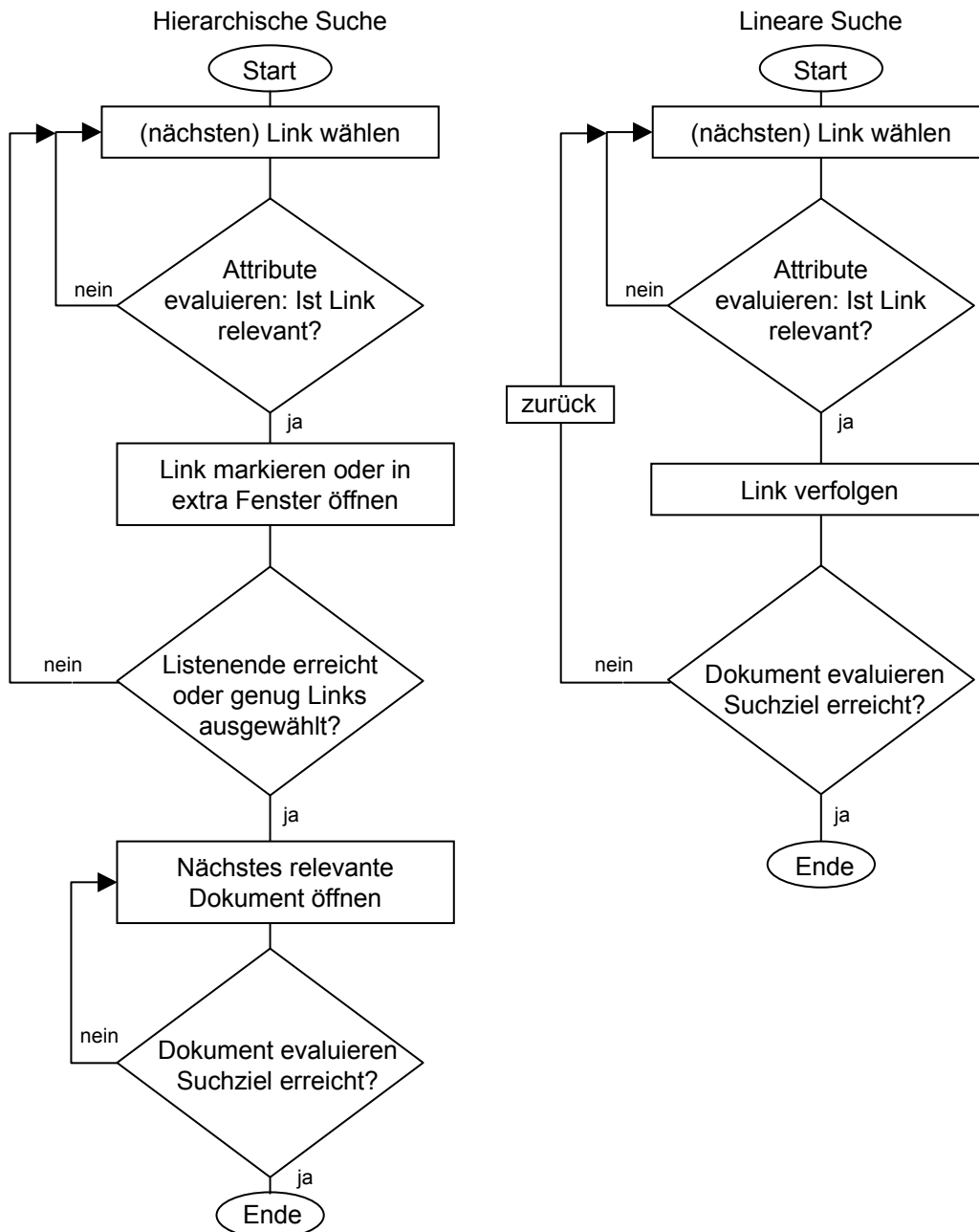
### 2.3.4 Anwendung von Task Description auf Ergebnisseiten

Auf der Ergebnisseite stehen Anwender vor dem Problem, diejenigen Treffer auszuwählen, die für ihr Informationsbedürfnis relevant erscheinen. Prinzipiell sind zwei extreme Verhaltensweisen möglich (vgl. Abbildung 6), wobei in der Praxis jedoch allerlei Mischformen vorkommen, da wie schon gezeigt das Suchverhalten keineswegs linear ist. Trotzdem können Anwender zur Evaluation der Ergebnisseite streng linear vorgehen. Nach der Entscheidung für einen relevanten Treffer rufen sie umgehend das entsprechende Dokument auf, prüfen, ob es eine genügende Antwort enthält, und kehren, sollte dies nicht der Fall sein, wieder zur Ergebnisseite zurück, um mit dem Evaluieren des nächsten Links fortzufahren. Die zweite Möglichkeit, die hierarchische Vorgehensweise, zerlegt den Suchvorgang in zwei Teile. Zunächst wird jeder Link geprüft. Erscheint dieser relevant, wird er markiert. Dies kann auch durch tatsächliches Anklicken geschehen, wie dies häufig bei spezialisierten Suchinterfaces zur Bibliografiercherche möglich ist. Ist dies nicht möglich, wird der Treffer im Gedächtnis behalten oder in einem neuen Fenster zur späteren Verwendung geöffnet. Nachdem die gesamte Ergebnisliste auf diese Weise evaluiert wurde, werden die markierten Dokumente auf Antworten durchsucht.

Beide Varianten werden in dieser Reinform kaum vorkommen. Beim linearen Vorgehen können Anwender häufig auf weiterführende Verweise stoßen und diese verfolgen, inklusive der oben beschriebenen negativen Auswirkungen wie dem Serendipity-Effekt (Kap. 2.3.2). Um dem entgegenzuwirken werden möglicherweise wie beim hierarchischen Vorgehen eigene Browserfenster mit den Suchergebnissen offen gehalten. Beim hierarchischen Vorgehen wiederum können Anwender jederzeit die Evaluation der Linkliste abbrechen, sobald dies aufgrund der Anzahl an ausgewählten Dokumenten sinnvoll erscheint. Einen Link sofort auszuwählen, wenn dieser akzeptabel erscheint, wird von Lergier und Resnick (2001) als selbst-terminierende Strategie (self-terminating strategy) bezeichnet. Behalten Anwender im Gedächtnis, welcher Treffer am erfolgversprechendsten ist, und wird dann der erfolgversprechendste nach Abarbeitung der gesamten Liste gewählt, sprechen sie von einer erschöpfenden Strategie (exhaustive strategy). Anwender würden daher bei hoher Aufgabenspezifität die selbst-terminierende Strategie wählen, bei unspezifischen Suchaufgaben sei das Vorgehen gemischt.

Trotz zahlreicher Arbeiten über das Suchverhalten allgemein wurde das Suchverhalten auf Ergebnisseiten selbst bisher vernachlässigt. Lergier und Resnick (2001) bestätigen:

It is currently unknown which of these parsing strategies describes behavior when searching online. The selection of parsing strategies may depend on many factors, such as the experience or personal preferences of the user, or the nature of the task. (S. 3)



**Abbildung 6:** Flussdiagramm zu linearem und hierarchischem Suchverhalten auf Ergebnisseiten.

### 2.3.5 Zusammenfassung und Ableitungen

In diesem Kapitel kristallisierte sich deutlich die Spezifität der Aufgabe als entscheidendste Dimension bei der Klassifikation von Suchanfragen heraus. Der Suchprozess wurde in acht Teilschritte gegliedert, wobei die Evaluation von Ergebnisseiten einer genaueren Betrachtung unterzogen wurde. Bei dieser Selektionshandlung lassen sich prinzipiell ein lineares und ein hierarchisches Vorgehen unterscheiden.

Daraus ergeben sich zusammen mit dem Faktor System einige Folgen für die Gestaltung von Ergebnisseiten. Existiert eine hohe Spezifität, wie dies zum Beispiel bei Fahrplanauskünften und in klassischer Bibliografierecherche mit einem OPAC der Fall ist, so lassen sich speziell angepasste Ergebnisseiten konstruieren, welche die gewünschten Informationen wie Zeilen



und Namen der Bahnhöfe oder Bibliografieangaben enthalten. Durch die richtige Auswahl der anzuzeigenden zusätzlichen Informationen (Kap. 2.2.4) kann der Anwender sein Ziel bereits direkt durch die Ergebnisseite erreichen, ohne dass diese nur eine Zwischenstation auf dem Weg zum Ziel darstellt, wie dies meistens bei Suchmaschinen der Fall ist. Jedoch kommt es auch bei Suchmaschinen vor, dass diese die Antwort bereits auf der Ergebnisseite präsentieren. Bei einem sehr spezifischen Informationsbedürfnis, beispielsweise nach einem Geburtsdatum, kann dieses bereits in der Beschreibung vorkommen. Solche Fälle sind jedoch nur bei stark konventionalisierter Schreibweise möglich, wie der Ausdruck „Burrhus Frederic Skinner (1904-1990)“ verdeutlicht. Ein lineares Vorgehen findet sich am ehesten bei spezifischen Aufgaben. Je unspezifischer eine Aufgabe jedoch wird und dadurch auch die Komplexität und Dauer des gesamten Suchverhaltens zunimmt, desto eher wird das hierarchische Vorgehen gewählt werden, um den cognitive overload, d.h. die zur Koordinierung des Suchverhaltens extra aufzuwendende Energie, so gering wie möglich zu halten. Ansätze zur Unterstützung der Selektionsentscheidungen und einer Reduzierung des cognitive overload zeigen sich beispielsweise in der Dokumenten-Vorschau-Funktion verschiedener Suchmaschinen (z.B. bei Fireball<sup>10</sup> und Excite<sup>11</sup>).

Nach der Task-Description soll nun ebenfalls auf die dahinterliegenden kognitiven Prozesse eingegangen werden, wie sie im Folgenden unter der Ebene Mensch subsumiert werden.

## 2.4 Mensch: Entscheidungspsychologie und Ergonomie

Der Ebene Mensch zugeordnet werden Kognitionen und allgemeine konstante Gesetzmäßigkeiten, die für alle Menschen gleichermaßen Gültigkeit besitzen. Der Begriff Mensch soll dabei deutlich machen, dass hier Anwender im Mittelpunkt stehen und nicht die Technik oder Aufgabenstellung. Bei der Beschreibung von Menschen können einerseits die Unterschiede zwischen verschiedenen Menschen herausgestellt (differentielle Psychologie) oder aber ihre Gemeinsamkeiten (allgemeine Psychologie) hervorgehoben werden. Im Folgenden werden die Unterschiede in den Kognitionen zwischen Anwendern unter entscheidungspsychologischen Gesichtspunkten betrachtet, danach wird auf die Bedeutung von Wissen als moderierende Variable im Suchprozess eingegangen. Den Gemeinsamkeiten zwischen Anwendern wird durch die Explizierung von Ergonomierichtlinien über Ergebnisseiten Rechnung getragen.

### 2.4.1 Die Selektionsentscheidung

Handelte das vorherige Kapitel vom direkt beobachtbaren Verhalten auf Ergebnisseiten, soll nun der Frage nach den zugrundeliegenden Kognitionen nachgegangen werden. Treffer einer Ergebnisseite dienen dem Zweck, den Anwender zu redaktionellen Inhalten zu führen. Wirth und Schweiger (1999) analysieren solche Links der Ergebnisseite im Kontext der Entscheidungspsychologie:

---

<sup>10</sup> <http://www.fireball.de/>

Damit bilden sie im entscheidungspsychologischen Sinne *Optionen*, d.h. Selektionsalternativen. Jede Erläuterung, die eine Option – einen Link – näher beschreibt, ist als Zusatzinformation oder *Attribut* aufzufassen, und erleichtert die Selektion potenziell. [...] Attribute ermöglichen die Evaluation eines Links, ohne dass das Ziel eines Links angesteuert werden muss. (S. 55f.)

Der Kern einer Ergebnisseite lässt sich demnach anhand der Anzahl an Optionen, also der Anzahl der Treffer und ihrer Attribute, charakterisieren. Die Attribute entsprechen hierbei den Eigenschaften eines Treffers: Position, Titel, Beschreibung, zusätzliche Informationen und zusätzliche Funktionen (vgl. Kap. 2.2.4). Anhand der Attribute entscheidet der Anwender, welche Links er weiterverfolgen möchte. Nach Saito und Ohmura (1998) wird die Information aus den Attributen mit einem mentalen Modell des Suchziels verglichen. Eine Entsprechung der Wörter in der Suchaufgabe mit den Beschreibungen von Links verkürzt nach Khan und Locatis (1998a) die Entscheidungsdauer und beschleunigt damit die Suche, erhöht jedoch nicht deren Genauigkeit. Eine ähnliche Unterscheidung in Zielbegriff (Suchaufgabe) und Linkbegriff (Beschreibung) treffen Wandke und Hurtienne (1999). Die Autoren unterscheiden verschiedene Relationen der Begriffe zueinander (z.B. Hierarchien, Synonyme und Nebenordnungen) und gehen davon aus, dass eine höhere Übereinstimmung der beiden Begriffe zu einer höheren Wahrscheinlichkeit in der Linkauswahl führt.

Auf welche Weise die Attribute bewertet werden, hängt maßgeblich vom zugrundegelegten Entscheidungsmodell ab. Payne, Bettman und Johnson (1993) beschreiben das Modell eines adaptiven Nutzers. Demnach passen Anwender die Höhe an investierter Zeit, den kognitiven Aufwand und die gewünschte Qualität der gesuchten Antwort der jeweiligen Situation an. Bei geringer Motivation und persönlicher Bedeutung der gesuchten Information würden daher auch suboptimale Ergebnisse akzeptiert und der Suchprozess entsprechend früher beendet. Wird eine hohe Anzahl relevanter Informationen bzw. eine umfassende Recherche gewünscht, ist ein entsprechend höherer Einsatz seitens der Anwender notwendig. Anwender handeln adaptiv, indem sie Kompromisse machen zwischen dem erforderlichen Aufwand und der Genauigkeit der Ergebnisse. Sie passen ihr Entscheidungsmodell also den Umständen an.

Komplexe rationale Entscheidungsstrategien sind möglich, bei denen eine vollständige Analyse aller Attribute erfolgt, wobei jeder Treffer anhand einer gewichteten Additionsregel bewertet wird (Wirth & Schweiger, 1999a). Dabei wird jedem Attribut ein Wert zugewiesen, der angibt, wie wichtig es für die Suche ist. Außerdem wird für jede konkrete Ausprägung eines Attributes bei jedem Treffer ein Wert festgestellt, der angibt, wie hoch der Deckungsgrad der Ausprägung dieses Attributes mit dem Informationsbedürfnis ist. Die persönliche Relevanz  $R$  eines Treffers kann daher mit Formel 1 beschrieben werden.

---

<sup>11</sup> <http://www.excite.de/>

Formel 1: Gewichtete Additionsregel zur Relevanzbestimmung.

$$R = \sum_{m=1}^n i_m \cdot j_m$$

n = Anzahl der Attribute, m = Zähler über alle Attribute,

j = Deckungsgrad der Ausprägung des Attributes m mit dem Informationsbedürfnis,

i = partielle Bedeutung des Attributes m für die Suche.

Die Frage ist nun, welche Treffer tatsächlich verfolgt werden. Gemäß Schwellenwertmethoden oder Satisficing-Regeln (Jungermann, Pfister & Fischer, 1998) wird jeder Treffer verfolgt, der eine bestimmte Relevanz überschreitet. Anwender haben also einen internen Cut-Off-Wert festgelegt, der überschritten werden muss. Ein Link wird sofort gewählt, wenn der Anwender eine ausreichende Nähe zum Informationsbedürfnis vermutet, also der mentale Vergleich eine hinreichend hohe Akzeptanz erwarten lässt. Dies entspricht der selbst-terminierenden Strategie nach Lergier und Resnick (2001). Würde nur der Treffer mit der höchsten Relevanz gewählt, so wäre das ein Beispiel für eine erschöpfende Strategie. Dies kann auch als Optimizing-Regel im Gegensatz zur Satisficing-Regel interpretiert werden. Der Schwellenwert erscheint dabei umso höher, je weniger Treffer der Anwender bereit ist zu verfolgen und je spezifischer die Aufgabe ist. Die exakte Höhe lässt sich im jeweiligen Suchkontext schätzen, sofern die Einschätzungen der Anwender bezüglich der Attribute (Variablen i und j) sowie ihre Entscheidungen bekannt sind.

Geringere kognitive Anforderungen entstehen bei Entscheidungsvereinfachungen. Dazu kann die Gewichtung weggelassen oder nur ein Teil der Attribute bewertet werden. Bei einer nur rudimentären Analyse mit minimalem kognitiven Aufwand bewertet der Anwender nur noch ein Attribut, das er für besonders wichtig hält und ignoriert sonstige Attribute (lexigrafische Regel). Solch eine Vereinfachung ist besonders für das Attribut Zusammenfassung sinnvoll. Aber auch beim Wiederfinden eines Dokumentes genügt gelegentlich schon ein Blick auf die URL für eine schnelle Entscheidung.

Jungermann, Pfister und Fischer (1998) berücksichtigen die Art und den Umfang kognitiven Aufwandes von Entscheidungen, indem sie vier Ebenen unterscheiden. Dabei nimmt das Niveau kognitiver Anstrengung von routinisierten über stereotype hin zu reflektierten und konstruktiven Entscheidungen zu, wobei sich Entscheidungen auf Ergebnisseiten zwischen stereotypen und reflektierten Entscheidungen bewegen. Routinisierte Entscheidungen zeichnen sich durch eine stets gleiche Anzahl und Art von Optionen und Attributen in immer wiederkehrenden Situationen aus. Zwischen ihnen wird weitgehend automatisch entschieden und sie benötigen den geringsten kognitiven Aufwand. Wählt ein Anwender für seine Suche stets dieselbe Suchmaschine als Startpunkt aus, so wäre ihr Aufruf im Webbrowser eine routinisierte Entscheidung. Stereotype Entscheidungen sind auf Ergebnisseiten zu erwarten, wenn nach erlerntem Bewertungsschema und durch ein „holistisches Affekturteil“ (Jungermann et al., 1998) entschieden wird. Darunter verstehen die Autoren, dass ein Link durch eine globale Bewertung ohne die detaillierte Analyse einzelner Attribute intuitiv selektiert wird. Wendet ein Anwender häufig, fast automatisiert die lexigrafische Regel auf ein einfaches Attribut an, so wären dies ebenfalls stereotype

Entscheidungen. Ein Link erscheint dann für den Anwender unmittelbar relevant oder irrelevant. Demgegenüber denkt der Anwender bei der reflektierten Entscheidung explizit über die Attribute nach und versucht diese mit dem Informationsbedürfnis in Beziehung zu setzen. Dies erfordert erheblichen kognitiven Aufwand und entspricht den oben angesprochenen rationalen Entscheidungsstrategien. Konstruktive Entscheidungen sind auf Ergebnisseiten nicht zu beobachten. Sie sind durch das Fehlen von Optionen gekennzeichnet, so dass diese erst gefunden werden müssen, oder dadurch, dass Anwender die Werte, anhand derer entschieden werden soll, erst erstellen müssen. Beides trifft auf Ergebnisseiten nicht zu.

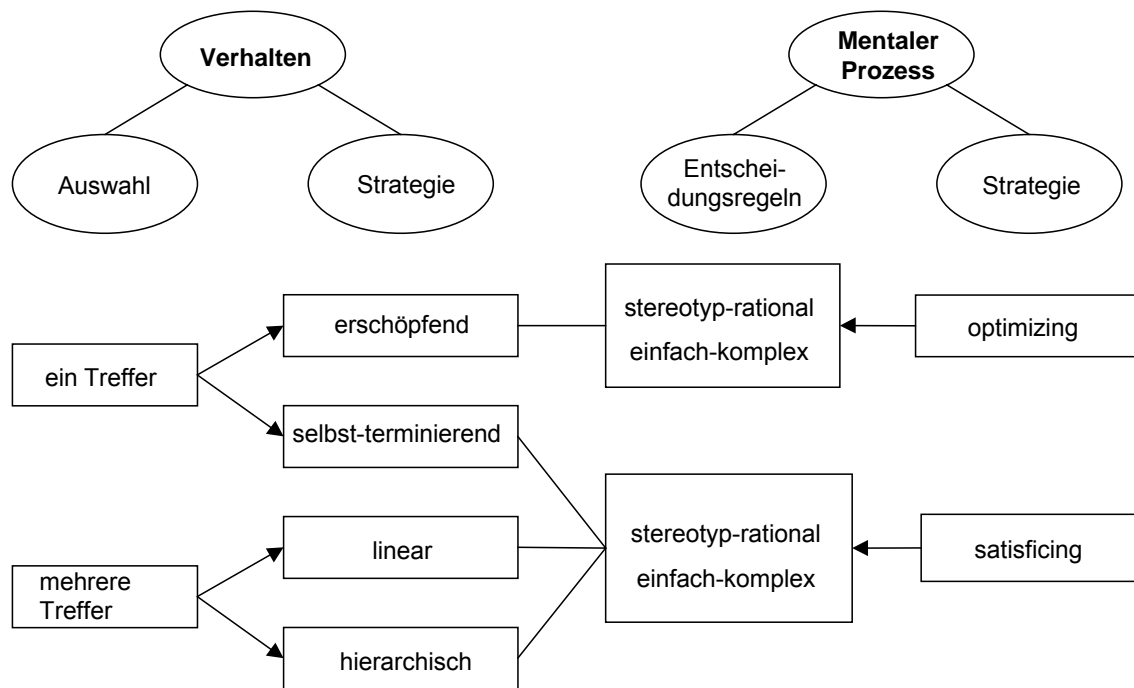
Zwar legt die vorliegende Arbeit ihren Schwerpunkt auf rationale Selektionsstrategien, jedoch soll an dieser Stelle auch kurz erwähnt werden, dass durchaus nicht-rationale Selektionsmodelle existieren. Diese berücksichtigen stochastische Regeln, affektive und habituelle Entscheidungen. Sie sollen jedoch eher das Browsen im Internet erklären und sind deshalb für das besonders zielgerichtete Suchen mit Suchmaschinen weniger geeignet und werden daher hier nicht näher erläutert (vgl. Wirth & Schweiger, 1999a).

#### **2.4.2 Zum Zusammenhang von Verhalten und Kognitionen**

Bisher wurden unterschiedliche Ansätze zum Selektionsverhalten beschrieben, sowohl im beobachtbaren Verhalten wie auch bei den zugrundeliegenden mentalen Prozessen. Diese stehen in einer bestimmten Beziehung zueinander (Abbildung 7), wie die folgenden Ausführungen zeigen werden. In ihrer Untersuchung legten Lergier und Resnick (2001) ihr Design so an, dass Anwender nur einen Treffer auf Ergebnisseiten auswählen konnten. Daraus ergaben sich die beiden von ihnen beobachteten Strategien, das erschöpfende und das selbst-terminierende Vorgehen. Stellt man das Verhalten in einen größeren Zusammenhang, so lässt sich feststellen, dass Anwender in der Mehrzahl der Fälle durchaus mehrere Treffer auswählen werden. Hier kann eine lineare und eine hierarchische Strategie unterschieden werden, wie sie vom Autor der vorliegenden Arbeit vorgeschlagen werden.

Bei mentalen Prozessen lassen sich ebenfalls zwei grundlegende Strategien postulieren. Anwender verfolgen eine optimierende Strategie (optimizing), wenn sie nur den relevantesten Treffer auf einer Ergebnisseite auswählen. Bei der Schwellenwertmethode (satisficing) wählen sie den- oder diejenigen Treffer aus, die eine gewisse Nähe zum Informationsbedürfnis unterschreiten. Bei beiden Strategien können wiederum sehr unterschiedliche Entscheidungsregeln zur Anwendung kommen, die sich nach ihrer Komplexität und in das Schema nach Jungermann, Pfister und Fischer (1998) einordnen lassen.

Da Anwender jedoch adaptiv handeln (Payne et al., 1993), sind sie sehr flexibel in ihren Strategien. So mag es vorkommen, dass auch innerhalb der linearen Strategie bei der Auswahl des nächsten in Frage kommenden Treffers eine Optimizing- anstelle der Satisficing-Strategie angewendet wird.



**Abbildung 7:** Integratives Modell zum Selektionsverhalten auf Ergebnisseiten.

Dann würden Anwender nicht nur den nächsten sinnvoll erscheinenden Treffer auswählen, sondern innerhalb jeder Schleife (vgl. Abbildung 6) die erschöpfende Methode auf die gesamte Ergebnisliste anwenden, wobei bei jedem Durchgang, die bereits ausgewählten Treffer eliminiert würden und der jeweils nächstbeste Treffer gewählt würde. Nach der Theorie des adaptiven Anwenders ist zu vermuten, dass solch ein komplexes Verhalten aufgrund der hohen erforderlichen Kontrolle und Gedächtniskapazität nicht oder nur selten vorkommt, da die anderen Strategien mit weniger Aufwand zum selben Ergebnis führen. So lassen sich die verschiedenen Strategien häufig miteinander verknüpfen, eine abschließende Festlegung, *wann genau* welche Strategien angewendet und wie sie verknüpft werden, muss hier jedoch offen bleiben, denn „generell ist festzuhalten, dass die Forschungslage in Bezug auf Selektionsentscheidungen im Internet bislang sehr dürftig ist“ (Wirth & Schweiger, 1999a, S. 65). Die in diesem Abschnitt beschriebene Integration kann deshalb nur als ein Ansatz zu einer Klärung verstanden werden, welche Strategien denkbar sind und worauf bei einem Untersuchungsdesign geachtet werden kann.

### 2.4.3 Die Bedeutung von Wissen

Eine bedeutende Rolle für erfolgreiches und effektives Suchen spielt das Wissen der Anwender. Das Phänomen des Lost-in-Hyperspace verliert verständlicherweise für diejenigen Anwender an Bedeutung, die mit dem Internet und insbesondere den Navigationsmöglichkeiten besonders vertraut sind. Unterschiede im Wissen sind demnach ein wichtiger Mediator beim Such- und Entscheidungsverhalten (Weber & Groner, 1999). Für die Navigation im Internet relevante Wissensstrukturen lassen sich dabei besonders gut durch die Theorie der mentalen Modelle beschreiben (Borgman, 1986; Noller, 2000). Diese beschreibt Wissen als ein die Realität abbildendes mentales Modell, das sich Personen von einem System konstruieren (Dutke, 1994). Es enthält sowohl Strukturen als auch Inhalte und

erlaubt Schlussfolgerungen für eigenes Verhalten. Bei den Wissensinhalten zeichnet sich besonders die Einteilung von Weber und Groner (1999) durch besondere Relevanz für den Untersuchungsgegenstand Suchmaschinen im Internet aus. Dieser Ansatz wird hier daher als Grundlage gewählt, um darin andere Befunde aus dem Kontext der Suche einzugliedern.

Weber und Groner unterscheiden drei Formen von Wissen, die für einen effizienten und erfolgreichen Suchprozess im Internet von Bedeutung sind (Weber & Groner, 1999, S. 182):

1. Wissen über das Web
2. Wissen über Suchinstrumente
3. Wissen über den Suchgegenstand

Wissen über das Web umfasst die Kenntnis von Strukturen und möglichen Inhalten des Internets. Darin enthalten ist die Kenntnis abstrakter möglicher Strukturmerkmale, beispielsweise dass Knoten (Webdokumente) mit Links verbunden werden (vgl. Nielsen, 1990), aber auch Annahmen über Aufbau und Verbund konkreter Webseiten, die zwischen Anwendern erheblich differieren können (Panero, 1995). Pollock und Hockley (1997) machen deutlich, wie eine Unterschätzung der Globalität des Internets zu einer unüberschaubaren Treffermenge führen können. Eine hohe Ähnlichkeit zwischen der abgebildeten Wirklichkeit und dem mentalen Modell sowie sinnvolle Metaphern (vgl. Noller, 2000) helfen damit bei einer effektiven Navigation während der Suche.

Zum Wissen über Suchinstrumente gehört die Kenntnis, zu welchem Zweck welche Suchmaschine eingesetzt werden sollte, wie diese zu bedienen ist sowie welche Suchstrategien eingesetzt werden können. Obwohl bei Weber und Groner (1999) keine Angaben darüber zu finden sind, bilden sich Anwender auch über Suchinstrumente ein mentales Modell, das Vorhersagen darüber erlaubt, wie Suchmaschinen auf Anwendereingaben reagieren werden (Borgman, 1986). Zur passenden Auswahl einer Suchmaschine ist die Kenntnis ihres Deckungsgrades notwendig, d.h. welchen Bereich des Internets eine Suchmaschine indexiert (vgl. dazu Lawrence & Giles, 1998b) und inwiefern sich dieser mit dem eigenen Informationsbedürfnis überschneidet. Ein falsches Modell kann dazu führen, dass Anwender mit einer Suchmaschine, die nur ein bestimmtes Webangebot indexiert (z.B. bei Die Suchfibel), fachfremde Suchanfragen durchführen und dadurch keine relevanten Ergebnisse erhalten. Zur Interpretation von Ergebnisseiten hilft Wissen über die Rankingalgorithmen von Suchmaschinen (vgl. dazu Grossan, 1997; Kleinberg, 1998; Lawrence & Giles, 1998a) und selbstverständlich ist auch Wissen über das Funktionieren einer Suchmaschine allgemein von Vorteil (vgl. dazu Kap. 2.2.2). Der effektive Einsatz von Suchmaschinen scheint für viele Anwender allerdings ein größeres Problem darzustellen. Das Wissen über die korrekte Suchsyntax und über die erfolgreiche Wortwahl bei der Sucheingabe scheint gering. Häufig verwenden Anwender nur wenige Wörter für eine Suchanfrage und erhalten so eine hohe und zu unspezifische Treffermenge (Lawrence & Giles, 1999; Spink, Bateman & Jansen, 1998). In einer Untersuchung von Jansen, Spink und Pfiff (2000) betrug die durchschnittliche Wortanzahl von Suchanfragen nur 2,21. Bei Hoelscher und Strube (1999) lag dieser Wert für durchschnittliche Anwender bei 1,66 und für Experten bei 3,64 Worte. Auch die Verwendung von booleschen Verknüpfungen ist für

viele Anwender zu komplex, und diese werden häufig falsch angewendet (Greene et al., 1990; Jansen, Spink & Saracevic, 1998; Nielsen, 1997). Diese Wissenslücken im Bereich der Suchinstrumente und verwendbarer Strategien wollen zahlreiche Ratgeber (vgl. Hartmann, Näf & Schäuble, 2000; Pior & Tiedemann, 2000) und Informationsangebote über erfolgversprechende Denkmethode(n) (vgl. Auble & Delongchamp, 2001; Brem & Boyes, 2000) schließen. Zu den möglichen Suchstrategien (vgl. Notess, 1997, 1998) zählen zum Beispiel die Verwendung mehrerer Suchmaschinen, der Einsatz unterschiedlicher Schreibweisen eines Wortes und die Beschränkung der Suche auf bestimmte Felder, wie das Datum oder die Sprache eines Dokuments (eine Liste von Artikeln über Suchstrategien findet sich bei Gunn, 2002). Hsieh-Yee (1993) zeigt, dass sich die Menge an Sucherfahrung in der Auswahl effektiver Strategien niederschlägt. Die Erfahrung im Umgang mit Computern, wie sie von mehreren Autoren als ebenfalls bedeutend für die Suche eingeschätzt wird (Drori, 2001; Hassenzahl & Prümper, 1999; Navarro-Prieto et al., 1999; Unz et al., 1999), lässt sich zusätzlich unter den Bereich der Suchinstrumente fassen, da Computer in diesem Sinne ebenfalls ein Suchinstrument darstellen. Eine Differenzierung erhält der Bereich Computererfahrung, indem zwischen deklarativem und prozeduralem Wissen unterschieden wird (Müsseler & Prinz, 2002). Diese Unterteilung erlaubt die Berücksichtigung unterschiedlicher durch Computerwissen vermittelter Effekte. Erstens zeigt sich deklaratives Wissen zu Computererfahrung in der richtigen Verwendung der Browserfunktionen (z.B. History- bzw. Verlauffunktion, Zurückschaltfläche, Anklicken von Links). Zweitens enthält das prozedurale Wissen die motorischen Fähigkeiten in der Bedienung des Computers. Dies beinhaltet die bekannte Beobachtung, dass erfahrene Computer-Anwender schneller und sicherer im Umgang mit der Bedienung der Maus sind als unerfahrene.

Wissen über den Suchgegenstand wird auch als Domänen- oder Bereichswissen bezeichnet (Marchionini, 1995). Es ist das vom Internet unabhängige Wissen im gesuchten Sachgebiet und hilft bei der Erstellung der Suchanfragen und der Bewertung der Ergebnisse. Erfahrene Anwender verwenden mehr Synonyme, die sie mit Hilfe von Thesauri erhalten, sowie speziellere und allgemeinere Suchwörter, um die Suche einzuschränken oder auszuweiten (Hsieh-Yee, 1993). Durch eine bereits vorhandene Wissensstruktur im gesuchten Gebiet können Angaben auf Ergebnisseiten leichter in relevante und irrelevante Informationen unterschieden und die wichtigsten Konzepte im Informationsbedürfnis erkannt werden (Marchionini, 1995). Eine Erweiterung erfährt das Konzept ‚Wissen über den Suchgegenstand‘ durch die Hinzunahme der Allgemeinbildung, wie besonders Pollock und Hockley (1997) betonen. Da bekannt ist, dass beispielsweise Rechtschreibschwächen eine Suche behindern können (Fidel et al., 1999), wirkt sich Allgemeinbildung in der korrekten Schreibweise von Suchwörtern aus. Nach der differentialpsychologischen Betrachtung der Anwender hinsichtlich Selektionsverhalten und Wissen folgt nun der allgemeinspsychologische Aspekt in Form von Betrachtungen zur Ergonomie.

## 2.4.4 Ergonomie

Trotz der vielen Unterschiede zwischen den Anwendern sind auch viele allgemeinspsychologische Gemeinsamkeiten vorhanden, die in Anweisungen zur Gestaltung von Ergebnisseiten bzw. Webseiten und Programmen allgemein umgesetzt werden können. Zahlreiche Style Guides und Designrichtlinien (vgl. für einen Überblick Ohnemus, 1997) mündeten daher in der europäischen Norm EN ISO 9241, die Kriterien für „Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten“ definiert und damit einen allgemeinen, gesetzlichen Standard geschaffen hat. Das Gesetz zur Bildschirmarbeitsverordnung (BildscharbV) vom 04.12.1996 bietet die Grundlage zur Anwendung der Norm in deutsches Recht. Für die Gestaltung von Ergebnisseiten interessant sind hierbei der Teil zehn mit sieben Grundsätzen der Dialoggestaltung und der elfte Teil mit drei Anforderungen an die Gebrauchstauglichkeit. Kommentare und Erläuterungen zu Teil 10 finden sich bei Schneider (1998). Zwar existieren aufzählende, spezielle Leitfäden zur Ergebnisseitengestaltung (Rappoport, 1999; Schulz, 2001), ein theoretischer Bezug der allgemeineren Norm auf Ergebnisseiten erscheint jedoch notwendig, um die Vergleichbarkeit von Forschungsergebnissen zu sichern. Der elfte Teil beschreibt, wie die sieben Grundsätze anhand der Anforderungen Effektivität, Effizienz und Zufriedenheit gemessen werden können. Auf die Anforderungen wird daher im Methodenteil der Arbeit näher eingegangen (siehe Kap. 3.1.6).

Da die Teile zehn und elf der EN ISO 9241 eine zentrale Stellung für die vorliegende Arbeit einnehmen und die Übertragung der Norm auf Ergebnisseiten bisher ausblieb, erscheint es angebracht, diese näher zu beleuchten. Von Suchmaschinen weitestgehend unabhängige Ausführungen zur Norm findet der Leser beim Beuth Verlag<sup>12</sup> und dem Informationsdienst Arbeit und Gesundheit – Schwerpunkt Bildschirmarbeit,<sup>13</sup> sowie bei KommDesign.<sup>14</sup>

### 2.4.4.1 Aufgabenangemessenheit

Definition: „Ein Dialog ist aufgabenangemessen, wenn er den Benutzer unterstützt, seine Arbeitsaufgabe effektiv und effizient zu erledigen.“

Ergebnisseiten sollten so gestaltet sein, dass der Anwender bei seiner Selektionsentscheidung zwischen den Links optimal unterstützt wird. Damit ist Aufgabenangemessenheit einer der zentralen Aspekte der vorliegenden Arbeit. Der Erreichung dieses komplexen Zieles dienen die Eigenschaften (Kap. 2.2.4) bzw. Attribute (Kap. 2.4.1) der Ergebnisseiten. Eine sinnvolle Beschreibung der repräsentierten Dokumente hat hierbei die größte Bedeutung, ist jedoch keineswegs trivial (vgl. Kap. 2.2.3). Zusätzliche Funktionen sollen die Effektivität und Effizienz erhöhen. Dazu gehört zum Beispiel die Möglichkeit, das repräsentierte Dokument in einem anderen Fenster oder einer Vorschau zu öffnen oder Vorschläge zu veränderten Suchausdrücken für die weitere Suche zu

---

<sup>12</sup> <http://www2.beuth.de/>

<sup>13</sup> <http://www.sozialnetz-hessen.de/Ergo-Online/>

<sup>14</sup> <http://www.kommdesign.de/texte/din.htm>



präsentieren, sogenanntes *query refinement* (McArthur & Bruza, 2000). Aus den Anfängen der Information Retrieval stammt eine Reformulierungsstrategie zur Suchanfrage, die es ermöglicht, Dokumente anzeigen zu lassen, die hohe Ähnlichkeiten mit einem der schon gefundenen Dokumente aufweisen (*relevance feedback*, Baeza-Yates & Ribeiro-Neto, 1999). Der große Vorteil dieses Verfahrens ist, dass Anwender lediglich eine Entscheidung darüber fällen müssen, welches Dokument für sie relevant ist und nicht einen neuen Suchausdruck formulieren müssen. Diese Funktionen werden in Suchmaschinen mit einem Link zu jedem Treffer angeboten, der mit „weitere Treffer von diesem Angebot zeigen“ (z.B. bei Fireball<sup>15</sup>), „Ähnliche Seiten“ (z.B. bei Google<sup>16</sup>) oder Analogem beschrieben werden kann. Aufgabenangemessene Suchmaschinen belasten den Anwender nicht mit unnötigen Informationen wie Werbung, auf die aus kommerziellen Gründen jedoch die wenigsten Suchmaschinen verzichten können.

#### **2.4.4.2 Selbstbeschreibungsfähigkeit**

Definition: „Ein Dialog ist selbstbeschreibungsfähig, wenn jeder einzelne Dialogschritt durch Rückmeldung des Dialogsystems unmittelbar verständlich ist oder dem Benutzer auf Anfrage erklärt wird.“

Dem Anwender sollen alle Informationen auf der Ergebnisseite dargeboten werden, die er für ihre Interpretation benötigt. Zunächst sind die das Suchergebnis beschreibenden Informationen anzugeben. Dazu gehören die Anzahl der gefundenen Treffer, der von der Suchmaschine tatsächlich verwendete Suchausdruck sowie weitere Angaben (s. Kap. 2.2.4). Einzelne Metaangaben enthalten eine Beschriftung (z.B. Datum: 15-1-03), sodass Anwender nicht aufgrund ihres Wissens schlussfolgern müssen, ob es sich bei der Zahlenangabe um ein Datum, eine Dateigröße oder eine andere technische Angabe handelt. Einige Suchmaschinen bieten eine Prozentangabe für jeden Treffer, die die Relevanz des gefundenen Dokuments widerspiegeln soll. Dies ist allerdings auch ein Negativbeispiel für Selbstbeschreibungsfähigkeit, da Anwendern die genaue Bedeutung der Zahlen verborgen bleibt, denn die genaue Berechnungsgrundlage wird von Suchmaschinen nicht veröffentlicht. Höhere Zahlen bedeuten demnach lediglich eine höhere Relevanz. Soweit ist diese Information jedoch bereits in der Rankingreihenfolge enthalten, womit die Prozentangabe für die meisten Anwender keine zusätzliche Information bietet.

Neben den Treffern sollen auch die Zusatzfunktionen selbstbeschreibungsfähig sein. So wird beispielsweise Relevanzfeedback, eine Technik, die kaum jemand aufgrund des Namens auf Anhieb verstehen dürfte, durch eine Umschreibung des zu erwarteten Effekts implementiert (vgl. Kap. 2.4.4.1). Auch die Funktion, ein repräsentiertes Dokument zu übersetzen, kann auf einfache Weise mit dem Text „Funktionen: Diese URL übersetzen“ beschrieben werden (z.B. bei Abacho<sup>17</sup>).

---

<sup>15</sup> <http://www.fireball.de/>

<sup>16</sup> <http://www.google.de/>

<sup>17</sup> <http://www.abacho.de/>

#### **2.4.4.3 Steuerbarkeit**

Definition: „Ein Dialog ist steuerbar, wenn der Benutzer in der Lage ist, den Dialogablauf zu starten sowie seine Richtung und Geschwindigkeit zu beeinflussen, bis das Ziel erreicht ist.“

Während des gesamten Suchprozesses sollten Anwender die vollständige Kontrolle über die Navigation von und zur Ergebnisseite behalten sowie auch Einfluss auf die Gestaltung der Ergebnisseite selber haben. Kontrolle über die Navigation wird darin deutlich, dass der Anwender jederzeit in der Lage ist, einen Link zu besuchen, eine neue Suchanfrage zu formulieren oder wieder zu vorherigen Ergebnissen zurückzukehren. Dies leisten in der Regel die Browser mit ihrer Adresszeile und den Vor- und Zurückschaltflächen (History-Funktion). Die Ergebnisseite lässt sich beispielsweise dadurch gestalten, dass festgelegt werden kann, welche Informationen oder wie viele der Treffer auf einer Ergebnisseite angezeigt werden sollen (z.B. bei Altavista<sup>18</sup>).

#### **2.4.4.4 Erwartungskonformität**

Definition: „Ein Dialog ist erwartungskonform, wenn er konsistent ist und den Merkmalen des Benutzers entspricht, z.B. seinen Kenntnissen aus dem Arbeitsgebiet, seiner Ausbildung und seiner Erfahrung sowie den allgemein anerkannten Konventionen.“

Verschiedene Ergebnisseiten innerhalb derselben und zwischen verschiedenen Suchmaschinen sollten stets gleich strukturiert sein. Darüber hinaus sollten Standards zur Webseitengestaltung auch auf Ergebnisseiten Gültigkeit haben, z.B. dass sich hinter unterstrichenen Wörtern Hypertext-Links verbergen. Die Tatsache, dass die verschiedenen Eigenschaften eines Treffers in einer festen Reihenfolge (Titel, Beschreibung, Zusatzinformationen, Zusatzfunktionalität) präsentiert werden, erleichtert die Verwendung mehrerer Suchmaschinen, da sich Anwender nicht für jede Suchmaschine auf einen eigenen Ergebnisdialog einstellen müssen.

#### **2.4.4.5 Fehlertoleranz**

Definition: „Ein Dialog ist fehlertolerant, wenn das beabsichtigte Arbeitsergebnis trotz erkennbar fehlerhafter Eingaben entweder mit keinem oder mit minimalem Korrekturaufwand seitens des Benutzers erreicht werden kann.“

Falsche Selektionsentscheidungen auf Ergebnisseiten dürfen den Anwender in seiner weiteren Suche nicht behindern. Er sollte ohne Anstrengung fähig sein, mit der Zurückfunktion seines Browsers jederzeit den Zustand vor der fehlerhaften Auswahl wiederherzustellen. Im größeren Zusammenhang des Suchens erweist sich die Fehlertoleranz auch darin, dass orthografische Fehler im Suchausdruck automatisch korrigiert werden (z.B. bei Google<sup>19</sup>).

---

<sup>18</sup> <http://de.altavista.com/>

<sup>19</sup> <http://www.google.de/>

#### **2.4.4.6 Individualisierbarkeit**

Definition: „Ein Dialog ist individualisierbar, wenn das Dialogsystem Anpassungen an die Erfordernisse der Arbeitsaufgabe sowie an die individuellen Fähigkeiten und Vorlieben des Benutzers zulässt.“

Aufbau und Aussehen von Ergebnisseiten und darzustellende Attribute einzelner Treffer lassen sich auf die Bedürfnisse des Anwenders hin definieren. Das bedeutet, dass Anwender festlegen können, wo welche Informationen wie angezeigt werden. Dieser auch mit dem Begriff des *Customization* bezeichnete Grundsatz ermöglicht beispielsweise Einstellungen der Farbgebung, die Anzahl anzuzeigender Treffer pro Ergebnisseite und die Sprache, in der die Dialoge geführt werden sollen (vgl. Notess, 2001). Die Individualisierung des Aufbaus, zum Beispiel die Verlegung der Werbung vom Anfang der Seite an das Ende, wird bei Suchmaschinen üblicherweise nicht unterstützt. Im größeren Rahmen des Suchprozesses lässt sich feststellen, dass Individualisierbarkeit in getrennten Eingabedialogen für Experten-/Profisuche und Anfänger-/normale Suche etabliert wurde. Im Expertendialog finden sich mehr Informationen und mehr Möglichkeiten, um die Suche zu steuern, die einen in der Suche Unerfahrenen überfordern würden. Die Individualisierbarkeit überschneidet sich teilweise mit der Steuerbarkeit, wie am Beispiel der Anzahl anzuzeigender Treffer zu sehen ist, zielt jedoch prinzipiell auf allgemeinere Einstellungen ab, die für alle Suchvorgänge gelten und sich auch langfristig speichern lassen (üblicherweise mit Hilfe von sogenannten Cookies, kleinen Textdateien, die lokal beim Anwender abgelegt werden). Im Sinne der Steuerbarkeit könnte also die Anzahl anzuzeigender Treffer für die aktuelle Suche auf dreißig erhöht werden, aber generell im Sinne der Individualisierbarkeit auf zehn festgelegt sein.

#### **2.4.4.7 Lernförderlichkeit**

Definition: „Ein Dialog ist lernförderlich, wenn er den Benutzer beim Erlernen des Dialogsystems unterstützt und anleitet.“

Ergebnisseiten sollen den Anwender darin unterstützen und anleiten, anhand welcher Informationen er seine Selektion treffen kann, und Hinweise über Aufbau und Eigenschaften der Ergebnisseiten liefern. Dies kann zum Beispiel in einer *Guided Tour* geschehen, die die Anwender mit besonderen Tricks in der Bedienung vertraut macht und die einzelnen Bereiche der Ergebnisseiten in der Art „wo finde ich was“ erklärt. Eine besondere Berücksichtigung sollte die Lernförderlichkeit im Bezug auf ältere Menschen erfahren, da sich bei ihnen Gedächtnisschwierigkeiten häufiger bemerkbar machen als bei jungen Menschen (vgl. Gourmelon, 1994). Überschneidungen mit der Selbstbeschreibungsfähigkeit sind dadurch gegeben, dass gute Dialogbeschreibungen auch zu einer erhöhten Lernförderlichkeit führen, indem sie dem Anwender das Erlernen neuer Dialoge durch Versuch und Irrtum erleichtern, ohne dass dieser eine Anleitung konsultieren muss.

#### **2.4.4.8 Kritische Würdigung**

Wie oben bereits dargelegt existieren auch andere Einteilungen zu Ergonomiefaktoren (vgl. Ohnemus, 1997). Beispielsweise definiert Jakob Nielsen fünf Designprinzipien (1993):

Learnability, Efficiency, Memorability, Errors und Satisfaction. Aufgrund des allgemeingültigen Anspruchs und der systematischen Struktur der EN ISO 9241 lassen sich andere Leitlinien jedoch oftmals vollständig integrieren. Beispielfhaft seien hier nur die größten Überschneidungen zwischen Jakob Nielsens Prinzipien und der EN ISO 9241 genannt, um die Allgemeingültigkeit der Norm zu unterstreichen. *Learnability* überschneidet sich in der EN ISO 9241 mit der Lernförderlichkeit und *Efficiency* mit der Aufgabenangemessenheit und Effizienz. *Memorability* findet sich in der Lernförderlichkeit und Erwartungskonformität wieder. *Errors* spiegelt sich in der Fehlertoleranz und *Satisfaction* in der Zufriedenheit.

Nach der Darstellung der Norm ist nun auch auf ihre Problempunkte einzugehen. Erstens überschneiden sich die Anforderungen teilweise untereinander, wie oben schon bei den Paaren Steuerbarkeit, Individualisierbarkeit und Selbstbeschreibungsfähigkeit, Lernförderlichkeit beschrieben. Die Norm selbst weist darauf hin:

Die Grundsätze der Dialoggestaltung sind nicht unabhängig voneinander, und es kann notwendig sein, die Vorteile eines Grundsatzes gegenüber denen eines anderen abzuwägen. (Schneider, 1998, S. 12)

Daraus folgt, dass die Anforderungen in einem widersprüchlichen Spannungsfeld zueinander liegen. Wird Selbstbeschreibungsfähigkeit beispielsweise durch Textbeschriftungen erreicht, so verlängert dies zwangsläufig die Ergebnisseite und gerät damit in Konflikt zum Ziel der Aufgabenangemessenheit, nach der ein Anwender nur die zur Aufgabenerfüllung notwendigen Informationen erhalten soll. Auch die Aufgabenangemessenheit selbst ist schwierig zu bestimmen. In Kapitel 2.2.3 zu den Beschreibungsmöglichkeiten einzelner Treffer wurde deutlich, dass keineswegs klar ist, welche Beschreibung für die repräsentierten Dokumente vorzuziehen ist. Einzelne Grundsätze müssen daher unterschiedliches Gewicht erhalten und können für verschiedene Anwender mit unterschiedlichem Wissenshintergrund auch zu unterschiedlichen Antworten führen. Bei widersprechenden Anforderungen bietet daher die Individualisierbarkeit ansatzweise einen Ausweg, indem sie durch unterschiedliche individuelle Gewichtungen der Grundsätze bzw. verschiedener Einstellungen eine individuelle Erfüllung mehrere Grundsätze ermöglicht, wie dies durch eine zusätzliche unabhängige Implementierung von Suchdialogen für Experten realisiert ist. Zweitens bietet die Norm lediglich Hinweise, auf welche Grundsätze zu achten ist, macht jedoch keine Angaben zu Mindestmaßen, die eine Erfüllung der Norm gewährleisten würden.

#### **2.4.5 Zusammenfassung**

Anwender verwenden die Attribute einzelner Treffer, um zu entscheiden, welche der durch die Treffer repräsentierten Dokumente sie sich genauer ansehen wollen. Je nach Spezifität der Aufgabe und persönlicher Bedeutung der Suche werden mehr oder weniger ressourcenintensive Entscheidungsregeln verwendet. Anwender passen ihr Entscheidungsverhalten demnach adaptiv den aktuellen Erfordernissen ihrer Suche an, wobei sie bestrebt sind, einen möglichst hohen Nutzen mit geringem Einsatz zu erreichen.

Ihr Wissen über das Internet, verschiedene Suchinstrumente und über ihren Suchgegenstand helfen ihnen bei einer effektiven Suche. Bisher ist jedoch festzustellen, dass Wissen über Suchinstrumente bei vielen Anwendern nur schwach ausgeprägt ist und zu fehlerhaften und ungenauen Suchausdrücken führt. Die Norm EN ISO 9241 Teil zehn erlaubt bei einer Auslegung auf Ergebnisseiten hin eine Einordnung der verschiedenen Attribute von Treffern und ihrer jeweiligen Bedeutungen zueinander. Insbesondere können verschiedene Zusatzfunktionen auf Ergebnisseiten bewertet werden. Im empirischen Teil der Arbeit (Kap. 3) wird deutlich werden, wie sich aus der Norm heraus systematisch Fragestellungen ableiten lassen und wie diese operationalisiert werden können. Vorher werden jedoch verschiedene bisherige Forschungsvorhaben zu Ergebnisseiten geschildert, um mögliche Methodenansätze dem der vorliegenden Arbeit gegenüberstellen zu können.

## 2.5 Aktuelle Forschungsparadigmen

In diesem Kapitel erfolgt eine auf seine wesentlichen Bestandteile konzentrierte Darstellung des Rahmenmodells. Die daraus gewonnene Taxonomie macht deutlich, welche Aspekte durch die bisherige Forschung untersucht wurden. Weiterhin werden die verschiedenen Designprinzipien von Untersuchungen einander gegenübergestellt (Tabelle 1). Dadurch wird im empirischen Teil der vorliegenden Arbeit erkennbar, wie sich der neu entwickelte methodische Ansatz von den bisherigen unterscheidet und welche Aspekte untersucht wurden. Anhand dreier Untersuchungen soll nun das aktuelle Forschungsparadigma zur Ergebnisseitengestaltung mit dem Schwerpunkt auf dem Attribut Beschreibung und den zugehörigen Methoden bzw. Arten der Zusammenfassungen der repräsentierten Dokumente dargestellt werden.

Unz, Capstick und Erbach (1999) stellten ihren Teilnehmern in Saarbrücken verschiedene Suchaufgaben und ließen mit einer Suchmaschine passende Ergebnisse finden und notieren (vgl. Tabelle 1). Dazu verwendeten die Autoren eine Suchmaschine, die einen bekannten Dokumentenbestand indexierte. Sie erreichten dadurch eine Untersuchungssituation, die hohe Ähnlichkeit mit tatsächlichem Suchverhalten hatte. Hintergrund der Untersuchung war die Entwicklung der Suchmaschine MULINEX (vgl. Erbach, Neumann & Uszkoreit, 1997). Insgesamt benötigte ein Teilnehmer drei Stunden. Ihr Untersuchungsdesign entsprach einer zweifaktoriellen Varianzanalyse mit einem Messwiederholungsfaktor, der über zwei Suchaufgaben variiert wurde. Im ersten Faktor variierten die Autoren die Art der Organisation der Suchergebnisse (kategorisch vs. nicht kategorisch, vgl. Kap. 2.2.3.5), im zweiten die Art der Zusammenfassung. Bei den Zusammenfassungen waren dabei drei Variationen vorhanden: Die „ersten 200 Zeichen“, „die Kernaussagen“ und „Wörter im Kontext“. Als abhängige Variablen wurden Performanz, Akzeptanz und Variablen des Nutzungsverhaltens erhoben. 84 Teilnehmer wurden auf sechs Gruppen verteilt, wodurch sich in einer Post-hoc-Analyse<sup>20</sup> eine Teststärke von 0,36 errechnen lässt (ANOVA,

---

<sup>20</sup> Die in der vorliegenden Arbeit nachträgliche Berechnung der Teststärken wurde mit GPOWER von Faul und Erdfelder (1992) durchgeführt.

Effektstärke Medium,  $\alpha$ -Fehler 0,05). Dies und die fehlende Kontrolle des Faktors Aufgabe könnte insgesamt zu dem uneinheitlichen Ergebnisbild der Untersuchung beigetragen haben. Performanz- oder Akzeptanzunterschiede konnten zwischen den Arten der Zusammenfassungen nicht festgestellt werden.

**Tabelle 1:** Vergleich von Untersuchungen über Ergebnisseiten mit einer Einordnung in die Ebenen Mensch, Maschine und Aufgabe.

<b>Autoren</b>	<b>Unz, Capstick &amp; Erbach, 1999</b>
<b>Methode</b>	2 Suchaufgaben. Teilnehmer suchen Antworten mit Suchmaschine.
<b>Design</b>	Zweifaktorielle ANOVA mit Messwiederholung über 2 Aufgaben. Faktor 1: Art der Zusammenfassung, dreistufig. Faktor 2: Art der Darstellung, zweistufig. gemessene Variablen: Verschiedene Zählungen zur Anzahl aufgerufener Dokumente, Quote richtiger Aussagen, Akzeptanz
<b>Teilnehmerzahl</b>	84
<b>Teststärke</b>	0,36
<b>Maschine</b>	Vorhandene Attribute: Position, Titel, Angaben zu URL, Dokumentgröße, Datum; Beschreibung in Faktor 1 operationalisiert.
<b>Aufgabe</b>	Spezifität in Faktor 2 operationalisiert.
<b>Mensch</b>	Computerwissen als Mediator erfasst.
<b>Autoren</b>	<b>Drori, 2001</b>
<b>Methode</b>	4 Suchaufgaben unterschiedlicher Spezifität. Teilnehmer suchen Antworten mit Suchmaschine.
<b>Design</b>	Zweifaktorielles Messwiederholungsdesign. Faktor 1 (Messwiederholung): Art der Zusammenfassung, dreistufig Faktor 2 (Messwiederholung): Suchaufgabe, vierstufig gemessene Variablen: Maße zur Performanz und Zufriedenheit
<b>Teilnehmerzahl</b>	128
<b>Teststärke</b>	0,94 für die in der Untersuchung durchgeführten abhängigen T-Tests.
<b>Maschine</b>	Vorhandene Attribute: Position; Titel und Beschreibung in Faktor 1 operationalisiert.
<b>Aufgabe</b>	Spezifität in Faktor 2 operationalisiert.
<b>Mensch</b>	Computerwissen, Menge an Sucherfahrung
<b>Autoren</b>	<b>Lergier &amp; Resnick, 2001</b>
<b>Methode</b>	6 Suchaufgaben. Suchmaschine wurde simuliert. Aufgabe nach erstem Klick beendet.
<b>Design</b>	Zweifaktorielle Varianzanalyse mit zwei zusätzlichen Messwiederholungsfaktoren Faktor 1: beste Linkposition, dreistufig Faktor 2: Wurde das für die Suche benötigte Attribut angezeigt?, zweistufig Faktor 3 (Messwiederholung): Spezifität der Suchaufgabe, zweistufig Faktor 4 (Messwiederholung): Welches Attribut ist für die Suche wichtig?, dreistufig gemessene Variablen: gewählter Link und Sicherheit das selektierter Link zu einer für die Suche passenden Antwort führt (pre-click-confidence).
<b>Teilnehmerzahl</b>	60
<b>Teststärke</b>	Nicht berechenbar aufgrund fehlender Angaben zur Korrelation innerhalb der Messwiederholungsfaktoren.
<b>Maschine</b>	Vorhandene Attribute: Position, Titel, Beschreibung, Angaben zu URL und Datum
<b>Aufgabe</b>	Spezifität in Faktor 3 operationalisiert. Selektionsstrategie untersucht.
<b>Mensch</b>	Internetserfahrung, Sucherfahrung

Drori (2001) ließ in Jerusalem mit einer Suchmaschine vier Suchaufgaben unterschiedlicher Spezifität bearbeiten (vgl. Tabelle 1). Auch hier war der indexierte Datenbestand vorab bekannt. Sein Untersuchungsdesign entsprach einem kompletten Messwiederholungsdesign mit zwei Faktoren. Im ersten Faktor wurde die Art der Zusammenfassung in drei

Bedingungen variiert: „Nur Dokumenttitel ohne weitere Beschreibung“, „einige erste Zeilen“ und die „Zeilen mit den für die Suche relevanten Inhalten“. Die letztere Beschreibung eines Treffers beruhte auf einem von ihm entwickelten Algorithmus und kann als Motivator der Untersuchung angesehen werden. Der zweite Faktor entsprach den vier Suchaufgaben, die sich in ihrer Spezifität unterschieden. Erhoben wurden Maße der Performanz und Zufriedenheit. 128 Teilnehmer bearbeiteten die vier Aufgaben und wurden dabei für jede Aufgabe einer Bedingung des ersten Faktors in zufälliger Reihenfolge zugewiesen. Daraus folgt, dass eine Art der Zusammenfassung für jeden Teilnehmer zweifach dargeboten wurde. Mittels Post-hoc-Analyse ergibt sich eine Schätzung der Teststärke von 0,94 (abhängiger t-Test, Effektstärke Medium,  $\alpha$ -Fehler 0,05, unspezifische Hypothese). Für das Performanzmaß „aufgewendete Zeit“ ergab sich ein Interaktionseffekt mit der Aufgabenspezifität. Bei unspezifischer Aufgabe war die Wörter-in-Kontext-Bedingung am praktikabelsten, wohingegen bei sehr spezifischer Aufgabe der Titel allein für das schnellste Ergebnis sorgte. Insgesamt war die Wörter-in-Kontext-Bedingung den anderen Bedingungen überlegen. Als problematisch bei dieser Untersuchung ist jedoch das Messwiederholungsdesign einzuschätzen, das jedem Teilnehmer alle Arten von Zusammenfassungen und unkontrolliert eine Art zweifach präsentierte. Im Zusammenhang mit der Abschlussbefragung nach allen Suchaufgaben wurde möglicherweise Fehlervarianz erzeugt, die durch die Methode der erzwungenen Differenzierung hervorgerufen wird.

Die beiden vorher genannten Studien verlangten von den Teilnehmern das Durchlaufen von mehreren kompletten Suchprozessen. Solche Untersuchungen sind sehr aufwendig durchzuführen, da nicht nur die Teilnehmer über einen langen Zeitraum gebunden sind, sondern häufig auch einen bestimmten Untersuchungsraum aufsuchen müssen. Darüber hinaus sind umfangreiche technische Vorbereitungen zu treffen, die mit dem Untersuchungsgegenstand selber wenig gemeinsam haben: Eine Datenbank muss aufgebaut und eine Suchmaschine bereit gestellt werden. Mag dies für eine Modellierung des gesamten Suchprozesses noch notwendig erscheinen, so lässt sich das relevante Verhalten auf Ergebnisseiten deutlich einfacher untersuchen, wie Lergier und Resnick (2001) und die noch darzustellende Untersuchung zeigen.

Zwar variierten Lergier und Resnick nicht das Attribut Beschreibung, jedoch nahmen sie einige Änderungen im bisher üblichen Untersuchungsdesign vor (vgl. Tabelle 1). Ihr Ziel lag in einer Beschreibung des Suchverhaltens auf Ergebnisseiten. Dazu variierten sie die Aufgabenspezifität zweistufig (spezifisch-unspezifisch) und konstruierten Aufgaben (dreistufiger Faktor), für deren Lösung die drei Attribute URL, Datum oder Dokumentgröße bedeutend waren. Dieses 2x3 Design wurde vollständig im Sinne eines Messwiederholungsdesign von allen 60 Teilnehmern durchlaufen. Zusätzlich wurde die Position des besten Links dreistufig (erste, fünfte, zehnte Position) und die Tatsache, ob das notwendige Attribut angezeigt wurde, zweistufig (ja/nein) variiert und jeder Stufe zehn Teilnehmer zugewiesen, wodurch das gesamte Design einer zweifaktoriellen Varianzanalyse (‚beste Linkposition‘ und ‚Wurde das für die Suche wichtige Attribut angezeigt?‘) mit zwei

zusätzlichen Messwiederholungsfaktoren (‚Spezifität‘ und ‚Welches Attribut ist für die Suche wichtig?‘) entsprach.<sup>21</sup> Teilnehmer gaben ihre Suchworte ein, erhielten dann jedoch vordefinierte Ergebnisseiten unabhängig von ihrer Eingabe. Sobald sie einen Link ausgewählt hatten, war die Suchaufgabe beendet und die Abschlussfragen folgten. Die Dokumente hinter den Links wurden nicht dargeboten. Ihre Ergebnisse zeigen, dass als Entscheidungsregel die selbst-terminierende Strategie (vgl. Kap. 2.3.4) häufiger bei spezifischen als unspezifischen Aufgaben angewendet wird. Bei unspezifischen Suchaufgaben ist das Vorgehen hingegen gemischt.

Die dargestellten Untersuchungen zeigen, welche theoretische und praktische Unterstützung das Rahmenmodell Mensch, Maschine, Aufgabe für eine erfolgreiche, systematische Forschung über Ergebnisseiten haben kann. Im nun folgenden empirischen Teil der Arbeit wird eine Untersuchung dargestellt, die sich an der Systematik des Rahmenmodells orientiert und neue Untersuchungsmethoden implementiert.

## 3 Methoden der empirischen Untersuchung

### 3.1 Untersuchungsdesign

Wie die vorherigen Kapitel deutlich zeigten, existieren zahlreiche Einflussfaktoren auf die hier untersuchten Effekte. In diesem Zusammenhang sind von Beobachtungen im Feld bis hin zu echten Experimenten verschiedene Untersuchungsdesigns denkbar. Der starke Einfluss von Störvariablen verhindert hierbei allerdings leicht die reliable Messung der gewünschten Effekte. Die für diese Arbeit durchgeführte Untersuchung implementiert daher eine möglichst hohe Kontrollmöglichkeit im Design. Im Folgenden werden die verschiedenen Facetten der bisherigen Untersuchungen miteinander verknüpft, wobei gleichzeitig möglichst viele Einflussfaktoren beachtet werden.

#### 3.1.1 Untersuchungsparadigma

Bei einer Untersuchung stellt sich zunächst die Frage, welche Variablen aktiv manipuliert werden können und welche Variablen sich nicht beeinflussen lassen. Dabei lässt sich feststellen, dass Variablen, die Anwender betreffen, nicht manipulierbar sind, sondern nur statistisch kontrolliert werden können. Sehr wohl lassen sich jedoch die Dimensionen in der Ebene Maschine mit nahezu beliebigen Ausprägungen realisieren. In einem festen Untersuchungsrahmen können sogar Elemente der Ebene Aufgabe, respektive die Suchaufgabenstellung, vorgegeben werden. Aus diesen Überlegungen heraus wird ein mehrfaktorielles varianzanalytisches Design mit Messwiederholungen konstruiert. Den ersten Faktor bildet das Attribut Beschreibung der repräsentierten Seite, im Folgenden

---

<sup>21</sup> Eine Teststärkeanalyse kann aufgrund fehlender Angaben zur Korrelation innerhalb der Messwiederholungsfaktoren für dieses Design nachträglich nicht berechnet werden.



Faktor Beschreibung genannt. Der zweite Faktor wird durch die Aufgabenspezifität gebildet und im Folgenden mit Faktor Spezifität bezeichnet. In der Ebene Maschine wird nur ein Attribut operationalisiert, da sich die Anzahl benötigter Versuchspersonen bei der Berücksichtigung weiterer Attribute mit jedem weiteren Faktor vervielfacht, um einen Verlust an Teststärke auszugleichen (Cohen, 1992). Alternativ können weitere Attribute auch als Messwiederholungsfaktoren konzipiert werden, was jedoch zu einer Untersuchungsverlängerung führen würden. Um möglichst viele Teilnehmer zu gewinnen, soll jedoch die gesamte Untersuchung nicht länger als durchschnittlich zwanzig Minuten in Anspruch nehmen, weshalb auch diese Möglichkeit hier ausscheidet. So befragten Batinic und Bosnjak (2000) Personen nach ihrer Bereitschaft an einer Online-Befragung teilzunehmen und empfehlen auf dieser Grundlage eine maximale Zeitdauer von 15 Minuten (vgl. dazu auch Bosnjak & Batinic, 1999). Eine Untersuchung zum Einfluss der Zusatzinformationen (z.B. URL, Datum) wurde nicht durchgeführt, da einfache Überlegungen zeigen, dass für jede Zusatzinformation Suchanfragen konstruierbar wären, bei der gerade diese Zusatzinformation für die Suche besonders relevant ist (vgl. Hock, 1998; Lergier & Resnick, 2001). Eine spezielle Zusatzinformation ist also je nach Suchanfrage mehr oder weniger wichtig. Weitere Variablen, die im Rahmenmodell als wichtige Einflussgrößen angesehen werden, werden als Kontrollvariablen erhoben. Um die Anzahl benötigter Versuchspersonen zu reduzieren, wurde der Faktor Spezifität als Messwiederholungsfaktor realisiert.

Ein Hauptziel der Arbeit ist die Konstruktion eines Untersuchungsdesigns, mit dem sich Ergebnisse leicht replizieren lassen und das mit geringem Aufwand in bestehende Suchmaschinen eingebaut werden kann. Um das zu erreichen, findet die Untersuchung über das Internet im Stile eines Web-Experiments statt. Die Zuweisung zu den einzelnen Bedingungen wird randomisiert (Reips, 2000). Eine Zufallsstichprobe konnte nicht gezogen werden, da die Population der Internetanwender unbekannt ist. Die Stichprobe kann daher als nicht-probabilistische Gelegenheitsstichprobe bezeichnet werden (Batinic & Bosnjak, 2000). Durch die Randomisierung ergibt sich eine hohe Aussagekraft innerhalb des Experiments mit beschränkter Generalisierbarkeit auf die Population (vgl. Reips, 2000).

Eine Besonderheit dieses Untersuchungsdesigns ist die Tatsache, dass Teilnehmer nicht die Möglichkeit erhalten, während der Untersuchung die Webseiten der repräsentierten Dokumente zu lesen. Dies ist deshalb nicht notwendig, weil das Ziel eine Evaluation von Ergebnisseiten und nicht der repräsentierten Dokumente oder der Qualität des Suchmaschinenrankings ist. Die Teilnehmer geben lediglich an, welche Treffer sie auswählen würden. Dies bietet zwei große Vorteile gegenüber einem Design, in dem eine gesamte Suche mit zu beantwortenden Fragen durchgeführt würde: Die Untersuchungsdauer wird erheblich verkürzt (z.B. von drei Stunden wie bei Unz, Capstick & Erbach, 1999, auf 20 Minuten in der vorliegenden Untersuchung), wodurch eine höhere Teilnehmerzahl erreicht werden kann. Außerdem reduziert dieses Vorgehen die Fehlervarianz, da die Messergebnisse nicht durch die unterschiedlichen Fähigkeiten von Teilnehmern, die gesuchten Informationen aus den repräsentierten Dokumenten zu extrahieren, beeinflusst werden. Der zweite Unterschied zu den oben dargestellten Untersuchungen (vgl. Kap. 2.5)

besteht darin, dass die für eine Suche verwendeten Suchwörter vorgegeben werden. Dies erscheint besonders sinnvoll vor dem Hintergrund, dass Teilnehmer häufig ohnehin Suchwörter aus dem Instruktionstext verwenden (Körper, 2000). Darüber hinaus werden die verwendeten Ergebnisseiten zur Minimierung der Fehlervarianz während der Untersuchung konstant gehalten. Es hätte also eine Coverstory konstruiert werden müssen, um die Suchanfrage real erscheinen zu lassen. Solch eine Täuschung der Teilnehmer über den Ablauf der Untersuchung bei gleichzeitiger Verlängerung der Untersuchungsdauer ist für die vorliegende Fragestellung nicht notwendig, da das Suchverhalten der Teilnehmer nicht Untersuchungsgegenstand ist.

Zusammenfassend handelt es sich bei diesem Design um ein randomisiertes Webexperiment bei Selbstselektion der Teilnehmer. Es wird nicht das komplette Suchverhalten beobachtet, sondern das Selektionsverhalten auf Ergebnisseiten bei vorgegebenen Suchaufgaben und verwendeten Suchwörtern. Die operationalisierten Faktoren betreffen die Attribut Beschreibung und die Aufgabenspezifität.

### 3.1.2 Operationalisierung der Faktoren Beschreibung und Spezifität

Im Faktor Beschreibung werden vier Stufen realisiert. Die im Folgenden beschriebenen beiden ersten Bedingungen spiegeln dabei die von Suchmaschinen am häufigsten verwendeten Beschreibungsvarianten wieder. Bei einem Treffer in der Untersuchung wird jeweils der Titel des repräsentierten Dokuments sowie eine der vier Beschreibungsvarianten dargestellt. Weitere Attribute (vgl. Kap. 2.2.4) werden nicht mit angegeben, um die zu entdeckenden Unterschiede zwischen den Bedingungen auf die Beschreibung und den Titel zurückführen zu können.

In der ersten Bedingung werden die im Suchausdruck verwendeten Suchwörter zusammen mit den sie umgebenden Wörtern, wie sie im repräsentierten Dokument vorkommen, zur Beschreibung der Treffer verwendet. Diese Methode wird als ‚Wörter-im-Kontext-Methode‘ bezeichnet (keywords-in-context, kurz KWIC, z.B. bei Google<sup>22</sup>). Dies entspricht der Extraktion signifikanter Sätze, wobei diese jedoch gekürzt werden (z.B. für die Suchwörter Euro und Größe: „... 5 Euro - Größe 120 mm x 62 mm. 10 Euro - Größe 127 mm x 67 mm. 20 Euro - Größe 133 mm x 72 mm. ...“).

Auch bei der zweiten Bedingung werden signifikante Sätze verwendet. Allerdings werden hierzu die ersten Zeichen aus einem Dokument extrahiert (z.B. bei Abacho). Diese Methode wird daher als ‚Textanfang‘ bezeichnet. Die reine Darstellung der ersten Wörter, wie sie im Dokument vorkommen und wie dies von einigen Suchmaschinen durchgeführt wird, ist jedoch nicht ohne weiteres sinnvoll. So finden sich Tabellenüberschriften und Bezeichnungen von Navigationsleisten eines Dokumentes häufig zu Beginn des Dokumentes. Ohne weitere Filter würden diese dann zu einem wenig aussagekräftigen Konglomerat von zufällig erscheinenden Wörtern führen. Erschwerend kommt hinzu, dass üblicherweise nur die ersten 200 Zeichen Text dargestellt werden, um die Beschreibung kurz

---

<sup>22</sup> <http://www.google.de/>

zu halten. In der vorliegenden Arbeit werden daher zwei Verbesserungen vorgeschlagen. Zunächst wird der beschreibende Anfangstext auf 400 Zeichen erweitert. Als nächstes wird ein verbesserter Algorithmus zur Extraktion angewendet. Dieser verhindert sinnleere Zeichenansammlungen dadurch, dass derjenige Text extrahiert wird, der sich bei einer Darstellung des Dokumentes durch einen Browser als zusammenhängender Text zeigt. 400 Zeichen zusammenhängender Text entsprechen damit redaktionellem Inhalt und geben sinnhafte Aussagen des Dokumentes wieder, ohne sich in Schlagworten oder Navigationswörtern zu verlieren. Für die Untersuchung findet die Extraktion auf diesem optischen Wege durch den Versuchsleiter statt. Als Ansatz zur maschinellen Umsetzung kann der Algorithmus wie folgt beschrieben werden:

1. Extrahiere die Zeichen zwischen den beiden Tags `<body>` und `</body>`
2. In diesen Zeichen suche nach  $n = 400$  zusammenhängenden Zeichen. Zeichen gelten als getrennt, wenn zwischen ihnen Tags vorkommen, die die Struktur der Seite bestimmen, wie z.B. alle Tags zu Tabellen und Frames.
3. Wurden  $n$  Zeichen gefunden, stelle diejenigen Zeichen dar, die auch von einem Browser dargestellt würden und beende den Algorithmus. Ansonsten reduziere  $n$  um 10 Zeichen und fahre mit Punkt 2 fort.

In der dritten Bedingung wird eine Extraktionsmethode der signifikanten Wörter realisiert. Die sieben häufigsten, semantisch sinnvollen Substantive im Dokument werden dargestellt (Bedingung Schlüsselwörter). Der HTML-Code wird dazu herausgefiltert. Identifiziert werden die Wörter mit Hilfe eines Programms zur Textanalyse, das in der Lage ist, die Worthäufigkeiten innerhalb eines Dokumentes zu zählen (Romppel, 2000). Aus der resultierenden Liste werden dann vom Versuchsleiter die sieben häufigsten Substantive ausgewählt, die für das Dokument eine inhaltliche Bedeutung haben. Für eine maschinelle Umsetzung können Stoppwörter definiert werden, die keine inhaltliche Bedeutung haben, also z.B. alle HTML-Tags, aber auch die bestimmten und unbestimmten Artikel in einer Sprache. Solche Stoppwörter werden aus der Häufigkeitsliste herausgefiltert. Dargestellt werden dann die ersten sieben verbleibenden Wörter.

Die vierte Bedingung stellt ein Konglomerat aus den vorherigen drei Bedingungen dar, womit festgestellt werden soll, inwieweit ein Maximum an Information bei der Selektionsentscheidung hilft (Bedingung MaxInfo). Die Angaben erfolgen untereinander in der Reihenfolge Schlüsselwörter, Wörter im Kontext, Textausschnitt des Anfangs.

Im Faktor zur Spezifität der Suchaufgabe werden zwei Stufen realisiert. Die spezifische Suchanfrage wird mit der Aufgabe „Welche Größe haben Euroscheine?“ operationalisiert, die unspezifische Aufgabe mit der Frage „Welche Folgen hat die Währungsstellung?“. Das theoretische Konzept der Spezifität wird dadurch umgesetzt, dass Informationen zur Euroscheingröße durchaus auf einer Webseite gefunden werden können und die Antwort zur Frage eindeutig in Millimeterangaben gegeben werden kann. Wer die Antwort gefunden hat wird sich daher über die Richtigkeit der Antwort sehr sicher sein. Anders ist die Situation mit der unspezifischen Frage nach den Folgen der Währungsstellung. Eine eindeutige Antwort auf diese Frage ist selbst bei einer umfassenden Recherche kaum zu

erwarten, da selbst vor der Einführung des Euro in Fachkreisen keine Einigkeit herrschte, wie sich die Währungsumstellung auswirken würde. So sieht sich der Suchende mit einer Fülle von möglichen Antworten konfrontiert, die je nach politischer und wirtschaftlicher Ausrichtung variieren.

Zusammenfassend und für einen einfacheren Wortgebrauch werden die Faktoren folgendermaßen bezeichnet: Der Faktor Beschreibung enthält die Bedingungen Kontext, Textanfang, Schlüsselwörter und MaxInfo. Der Faktor Spezifität (Aufgabe) enthält die Bedingungen spezifisch und unspezifisch.

### **3.1.3 Aufbau und Ablauf der Studie**

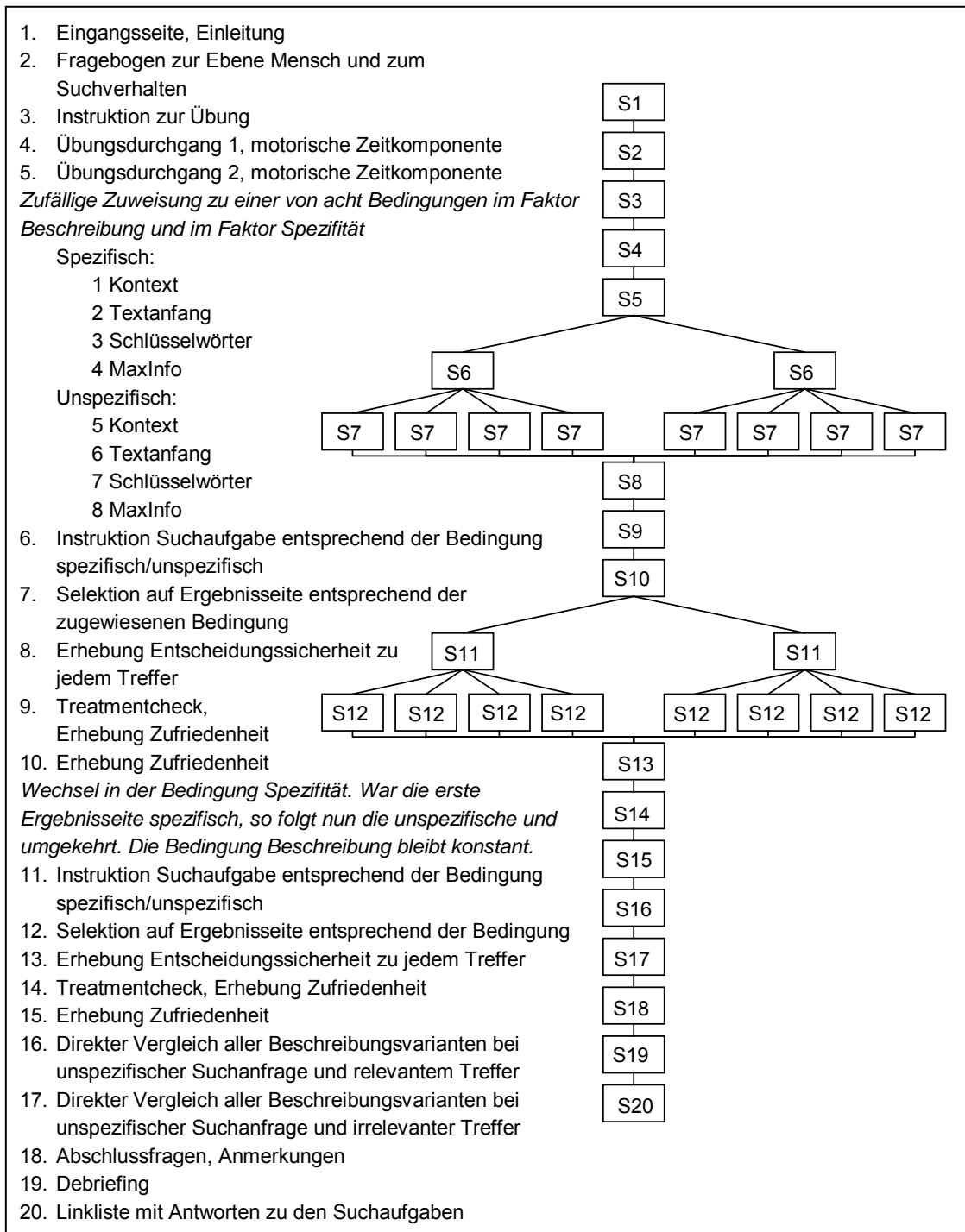
Zu Beginn des Experiments werden die Teilnehmer randomisiert einer der vier Stufen des Faktors Beschreibung zugewiesen und durchlaufen beide Stufen des Faktors Spezifität, wobei die Sequenz permutiert wird, um Reihenfolgeeffekte zu kontrollieren. Daraus ergeben sich acht unterschiedliche Bedingungen mit ebenso vielen Ergebnisseiten. Die einzelnen Webseiten werden nach der Darstellung des kompletten Designs und des Ablaufs der Untersuchung in Abbildung 8 mit den zugehörigen Designüberlegungen verbunden und erläutert.

Die verwendeten Webseiten finden sich in Anhang I. Zur leichteren Auffindbarkeit werden Bezüge zu Experimentalseiten in Klammern mit einem ‚S‘ gefolgt von der Webseitenzahl angegeben (z.B. S1). Bezug auf einzelne Items geschieht entsprechend mit dem Buchstaben ‚I‘ gefolgt von der Position auf der Webseite (z.B. S13, I1).

#### **3.1.3.1 Allgemeine Designprinzipien**

In diesem Abschnitt werden diejenigen Überlegungen im Untersuchungsdesign diskutiert, die für alle Webseiten der Untersuchung Gültigkeit besitzen (vgl. Dillman, 2000). Jede Untersuchungsseite beginnt mit einer Fortschrittsanzeige am oberen Bildschirmrand. Diese zeigt dem Anwender stets, wie weit die Untersuchung vorangeschritten ist und soll der Motivation der Teilnehmer dienen (Couper, Traugott & Lamias, 2001). Ein deutlicher Pfeil am Ende jeder Seite soll den Teilnehmern ermöglichen, leicht und schnell zur nächsten Untersuchungsseite zu gelangen.

Bei der Konstruktion von Frage-Items wird für die Skalierung von Antworten eine fünfstufige Skala gewählt, da sich diese auf den meisten Bildschirmseiten ohne Verzerrungen in der Metrik mit Beschreibungen zu jeder Ausprägung darstellen lassen. Dies ist bedeutend, damit die Abstände zwischen einzelnen Abstufungen stets gleich sind, weil sonst einzelne Ausprägungen der Skala salienter erscheinen als andere (Dillman, 2000). Die Bezeichnungen entsprechen dabei den üblichen Sprachregeln einer Likertskalierung (vgl. Anhang I). Um bei der Auswertung eine Entscheidung zwischen zufälliger Itemauslassung und von Teilnehmern beabsichtigtem Nicht-Antworten zu ermöglichen, wird bei jedem Item eine Kategorie „weiß nicht“ hinzugefügt. Items werden entsprechend den üblichen Empfehlungen (z.B. möglichst klare Aussagen, keine Verneinungen, ausgeglichene Polung) konstruiert (Dillman, 2000; Lienert & Ratz, 1994).



**Abbildung 8:** Ablauf der Untersuchung.

Weiterhin werden die Webseiten so erstellt, dass sie eine Breite von maximal 550 Bildpunkten haben, damit sie auf allen gängigen Bildschirmauflösungen gleich dargestellt werden und keine Effekte aufgrund von unterschiedlichen Bildschirmauflösungen auf Seiten der Teilnehmer zu erwarten sind. Da nicht bekannt ist, wie die Verteilungskurven des Alters der tatsächlichen Teilnehmer und ihrer Sucherfahrungen aussehen, werden diese Angaben über offene Textfelder erhoben, um die Messgenauigkeit zu maximieren. Zur besseren Abgrenzung und Übersichtlichkeit einzelner Items und der Instruktionstexte werden diese

mit unterschiedlichen Farbtönen unterlegt (Instruktionstexte mit türkisen, Fragebogenitems mit hell- und dunkelorange Farbtönen).

### **3.1.3.2 *Eingangsseite und Eingangsfragebogen***

Die Eingangsseite (S1) begrüßt die Teilnehmer, bietet eine kurze Einführung in den Untersuchungsgegenstand und zeichnet ein Bild vom Ablauf der Untersuchung. Des Weiteren stehen hier wichtige Bedienungshinweise sowie eine Erklärung zum Datenschutz.

Demografische und andere persönliche Merkmale der Teilnehmer können am Ende oder zu Beginn der Untersuchung erhoben werden. Damit Variablen jedoch als unabhängige Kontrollvariablen in die statistische Auswertung aufgenommen werden können, müssen diese entweder unbeeinflussbar durch den Untersuchungsgegenstand sein, wie z.B. das Geschlecht, oder vor Beginn der eigentlichen Untersuchung erhoben werden. Selbsteinschätzungen könnten durch den subjektiv wahrgenommenen vorherigen Erfolg oder Misserfolg der Selektionsentscheidungen beeinflusst werden. Zu Beginn des Eingangsfragebogens (S2) werden deshalb Informationen über die Verwendung von Suchmaschinen gesammelt. Danach folgen Items, die durch die Angabe von konkretem Verhalten die Erfahrung im Umgang mit Suchmaschinen, spezifischen und unspezifischen Suchanfragen, dem Computer und dem Internet erheben. Abschließend werden Fragen zum Bereichswissen gestellt. Wissen wird als Selbsteinschätzung erhoben, um die Untersuchungsdauer kurz zu halten, wobei nach Empfehlungen von Bortz und Döring (2002) für reliable und valide Ergebnisse die Items größtenteils auf konkrete Verhaltensweisen abzielen (z.B. „An wie vielen Tagen saßen Sie letzte Woche an einem Computer“). Ähnliche Formulierungen verwenden Hassenzahl und Prümper (1999) zur Erfassung der Computererfahrung. Wanger, McDonald und Berger (1980, zitiert nach Hsieh-Yee, 1993) definieren Sucherfahrung durch Menge an durchgeführten Suchanfragen pro Monat.

Werden Angaben erhoben, die mit Gedächtnisleistungen der Teilnehmer verbunden sind, so kann dies zu Verzerrungen und Fehleinschätzungen der eigenen Angaben führen (Bortz & Döring, 2002). Aufgrund dieser Gedächtnisproblematik wird in den Items daher nach ‚Anzahl der Tage‘ und nicht nach ‚Anzahl der Stunden‘ gefragt. Eine Angabe von Stunden ist nur dann genauer, wenn diese auch korrekt erinnert werden können, wobei Teilnehmer in den Voruntersuchungen hierbei jedoch große Schwierigkeiten berichteten. Mit den gewählten Formulierungen soll daher die Messung so genau wie möglich sein, ohne jedoch eine Exaktheit der Werte vorzutäuschen, die nicht existiert.

### **3.1.3.3 *Übungsdurchgänge***

Um die Selektionsentscheidung auf Ergebnisseiten zu messen, wird eine für diese Arbeit speziell entwickelte Methode angewendet. Das normale Selektionsverhalten auf Ergebnisseiten zeigt sich darin, dass Anwender in der Selektionsentscheidung einen Treffer als relevant auswählen und den zugehörigen Link auf das repräsentierte Dokument anwählen. Damit wird üblicherweise die Ergebnisseite verlassen und das repräsentierte Dokument geladen. Für die Hypothesen dieser Untersuchung spielen die tatsächlichen

repräsentierten Dokumente jedoch keine Rolle, da nur die durch die Beschreibung der Treffer vermittelten Effekte in den abhängigen Variablen entscheidend sind. Je mehr Treffer evaluiert werden, desto mehr mögliche Messergebnisse resultieren daraus und umso höher ist die aus den aggregierten Messungen entstehende Reliabilität (Wittmann, 1985). Die im Theorieteil durchgeführte Aufgabenanalyse (siehe Kap. 2.3.4) zeigt, dass bei einer Suche mehrere Treffer evaluiert und auch ausgewählt werden können. Daraus ergibt sich im Gegensatz zum Ansatz von Lergier und Resnick (2001), dass Teilnehmer jeden Link anwählen sollen, den sie in einer tatsächlichen Suche ebenfalls weiterverfolgen würden.

Die Messung wird dadurch weiter präzisiert, dass für jeden Link eine Selektionsentscheidung gefordert wird. Da dieses Vorgehen – für jeden Link anzugeben, ob er ausgewählt wird oder nicht – für Anwender ungewohnt und neu ist, werden zwei Übungsdurchgänge vorangestellt. Damit wird zunächst das Anwählen der „radio buttons“ geübt, damit die Zeitmessungen auf den Ergebnisseiten möglichst wenig dadurch verfälscht werden, dass sich Teilnehmer je nach Vorerfahrung mehr oder weniger stark an die ungewohnte Selektionsaufgabe gewöhnen müssen. Als Zweites erlauben diese Übungsdurchgänge jedoch auch eine Zeitmessung und damit einen Indikator, wie hoch die motorische Geschwindigkeit der Teilnehmer ist.

Generell wird bei quantitativen Ansätzen angenommen (z.B. GOMS-Analyse, Banke, 2000; Körndle, 1993), dass sich die Zeit, die für einen gesamten Verhaltensprozess benötigt wird, aus mehreren Zeitkomponenten zusammensetzt: Im vorliegenden Fall besteht der gesamte Selektionsprozess aus der benötigten Zeit zum Lesen des Textes, der Evaluation eines Treffers sowie einer motorischen Bewegung zum Anwählen des Links. Die in der vorliegenden Arbeit als Fehlervarianz aufzufassenden Komponenten ‚motorische Bewegung‘ und ‚Lesedauer‘ können damit in die Auswertung einfließen, um eine exaktere Messung der Entscheidungszeit zu erreichen.

Nach einer Instruktion (S3) ist im ersten Übungsdurchgang (S4) für sieben Treffer direkt angegeben, ob diese ausgewählt werden sollen oder nicht. Die Zeitkomponenten Lesedauer sowie motorische Bewegung fallen in dieser Messung zusammen, da eine Einzelmessung beider Komponenten über das Internet nicht ohne die Einrichtung eines weiteren Zusatzprogramms beim Teilnehmer möglich ist. Solch ein Zusatzprogramm könnte z.B. mit Hilfe von Java realisiert werden, was jedoch die Teilnahmeakzeptanz reduzieren und den Teilnehmerkreis auf Personen mit javafähigem Browser einschränken würde. Daher wurde auf ihren Einsatz verzichtet. Um die Zeitmessungen als motorische Komponente mit Lesedauer interpretieren zu können, muss weiterhin angenommen werden, dass die Selektionskomponente durch den unmittelbar verständlichen Text vernachlässigbar gering ist. Ein zweiter Durchgang (S5), bei dem andere Links ausgewählt werden sollten als im ersten Übungsdurchgang, wird ebenfalls durchgeführt, um die Reliabilität der Zeitmessung zu erhöhen.

### 3.1.3.4 Ergebnisseiten

Zunächst wird das Suchszenario beschrieben (S6, S11). Die Teilnehmer werden aufgefordert, sich ein bestimmtes Informationsbedürfnis vorzustellen, indem sie auf eine Frage, die im Folgenden spezifiziert wird, eine Antwort suchen. Je nachdem, welche Bedingung der Spezifität dargeboten wird, geschieht dies mit den Fragen „Welche Größe haben die deutschen Euroscheine in Millimeter?“ oder „Welche wirtschaftlichen Folgen hat die Währungsumstellung?“. Den Teilnehmern wird mitgeteilt, welche Information sie bei den Treffern erhalten werden (Erklärung zum Faktor Beschreibung). Zusätzlich wird die Erklärung, wie Treffer ausgewählt werden können, erneut wiederholt und die Teilnehmer zu zügigen Entscheidungen angehalten, um die Fehlerstreuungen in den Zeitmessungen innerhalb der Teilnehmer zu reduzieren.

Die Ergebnisseite mit der Trefferliste (S7, S12) beginnt mit einer Wiederholung der Suchaufgabe, den verwendeten Suchwörtern („Euroscheine Größe“ bzw. „Euro wirtschaftliche Folgen“) gefolgt von sieben Treffern, bei denen jeweils der Titel sowie eine Beschreibung des repräsentierten Dokumentes entsprechend der Faktorbedingung angegeben ist. Für jede Selektion oder Nicht-Selektion eines Treffers wird die absolute Systemzeit im Computer der Teilnehmer erfasst. Daraus lässt sich die relative Zeit errechnen, die zwischen zwei Treffer-Selektionen verstreicht. Sobald die Seite bei dem Teilnehmer auf dem Bildschirm aufgebaut ist, beginnt die Zeitmessung. Die relativen Zeiten werden aggregiert als Maß für die Entscheidungsdauer interpretiert (siehe Kap. 3.1.6.2). Außerdem wird die Selektionsentscheidung gemessen, woraus sich verschiedene Kriterien zur Güte der Entscheidung errechnen lassen (siehe Kap. 3.1.6.1).

In der Untersuchung wurde vom Standard angezeigter Treffer bei Ergebnisseiten (in der Regel zehn) abgewichen. Es werden sieben Treffer dargestellt, da diese bei einer Bildschirmauflösung von 1024\*786 Bildpunkten alle auf einmal dargestellt werden können und damit ein Scrollen des Bildschirms seitens der Teilnehmer unnötig machen. Solch ein Scrollvorgang würde die Aussagekraft der Zeitmessung erheblich reduzieren, da nicht ohne Zusatzprogramm erfasst werden kann, wann dieser beginnt und endet und inwieweit die Evaluierung der Treffer durch den Scrollvorgang beeinflusst wird.

Die einzelnen Treffer wurden vor der Untersuchung in einer tatsächlichen Internetrecherche mit den oben angegebenen Wörtern ausgewählt und zusammengefügt, wobei der vierte Treffer so gestaltet ist, dass er eindeutig als unzutreffend identifiziert werden kann. Der vierte Treffer erhöht dadurch die Ähnlichkeit zu tatsächlichen Suchergebnissen, die ebenfalls gelegentlich auf für den Suchenden völlig abwegige Dokumente verweisen. Das Muster, welcher Treffer ein für die Suche relevantes Dokument repräsentiert, ist für alle Probanden gleich, um eventuelle Effekte, die durch unterschiedliche Güte von Ergebnisseiten erzeugt werden könnten, zu minimieren. Dies ist ebenfalls deshalb sinnvoll, weil das Ranking einen erheblichen Einfluss auf die Auswahl einzelner Treffer hat (vgl. Ergebnisse bei Abseits.de, 2001) und dadurch einen Störeinfluss in der Untersuchung darstellt. Vollständig korrekte Selektionsentscheidungen bei allen sieben Treffern führen daher zu einer festen Abfolge von ja-nein Entscheidungen (ja, nein, ja, nein, nein, ja, nein). Zu prüfen ist jedoch, ob die



wahrgenommene Relevanz seitens der Teilnehmer dieser vorab bestimmten Abfolge von korrekten Entscheidungen entspricht (siehe Kap. 4.1.1).

### **3.1.3.5 Einschätzungen der Treffer**

Um eine genauere Einschätzung der einzelnen Treffer zu erlauben, werden die vorhergehenden Treffer erneut präsentiert (S8, S13). Bei der Auswahl mehrerer Treffer sind sich Anwender unterschiedlich sicher, wie sehr die einzelnen Dokumente tatsächlich die gesuchte Information enthalten. Lergier und Resnick (2001) erheben die ‚Pre-Click-Confidence‘ und verstehen darunter die Sicherheit, mit der Anwender annehmen, dass ihre Auswahlentscheidung zu einer richtigen Antwort führen würde. In ihrem Design wurde nur ein Link ausgewählt und bewertet. Solch eine Frage an die Teilnehmer wird aber problematisch, wenn mehrere Entscheidungen bewertet werden sollen. Dies ergaben die Voruntersuchungen. Dort wurden Teilnehmer aufgefordert anzugeben, wie sicher sie in ihrer Entscheidung sind. Dabei zeigte sich, dass Teilnehmer bei den sieben Treffern Schwierigkeiten hatten, sich für jeden einzelnen zu erinnern, welchen sie für relevant befunden hatten. Die Frage „Wie wahrscheinlich führt dieser Link zu den gesuchten Informationen“ eignet sich zur Bewertung der Treffer. Extremere Einschätzungen seitens der Teilnehmer werden als höhere Entscheidungssicherheit interpretiert, Angaben in der Mitte der Skala als Unsicherheit (siehe Kap. 3.1.6.3).

### **3.1.3.6 Bewertungen der Ergebnisseiten**

Auf zwei weiteren Webseiten (S9-10, S14-15) werden Zufriedenheits- und andere Einstellungsmaße zur Ergebnisseite erhoben (siehe Kap. 3.1.6.3). Zu Beginn findet ein Treatmentcheck statt, womit festgestellt werden kann, ob die Bedingung der Spezifität von den Teilnehmern korrekt wahrgenommen wird (z.B. „Die Suche finde ich komplex – einfach“). Danach behandeln Items die Trefferliste als Ganzes (z.B. „Es war schwierig zwischen den Links auszuwählen“, „Mit den angebotenen Links war ich zufrieden“). Anschließend folgen Items zu den Beschreibungstexten (z.B. „Die Texte waren unverständlich“, „Die Texte unterstützten schnelle Entscheidungen“, siehe Anhang I). Die Items wurden erstellt, indem Fragen aus vorhergehenden Untersuchungen (Drori, 2001; Lergier & Resnick, 2001; Unz et al., 1999) verwendet und durch neu erstellte ergänzt wurden.

### **3.1.3.7 Direkter Vergleich der Beschreibungsvarianten**

Die Messungen in den Bedingungen zum Faktor Beschreibung auf den vorhergehenden Webseiten erfolgen zwischen verschiedenen Teilnehmern. Auf den Seiten zum direkten Vergleich aller Beschreibungsvarianten (S16-17) geben Teilnehmer Einschätzungen zu allen Beschreibungsvarianten ab. Diese voneinander abhängigen Messungen werden durchgeführt, um die Ergebnisse mit der Untersuchung von Drori (2001), der ein vollständiges Messwiederholungsdesign anwandte, vergleichen zu können. Auf der ersten Seite (S16) wird ein für die Suche relevanter Treffer bei unspezifischer Suchaufgabe in allen vier Beschreibungsvarianten dargestellt und bewertet. Eine zweite Seite (S17) überprüft einen irrelevanten Treffer bei ebenfalls unspezifischer Fragestellung. Um die Untersuchungsdauer kurz zu halten und weil bei spezifischer Frage die gesuchten

Antworten teilweise in der Beschreibung enthalten sind und dadurch eine eindeutige Präferenz nahe legen, wurde auf den Vergleich bei spezifischer Suchaufgabe verzichtet.

### **3.1.3.8 Abschluss der Untersuchung**

Die Untersuchung wird durch Fragen abgeschlossen, in denen die Teilnehmer ihre ernsthafte Teilnahme bestätigen, ihre Daten von einer Auswertung ausschließen können und angeben, wie sie auf die Untersuchung aufmerksam geworden sind (S18). Danach folgen drei offene Fragen, wie gute Beschreibungen von Treffern aussehen, welche Informationen die Teilnehmer für ihre Selektionsentscheidung heranziehen und ein Bereich für weitere Anmerkungen zur Untersuchung. Die nächste Webseite enthält das Debriefing mit Erläuterungen zur Forschungsfrage der Untersuchung und der Möglichkeit, die Emailadresse zum Erhalt der Ergebnisse der Untersuchung anonym und unabhängig von den erhobenen Daten anzugeben (S19). Die zwanzigste und letzte Webseite enthält Links, die Antworten zu den simulierten Suchanfragen liefern (S20).

### **3.1.4 Exkurs Technische Umsetzung**

Die Untersuchung wurde über das Internet durchgeführt und umfasste je Bedingung 20 Webseiten. Da die Konstruktion der Webseiten, die Programmierung und Anpassung von Programmen zur Datenspeicherung und Aufbereitung vom Autor der vorliegenden Arbeit durchgeführt wurden, soll dieser Abschnitt die technische Seite der Untersuchung umreißen.

Jede für die Untersuchung notwendige Webseite lag fast vollständig auf dem Server vor, um die Antwortzeiten des Servers zu reduzieren und den Seitenaufbau bei den Teilnehmern zu beschleunigen. Um Fehler im Code der Webseiten zu vermeiden, wurde ein Programm geschrieben, mit dem sich aus prototypischen Dateien (diese enthalten den Text zwischen den beiden Body-Tags) alle 160 benötigten Webseiten erstellen lassen. Die Seiten wurden so konzipiert, dass Darstellungen mit verschiedenen Versionen des Internetexplorer und Netscape gleich aussahen. Dazu wurde die Mehrheit der Strukturmerkmale direkt in jede Seite kodiert und die Verwendung von Cascading-Style-Sheets auf Elemente beschränkt, die sowohl vom Internetexplorer als auch von Netscape korrekt interpretiert werden. Die Zeitmessungen, der Fortschrittsbalken und die Identifizierung des verwendeten Browsers und der Bildschirmauflösung wurden mit Javascript realisiert (zur Problematik und Durchführung von Zeitmessungen vgl. Eichstaedt, 2001; Janetzko, 1999).

Damit Teilnehmer eindeutig über alle Untersuchungsseiten identifiziert werden können, erhalten sie auf der Startseite (S1) eine eindeutige ID in Form eines Zeitstempels. Diese ID wird durch ein CGI-Skript (common gateway interface) in jede Webseite kodiert, bevor sie an die Teilnehmer versendet wird. Manipulationen an der ID werden dadurch ausgeschlossen. Nachdem eine Seite vom Teilnehmer bearbeitet wurde, wird sie durch den Teilnehmer-Klick auf den grafischen Pfeil am Ende jeder Seite an ein CGI-Skript auf dem Server gesendet, das die Speicherung der Daten sicherstellt und die nächste Untersuchungsseite an die Teilnehmer sendet. Um durch gleichzeitige Schreibvorgänge verursachte Zugriffsverletzungen beim Speichern der Daten zu vermeiden, was zu Datenverlust führen könnte, werden die gewonnenen Daten für jede Webseite getrennt in

eine einzelne Datei gespeichert. Diese sukzessive Speicherung der Daten ermöglicht zudem eine Dropoutanalyse (siehe Kap. 4.1.3).

Nach Ende der Untersuchung werden die gewonnenen 160 Dateien zu einer Datei zusammengefügt. Diese Datei enthält neben den interessierenden auswertbaren Daten auch viel irrelevantes, durch Abbrecher verursachtes Material. Zudem liegen die Daten in einer unsortierten und für Statistikprogramme nicht lesbaren Form vor. Durch die gleichzeitige Teilnahme mehrerer Personen kommt es außerdem vor, dass Daten von einer Person nicht linear hintereinander vorliegen, sondern sich mit Daten von anderen Teilnehmern überschneiden. Es wurde daher ein Programm in Java geschrieben, das die zu einem Teilnehmer gehörenden Daten anhand der ID identifiziert und sortiert. Gleichzeitig werden Dubletten, die durch mehrfaches Absenden einer Webseite entstehen, herausgefiltert und die Daten für den nächsten Schritt aufbereitet. Danach wird ein Programm (Parser) eingesetzt, das die zur Auswertung gewünschten Variablen ausliest und in ein für Statistikprogramme lesbares Format umschreibt. Besonders gute Ergebnisse bei der Identifizierung verwendbarer Daten wurde mit der Abschlusssaufforderung „Wählen Sie deshalb bitte  Ich habe ernsthaft an der Untersuchung teilgenommen“ erzielt. Konnte bei einem Datensatz aufgrund dieser Variable auf eine ernsthafte Teilnahme geschlossen werden, dann ergaben auch die anderen Filtermethoden kein widersprechendes Bild.

Mehrere Probleme im Datensatz ergeben sich typischerweise bei Webuntersuchungen. Teilnehmer müssen möglichst eindeutig identifiziert werden und Dubletten, die durch mehrfache Teilnahme oder falsche Bedienung entstehen, herausgefiltert werden. Abbrüche während der Untersuchung (Dropouts) kommen ebenfalls sehr häufig vor, und auch die nicht ernsthafte Teilnahme (Lurking, vgl. Bosnjak & Tuten, 2001), bei der sinnleere Antworten gegeben werden, sollten erkannt und gefiltert werden. Diese Probleme werden durch die Vergabe einer eindeutigen ID reduziert. Bei mehrfachem Aufruf der Eingangsseite zur Untersuchung würden Teilnehmer jeweils eine neue ID erhalten. Daher wird auch die IP-Adresse ausgelesen, eine Nummer, die zeitgleich nur einmal im Internet vorkommen kann und einen Computer im Internet eindeutig identifiziert. Zwar können Teilnehmer mit der ID eindeutig innerhalb eines Untersuchungsdurchlaufs identifiziert werden, aber es ist denkbar, dass Probanden mehrfach teilnehmen. Dies würde sich an einer Wiederholung der IP-Nummer im Datensatz bemerkbar machen. Zusätzlich ist es theoretisch möglich, dass Teilnehmer an verschiedenen Computern mehrfach teilnehmen. Dies kann nicht entdeckt werden. Da die Untersuchung jedoch keinerlei monetäre Belohnung bot, ist ein solcher Fall sehr unwahrscheinlich. Das zeitlich direkt aufeinander folgende Auftreten derselben IP bei mehr als einem Teilnehmer hintereinander kam in der Untersuchung nicht vor und ist ebenfalls als Indiz gegen Mehrfachteilnahme zu interpretieren. Um die Menge an Mehrfachteilnehmern einschätzen zu können, wurden manuelle Vergleiche beim mehrfachen Auftreten gleicher IPs von Datum, Anfangszeit der Untersuchung und der zur IP-Nummer gehörenden Webadresse durchgeführt. Fand sich dieselbe IP-Nummer mehrfach im Datensatz, so wurde sichergestellt, dass diese IP-Nummer von Internet Providern oder Proxies (z.B. Uni München) stammen, was anhand der Hostnamen zu erkennen ist. Andernfalls wurden die zugehörigen Daten entfernt. Internetprovider

verwalten weniger IP-Nummern als sie Kunden haben. Daher vergeben sie jeweils dynamisch eine IP-Nummer an ihre Kunden, wenn diese sich ins Internet einwählen. Proxies und auch Firewalls bieten ähnliche Zugänge für Internetnutzer. Beide stellen einen Durchgang ins Internet dar, hinter dem sich viele verschiedene Anwender befinden können. Bei IP-Nummern von Internet Providern, Proxies und Firewalls sollte daher davon ausgegangen werden, dass es sich um verschiedene Teilnehmer handelt. Die tatsächliche Mehrfachteilnahme eines Teilnehmers lässt sich jedoch nicht feststellen.

Weitere durchgeführte technische Maßnahmen betrafen das Ausschalten der Backbuttonfunktionalität, um Mehrfacheingaben zu vermeiden, und ein automatisches Maximieren des Browserfensters auf die maximale Fenstergröße zu Beginn der Untersuchung, um Fehlervarianz durch unterschiedliche Fenstergrößen zu minimieren. Dies ist insbesondere für die Zeitmessungen auf den Ergebnisseiten erforderlich, um ein unerwünschtes Scrollen zu verhindern (vgl. Kap. 3.1.3.4).

Die Hauptziele in der technischen Umsetzung sind damit Messung, Speicherung und Aufbereitung der Daten. Weitere Maßnahmen sollen störende Einflüsse reduzieren, die Fehlervarianz in den Daten erzeugen können.

### **3.1.5 Vortests, Durchführung und Ableitungen**

Vor der Konstruktion der Webseiten wie sie bisher beschrieben wurden fanden Vortests mit Prototypen der Untersuchungsseiten statt. Diese dienten zur Prüfung, ob die verwendeten Erhebungs- und Auswertungsmethoden funktionierten und ob die Texte der Webseiten von Teilnehmern so aufgefasst wurden, wie dies beabsichtigt war. Insgesamt fanden drei Vortests statt. Zunächst werden die eingesetzten Methoden beschrieben, um anschließend auf die gewonnenen Erkenntnisse und daraus resultierenden Designänderungen einzugehen. Die in den Vortests verwendeten Webseiten finden sich im Anhang H.

#### **3.1.5.1 Beobachtete Teilnahme mit lautem Denken**

Im Januar 2002 wurden sechs Personen bei der Teilnahme an der Voruntersuchung direkt beobachtet, wobei die Methode des lauten Denkens (MLD) eingesetzt wurde. „[D]as sofortige, reflexhafte Aussprechen präsender Gedanken während der Navigation im WWW ist der methodische Vorteil der MLD und erlaubt ‚Einblicke‘ in Informationsaufnahme und -verarbeitung.“ (Wirth & Brecht, 1999, S. 159). Dadurch kann besonders gut erkannt werden, wo Schwierigkeiten in der Untersuchung auftauchen und welche Texte problematisch sind (Dillman, 2000). Entsprechend konnten die Instruktionstexte durch die Methode präzisiert werden.

#### **3.1.5.2 Gruppendiskussion**

Im Februar 2002 nahmen im Rahmen eines Seminars zur Methodologie von Webuntersuchungen eine Gruppe von acht Personen am Vortest teil. Anschließend wurde die Untersuchung diskutiert. Der Vorteil bestand hierbei darin, dass konkret methodische Schwierigkeiten im Design aufgedeckt werden konnten.

### **3.1.5.3 Feldtest**

Um die reibungslose Durchführung in der Hauptuntersuchung sicherzustellen, wurde abschließend Anfang März 2002 ein Probelauf durchgeführt, der die Designverbesserungen der ersten beiden Vortests enthielt. Dabei sollte vor allem die technische Umsetzung auf Korrektheit geprüft werden. 19 Personen kamen der Aufforderung per E-Mail, an der Untersuchung teilzunehmen, nach. Ihre Rückmeldungen sorgten für eine Überarbeitung der offenen Fragen am Ende der Untersuchung (S18).

### **3.1.5.4 Designänderungen**

Die Vortests erwiesen sich als besonders geeignet, um unklare Textformulierungen aufzudecken. Davon profitierten besonders die Instruktionstexte und die Items zur Zufriedenheit. Fragen wurden mit einem deutlicheren inhaltlichen und zeitlichen Bezug versehen (vgl. Bortz & Döring, 2002). Die Verwendung von Matrixfragen (Fragen und Antwortkategorien blockweise angeordnet) und dadurch verursachte erhöhte Korrelationen zwischen den Items wurden reduziert.

Es wurden zwei wesentliche Änderungen im Design vorgenommen. Diese betrafen die Bewertung einzelner Treffer (S8, S13) und die Gesamtbewertung der Beschreibungstexte der Treffer (S9-10, S14-15). Im Vortest erfolgte die Bewertung der einzelnen Treffer dadurch, dass Teilnehmer für die selektierten Treffer eine Rangfolge bilden sollten. Solch ein persönliches Ranking erlaubt Aussagen darüber, welche Treffer gegenüber anderen der gleichen Ergebnisseite relevanter erscheinen. Keine Aussagen können hingegen zwischen verschiedenen Ergebnisseiten getroffen werden. Das Ranking wurde daher durch ein Bewertungsverfahren ersetzt, bei dem für jeden Treffer angegeben wird, wie wahrscheinlich der Treffer zu den gesuchten Informationen führt. Durch diese Angabe für jeden einzelnen Treffer können sowohl die Treffer innerhalb einer Ergebnisseite verglichen werden, wie auch die Mittelwerte zwischen verschiedenen Ergebnisseiten.

Die zweite wesentliche Änderung im Design betraf die Erhebungsmethode zur Bewertung der Trefferbeschreibungen. Im Vortest fand eine Bewertung getrennt nach ausgewählten und nicht-ausgewählten Treffern statt. Diese Erhebungsmethode wies einige Schwächen auf, was auf mehrere Ursachen zurückzuführen ist: Teilnehmer merkten an, dass sie sich nicht ausreichend erinnern konnten, um tatsächlich einen Unterschied zwischen ausgewählten und nicht-ausgewählten Treffern festzustellen. Ein weiterer Problempunkt war, dass es kaum möglich schien, von der jeweiligen Relevanz der Treffer auf die Beschreibungsqualität im Allgemeinen zu abstrahieren, wie es für die Untersuchung nötig gewesen wäre. Teilnehmer bewerteten nach eigener Aussage, ob relevante Treffer besser waren als nicht relevante, anstatt wie von der Instruktion gefordert anzugeben, wie gut die Beschreibung der Treffer zur Entscheidung beitrug. Demnach sollte ein Treffer, der eindeutige Hinweise darauf gibt, dass er für die Suchanfrage nicht relevant ist, eine positive Bewertung erhalten, da die Beschreibung ihren Zweck optimal erfüllt. Dies war jedoch nicht der Fall. Nicht-relevante Treffer erhielten eine schlechtere Bewertung als relevante, womit sich keine zusätzliche Information für die Fragestellung dieser Arbeit ergab. Zwar mag argumentiert werden, dass eine längere Instruktion den Sachverhalt entschärft hätte, jedoch ergäbe sich

auch dann eine Formulierung, in der Teilnehmer von der sonst üblichen Sichtweise auf die Relevanz von Treffern abstrahieren müssten. Da im Webexperiment nicht kontrolliert werden kann, inwieweit dies tatsächlich gelingt und die oben angesprochene Gedächtnisproblematik ein deutlicher Nachteil ist, wurde auf diese Differenzierung zugunsten eines kürzeren Fragebogens verzichtet.

### 3.1.6 Anforderungen an die Gebrauchstauglichkeit

Dieses Kapitel beschreibt, welche abhängigen Variablen zum Vergleich der verschiedenen Faktorstufen herangezogen werden. Die ISO 9241-11 definiert die Anforderungen Effektivität, Effizienz und Zufriedenheit. Diese werden als Hauptkategorien für die im Folgenden darzustellenden Kriterien verstanden.

#### 3.1.6.1 Effektivität

In der ISO 9241-11 wird Effektivität als die Genauigkeit und Vollständigkeit definiert, mit der das Ziel des Benutzers erreicht wird. In der Information Retrieval Literatur werden in diesem Zusammenhang die Begriffe *Genauigkeit* (precision) und *Vollständigkeit* (Recall) verwendet. Sie sind die gebräuchlichsten Maße für die Leistungsfähigkeit von Suchmaschinen bezogen auf die erhaltenen Ergebnisseiten (Harter, 1992; Konrad, 1992). Dazu muss für eine gegebene Suchanfrage vorab bekannt sein, welche Dokumente für diese Suchanfrage relevant und welche irrelevant sind. Diese Festlegung geschieht normalerweise durch Experten für eine gegebene Fragestellung bei vorgegebenem Datenbestand. Genauigkeit oder Präzision bezeichnet die Menge an erhaltenen relevanten Dokumenten, die eine Suchmaschine in Bezug zur Gesamtzahl der gelieferten Dokumente liefert. Vollständigkeit bezeichnet die Menge an erhaltenen relevanten Dokumenten in Bezug zur Gesamtzahl aller relevanten Dokumente. In der Literatur über Entscheidungstheorie sind die Begriffe Sensitivität, Teststärke,  $\alpha$ -Fehler und  $\beta$ -Fehler gebräuchlich. Diese Begriffe lassen sich mathematisch ineinander überführen wie Tabelle 2 zur Entscheidungsmatrix veranschaulicht. Sind jeweils die Inhalte zweier Zellen in einer Zeile oder Spalte bekannt, so kann der Inhalt der dritten Zelle errechnet werden. So gilt z.B.  $Teststärke = 1 - \beta\text{-Fehler}$ .

Tabelle 2: Entscheidungsmatrix.

	Treffer ist relevant	Treffer ist irrelevant	Randsummen
selektiert	Genauigkeit Vollständigkeit (Sensitivität)	$\beta$ -Fehler	Anzahl selektierter Treffer
nicht selektiert	$\alpha$ -Fehler	Teststärke	Anzahl nicht selektierter Treffer
Randsummen	Anzahl relevanter Treffer	Anzahl irrelevanter Treffer	Anzahl Treffer auf Ergebnisseite

Aus der Entscheidungsmatrix lassen sich die möglichen Ausgänge von Entscheidungen ablesen. Ist ein Treffer auf einer Ergebnisseite relevant und wird er von Anwendern selektiert, so ergeben sich daraus die beiden Kriterien Genauigkeit und Vollständigkeit (letzteres entspricht der Sensitivität), die jeweils auf die Zeilen- bzw. Spaltensumme bezogen werden. Jede der vier Zellen in der Entscheidungsmatrix könnte als Kriterium für die Effektivität herangezogen werden. Da die verschiedenen Kriterien jedoch alle direkt mathematisch voneinander abhängen, genügt es, die beiden gebräuchlichsten Kriterien Genauigkeit und Vollständigkeit zu prüfen. Die Selektionsentscheidungen der Anwender unter dem Gesichtspunkt der Effektivität werden daher in der vorliegenden Arbeit mit Hilfe dieser beiden Kriterien untersucht.

Inwieweit der Anwender in der Lage ist, alle für ihn relevanten Treffer zu erkennen und auszuwählen und damit alle ihn interessierenden Informationen zu erreichen, entspricht der Vollständigkeit (recall). Genauigkeit (precision) ist ein Maß dafür, inwieweit ausschließlich relevante Treffer ausgewählt werden. Würde der Anwender alle dargebotenen Treffer wählen, ergäbe sich eine Vollständigkeit von hundert Prozent, jedoch nur eine geringe Genauigkeit. Würden umgekehrt zwei relevante Dokumente ausgewählt und damit hundert Prozent Genauigkeit erreicht, jedoch viele andere relevante Dokumente ignoriert, führte dies zu einer geringen Vollständigkeit. Genauigkeit ist die Anzahl selektierter und gleichzeitig relevanter Treffer dividiert durch die Gesamtanzahl der selektierten Treffer (Formel 2). Vollständigkeit ist die Anzahl selektierter und gleichzeitig relevanter Treffer dividiert durch die Gesamtanzahl aller relevanten Treffer (Formel 3).

**Formel 2:** Genauigkeit der Selektionsentscheidungen.

$$\text{Genauigkeit} = \frac{\text{relevant} \cap \text{selektiert}}{\text{selektiert}}$$

**Formel 3:** Vollständigkeit der Selektionsentscheidungen.

$$\text{Vollständigkeit} = \frac{\text{relevant} \cap \text{selektiert}}{\text{relevant}}$$

Relevant ist ein Treffer in der vorliegenden Arbeit dann, wenn eine eindeutige Mehrheit der Teilnehmer den Treffer für relevant hält (siehe Kap. 4.1.1).

### **3.1.6.2 Effizienz**

Effizienz wird in der ISO 9241-11 definiert als das „Verhältnis von Aufwand zu erreichter Effektivität bei der Zielerreichung“. Je höher der kognitive und zeitliche Aufwand im Entscheidungsprozeß bei gleicher Effektivität, desto geringer ist die Effizienz. Entsprechend verringert sich die Effizienz, wenn viele irrelevante Treffer ausgewählt würden (geringe Genauigkeit), durch den höheren Zeitaufwand, den der Anwender bei der Betrachtung irrelevanter Seiten vergeuden würde. Die Effizienz enthält damit den zeitlichen Aspekt von Messungen. In der vorliegenden Untersuchung wird die Zeitspanne in Millisekunden gemessen, die die Teilnehmer für jede Selektion benötigen. Die aggregierten relativen Zeiten

(siehe Kap. 3.1.3.4) zwischen den einzelnen Selektionen werden als *Entscheidungszeit* für die Selektionsentscheidung interpretiert.

Die *Erfolgsrate* (success rate) von Wolfram und Dimitroff (1997) entspricht der von der ISO Norm vorgegebenen Definition von Effizienz dadurch, dass die Effektivität in Relation zur benötigten Zeit gesetzt wird. Die Autoren definieren die Erfolgsrate als die Genauigkeit gewichtet mit der Vollständigkeit in Relation zur für die Suche benötigten Zeit. Daran angelehnt wird die Erfolgsrate auf den vorliegenden Untersuchungsgegenstand angepasst: Die Zeitdauer der Suche wird durch die Entscheidungszeit ersetzt. Die Erfolgsrate in der vorliegenden Arbeit entspricht damit der Genauigkeit gewichtet mit der Vollständigkeit in Relation zur Entscheidungszeit (Formel 4).

**Formel 4:** Erfolgsrate der Suche nach Wolfram und Dimitroff (1997).

$$\text{Erfolgsrate} = \frac{\text{Genauigkeit} \times \text{Vollständigkeit}}{\text{Entscheidungszeit}}$$

### 3.1.6.3 Zufriedenheit

Zufriedenheit wird in der ISO 9241-11 definiert als die „Freiheit von Beeinträchtigung und positive Einstellung des Benutzers gegenüber der Nutzung des Produktes“. Zufriedenheit umfasst die Einstellung des Anwenders gegenüber der Ergebnisseite: Ist sie hilfreich, sind die Beschreibungen verständlich, fühlt sich der Anwender durch zu viel Information beeinträchtigt oder kann er aufgrund von zu wenig Informationen eine Entscheidung nur schwer fällen. *Zufriedenheit* wird in der vorliegenden Studie mit vierzehn Items erfasst, die Einschätzungen zur Trefferliste und den Beschreibungstexten erlauben (siehe Kap. 3.1.3.6). Darüber hinaus sollen unter die Anforderung Zufriedenheit auch weitere Einschätzungen von Teilnehmern gefasst werden: Die Sicherheit, mit der Teilnehmer Treffer auswählen, wird als *Entscheidungssicherheit* bezeichnet (siehe Kap. 3.1.3.5) und weist Ähnlichkeit zur ‚Pre-Click-Confidence‘ von Lergier und Resnick (2001) auf (vgl. Kap. 3.1.3.5). Teilnehmern werden ähnlich wie bei Drori (2001) ebenfalls alle Beschreibungsvarianten präsentiert, wodurch erfasst wird, welche Variante sie im *direkten Vergleich* präferieren (siehe Kap. 3.1.3.7).



### 3.1.7 Fazit

Wie die verschiedenen Elemente des Rahmenmodells operationalisiert und welche abhängigen Variablen zur Messung herangezogen werden, wird nun in Tabelle 3 zusammengefasst.

Tabelle 3: Variablen im Untersuchungsdesign.

<b>Theoretische Konstrukte und ihre Teilaspekte</b>	<b>Operationalisierungen, gemessene Variablen</b>
<b><i>Mensch</i></b>	
Demografische Merkmale	Alter, Geschlecht, Rekrutierungsweg
Wissen	Acht Teilfragen zu: Welche Suchmaschinen wurden in der letzten Woche wie lange, bei wie vielen Suchanfragen verwendet? Wie viele Tage wurde in der letzten Woche am PC und im Internet gearbeitet? Drei Fragen zur Selbsteinschätzung: Wissen über Suchmaschinen, Währungsumstellung und Wirtschaft.
Motorische Zeitkomponente bei der Trefferauswahl und Lesedauer	Zeitmessung in den Übungsdurchgängen zur Trefferauswahl.
<b><i>Maschine</i></b>	
Attribut Beschreibung	Vier Beschreibungsvarianten: Kontext, Textanfang, Schlüsselwörter, alle drei vorherigen zusammen (MaxInfo)
Andere Attribute	Der Titel des repräsentierten Dokuments wird angezeigt
<b><i>Aufgabe</i></b>	
Spezifität	Zwei Aufgaben mit unterschiedlicher Spezifität: „Welche Größe haben Euroscheine?“ und „Welche Folgen hat die Währungsumstellung?“.
Verhalten	Es sind Schlussfolgerungen anhand der Selektionsreihenfolge möglich.
<b><i>Abhängige Variablen</i></b>	
Effektivität	Genauigkeit
	Vollständigkeit
Effizienz	Erfolgsrate
	Entscheidungszeit
Zufriedenheit	Entscheidungssicherheit
	Vierzehn Items zur Zufriedenheit
	Direkter Vergleich der vier Beschreibungsvarianten

## 3.2 Forschungsleitende Fragen

Nach der Darstellung der Kriterien können nun die forschungsleitenden Fragen genauer betrachtet werden.

### 3.2.1 Maschine

Aus der Leitfrage „Wie sollte das Attribut Beschreibung gestaltet sein, damit Anwender optimal bei ihrer Selektionsentscheidung unterstützt werden?“ ergibt sich für die Untersuchung die Fragestellung: „Existieren Unterschiede zwischen den verschiedenen Bedingungen im Faktor Beschreibung?“. Diese lassen sich weiter unterteilen in die Einzelfragen:

Frage 1.1: Welche Beschreibung führt zur höchsten Genauigkeit?

Frage 1.2: Welche Beschreibung führt zur höchsten Vollständigkeit?

Frage 1.5: Welche Beschreibung führt zur höchsten Erfolgsrate?

Frage 1.4: Welche Beschreibung erlaubt die geringste Entscheidungszeit?

Frage 1.3: Welche Beschreibung bietet die höchste Entscheidungssicherheit?

Frage 1.6: Welche Beschreibung führt zur größten Zufriedenheit?

Frage 1.7: Welche Beschreibung präferieren Anwender im direkten Vergleich?

### 3.2.2 Aufgabe

Die Leitfrage in der Ebene Aufgabe lautet „Welchen Einfluss hat die Spezifität der Aufgabe auf die Suche?“. Aus der Theorie zur Aufgabenspezifität (siehe Kap. 2.3.1) ergibt sich die Hypothese, dass für jedes Kriterium, das sich auf den Erfolg der Suche bezieht, Anwender bei spezifischer Suchaufgabe erfolgreicher sind als bei unspezifischer Suchaufgabe. Im Einzelnen wird eine höhere Genauigkeit, Vollständigkeit und Erfolgsrate erwartet. Anwender bei spezifischen Suchaufgaben benötigen weniger Entscheidungszeit, die Entscheidungssicherheit ist ebenso wie die Zufriedenheit höher.

### 3.2.3 Mensch

Aus der Leitfrage „Welchen Einfluss haben die Unterschiede zwischen den Anwendern auf den Sucherfolg?“ lässt sich die Fragestellung formulieren „Existiert ein moderierender Einfluss der Eigenschaften der Anwender auf die abhängigen Variablen?“.

## 3.3 Auswertungsverfahren

Um die Ergebnisdarstellungen nicht mit Ausführungen zu Voraussetzungsüberprüfungen und Erläuterungen zu Test-Statistiken zu überfrachten, werden an dieser Stelle die eingesetzten Auswertungsverfahren beschrieben. Das Design der vorliegenden Untersuchung lässt sich mit Hilfe von einfaktoriellen Varianzanalysen mit Messwiederholungsfaktoren auswerten. Eine Varianzanalyse mit Messwiederholung ist als Erweiterung des t-Tests für abhängige Stichproben anzusehen (Bortz, 1999). Hierbei soll die englische Terminologie nach Brace, Kemp und Snelgar (2000) verwendet werden, da diese

einerseits einfacher im Wortgebrauch und andererseits eindeutiger in der Bezeichnung der verschiedenen verwendeten Auswertungsdesigns ist. Der Faktor Beschreibung entspricht demnach einem between-subjects Faktor, der Faktor Spezifität einem within-subjects Faktor. Der Faktor Spezifität ist damit stets ein Messwiederholungsfaktor. Bei einer Auswertung über das Kriterium Genauigkeit kann daher von einer mixed ANOVA mit einem within-subjects und einem between-subjects Faktor gesprochen werden.

Die Voraussetzungen der Varianzanalyse ohne Messwiederholung, wie sie auch für Varianzanalysen mit Messwiederholung gelten, können, da die Varianzanalyse ein relativ robustes Verfahren ist und angesichts der großen Stichprobe, als gegeben bzw. als vernachlässigbar betrachtet werden (Bortz, 1999). Eine Besonderheit bei Messwiederholungen stellt jedoch die geforderte Unabhängigkeit zwischen den verschiedenen within-subjects Faktorstufen dar. Da diese üblicherweise miteinander korreliert sind, führt der F-Test nur dann zu richtigen Entscheidungen, wenn eine weitere Voraussetzung erfüllt ist, die Sphäritätsannahme. Diese fordert, dass (a) die Varianzen unter den einzelnen Faktorstufen und (b) die Korrelationen zwischen den Faktorstufen homogen sind. Dies kann mit Hilfe des Mauchly-Tests überprüft werden (Brace et al., 2000). Ist die Homogenitätsannahme verletzt, so entscheidet die Varianzanalyse mit Messwiederholung progressiv (Bortz, 1999). Um dies zu vermeiden, wird in den folgenden Varianzanalysen zunächst mit Hilfe des Mauchly-Tests die Voraussetzung überprüft. Ist er signifikant, wird die Test-Statistik für den multivariaten Ansatz mit Wilks  $\lambda$  angewendet. Sie ist an weniger Voraussetzungen geknüpft und entscheidet konservativer als die Varianzanalyse. Der Mauchly-Test selbst wird nicht berichtet, sondern jeweils die erforderliche Test-Statistik zur Auswertung herangezogen. Zur Identifizierung, welche Bedingungen im Faktor Beschreibung signifikant voneinander abweichen, werden Post-hoc-Mehrfachvergleiche mit Hilfe des Tukey-Tests durchgeführt. Die ebenfalls eingesetzten Tests Scheffé und Bonferroni ergeben in der vorliegenden Arbeit weitestgehend dieselben Effekte, scheinen jedoch konservativer zu sein. Da die Relevanz für die einzelnen Treffer relativ eindeutig ist (siehe Kap. 4.1.1), können die Abweichungen in den Selektionsentscheidungen zwischen den Teilnehmern nur gering sein. Um die damit verbundene geringe Effektgröße zu entdecken, ist der Einsatz eines progressiven Post-hoc-Tests sinnvoll. Es wird daher darauf verzichtet, die Post-hoc Tests einzeln aufzuführen und jeweils nur angegeben, welche Unterschiede sich als signifikant erweisen. Zur Interpretation der Effekte werden Abbildungen herangezogen, die die von der Varianzanalyse geschätzten Mittelwerte darstellen. In Anhang A finden sich die jeweils zugehörigen in der Stichprobe berichteten Mittelwerte.

## 4 Ergebnisse

### 4.1 Überprüfung der Daten und der Implementation

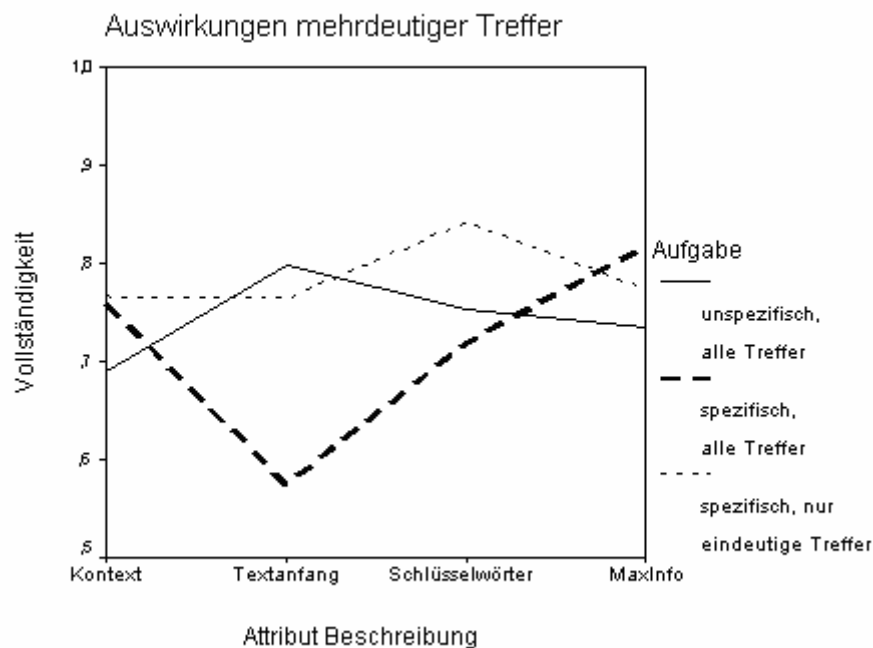
Um in der Interpretation der Ergebnisse zulässige kausale Schlüsse zu ziehen, ist es unumgänglich, die erhobenen Daten und das Design verschiedenen Überprüfungen zu unterziehen. Diese sollen eine Gefährdung der Validität möglichst minimieren (Cook & Campbell, 1979). Alle in diesem Kapitel dargestellten Überprüfungen dienen dem Ziel festzustellen, ob störende Effekte vorliegen, die die Aussagekraft der Ergebnisse reduzieren können.

#### 4.1.1 Relevanz, Prüfen der Treffertauglichkeit

Treffer fungieren in der vorliegenden Arbeit als Testitems und müssen daher auch gewissen Eigenschaften genügen, um sinnvoll in die Auswertung einfließen zu können. Da die Treffer einzeln für die Untersuchung erzeugt wurden, kann es nicht das Ziel der Arbeit sein, die zugrundeliegenden Algorithmen, die die Beschreibungen erzeugen, gegeneinander abzuwägen. Um eine Aussage über die Effektivität der Algorithmen treffen zu können, wäre eine weit größere Menge an Treffern notwendig. Entsprechend ist es denkbar, dass einzelne Treffer der vorliegenden Arbeit aufgrund irreführender Texte einen zu hohen Einfluss auf die Selektionsentscheidung haben und dadurch irrtümlich zu einem Effekt durch die Beschreibung führen. Treffer, bei denen die Mehrzahl der Teilnehmer nicht eindeutig feststellen kann, ob das repräsentierte Dokument relevant ist oder nicht, müssen daher von der Auswertung ausgeschlossen werden. Ein Treffer wird in diesem Sinne als relevant definiert, wenn eine eindeutige Mehrheit an Teilnehmern ihn für relevant hält. „By definition, whatever a user says is relevant, is relevant, and that is about all that can be said.“ (Harter, 1992, S. 603) Die Auswirkungen von nicht eindeutig vorhandener Relevanz eines Treffers wird an Abbildung 9 deutlich. Für das Kriterium Vollständigkeit ergeben sich unterschiedliche Effekte je nachdem, ob alle Treffer bei spezifischer Aufgabe in die Auswertung eingehen oder nur diejenigen Treffer, bei denen eine Relevanzzuordnung von der Mehrheit der Teilnehmer eindeutig vorgenommen wird. Der beobachtete Einbruch an Vollständigkeit bei der Beschreibung Textanfang ergibt einen signifikanten Interaktionseffekt mit der Spezifität der Aufgabe ( $F=8,43$ ;  $df=3$ ;  $p=0,00$ ), wobei die Haupteffekte nicht signifikant werden (Beschreibung:  $F=1,81$ ;  $df=3$ ;  $p=0,15$ ; Spezifität:  $F=1,08$ ;  $df=1$ ;  $p=0,30$ ). Dies kommt dadurch zustande, dass bei Treffer Nr. 6 im Textanfang keine Hinweise auf die Relevanz zu finden sind, die Beschreibungen Kontext und Schlüsselwörter jedoch eindeutige Hinweise enthalten (vgl. Anhang I). Dadurch wird der Treffer Nr. 6, obwohl er vor der Untersuchung als relevant klassifiziert wurde, von Teilnehmern in der Bedingung Textanfang als irrelevant identifiziert. Der Effekt beruht dadurch allein auf einer ungeeigneten Beschreibung eines Treffers. Die Wahl eines anderen Webdokuments als repräsentiertes Dokument hätte in diesem Fall leicht andere Effekte herbeiführen können. Der Effekt muss daher als Artefakt bezeichnet werden und der Treffer wird dementsprechend aus der Auswertung ausgeschlossen. Nach dem Ausschließen der in

der spezifischen Bedingung als ungeeignet identifizierten Treffer Nr. 6 und Nr. 7 verschwindet der beschriebene Interaktionseffekt und die Haupteffekte bleiben nicht signifikant (Interaktion:  $F=1,08$ ;  $df=3$ ;  $p=0,36$ ; Beschreibung:  $F=1,15$ ;  $df=3$ ;  $p=0,33$ ; Spezifität:  $F=2,51$ ;  $df=1$ ;  $p=0,12$ ).

Ungeeignete Treffer können statistisch mittels t-Test für eine Stichprobe gegen den erwarteten Mittelwert getestet werden, der bei zufälliger Selektionsentscheidung entstehen würde. Da eine Selektion mit Eins und eine Nichtselektion mit Null kodiert werden, liegt der erwartete Mittelwert für zufällige Entscheidungen bei 1,5. In die Auswertung werden nur solche Treffer aufgenommen, bei denen die Entscheidungen sich signifikant von zufälligen Selektionen unterscheiden.



**Abbildung 9:** Auswirkungen mehrdeutiger Treffer auf die Kriterien.

Die Durchführung mehrerer solcher Tests auf denselben Effekt hin führt zu einer  $\alpha$ -Fehler-Kumulierung. Aus diesem Grund wird die Bonferoni-Korrektur nach Formel 5 eingesetzt (Bortz, 1999).

**Formel 5:** Bonferoni-Korrektur

$$\text{Neue Irrtumswahrscheinlichkeit } \alpha' = \frac{\alpha}{\text{Anzahl Tests}} = \frac{0,05}{14} = 0,00357$$

Effekte sind demnach erst ab einer Irrtumswahrscheinlichkeit von gleich oder unter 0,00357 signifikant. Tabelle 4 zeigt die statistische Prüfung der Treffer für die spezifische und unspezifische Suchaufgabe. Daraus ist ersichtlich, dass sich die Treffer Nr. 6 und Nr. 7 der spezifischen Aufgabe deutlich von den restlichen Treffern unterscheiden und daher von der Auswertung ausgeschlossen werden.

**Tabelle 4:** T-Test für eine Stichprobe zur Prüfung der Treffertauglichkeit.

Aufgabe	Treffer	N	Mittelwert	df	Sig. (2-seitig)
spezifisch	L1	189	1,08	188	,000
	L2	187	1,98	186	,000
	L3	188	1,34	187	,000
	L4	186	1,99	185	,000
	L5	185	1,97	184	,000
	L6	188	1,41	187	,013
	L7	188	1,57	187	,041
unspezifisch	L1	187	1,28	186	,000
	L2	181	1,88	180	,000
	L3	187	1,26	186	,000
	L4	179	1,98	178	,000
	L5	180	1,88	179	,000
	L6	184	1,17	183	,000
	L7	176	1,98	175	,000

#### 4.1.2 Unvollständige Angaben (Missings)

Fehlende Werte (Missings) im Datensatz entstehen aus den unterschiedlichsten Gründen heraus. Teilnehmer können bewusst die Antwort auf eine Frage verweigern oder eine Frage zufällig übersehen. Problematisch werden fehlende Angaben bei einzelnen Variablen, wenn diese systematisch mit dem Untersuchungsgegenstand im Zusammenhang stehen. Little und Rubin (1987) unterscheiden drei mögliche Ausfallmuster: ‚missing completely at random‘ (MCAR), ‚missing at random‘ (MAR) und ‚non-ignorable‘ (NI). Sind die fehlenden Werte rein zufälliger Natur und findet sich daher kein systematischer Zusammenhang zwischen ihnen und den vorhandenen Werten, so sprechen die Autoren von ‚missing completely at random‘. In diesem Fall besteht keine Gefahr für die Validität. Bei ‚missing at random‘ findet sich zwar ein systematischer Zusammenhang, dieser kann jedoch mit Hilfe bekannter anderer Variablen ausreichend vorhergesagt und damit auch statistisch kontrolliert werden. In die Kategorie ‚non-ignorable‘ fallen fehlende Werte dann, wenn das Ausfallmuster systematisch ist und nicht kontrolliert werden kann. In diesem Fall besteht eine ernsthafte Gefährdung der Validität in den Daten.

Systematische Zusammenhänge zwischen unvollständigen und vollständigen Daten lassen sich im bivariaten Fall mit Hilfe von t-Tests berechnen. Dazu wird die Variable mit den unvollständigen Werten in eine ‚Dummyvariable‘ umkodiert: Fehlende Werte erhalten den Wert Null, vorhandene Werte werden in Eins umgewandelt. Anschließend kann diese Variable als Gruppenfaktor im t-Test verwendet werden, um Mittelwertsunterschiede in beliebigen anderen Variablen festzustellen. Sind diese Unterschiede nicht signifikant, kann das Ausfallmuster der Variable als ‚missing completely at random‘ betrachtet werden. Im multivariaten Fall, wie in der vorliegenden Arbeit, kann der Little’s MCAR-Test verwendet

werden. Er beruht auf einem Vergleich der Verteilung der vollständigen Werte mit der Verteilung der geschätzten fehlenden Werte. Bei signifikanter Abweichung der beiden Verteilungen voneinander liegt kein zufälliges Ausfallmuster vor.

Die wesentlichen Aussagen in der vorliegenden Arbeit betreffen Unterschiede, die mit den Faktorstufen im varianzanalytischen Design zusammenhängen. Gefährdungen sind daher besonders durch systematische Ausfallmuster gegeben, die mit den Faktorstufen kovariieren. Solch ein Zusammenhang zwischen fehlenden Angaben und den Faktorstufen würde eine kausale Interpretation verhindern. Um die Art der Ausfallmuster festzustellen, wurden die kriterienrelevanten Variablen daher mit Hilfe des Little's MCAR-Tests auf systematische Ausfälle bezüglich der Faktorstufen hin untersucht. Diese Tests waren für die bedeutsamen Kriterien nicht signifikant (Tabelle 5), sodass die fehlenden Werte keine Gefährdung der Validität darstellen und in der Auswertung keiner besonderen Beachtung über die normalen Verfahren hinaus bedürfen. Fehlende Werte werden daher in der folgenden Auswertung mit den üblichen Verfahren des fallweisen Ausschlusses (listwise deletion) und des paarweisen Ausschlusses (pairwise deletion) behandelt (für eine Übersicht vgl. Bosnjak, 2002). Bei den Übungsdurchgängen zeigt sich ein signifikantes Ausfallmuster, diese werden jedoch ohnehin aufgrund der misslungenen Operationalisierung von der Auswertung ausgeschlossen (siehe Kap. 4.1.4).

**Tabelle 5:** Little's MCAR Test zur Prüfung auf zufällige Ausfallmuster und prozentuale Anzahl fehlender Werte.

Variablen	Prozent fehlender Werte	Little's MCAR Wert	df	p
Eingangsforschfragebogen S1	7,05%	442,996	430	0,322
Übung S3; S4	11,67%	3219,786	198	0,000
Selektion von Links S7; S12	3,33%, 1,94% bei spezifischer, 4,71% bei unspezifischer Aufgabe	195,910	195	0,468
Entscheidungssicherheit S8, S13	0,90%	87,402	119	0,987
Treatmentcheckfragen S9, I1-3; S14, I1-3	4,38%	60,705	48	0,103
Einstellungsfragen S9, I6-11; S10; S14, I6-11; S15	1,61%	531,980	575	0,900
Direkter Vergleich S16; S17	1,31%	66,961	56	0,150

Fehlende Angaben unter fünf Prozent können als wenige Missings aufgefasst werden (o.V., 2000). Das Verhältnis von vorhandenen Daten zu fehlenden Werten kann daher als gut bis sehr gut bezeichnet werden.

#### 4.1.3 Dropouts, Prüfen der Randomisierung

Datenverluste entstehen nicht nur durch die Nicht-Beantwortung einzelner Items, sondern auch dadurch, dass Teilnehmer die Untersuchung vorzeitig abbrechen. Systematische Abbrüche führen zu einer Verzerrung der Stichprobe und würden das Experiment auf die

Aussagemöglichkeiten eines Quasi-Experiments herabstufen (Cook & Campbell, 1979). Um solch einen Effekt zu prüfen, wird eine Dropoutanalyse durchgeführt (Abbildung 10).

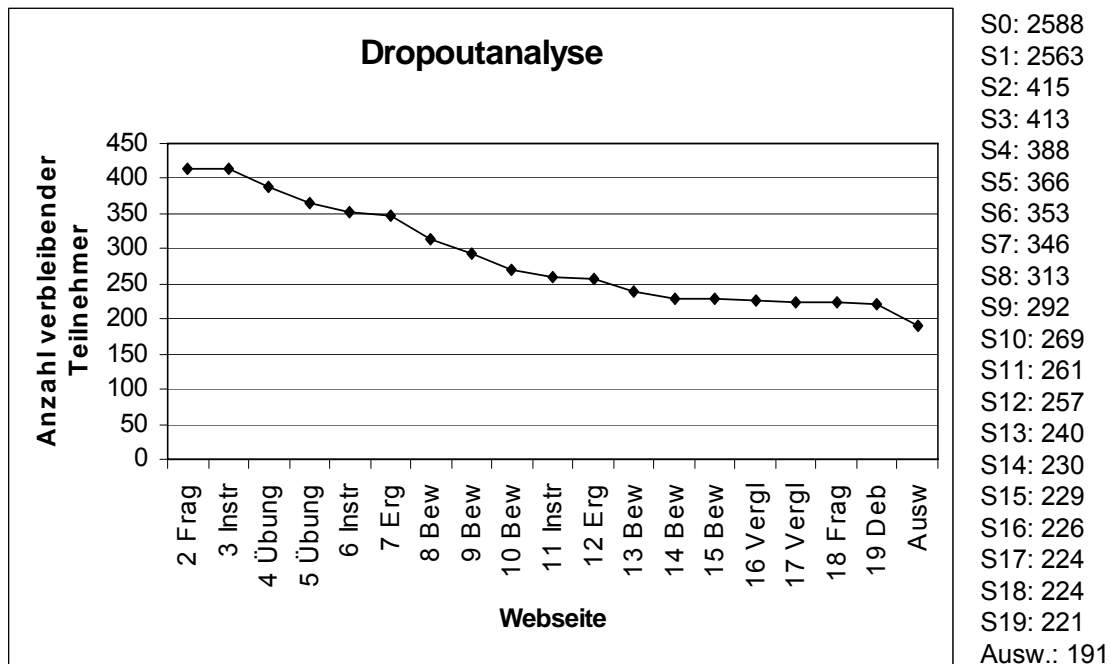


Abbildung 10: Dropoutanalyse.

Die Abnahme der Teilnehmer im Selbstselektionsprozess bis hin zur verbleibenden Teilnehmeranzahl wird im Folgenden beschrieben. Sobald sich der Teilnehmer für die Untersuchung interessiert, besucht er die Startseite. Hier besteht eine erste Hürde darin, dass Javascript für eine Teilnahme eingeschaltet sein muss (vgl. Kap. 3.1.4). War Javascript nicht eingeschaltet, erhielt der Teilnehmer eine entsprechende Meldung („Die Untersuchung funktioniert nur mit eingeschaltetem Javascript. Bitte aktivieren Sie Javascript in ihrem Browser, um an der Untersuchung teilzunehmen. Rufen Sie dann diese Seite erneut auf.“). Laut der Zählung mit Nedstat betraf dies 125 Personen von insgesamt ungefähr 2660 potenziellen Teilnehmern, die ihre Teilnahmebereitschaft durch Besuchen der ersten Untersuchungsseite bekundeten.<sup>23</sup> Bei aktiviertem Javascript wurde der Teilnehmer zufällig einer der acht Bedingungen zugewiesen. Auf jeder der ersten 19 Untersuchungsseiten beenden mehrere Teilnehmer die Untersuchung. Abbildung 10 zeigt, dass 415 Personen den Anfangsfragebogen ausfüllten.

<sup>23</sup> Nedstat (<http://www.nedstat.com/>) ist ein erweiterter Besucherzähler, der auf Webseiten eingebunden werden kann. Er wird auf der ersten Seite der Untersuchung verwendet, auf der die randomisierte Zuweisung zu einer der Experimentalgruppen und die automatische Weiterleitung zur ersten Untersuchungsseite geschieht. Außerdem wird er auf jeweils der ersten Untersuchungsseite eingesetzt. Die so ermittelten Besucherzahlen entsprechen einer max. Schätzung von potenziellen Teilnehmern, wobei jeder Seitenabruf vom Server gezählt wird. Die tatsächliche Besucherzahl ist vermutlich niedriger, kann jedoch aufgrund der Erhebungstechnik, die wiederum in der Technik des Internet begründet ist, nicht ermittelt werden.



Am Ende der Untersuchung hatten 221 Personen alle Seiten besucht. Um auswertbare Datensätze leichter identifizieren zu können, wurden Teilnehmer aufgefordert, das Item „Ich habe ernsthaft an der Untersuchung teilgenommen“ anzukreuzen. Dieses Item hat eine sehr gute Validität, da Personen, die den Fragebogen ernsthaft bearbeiten, auch dieses Item korrekt beantworteten, wohingegen „Durchklicker“ (Lurker, vgl. Bosnjak & Tuten, 2001) sich kaum die Mühe machen, hier eine Falscheingabe zu erzeugen. 28 leere und nicht aussagekräftige Daten wurden damit erfolgreich ausgefiltert. Zudem gab es aufgrund ethischer Überlegungen die Möglichkeit, die Daten von der Untersuchung auszuschließen, was 2 Personen in Anspruch nahmen.

Systematische Ausfälle können zwei Ursachen haben. Erstens können Ausfälle durch Unterschiede zwischen den Personen entstehen. Dies kann teilweise dahingehend geprüft werden, ob sich die Angaben im Anfangsfragebogen der Abbrecher (224 Personen) von denen unterscheiden, die in die Auswertung der Untersuchung eingehen (191 Personen). Ein systematischer Ausfall von Teilnehmern wird durch Mittelwertvergleiche der beiden Gruppen (Abbrecher, tatsächliche Teilnehmer) geprüft. Abhängige Variablen der eingesetzten t-Tests sind die Variablen zum Alter, dem Geschlecht, der Computererfahrung, der Interneterfahrung, dem Suchmaschinenwissen, dem Wissen über den Euro und dem Wirtschaftswissen. Hierbei zeigt sich bei berücksichtigter Bonferoni-Korrektur lediglich ein signifikanter Unterschied in der Selbsteinschätzung zum Suchmaschinenwissen (siehe Anhang C). Demnach unterscheiden sich die tatsächlichen Teilnehmer von den Abbrechern lediglich dadurch, dass Teilnehmer ein höheres Suchmaschinenwissen angeben ( $\bar{x}_{Teilnehmer} = 3,62$ ;  $\bar{x}_{Abbrecher} = 3,21$ ;  $F=3,30$ ;  $p=0,001$ ). Ausfälle bedingt durch Unterschiede zwischen den Personen betreffen die Generalisierbarkeit der Ergebnisse (vgl. Kap. 0)

Zweitens können Ausfälle durch die Bedingungen in den Faktorstufen verursacht werden. Dies würde die Aussagekraft von monokausalen Zusammenhängen der Faktorstufen mit den abhängigen Variablen einschränken. Um dies zu prüfen, wird ein  $\chi^2$ -Test über die Häufigkeiten in den einzelnen Faktorstufen durchgeführt (Tabelle 6). Der  $\chi^2$ -Test zeigt keine signifikanten Unterschiede in den Häufigkeiten ( $\chi_7 = 4,991$ ;  $p=0,661$ ). Daraus lassen sich mehrere Schlussfolgerungen ziehen: Ein systematischer Ausfall in einzelnen Faktorstufen (zwischen Abbrecher und Teilnehmer) kann nicht beobachtet werden. Ebenso existieren keine Unterschiede in den Häufigkeiten zwischen den einzelnen Faktorstufen. Die Zufallszuweisung zu den einzelnen Faktorbedingungen kann damit als gelungen betrachtet werden. Dadurch ist die interne Validität in dem Sinne, dass Schlussfolgerungen aufgrund der Faktorbedingungen auf die abhängigen Variablen möglich sind, nicht gefährdet (Cook & Campbell, 1979). Außerdem kann durch die gelungene Zufallszuweisung davon ausgegangen werden, dass eine Voraussetzung der Varianzanalyse (voneinander unabhängige Fehlerkomponenten) erfüllt ist (Bortz, 1999).

**Tabelle 6:** Häufigkeiten in den einzelnen Faktorstufen und zwischen Teilnehmern und Abbrechern.

<b>Experimentelle Bedingung</b>	<b>Abbrecher</b>	<b>Teilnehmer</b>	<b>Summe</b>
Kontext, spezifisch	39	29	68
Kontext, unspezifisch	28	29	57
Textanfang, spezifisch	23	26	49
Textanfang, unspezifisch	25	25	50
Schlüsselwörter, spezifisch	26	20	46
Schlüsselwörter, unspezifisch	25	18	43
MaxInfo, spezifisch	41	25	66
MaxInfo, unspezifisch	17	19	36
Summe	224	191	415

#### **4.1.4 Manipulationsüberprüfung (Treatmentcheck)**

Ein Manipulationscheck ist sinnvoll, wenn überprüft werden soll, ob die Operationalisierung der Faktorstufen, die Einfluss auf die Teilnehmer nehmen, gelungen ist. Die Spezifität der Aufgabe wird durch zwei unterschiedliche Suchaufgaben operationalisiert. Um überprüfen zu können, ob spezifische Suchaufgaben tatsächlich als einfacher und leichter empfunden werden, findet eine Bewertung der Suchaufgabe nach den Ergebnisseiten statt (S9 und S14, I1-3). Mittels einer Varianzanalyse mit den Items und der Spezifität als zwei Messwiederholungsfaktoren wird geprüft, ob die Manipulation im Sinne eines ‚Treatments‘ gelungen ist. Das Ergebnis zeigt, dass die Konstruktion der Suchaufgaben als gelungen betrachtet werden kann. Der erwünschte Haupteffekt der Spezifität ist höchst signifikant ( $F=51,60$ ;  $df=1$ ;  $p=0,00$ ).

#### **4.1.5 Reihenfolgeeffekte**

Da Teilnehmer zwei Suchaufgaben bearbeiten, ist es denkbar, dass sich die Reihenfolge, in der die Suchaufgaben bearbeitet werden, auf die Ergebnisse auswirkt. Solche unerwünschten Effekte können zum Beispiel durch Ermüdung entstehen. Um dem entgegenzuwirken, werden Teilnehmer zufällig einer der beiden Aufgabenabfolgen zugewiesen (spezifisch-unspezifisch vs. unspezifisch-spezifisch). Mittels einer Varianzanalyse mit dem unabhängigen Faktor Reihenfolge und dem Messwiederholungsfaktor Kriterien (vgl. Kap. 3.1.6) wird ein systematischer Reihenfolgeeffekt für die verschiedenen Kriterien überprüft. Das Ergebnis ist nicht signifikant ( $F=1,807$ ;  $df=1$ ;  $p=0,182$ ). Ein Reihenfolgeeffekt und die damit verbundene systematische Verzerrung der Ergebnisse kann daher ausgeschlossen werden.

#### **4.1.6 Probleme in den Übungsdurchgängen**

Die Übungsaufgaben sollten der Messung der motorischen Zeitkomponente bei vernachlässigbaren Entscheidungszeiten dienen (S3-4, vgl. Kap. 3.1.3.3). Dies gelang nicht. Tabelle 7 zeigt, dass der für einfach gehaltene Instruktionstext „Diesen Link auswählen“ keineswegs eindeutig ist. Erstens ist die hohe Zahl an Teilnahmeverweigerern deutlich zu

erkennen (11,67% fehlende Angaben) und zweitens unterscheiden sich sämtliche Selektionsmuster von den durch den Instruktionstest erwarteten. Der zur Überprüfung eingesetzte t-Test bei einer Stichprobe mit erwarteten Mittelwerten entsprechend dem Instruktionstext wird für alle Durchgänge höchst signifikant. Die Übungsaufgaben werden aus diesem Grund von der weiteren Auswertung ausgeschlossen.

**Tabelle 7:** Korrektheit der Angaben in den Übungsdurchgängen zur Selektion.

Beispiel- treffer Nr.	Anzahl ausgewählt	Anzahl nicht ausgewählt	Instruktions- ziel	T	df	p	N
1	162	13	Auswählen	3,737	174	0,000	175
2	147	29	Auswählen	5,876	175	0,000	176
3	24	136	nicht Ausw.	-5,297	159	0,000	160
4	20	138	nicht Ausw.	-4,770	157	0,000	158
5	141	29	Auswählen	5,896	169	0,000	170
6	18	142	nicht Ausw.	-4,489	159	0,000	160
7	147	26	Auswählen	5,516	172	0,000	173
8	27	139	nicht Ausw.	-5,661	165	0,000	166
9	157	21	Auswählen	4,866	177	0,000	178
10	20	145	nicht Ausw.	-4,756	164	0,000	165
11	150	27	Auswählen	5,628	176	0,000	177
12	16	147	nicht Ausw.	-4,199	162	0,000	163
13	14	150	nicht Ausw.	-3,900	163	0,000	164
14	148	29	Auswählen	5,873	176	0,000	177

Das Scheitern der Operationalisierung kann mehrere Ursachen haben, wobei der Interpretation hier vorgegriffen werden soll: Einerseits scheint der Zweck der Übungen trotz eigener Instruktionseite (S3) nicht deutlich genug. Darauf deuten Anmerkungen (S18, I6) hin, in denen Teilnehmer die Übungsdurchgänge als „verwirrend“ bezeichnen und anmerken, „der Sinn [sei] nicht auszumachen“. Außerdem ist es möglich, dass die Übungen für zu einfach gehalten wurden und dadurch in paradoxer Weise die Teilnehmer von einer Teilnahme abhielten. Dazu schrieb eine Teilnehmerin: „Bei den Linkübungen (Entscheidung ja/nein) waren leider keine Links zu lesen. Alles andere war aber vollständig und leicht verständlich und deshalb auch ernsthaft bearbeitet.“

## 4.2 Beschreibung der Stichprobe

Von April bis Juni 2002 konnte online an der etwa zwanzigminütigen Untersuchung teilgenommen werden. In die Auswertung gingen 191 Fälle ein. Die Teilnehmer waren 127 Männer (67,5%) und 61 Frauen (32,4%; drei ohne Angabe). Das Durchschnittsalter betrug etwa 31 Jahre. Die meisten Teilnehmer wurden durch einen Link von anderen Seiten (72 Personen), durch E-Mail (55) oder eine Suchmaschine (22) auf die Untersuchung aufmerksam (Abbildung 11). Es handelt sich um Anwender, die fast täglich mit Computern arbeiten. Durchschnittlich beschäftigen sie sich sechs Tage in der Woche mit Computern und

sind fast sechs Tage in der Woche im Internet. Ihr Wissen über Suchmaschinen wird von ihnen auf einer fünf-stufigen Skala leicht überdurchschnittlich, ihr Wissen über die Währungsstellung zum Euro und ihr Wirtschaftswissen durchschnittlich eingeschätzt. Die meisten Teilnehmer verwendeten den Internetexplorer von Microsoft und hatten eine Bildschirmauflösung von 1024x768 Bildpunkten. Einzelheiten der Stichprobe zeigen Tabelle 8 und Tabelle 9.

Obwohl die Maximalwerte bei der Anzahl durchgeführter Suchanfragen pro Woche recht hoch sind, werden die Teilnehmer nicht von der Auswertung ausgeschlossen, da die abhängigen Variablen trotzdem plausible Werte liefern. Dies gilt ebenso für den Teilnehmer im Alter von 78 und den Teilnehmer, der fast vier Stunden für die Untersuchung benötigte. Die relevanten Zeitmessungen auf den Ergebnisseiten unterscheiden sich nicht von denen anderer Teilnehmer.

**Tabelle 8:** Beschreibung der Stichprobe.

<b>Variable</b>	<b>N</b>	<b>Min</b>	<b>Max</b>	<b>M</b>	<b>SD</b>
Anzahl verwendete Suchmaschinen pro Woche	183	0	30	2,72	3,61
Dauer der Suchmaschinennutzung pro Woche in Stunden	160	0	40	3,81	5,20
Anzahl der gestellten Suchanfragen in einer Woche	161	0	500	43,75	70,11
Davon einfache Suchanfragen	153	0	250	15,71	29,22
Davon komplexe Suchanfragen	148	0	480	24,80	49,22
Anzahl Tage mit Computernutzung in einer Woche	189	2	7	6,17	1,13
Anzahl Tage mit Internetnutzung in einer Woche	189	1	7	5,72	1,45
Geschätztes Wissen über Suchmaschinen	187	1	5	3,64	0,91
Geschätztes Wissen über den Euro	185	1	5	3,37	0,93
Geschätztes Wirtschaftswissen	188	1	5	3,07	0,99
Alter	186	17	78	30,90	10,16
Benötigte Gesamtzeit (Minuten, Sekunden)	187	4,39	231	20,21	18,49

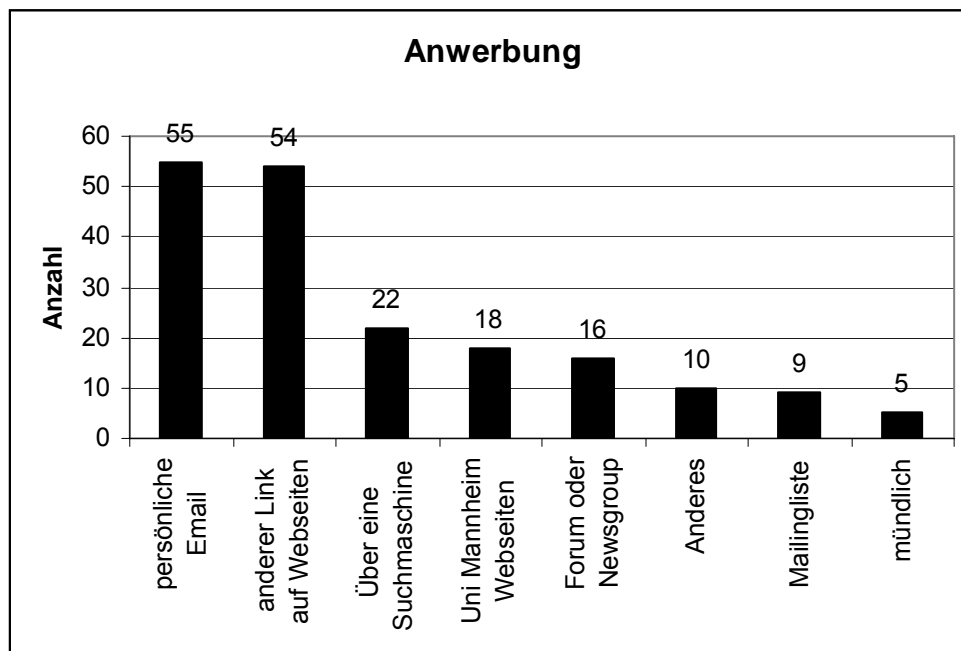
**Tabelle 9:** Verwendete Bildschirmauflösung und Browser.

<b>Auflösung</b>	<b>Häufigkeit</b>	<b>Prozent</b>	<b>Browser</b>	<b>Häufigkeit</b>	<b>Prozent</b>
Höhere	43	22,5	MS Explorer	158	82,7
640	1	0,5	Netscape	25	13,1
800	32	16,8	Andere	8	4,2
1024	115	60,2			
Gesamt	191	100,0	Gesamt	191	100,0

Die Verteilung der männlichen und weiblichen Teilnehmer entspricht recht genau der von der Academic Data (1997) berichteten Verteilung der Internet-Nutzer von 68% Männern und 32% Frauen (zitiert nach Hauptmanns, 1999). Das Durchschnittsalter der Teilnehmer liegt jedoch einige Jahre unter dem erwarteten Wert von 35,7 Jahren. Ähnliche Werte für die Internetpopulation finden sich bei Batinic und Bosnjak (2000), sowie bei Pitkow und Kehoe (1996)

Um möglichst viele Teilnehmer zu gewinnen, wurde auf verschiedenen Wegen auf die Untersuchung aufmerksam gemacht. In Anlehnung an Vogt (1999) wird zwischen potenziellen Teilnehmern (ungefähr 2660 Personen<sup>24</sup>), die die erste Seite der Untersuchung aufgerufen haben, und tatsächlichen Teilnehmer (191 Personen), die an der Untersuchung bis zum Abschluss teilnahmen und auswertbare Daten erzeugten, unterschieden. Am Ende der Untersuchung wurde erhoben, wie die Teilnehmer auf die Untersuchung aufmerksam geworden waren. Zwischen den Rekrutierungswegen kann daher nur für tatsächliche Teilnehmer differenziert werden (Abbildung 11).

Der Freundeskreis des Versuchsleiters wurde mit Emails (Anhang E) und mündlich gebeten, an der Untersuchung teilzunehmen und anderen von der Untersuchung zu berichten (55 tatsächliche Teilnehmer). Mit einem Service des Domainnamenverwalters Network Information Center<sup>25</sup> wurde die Startseite zur Indexierung bei mehreren Suchmaschinen angemeldet. Später konnte jedoch festgestellt werden, dass von den angekündigten 250 Suchmaschinen nur 16 eine erfolgreiche Anmeldung zurückmeldeten, wovon wiederum schließlich nur Lycos die Untersuchungsseite erfolgreich indexiert hatte.



**Abbildung 11:** Anwerbung.

Außerdem erfolgten auch manuelle Eintragungen in die laut Suchfibel<sup>26</sup> wichtigsten Suchmaschinen (Anhang F). Dabei wurde besonders auf die Besonderheiten von Suchmaschinen für erfolgreiche Indexierung, wie zum Beispiel die Angabe von Metainformationen (u.a. Schlüsselwörter), Rücksicht genommen (vgl. Karzauninkat, 1998a).

<sup>24</sup> Diese Schätzung geschieht anhand der Seitenabrufe der ersten Untersuchungsseite vom Server.

<sup>25</sup> <http://www.nic.de.vu/>

<sup>26</sup> <http://suchfibel.de/>

Diese Anmeldungen waren erfolgreich. Mit den Suchmaschinen Google<sup>27</sup> und Fireball<sup>28</sup> wurden Webseitenanbieter identifiziert, die ein Interesse am Thema haben und daher als kostenlose Werbeträger in Frage kamen. Solche Links von anderen Webseiten (vgl. Anhang F) brachten 54 tatsächliche Teilnehmer. Andere Teilnehmer erfuhren über Suchmaschinen, durch Foren, Newsgroups, Mailinglisten, mündlich oder auf nicht näher angegebene Weise von der Untersuchung. Für eine Teilnahme gab es keine materielle Belohnung.

### 4.3 Effekte in der Ebene Maschine und Aufgabe

Erläuterungen zu den verwendeten Auswertungsverfahren wurden in Kapitel 3.3 behandelt, weshalb sich die Beschreibung der Verfahren im Folgenden auf eine einfache Nennung des eingesetzten Verfahrens beschränkt. Da die Fragestellungen zu den Ebenen Maschine und Aufgabe jeweils im selben Testverfahren untersucht werden, werden sie im gleichen Abschnitt behandelt. Eine Tabelle gibt jeweils die Test-Statistik wider, eine Abbildung veranschaulicht anhand der Mittelwerte die Richtung der Effekte, um eine Interpretation zu erleichtern. Aufgrund der hohen Zahl an Untersuchungsergebnissen wird die Interpretation jeweils bei den berichteten Ergebnissen statt in einem eigenen Abschnitt vorgestellt.

#### 4.3.1 Genauigkeit

Zur Auswertung der Genauigkeit wurde eine mixed ANOVA verwendet. Hierbei ist lediglich der Hauptfaktor Spezifität höchst signifikant (Tabelle 10). Aus Abbildung 12 lässt sich erkennen, dass unspezifische Aufgaben zu einer geringeren Genauigkeit führen als spezifische Aufgaben. Dies entspricht dem erwarteten Zusammenhang. Unterschiede in den Beschreibungsvarianten lassen sich nicht feststellen.

Tabelle 10: Genauigkeit.

Effekt	F	df	Signifikanz
Beschreibung	0,584	3	0,626
Spezifität	16,941	1	0,000***
Beschreibung x Spezifität	2,187	3	0,091

<sup>27</sup> <http://www.google.de/>

<sup>28</sup> <http://fireball.de/>

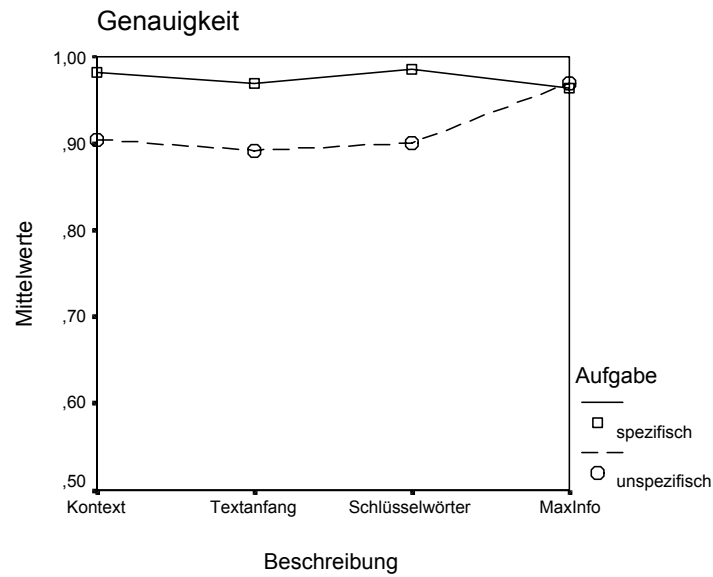


Abbildung 12: Genauigkeit hinsichtlich Beschreibung und Spezifität.

#### 4.3.2 Vollständigkeit

Tabelle 11: Vollständigkeit.

Effekt	F	df	Signifikanz
Beschreibung	1,153	3	0,329
Spezifität	2,510	1	0,115
Beschreibung x Spezifität	1,083	3	0,358

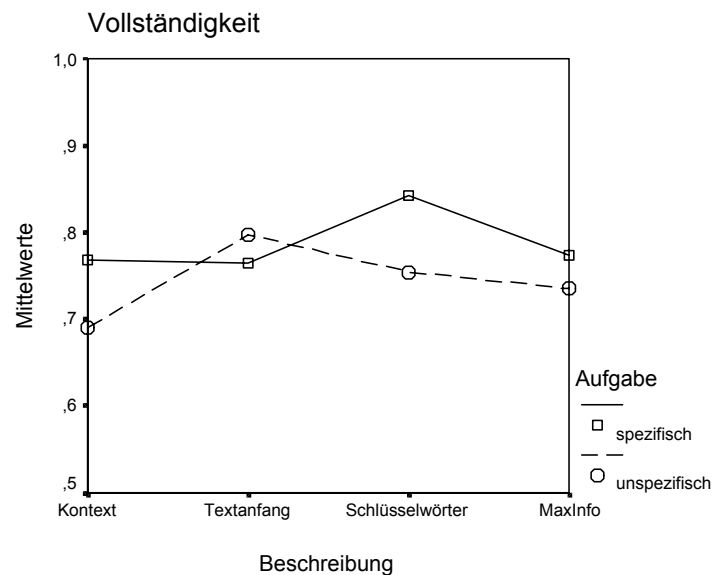


Abbildung 13: Vollständigkeit hinsichtlich Beschreibung und Spezifität.

Zur Auswertung der Vollständigkeit wurde eine mixed ANOVA verwendet. Hierbei sind keine signifikanten Effekte zu beobachten (Tabelle 11). Die in Abbildung 13 angedeuteten Unterschiede erhalten dadurch keinerlei inhaltliche Bedeutsamkeit.

### 4.3.3 Erfolgsrate

Zur Auswertung der Erfolgsrate wurde eine mixed ANOVA verwendet. Hierbei sind die Hauptfaktoren Spezifität und Beschreibung höchst signifikant (Tabelle 12). Post-hoc-Mehrfachvergleiche zeigen, dass sich die Bedingung Schlüsselwörter signifikant von den beiden Bedingungen Kontext und MaxInfo unterscheidet ( $p=0,31$ ;  $p=0,000$ ). Weiterhin ist der Unterschied zwischen den Bedingungen Textanfang und MaxInfo signifikant ( $p=0,005$ ). Aus Abbildung 14 lässt sich erkennen, dass eine Beschreibung der Treffer mit Schlüsselwörtern den Beschreibungsvarianten Kontext und MaxInfo überlegen ist. Außerdem ist eine Beschreibung durch Textanfang der Konglomeratvariante MaxInfo überlegen. Spezifische Suchaufgaben ergeben eine höhere Erfolgsrate als unspezifische Suchaufgaben.

Tabelle 12: Erfolgsrate.

Effekt	F	df	Signifikanz
Beschreibung	7,113	3	0,000***
Spezifität	12,490	1	0,001***
Beschreibung x Spezifität	0,472	3	0,702

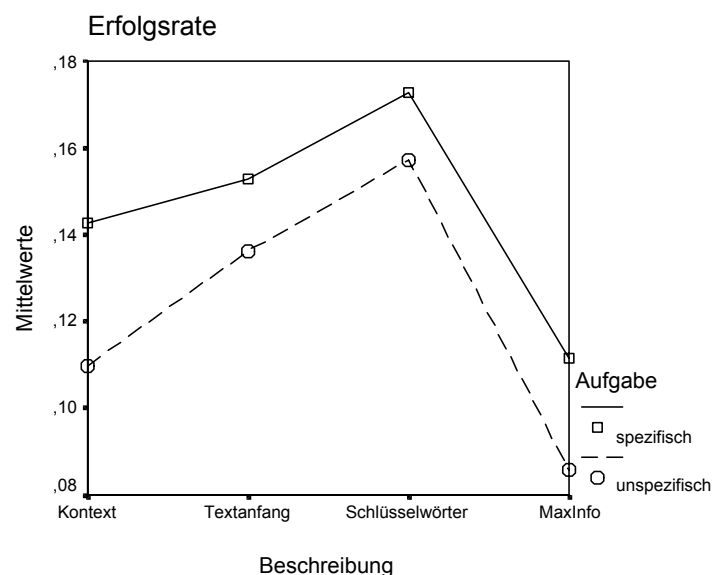


Abbildung 14: Erfolgsrate hinsichtlich Beschreibung und Spezifität.

### 4.3.4 Entscheidungszeit

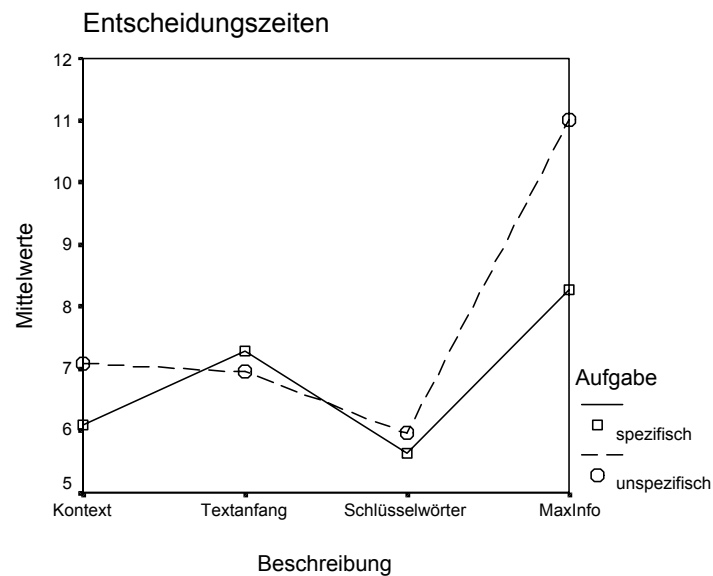
Zur Auswertung der Entscheidungszeit wurde eine mixed ANOVA verwendet. Hierbei ist der Faktor Spezifität signifikant und der Faktor Beschreibung höchst signifikant (Tabelle 13). Post-hoc-Mehrfachvergleiche zeigen, dass sich die Bedingung MaxInfo signifikant von den



anderen drei Bedingungen Kontext, Textanfang und Schlüsselwörter unterscheidet ( $p=,001$ ;  $p=0,014$ ;  $p=0,000$ ). Aus Abbildung 15 lässt sich erkennen, dass Anwender bei einem Konglomerat aller drei anderen Bedingungen (MaxInfo) deutlich mehr Zeit für ihre Entscheidungen benötigen als für jede der drei Beschreibungsvarianten allein. Spezifische Suchaufgaben erlauben schnellere Entscheidungen als unspezifische Suchaufgaben. Dieser Effekt beruht offenbar auf schnelleren Entscheidungen bei spezifischen Aufgaben in den Bedingungen Kontext und MaxInfo.

**Tabelle 13:** Entscheidungszeiten.

Effekt	F	df	Signifikanz
Beschreibung	7,078	3	0,000***
Spezifität	4,392	1	0,038*
Beschreibung x Spezifität	2,183	3	0,092



**Abbildung 15:** Entscheidungszeiten hinsichtlich Beschreibung und Spezifität in Sekunden pro Webseite.

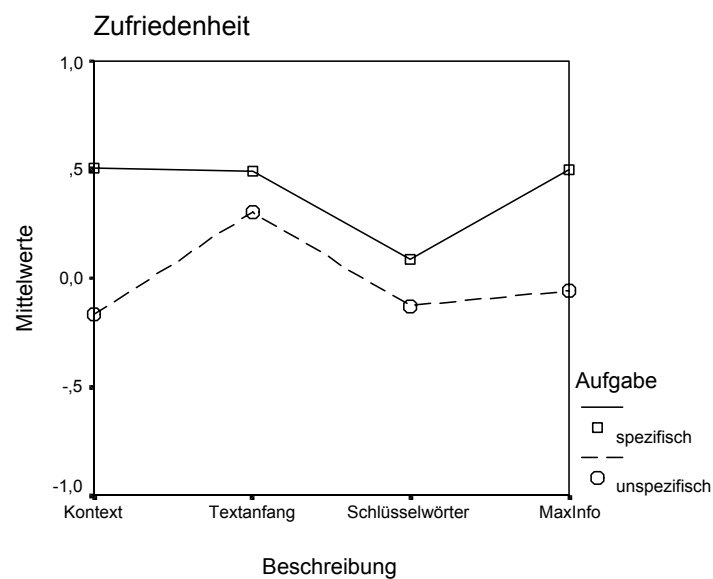
#### 4.3.5 Zufriedenheit

Bei den vierzehn Items zur Zufriedenheit kann nicht davon ausgegangen werden, dass sie den Kriterien einer Testskala genügen (Lienert & Ratz, 1994). Es ist daher streng genommen nicht zulässig, sie als Likert-Skala zu interpretieren und für eine Auswertung zu aggregieren. Aus diesem Grund geht jedes der vierzehn Items als eigene Faktorstufe im Sinne eines zweiten within-subjects Faktors in die Auswertung ein (14-stufig). Der erste within-subjects Faktor ist dabei die Spezifität (2-stufig) und der between-subjects Faktor die Beschreibung (4-stufig). Durch die nun insgesamt 28 verschiedenen within-subjects Bedingungen ist die Sphäritätsannahme verletzt: Die Inter-Item-Korrelationen sind nicht wie gefordert über alle

Items gleich. Daher wird zur Auswertung der Zufriedenheit der multivariate Ansatz verwendet (vgl. Kap. 3.3). Berichtet werden nur die interessierenden Effekte. Der Faktor Spezifität erweist sich als höchst signifikant, der Faktor Beschreibung als sehr signifikant (Tabelle 14). Auch der Interaktionseffekt zwischen Beschreibung und Spezifität ist höchst signifikant. Post-hoc-Mehrfachvergleiche zeigen, dass sich die Bedingung Schlüsselwörter von den Bedingungen Kontext ( $p=0,003$ ) und Textanfang ( $p=0,003$ ) signifikant unterscheidet.

**Tabelle 14:** Zufriedenheit. Multivariater Test mit Wilks  $\lambda$ .

Effekt	F	df 1	df 2	Signifikanz
Beschreibung	3,987	3	---	0,009**
Spezifität	61,004	1	154,000	0,000***
Beschreibung x Spezifität	5,850	3	154,000	0,001***



**Abbildung 16:** Zufriedenheit hinsichtlich Beschreibung und Spezifität.

Bei einer Interpretation muss jedoch der Interaktionseffekt mit berücksichtigt werden. Aus Abbildung 16 lässt sich erkennen, dass bei spezifischer Aufgabe die Beschreibungsvarianten Kontext und Textanfang der Beschreibungen mit Schlüsselwörtern überlegen sind. Bei unspezifischer Aufgabe ist lediglich eine Beschreibung mit Textanfang den Schlüsselwörtern überlegen. Insgesamt zeigt sich, dass Anwender mit Textanfang zufriedener sind als mit Schlüsselwörtern. Weiterhin eindeutig ist, dass Ergebnisseiten bei spezifischen Suchaufgaben zu einer höheren Zufriedenheit führen als bei unspezifischen Aufgaben. Welche Items im einzelnen für die Effekte verantwortlich sind, wird in einem eigenen Kapitel zu klären sein (siehe Kap. 4.5).

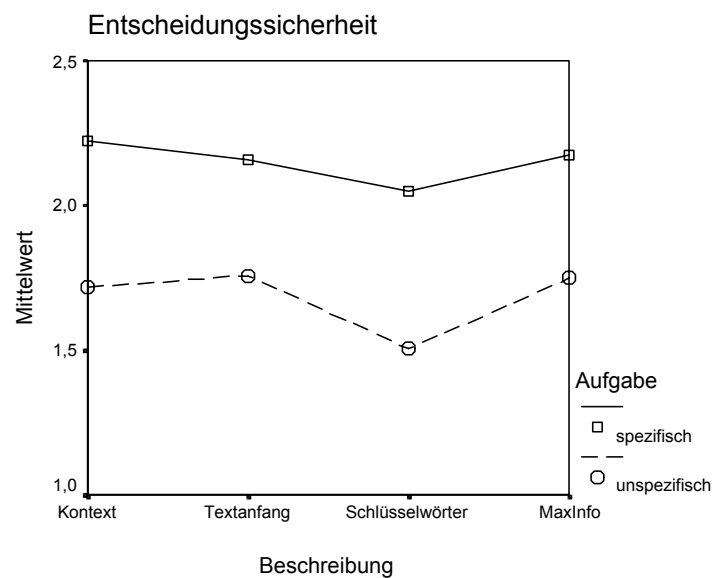
#### 4.3.6 Entscheidungssicherheit

Zur Auswertung der Entscheidungssicherheit wurde eine mixed ANOVA verwendet. Hierbei ist der Faktor Spezifität höchst signifikant und der Faktor Beschreibung signifikant (Tabelle 15). Post-hoc-Mehrfachvergleiche zeigen, dass der Unterschied zwischen den

Bedingungen Kontext und Schlüsselwörter die Signifikanzgrenze nur fast erreicht ( $p=0,052$ ). Aus Abbildung 17 lässt sich erkennen, dass Anwender der Tendenz nach bei Schlüsselwörtern in ihrer Selektion am unsichersten sind. Deutlich zeigt sich jedoch, dass spezifische Suchaufgaben sichere Entscheidungen erlauben als unspezifische Aufgaben.

**Tabelle 15:** Entscheidungssicherheit.

Effekt	F	df	Signifikanz
Beschreibung	2,772	3	0,043
Spezifität	157,438	1	0,000
Beschreibung x Spezifität	0,809	3	0,490



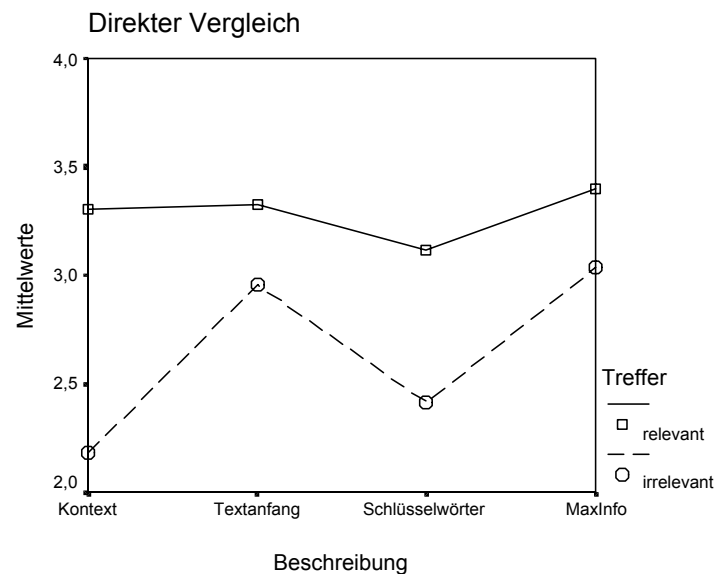
**Abbildung 17:** Entscheidungssicherheit hinsichtlich Beschreibung und Spezifität.

#### 4.3.7 Direkter Vergleich

Die vier verschiedenen Beschreibungsvarianten werden von Teilnehmern je einmal mit relevanten Treffern und mit irrelevanten Treffern miteinander verglichen („Wie gut oder schlecht ist die jeweilige Darstellung des Links?“). Diese beiden Varianten gehen daher als within-subjects Faktor ‚Treffer‘ in die Auswertung ein (2-stufig). Da alle vier Beschreibungsvarianten im direkten Vergleich von allen Teilnehmern bewertet werden, ist auch der Faktor Beschreibung in diesem Fall ein within-subjects Faktor (4-stufig). Aufgrund der sich daraus ergebenden acht within-subjects Bedingungen mit ihren verschiedenen Interkorrelationen ist die Spherizitätsannahme verletzt. Daher wird zur Auswertung des direkten Vergleichs ein multivariater Ansatz verwendet (vgl. Kap. 3.3).

**Tabelle 16:** Direkter Vergleich aller vier Beschreibungsvarianten. Multivariater Test mit Wilks  $\lambda$ .

Effekt	F	df 1	df 2	Signifikanz
Beschreibung	12,414	3	176	0,000
Treffer	67,669	1	178	0,000
Beschreibung x Treffer	10,160	3	176	0,000

**Abbildung 18:** Direkter Vergleich aller vier Beschreibungsvarianten getrennt nach Trefferrelevanz.

Die beiden Haupteffekte und der Interaktionseffekt zwischen Beschreibung und Trefferrelevanz sind höchst signifikant (Tabelle 16). Weil es sich beim Faktor Beschreibung in diesem Fall um einen within-subjects Faktor handelt, stehen keine Post-hoc-Mehrfachvergleiche zur Verfügung. Die im Folgenden interpretierten Unterschiede in der Beschreibung müssen daher mit Vorsicht behandelt werden. Eindeutig ist zu erkennen, dass relevante Treffer in jeder Bedingung eine bessere Bewertung erhalten als irrelevante Treffer (Abbildung 18). Dies kann inhaltlich damit erklärt werden, dass Anwender solche Treffer präferieren, die sie für relevant halten. Aufgrund des Interaktionseffekts ist die Interpretation des Effekts im Faktor Beschreibung weniger eindeutig. Hier sind am ehesten die Unterschiede bei irrelevanten Treffern bedeutsam. Die Beschreibungsvarianten Textanfang und MaxInfo werden bei irrelevanten Treffern gegenüber den Varianten Kontext und Schlüsselwörter präferiert. Dies würde darauf hindeuten, dass für eine Beurteilung von irrelevanten Treffern der Textanfang der repräsentierten Dokumente am geeignetsten ist.

#### 4.4 Moderierender Einfluss der Ebene Mensch

Für jede der in Kapitel 4.3 beschriebenen durchgeführten Auswertungen wurden die verschiedenen Variablen der Ebene Mensch einzeln als Kovariate geprüft, um ihren moderierenden Einfluss einschätzen zu können. Entgegen den Erwartungen zeigten sich keine weiteren Effekte durch die Hinzunahme moderierender Variablen. Die geprüften Variablen sind in Tabelle 8 (Seite 70) zur Stichprobenbeschreibung dargestellt (die Zeitdauer

wurde nicht als Kovariate eingesetzt). Auch Korrelationen zwischen den Variablen der Ebene Mensch und den abhängigen Variablen in Kapitel 4.3 sind weitgehend nicht signifikant (siehe Anhang B). Einzige Ausnahme bildet der Zusammenhang zwischen Alter der Teilnehmer und der Entscheidungssicherheit bei spezifischer Suchanfrage ( $r=-0.292$ ). Demnach geben jüngere Teilnehmer bei spezifischer Suchanfrage für ihre Selektionen eine höhere Entscheidungssicherheit an als ältere Teilnehmer. Die gefundene Korrelation ist jedoch gering und der Zusammenhang sehr speziell, so dass eine weitergehende Interpretation dieses Effektes nicht sinnvoll ist.

Dies kann zwei mögliche Ursachen haben. Einerseits ist es möglich, dass die Variablen in der Ebene Mensch einer zu geringen Streuung unterliegen. Je geringer die Varianz einer Variablen, desto geringer fällt auch der Zusammenhang mit anderen Variablen aus. Khan und Locatis (1998b) definieren Such-Experten danach, ob sie mehr als fünf Stunden in der Woche mit dem Browsen im Internet verbringen. Demnach würde die vorliegende Stichprobe fast ausschließlich aus Experten bestehen und wenig Raum für Zusammenhänge bieten, die sich aus Unterschieden zwischen erfahrenen und weniger erfahrenen Teilnehmern ergeben könnten. Andererseits ist es auch denkbar, dass das Untersuchungsdesign selbst so angelegt ist, dass moderierende Effekte nicht zum Tragen kommen. Dies wäre ein Zeichen für die Güte des Designs. Durch die Tatsache, dass nicht das gesamte Suchverhalten, sondern das Verhalten auf Ergebnisseiten untersucht wird, konnten viele Varianzquellen, die besonders durch die Ebene Mensch beeinflusst sind, schon im Vorfeld ausgeschaltet werden. Diese betreffen vor allem eine Vielzahl an kognitiven Problemlöseprozessen, die ansonsten während des gesamten Suchprozesses ablaufen (Fang & Salvendy, 2000).

## 4.5 Skalenbildung Zufriedenheit

Bei vierzehn verwendeten Items zur Zufriedenheit und signifikanten Unterschieden in den Beschreibungsvarianten (siehe Kap. 4.3.5) stellt sich die Frage, welche Items zur Differenzierung dieser Varianten beitragen und inwiefern sich dadurch eine Testskala mit weniger Items konstruieren lässt. Das Kriterium für eine solche Skala ist dementsprechend die Differenzierungsfähigkeit zwischen den verschiedenen Bedingungen des Faktors Beschreibung. Zur Identifizierung geeigneter Items werden daher Varianzanalysen mit dem Messwiederholungsfaktor Aufgabe für jedes Item durchgeführt. Items eignen sich für eine Skalenbildung, wenn diese eine signifikante Differenzierung innerhalb des Faktors Beschreibung ohne Interaktionseffekt mit der Aufgabenspezifität erlauben. Items mit Interaktionseffekt sollen nicht zur Skalenbildung verwendet werden, da diese eine Interpretation der Haupteffekte erschweren. Die Durchführung mehrerer solcher Tests auf denselben Effekt hin führt zu einer  $\alpha$ -Fehler-Kumulierung. Aus diesem Grund wird die Bonferoni-Korrektur nach Formel 5 (Seite 57) eingesetzt (Bortz, 1999). Effekte sind demnach erst ab einer Irrtumswahrscheinlichkeit von gleich oder unter 0,00357 signifikant. Aus Tabelle 17 ergibt sich die Eignung der vier Items 8, 9, 10, 11.

Ohne Bonferoni-Korrektur erscheinen auch die Items 2, 4 und 6 geeignet. Eine Plausibilitätsüberprüfung der Itemtexte zeigt, dass solche Items ungeeignet sind, die sich weder direkt auf die Selektionsentscheidungen (4, 10, 11) noch auf das Verständnis der Texte (2, 8, 9) beziehen. Eine Ausnahme stellt Item 6 dar, das direkt nach der Zufriedenheit fragt.

Soll die Skala einen gemeinsamen Faktor repräsentieren, so ist eine hohe Homogenität zwischen den Items wünschenswert. Die interne Konsistenz wird mit Hilfe von Cronbachs  $\alpha$  geprüft. Der sich hierbei ergebende Wert von Cronbachs  $\alpha=0,87$  kann als gut bezeichnet werden. Darauf aufbauend kann eine aggregierte Skala Zufriedenheit aus den sieben Items gebildet werden. In diesem Zusammenhang muss jedoch erwähnt werden, dass weitere Untersuchungen notwendig sind, um die Eignung der hier identifizierten Items zu validieren.

**Tabelle 17:** Differenzierungsfähigkeit der Items zur Zufriedenheit.

Effekt Item	Beschreibung			Spezifität			Interaktion			Eig- nung
	F	df	Sig.	F	df	Sig.	F	df	Sig.	
1	0,377	3	0,770	0,56	1	0,457	3,250	3	0,023	
2	2,653	3	0,050	33,06	1	0,000	1,995	3	0,116	X
3	0,951	3	0,417	12,75	1	0,000	2,242	3	0,085	
4	2,935	3	0,035	29,45	1	0,000	2,737	3	0,045	X
5	1,132	3	0,338	50,29	1	0,000	4,291	3	0,006	
6	3,649	3	0,014	46,68	1	0,000	4,353	3	0,005	X
7	1,752	3	0,158	1,26	1	0,262	0,959	3	0,413	
8	5,583	3	0,001	42,80	1	0,000	1,456	3	0,228	X
9	10,89	3	0,000	48,58	1	0,000	4,579	3	0,004	X
10	4,001	3	0,009	60,02	1	0,000	3,845	3	0,011	X
11	4,983	3	0,002	32,38	1	0,000	3,721	3	0,013	X
12	2,966	3	0,033	12,08	1	0,001	3,152	3	0,026	
13	13,460	3	0,000	5,49	1	0,020	7,564	3	0,000	
14	1,482	3	0,221	28,71	1	0,000	2,921	3	0,035	

## 4.6 Zusammenfassende Interpretation

In diesem Abschnitt werden die Einzelergebnisse zu zwei Hauptaussagen zusammengefasst. Zu Beginn werden Unterschiede in den Beschreibungsvarianten gegeneinander abgewogen, im zweiten Teil Ergebnisse zur Spezifität vorgestellt.

Bei den Beschreibungsvarianten werden zunächst die verschiedenen Ergebnisse zur Effektivität und Effizienz berichtet. Anschließend wird auf die Variablen zur Anforderung der Zufriedenheit eingegangen. In einem dritten Abschnitt wird eine Synthese aus beiden Bereichen gebildet.

Für die Kriterien Genauigkeit und Vollständigkeit ergeben sich keine signifikanten Unterschiede zwischen verschiedenen Beschreibungsvarianten. Sie gehen jedoch in das bedeutsame Kriterium der Erfolgsrate ein. Vergleiche der Erfolgsrate zeigen, dass Trefferbeschreibungen mit Schlüsselwörtern den beiden Beschreibungsvarianten Wörter-im-Kontext und der Konglomeratbedingung MaxInfo überlegen sind. Weiterhin ist Textanfang ebenfalls der Variante MaxInfo überlegen. Zwischen den Varianten mit Schlüsselwörtern und Textanfang zeigen sich keine signifikanten Unterschiede. Bezogen auf die Erfolgsrate kann also geschlussfolgert werden, dass Beschreibungstexte, die sich auf Schlüsselwörter der repräsentierten Dokumente beschränken, am vielversprechendsten sind. Eine Alternative könnte jedoch auch im Beginn der repräsentierten Dokumente liegen, da sich die Bedingung Textanfang nicht signifikant von der Bedingung mit Schlüsselwörtern unterscheidet. Wenig verwundert das Ergebnisbild bei den Entscheidungszeiten, die ebenfalls in das Kriterium Erfolgsrate eingehen. In der Konglomeratbedingung MaxInfo, die alle anderen drei Bedingungen in sich vereinigt, benötigen Anwender signifikant mehr Zeit als für jede der anderen drei Bedingungen. Unterschiede in den Entscheidungszeiten der Beschreibungsvarianten Kontext, Textanfang und Schlüsselwörter können hingegen nicht festgestellt werden. Für das Effizienzmaß Erfolgsrate können demnach die Beschreibungsvarianten Textanfang und Schlüsselwörter als Favoriten gesehen werden.

Bei einer Bewertung der Ergebnisseite im Hinblick auf die Anforderung Zufriedenheit kann ein anderes Bild beobachtet werden. Beschreibungen durch Schlüsselwörter erhalten bei den Items zur Zufriedenheit die schlechteste Bewertung durch Anwender und unterliegen signifikant der Variante mit Textanfang. Im Sinne eines Interaktionseffektes ist eine Beschreibung mit Wörtern im Kontext lediglich bei spezifischer Suchaufgabe der Schlüsselwortmethode überlegen. Ähnliches ergibt sich auch bei einer Betrachtung der Entscheidungssicherheit. Zwar zeigen sich keine signifikanten Unterschiede, aber tendenziell sind sich Anwender bei Treffern, die durch Schlüsselwörter beschrieben werden, am unsichersten. Beim direkten Vergleich aller vier Beschreibungsvarianten können aufgrund eines Interaktionseffektes mit der Aufgabe am sinnvollsten die irrelevanten Treffer interpretiert werden. Textanfang wird bei irrelevanten Treffern den Beschreibungen mit Schlüsselwörtern und Wörtern im Kontext vorgezogen. MaxInfo wird ebenfalls gegenüber Schlüsselwörtern und Wörtern im Kontext präferiert. Hier kann jedoch davon ausgegangen werden, dass dies durch den enthaltenen Textanfang verursacht wird. Ein eigener Effekt aufgrund der Ansammlung aller Beschreibungsvarianten wird nicht vermutet, da die Bedingung MaxInfo nicht besser bewertet wird als der Textanfang allein. Bei relevanten Treffern zeigt sich diese Tendenz in stark abgeschwächter Form. Offenbar möchten Anwender zur Identifizierung von irrelevanten Treffern möglichst zusammenhängende Informationen, wie sie der Textanfang bietet. Dies kann damit erklärt werden, dass alle Beschreibungsvarianten genug Informationen enthalten, um einen Treffer als relevant einzuschätzen. Für die Entscheidung, dass ein Treffer jedoch irrelevant ist, bieten die Beschreibungen mit Wörtern im Kontext und Schlüsselwörtern zuwenig Hinweise. Dieses Fehlen an Hinweisen kann auch die größere Entscheidungsunsicherheit bei der Variante mit Schlüsselwörtern erklären. Da Anwender bei einer Ergebnisseite zunächst davon ausgehen,

dass ein Treffer für die Suche relevant ist, sind bei irrelevanten Treffern Beschreibungen notwendig, die eine klare Abgrenzung zum Informationsbedürfnis bieten. Zusammenfassend kann für die Anforderung zur Zufriedenheit die Beschreibung durch Textanfang als Favorit festgehalten werden. In sehr eingeschränkter Form, d.h. bei spezifischen Suchaufgaben und für relevante Treffer, ist es erlaubt, auch die Wörter im Kontext Methode zur Beschreibung ohne Nachteile gegenüber der Variante Textanfang heranzuziehen.

Bei einer genauen Betrachtung der Anforderungen zur Effizienz und Zufriedenheit fällt auf, dass Textanfang als Beschreibungsvariante in beiden Anforderungen gute Ergebnisse erzielt, wohingegen die Wörter im Kontext Methode nur sehr eingeschränkt empfohlen werden kann. Die Variante Schlüsselwörter wiederum ergibt zwar ein gutes Effizienzmaß, wird aber durch die Anwender selbst als eher schlecht bewertet. Die Zufriedenheit ist allerdings eine wesentliche Anforderung der ISO 9241-11. Bei gleichwertiger Effizienz mehrerer Beschreibungsvarianten ist es angebracht, der Meinung der Anwender im Sinne freier Entscheidungswahl den Ausschlag zu geben. Da sich die Effizienz bei Schlüsselwörtern von Textanfang nicht signifikant unterscheidet, kann die aufgeworfene Fragestellung zu Beschreibungsvarianten (siehe Kap. 3.2.1) daher wie folgt beantwortet werden: Erstens existieren beobachtbare Unterschiede zwischen den verschiedenen Beschreibungsvarianten. Zweitens wird der Entscheidungsprozess bei der Bewertung von Treffern auf Ergebnisseiten am optimalsten durch redaktionellen Anfangstext aus den repräsentierten Dokumenten unterstützt.

Sehr viel eindeutiger ist die Befundlage zur Spezifität der Suchaufgabe. Bei allen Kriterien schneiden spezifische Suchaufgaben besser ab als unspezifische. Die einzige Ausnahme bildet die Vollständigkeit, bei der keinerlei signifikante Unterschiede festgestellt werden. Spezifische Suchaufgaben erlauben demnach eine höhere Effektivität und Effizienz der Suche als unspezifische Aufgaben. Zudem fühlen sich Anwender in ihren Selektionsentscheidungen sicherer und sind mit den Ergebnisseiten zufriedener als bei unspezifischen Aufgaben. Dieses Ergebnis entspricht dem von der Theorie her erwarteten Bild (siehe Kap. 2.3.1) und den Hypothesen der Fragestellung (siehe Kap.3.2.2).

## 5 Diskussion

Begonnen wird die Diskussion mit einer Zusammenfassung der Untersuchung. Danach wird auf die Vorteile des Designs eingegangen und Verbesserungsmöglichkeiten aufgezeigt. Der dritte Abschnitt bietet einen Einblick in die Teststärke des Designs. Der Generalisierbarkeit der Ergebnisse auf die Internetpopulation und der Problematik der Selbstselektion wird in einem eigenen Abschnitt Rechnung getragen. Abgeschlossen wird die Diskussion durch einen Beitrag zur Umsetzung der Ergebnisse in die Praxis von Suchmaschinen und Suchenden. Den Abschluss bietet der Ausblick, in dem mögliche zukünftige Forschungsvorhaben skizziert werden.



## 5.1 Gesamtfazit

Ein Ausgangspunkt der vorliegenden Arbeit war der Mangel an Konzepten, mit denen sich die Gebrauchstauglichkeit der Ergebnisseiten von Suchmaschinen untersuchen lassen. Anwender sind auf die Beschreibung der Treffer angewiesen, um zu entscheiden, ob diese für die Suche relevant sind und weiterverfolgt werden sollen. Die Trefferbeschreibung ist damit ein Kernelement im Entscheidungsprozess der Suche. Bisher ungelöst war die Frage, welcher Text aus den verlinkten Dokumenten zur Beschreibung am besten geeignet ist und den Entscheidungsprozess optimal unterstützt. Aus diesem Grund wurde ein neues Rahmenmodell für eine systematische und differenzierte Forschung über Ergebnisseiten von Suchmaschinen entwickelt. Dieses enthält die drei Ebenen Mensch, Maschine und Aufgabe. Jede Ebene wurde in ihre Bestandteile zerlegt und diejenigen Elemente im Einzelnen betrachtet, die einen Einfluss auf den Entscheidungsprozess nehmen können.

Darauf aufbauend wurde eine Untersuchung konzipiert. Als abhängige Variablen zum Vergleich von Beschreibungsvarianten von Treffern wurden Kriterien aus den Anforderungen zur Gebrauchstauglichkeit der EN ISO 9241-11 abgeleitet. Berücksichtigt wurden neben den im Information Retrieval gängigen Kriterien zur Effektivität (Präzision und Vollständigkeit) auch zwei Maße zur Effizienz (Erfolgsrate und Entscheidungszeiten) sowie zur Zufriedenheit (Fragen zur Zufriedenheit, Entscheidungssicherheit, Präferenz von Beschreibungsvarianten im direkten Vergleich). In einem Web-Experiment wurden die Trefferbeschreibungen zu den repräsentierten Dokumenten und die Spezifität der Suchaufgabe variiert. Die Trefferbeschreibungen entsprachen den aktuell konkurrierenden Ansätzen: (a) die Suchwörter im Kontext des Vorkommens im Dokument darzustellen, (b) Text vom Anfang des Dokuments darzustellen und (c) Schlüsselwörter zur Beschreibung zu verwenden. Teilnehmer trafen Selektionsentscheidungen auf Ergebnisseiten anhand zweier unterschiedlich spezifischer Suchaufgaben. Die Daten von 191 Personen erlaubten einen Vergleich der Gebrauchstauglichkeit verschiedener Varianten von Trefferbeschreibungen sowie Aussagen zu den Effekten durch unterschiedliche Aufgabenschwierigkeiten.

Die Ergebnisse zeigen, dass eine Linkbeschreibung durch Anfangstext mit 400 Zeichen redaktionellem Inhalt den anderen Bedingungen überlegen ist und die Selektionsentscheidung am besten unterstützt. Bei der spezifischen Suchaufgabe konnten theoriekonform bessere Selektionsentscheidungen als bei der unspezifischen festgestellt werden. Einflüsse aufgrund von Unterschieden zwischen Teilnehmern konnten nicht festgestellt werden.

## 5.2 Anmerkungen zum Design

Wie sich herausgestellt hat, bietet das Rahmenmodell eine sehr gute Basis, auf deren Grundlage sich die vorliegende Untersuchung aufbauen ließ. Die Operationalisierung kann als gelungen bezeichnet werden, wie die verschiedenen Prüfungen zu diesem Sachverhalt deutlich machten (siehe Kap. 4.1). Das gemischte Design ermöglichte die Erfassung von Effekten sowohl innerhalb wie auch zwischen verschiedenen Teilnehmern und war dadurch

besonders resistent gegen mögliche Störeinflüsse von Drittvariablen. Mit Hilfe des direkten Bezugs zu den Grundsätzen der Dialoggestaltung und den Anforderungen an die Gebrauchstauglichkeit in der EN ISO 9241 konnten sinnvolle Kriterien entwickelt werden. Diese haben sich als erfolgreich in der Differenzierungsfähigkeit zwischen verschiedenen Varianten von Trefferbeschreibungen und Suchaufgaben erwiesen.

Die Definition von Relevanz als Relevanz aus Anwenderperspektive hat den Vorteil, dass nicht im voraus für jede Suchaufgabe entschieden werden muss, welche Treffer mit dem Informationsbedürfnis in Zusammenhang stehen. Eine solche Relevanzdefinition könnte daher auch für laufende Suchanfragen bei Suchmaschinen eingesetzt werden (vgl. Kap. 5.6).

Schwächen der Untersuchung fanden sich in der Umsetzung der Übungsaufgaben. Anstelle der verwendeten inhaltsleeren Selektionsaufgabe könnte hier eine standardisierte Suchanfrage als Kontrollvariable für spätere Anfragen eingesetzt werden. Sie könnte helfen, ein (hier nicht näher definiertes) Konstrukt allgemeiner Verarbeitungsgeschwindigkeit auf Ergebnisseiten zu erfassen und statistisch konstant zu halten. Auch die Messung der Variablen in der Ebene Mensch kann verbessert werden. Die in der vorliegenden Arbeit verwendeten Items differenzieren möglicherweise nicht genug zwischen Teilnehmern und verhindern dadurch die Entdeckung von Zusammenhängen zwischen Teilnehmern mit hoher und niedriger Sucherfahrung (vgl. Kap. 4.4). Dies gilt entsprechend auch für die anderen erhobenen Wissensbereiche. Weiterhelfen würden hierbei möglicherweise Fragen nach ausgeübtem Beruf und erreichtem Bildungsabschluss, da diese Fragen weniger von der subjektiven Selbsteinschätzung abhängen als die verwendeten Wissensfragen. An der Dropoutanalyse lässt sich ablesen, dass viele Teilnehmer durch die Übungsaufgaben von einer weiteren Teilnahme abgeschreckt wurden und die Teilnehmerzahl in der zweiten Hälfte der Untersuchung einen zweiten Einbruch erleidet. Das lässt darauf schließen, dass die durchschnittlichen zwanzig Minuten, die für die Untersuchung benötigt wurden, zu lang sind. Gekürzt werden kann die Untersuchung, indem die vierzehn Items zur Zufriedenheit durch die noch zu validierende neue Zufriedenheitsskala ersetzt wird. Zusammenfassend kann also festgestellt werden, dass das Untersuchungsdesign klare Stärken aufweist, aber in einigen Punkten weitere Verbesserungen möglich sind.

### 5.3 Teststärkenanalyse

Die angenommene Irrtumswahrscheinlichkeit ( $\alpha$ -Fehler) zur Signifikanzprüfung von Tests bietet noch keine Aussage darüber, wie gut ein Test geeignet ist, auch kleine Effekte zu entdecken (Teststärke). Um die tatsächliche Stärke des vorliegenden Designs bewerten zu können, wird eine Poweranalyse mit Hilfe des Programms G-Power von Faul und Erdfelder (1992) durchgeführt. Teststärke,  $\alpha$ -Fehler, Effektgröße und Probandenzahl stehen dabei in direkter Beziehung zueinander (Cohen, 1992). Um eine kleinere Effektgröße zu entdecken oder eine höhere Teststärke zu erzielen, ist stets eine höhere Probandenzahl notwendig, wenn die anderen Größen konstant bleiben sollen.

Da nur solche Treffer in die Auswertung eingehen, für die die Mehrheit der Teilnehmer eine eindeutige Entscheidung bezüglich der Relevanz treffen können, sind die möglichen Streuungen in den Selektionsentscheidungen per Definition gering. Daraus folgt, dass kleine Effekte für die Effektivitäts- und Effizienzmaße zu erwarten sind, da sich diese aus den Selektionsentscheidungen errechnen. Cohen schreibt zu den erwarteten Effektstärken: „[E]ffect sizes in behavioral science are generally small [...]“ (Cohen, 1988, S. 284). Für die Zufriedenheitsmaße können daher ebenfalls kleine Effekte erwartet werden. An anderer Stelle schreibt Cohen zur Bedeutsamkeit mittlerer Effektgrößen: „My intent was that medium ES [effect size] represent an effect likely to be visible to the naked eye of a careful observer.“ (Cohen, 1992, S. 156). Die erwarteten Effekte bei Untersuchungen zur Gebrauchstauglichkeit der Ergebnisseiten sind demnach nicht für das bloße Auge zu erkennen und können als klein bezeichnet werden. Ein allgemeiner Ansatz zur Berechnung der Effektgrößen für das vorliegende Untersuchungsdesign einer mixed ANOVA (vgl. dazu Kap. 3.3) existiert nicht (Faul & Erdfelder, 1992). Die durchgeführten Berechnungen sind daher in Anhang G aufgeführt. Beispielhaft wird in Tabelle 18 die Teststärke des verwendeten Untersuchungsdesigns für kleine Effektgrößen bei den beiden extremen Ausprägungen der Vollständigkeit und der Erfolgsrate dargestellt. Die Teststärken für die Variablen Genauigkeit, Entscheidungszeit sowie Entscheidungssicherheit liegen zwischen diesen beiden Extremen. Wie hoch die Teststärke im Einzelfall ausfällt, hängt beim mixed ANOVA Design zum Teil von der Korrelation der Kriterien im within-subject Faktor ( $r_{\text{Erfolgsrate}} = 0,456$ ;  $r_{\text{Vollständigkeit}} = 0,08$ ), also der Aufgabe ab (Buchner, Erdfelder & Faul, 1997). Bei erwarteten kleinen Effekten führt die geringe Fallzahl zu einer geringen Teststärke für einige Kriterien. Zusammenfassend konnte jedoch gezeigt werden, dass sich bei Betrachtung aller Kriterien durchaus Ergebnisse und Empfehlungen ableiten lassen.

**Tabelle 18:** Teststärken bei  $\alpha = 0,05$ ;  $N=191$ ; Effektstärke  $f^2 = 0,01$ .

Effekt	Kriterium	Teststärke
Beschreibung	Erfolgsrate	0,24
	Vollständigkeit	0,31
Aufgabe	Erfolgsrate	0,75
	Vollständigkeit	0,53
Interaktion	Erfolgsrate	0,58
	Vollständigkeit	0,36

## 5.4 Validität der Untersuchung

In diesem Abschnitt wird die interne und externe Validität des Webexperiments diskutiert. Die interne Validität von Untersuchungen beschreibt, inwiefern kausale Zusammenhänge zwischen verschiedenen Variablen bestehen und zulässig sind. Die externe Validität beschreibt, auf welche Population sich der gefundene Kausalzusammenhang generalisieren lässt (Cook & Campbell, 1979).

Zur Sicherstellung der internen Validität wurden die Untersuchungsergebnisse zahlreichen Prüfungen unterzogen (Kap. 4.1). Als mögliche Gefährdung konnte ein Unterschied zwischen Abbrechern und Teilnehmern in Bezug auf das selbst eingeschätzte Suchmaschinenwissen identifiziert werden. Eine tatsächliche Gefährdung liegt jedoch nicht vor, weil das Wissen der Teilnehmer nicht mit den Kriterien der Untersuchung korreliert ist. Der kausale Zusammenhang zwischen der Beschreibungsvariante Textanfang und den verschiedenen Kriterien wird dadurch betont, dass 155 von 191 Teilnehmern angaben, Google<sup>29</sup> als häufigste Suchmaschine in der letzten Woche eingesetzt zu haben. Die von Google verwendete Beschreibungsmethode wurde jedoch in der Bedingung Kontext operationalisiert. Obwohl die Mehrheit der Teilnehmer offenbar mit der Beschreibungsvariante Kontext bestens vertraut sind, führt Textanfang zu zufriedeneren Teilnehmern.

Zwar lautet das Ergebnis der vorliegenden Arbeit, dass Textanfang am besten zur Beschreibung von Treffern geeignet ist, dies muss jedoch im Zusammenhang der gegeneinander abgewogenen Beschreibungsvarianten gesehen werden. Es ist anzunehmen, dass noch bessere Methoden zur Beschreibung von Treffern existieren (siehe Kap. 5.6).

Ein Problem in Webexperimenten stellt die Tatsache dar, dass keine Zufallsstichprobe aus der Population gezogen werden kann. Teilnehmer können dadurch nicht als repräsentativ für die Population angesehen werden (Couper, 2000). Dies schränkt die Generalisierbarkeit der Ergebnisse über die Stichprobe hinaus ein. Als zu untersuchende Population gelten in der vorliegenden Arbeit Internetnutzer, die Suchmaschinen verwenden. Das Problem der Repräsentativität wird daher etwas entschärft, weil das Medium der Datenerhebung dasselbe ist, in dem sich auch die zu untersuchende Population befindet. Zudem konnte im Kapitel zur Stichprobenbeschreibung (Kap. 4.2) gezeigt werden, dass sich die Geschlechterverteilung in der Internetpopulation mit der in der Untersuchung gefundenen deckt. Da zudem das Geschlecht nicht signifikant mit den Kriterien korreliert, können die gefundenen Ergebnisse als repräsentativ in Bezug auf die Geschlechterverteilung der Internetpopulation gelten. In Bezug auf das Alter zeigt sich jedoch eine Tendenz zu jüngeren Teilnehmern als in der Population. Ebenfalls berichten Teilnehmer eine höhere Einschätzung an Suchmaschinenwissen als Abbrecher. Die Allgemeingültigkeit der Ergebnisse ist daher dahingehend einzuschränken, dass sie für jüngere Menschen mit etwas überdurchschnittlicher Sucherfahrung gelten.

---

<sup>29</sup> <http://www.google.de>

Ein weiterer wichtiger Aspekt ist, dass durch das experimentelle Design zwar eine hohe interne Validität erreicht werden konnte, die externe Validität jedoch dadurch eingeschränkt wird. Der experimentelle Aufbau lässt berechtigterweise die Frage aufkommen, ob solche Effekte auch bei tatsächlicher Suche zu beobachten sind. Der letzte Abschnitt der Diskussion (Kap. 5.6) skizziert daher eine mögliche Umsetzung der Untersuchungsfragen in eine Felduntersuchung, die diese Vermutungen bestätigen könnte.

## 5.5 Implikationen für die Praxis

Gemäß der gefundenen Ergebnislage zu Beschreibungen von Treffern auf Ergebnisseiten und der Spezifität der Suchaufgabe lassen sich zwei Empfehlungen für die Praxis ableiten.

Redaktioneller Anfangstext von repräsentierten Dokumenten scheint anderen Methoden der Darstellung überlegen. Obwohl ausführlichere Informationen über das repräsentierte Dokument zu einer Verbesserung der Entscheidungen hinsichtlich Effizienz und Zufriedenheit führen, existiert ein Trade-Off mit der Entscheidungszeit. Je mehr Text den Anwendern zum Evaluieren von Treffern zur Verfügung steht, desto höher wird auch die Bearbeitungszeit der Ergebnisseite werden. Designer von Ergebnisseiten sehen sich daher in einem Konflikt. Sollen sie die Attribute der Links kurz und gering halten, um eine schnelle und oberflächliche Evaluation zu ermöglichen, oder sind ausführlichere Informationen sinnvoller, die die Ergebnisseite deutlich länger werden lassen, aber auch eine genauere Darstellung der zu erwarteten Inhalte erlauben?

Zu dieser Problematik konnte die vorliegende Arbeit einige Antworten geben, die im Folgenden für die Praxis weiter expliziert werden: Vor die Wahl gestellt, welche Beschreibungstechniken Suchmaschinenbetreiber implementieren sollen, können sie aussagekräftige und gebrauchstaugliche Ergebnisseiten erzeugen, indem sie redaktionellen Text vom Anfang der repräsentierten Dokumenten verwenden. Als Einstiegshilfe zur Implementation kann der in Kapitel 3.1.2 vorgeschlagene Algorithmus herangezogen werden.

Die in dieser Untersuchung festgestellte selbst gewählte Präferenz der Anwender widerspricht den Ergebnissen anderer Untersuchungen (vgl. Drori, 2001). Jedoch wurden mit 400 Zeichen vom Textanfang die sonst üblichen 50-200 Zeichen deutlich überschritten. Dadurch scheinen Anwender in der Lage zu sein, sich ein genaueres Bild vom verlinkten Dokument zu machen als dies mit den Methoden „Wörter im Kontext“, „Kernaussagen oder Schlüsselwörter“ oder „Anfangstext mit 200 Zeichen“ möglich ist. Bisher relativ ungelöst ist die Frage, ab wann der Vorteil längerer Beschreibungen sich aufgrund der höheren Bearbeitungszeit in einen Nachteil verwandelt. Das solch ein Fall nicht von der Hand zu weisen ist, zeigen die Ergebnisse in derjenigen Bedingung, die die drei Beschreibungsvarianten Wörter im Kontext, Textanfang und Schlüsselwörter in einem Konglomerat miteinander verbindet: Sie schneidet hinsichtlich der Effizienz deutlich schlechter ab als die anderen drei Bedingungen für sich genommen.

Im Sinne des Grundsatzes der Steuerbarkeit von Dialogen erscheint es wünschenswert, dem Anwender die Entscheidung zu überlassen, wie viele Zeichen aus dem Anfangstext dargestellt werden sollen. Die Menge ließe sich damit der jeweiligen Fragestellung anpassen. So wie bisherige Suchmaschinen die Anzahl der Treffer pro Seite einstellbar halten, könnte eine weitere Option eingebaut werden, die eine Auswahl mit sinnvollen Mengen Zeichen Anfangstext erlaubt (z.B. 100, 200, 400, 800). Im nächsten Abschnitt (Kap. 5.6) wird auf offene Fragestellungen in diesem Zusammenhang eingegangen.

Die zweite Empfehlung betrifft das Suchverhalten der Anwender. Für Anwender scheint es lohnenswert sich mit ihrem Informationsbedürfnis noch vor Beginn der Suche mit Suchmaschinen näher auseinanderzusetzen. Gelänge es Anwendern, ein relativ unspezifisches Informationsbedürfnis in mehrere spezifische Suchaufgaben zu zerlegen, so würden sie mit einer erfolgreicherer Suche und einer höheren Zufriedenheit mit den erzielten Ergebnissen belohnt werden. Beispielsweise kann die in der Untersuchung gestellte unspezifische Suchaufgabe „Welche wirtschaftlichen Folgen hat die Währungsumstellung?“ in mehrere spezifischere Suchaufgaben zerlegt werden. Dies wären in diesem Fall Fragen nach Teilaspekten, auf die sich die Währungsumstellung auswirkt. Zum Beispiel: „Wie wirkt sich die Währungsumstellung auf die Börse aus?“, „Welche Folgen hat die Währungsumstellung für den Im- und Export?“, „Wie wirkt sich die Währungsumstellung auf den Devisenmarkt aus?“ Eine andere Möglichkeit der näheren Spezifizierung stellt die Frage nach den Quellen dar, die die gewünschten Informationen bereitstellen könnten. Können solche Quellen identifiziert werden, schränkt dies den Suchraum weiter ein. Mögliche Anbieter von Informationen zur Währungsumstellung könnten die Bundesregierung, politische Parteien, wirtschaftliche Interessenverbände oder Finanzzeitschriften sein.

Vorüberlegungen zur Suche unter den Aspekten der Themenzerlegung und Quellenbestimmung helfen beim Bilden von spezifischeren Suchanfragen und führen dadurch zu einem effizienterem Suchprozess mit zufriedeneren Anwendern.

## 5.6 Ausblick

Aus dem Vorschlag, die Länge von Beschreibungen variabel einstellbar zu gestalten, ergibt sich die Fragestellung, inwieweit Anwender solch eine Option auch einsetzen würden. Wird der Gedanke zur Steuerbarkeit konsequent weitergeführt, so kann ebenfalls gefordert werden, dass es Anwendern möglich sein sollte, zwischen verschiedenen möglichen Beschreibungsvarianten wählen zu können. Möglicherweise existieren verschiedene Typen von Anwendern, wobei einige bessere Ergebnisse mit kurzen Beschreibungen erzielen, während andere wiederum mit ausführlicheren Texten besser zu Recht kommen.

Eine möglicherweise bessere Variante zur Beschreibung von Treffern als die hier untersuchten kann in einer Verbindung der Vorteile von Kontextinformationen mit zusammenhängendem Text vermutet werden. Dokumente beginnen nicht immer mit redaktionellem Text, der erfolgreich über die wesentlichen Inhalte des Dokuments aufklärt.

Ein Zusammenhang mit der Suchanfrage wird ebenfalls nicht zwangsweise deutlich. Daher kann es sinnvoll sein, die Methode Wörter im Kontext einzusetzen, jedoch dabei erheblich mehr Text neben den Wörtern aus dem Suchausdruck darzustellen. Dies würde den Nachteil der Methode Kontext, dass Beschreibungen gelegentlich unzusammenhängend erscheinen, aufheben können. Um nun die Trefferbeschreibungen nicht mit jedem weiteren Wort im Suchausdruck zu verlängern, könnte die Anzahl an darzustellenden Zeichen festgelegt sein. Die Länge der Passagen zu einzelnen Wörtern des Suchausdrucks könnten dann jeweils entsprechend auf den festgelegten Raum angepasst werden.

Die Integration der aufgeworfenen Forschungsfragen in einer Felduntersuchung mit echten Suchanfragen und Suchmaschinen kann wie folgt skizziert werden: Beim Betreten einer Suchmaschinenstartseite werden Personen für die Untersuchung angeworben. Der Nutzer beginnt dann seine durch ein reales Informationsbedürfnis ausgelöste Suche, während gleichzeitig verschiedene Variablen über die durchgeführte Suche erhoben werden. Bei Bedarf können Teilnehmer vor Beginn der Suche zu einigen Angaben aufgefordert werden, wie z.B. die Wahl einer Beschreibungsvariante. Während der Suche wird die Selektion eines Treffers auf einer Ergebnisseite registriert. Nach Beenden der Suche kann nach der Relevanz des Treffers und der Spezifität der Aufgabe gefragt werden. Diese Relevanzeinschätzung durch die Teilnehmer kann als ein wichtiges Kriterium herangezogen werden, um zu vergleichen, ob Unterschiede zwischen Untersuchungsbedingungen dahingehend existieren, wie viele irrelevante Treffer ausgewählt werden. Je weniger irrelevante Treffer eine Bedingung verursacht, desto überlegener wäre die zugehörige Beschreibungsvariante. Ergebnisse aus solch einem Vorgehen hätten eine hohe Aussagekraft mit direktem Bezug zu tatsächlichem Suchverhalten.

Es war das Anliegen des Autors der vorliegenden Arbeit, einen psychologischen und methodischen Beitrag zur Erforschung der Gebrauchstauglichkeit der Ergebnisseiten bei Suchmaschinen zu leisten. Es wäre erfreulich, wenn einige Gedanken aus dieser Arbeit in die weitere Forschung einfließen würden.

## 6 Literatur

- Abseits.de. (2001). *Wieviele Ergebnisseiten von Suchmaschinen werden gelesen?* [Online]. Verfügbar unter: <http://www.abseits.de/wievieleseiten.htm> [12.11.2002].
- Amitay, E. (2001). What lays in the Layout: Using anchor-paragraph arrangements to extract descriptions of Web documents. Unveröffentlichte Dissertation, Macquarie University Sydney, Australia.
- Amitay, E. & Paris, C. (2000). Automatically Summarising Web Sites - Is There A Way Around It? In Proceedings of the ninth international conference on Information and knowledge management (CIKM) in Washington, DC, USA, November 6-11, 2000 (S. 173-179). New York: ACM Press.
- Auble, R. S. & Delongchamp, D. L. (2001). *Review: Teaching Internet Search Strategies Using Critical Thinking Skills* [Online]. Verfügbar unter: <http://imet.csus.edu/imet1/scotty/research/litreview.htm> [14.10.2002].
- Baeza-Yates, R. & Ribeiro-Neto, B. (1999). *Modern Information Retrieval* (3. Aufl.). Harlow: Addison-Wesley Longman Limited.
- Banke, K. (2000). Wettklicken - Vergleich: Benutzbarkeit von Websites. *iX*, 12, 98-103. Online verfügbar unter: <http://www.heise.de/ix/artikel/2000/12/098/>.
- Bates, M. J. (1998). The design of browsing and berrypicking techniques for the on-line search interface. *Online Review*, 13 (5), 407-431. Online verfügbar unter: <http://www.gseis.ucla.edu/faculty/bates/berrypicking.html> [14.10.2002].
- Batinic, B. & Bosnjak, M. (2000). Fragebogenuntersuchungen im Internet. In B. Batinic (Hrsg.), *Internet für Psychologen* (S. 287-318). Göttingen: Hogrefe.
- Bekavac, B. (1999). Suche und Orientierung im WWW. Verbesserung bisheriger Verfahren durch Einbindung hypertextspezifischer Informationen. Konstanz: Universitätsverlag Konstanz.
- Bernard, M., Baker, R. & Fernandez, M. (2002). Paging vs. Scrolling: Looking for the Best Way to Present Search Results. *Usability News* [Online], 4 (1). Verfügbar unter: <http://psychology.wichita.edu/surl/newsletter.htm>.
- Bernard, M., Chaparro, B. & Thomasson, R. (2000). Finding Information on the Web: Does Whitespace Matter? *Usability News* [Online], 2 (1). Verfügbar unter: <http://psychology.wichita.edu/surl/newsletter.htm>.
- Bloom, B. S. (1972). *Taxonomie von Lernzielen im kognitiven Bereich* (2. Aufl.). Weinheim: Beltz.
- Borgman, C. L. (1986). The user's mental model of an information retrieval system: an experiment on a prototype online catalog. *International Journal of Man-Machine Studies*, 24, 47-64.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5. Aufl.). Berlin: Springer.
- Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (3. Aufl.). Berlin: Springer.
- Bosnjak, M. (2002). (Non)Response bei Web-Befragungen. Auswahl, Erweiterung und empirische Prüfung eines handlungstheoretischen Modells zur Vorhersage und Erklärung des



- Partizipationsverhaltens bei Web-basierten Fragebogenuntersuchungen. Aachen: Shaker-Verlag.
- Bosnjak, M. & Batinic, B. (1999). Determinanten der Teilnahmebereitschaft. In B. Batinic, A. Werner, L. Gräf & W. Bandilla (Hrsg.), *Online Research. Methoden, Anwendungen und Ergebnisse* (S. 145-158). Göttingen: Hogrefe.
- Bosnjak, M. & Tuten, T. L. (2001). Classifying Response Behaviors in Web-based Surveys. *Journal of Computer-Mediated Communication*, 6 (3). Online verfügbar unter: <http://www.ascusc.org/jcmc/vol6/issue3/boznjak.html> [09.11.2002].
- Brace, N., Kemp, R. & Snelgar, R. (2000). *SPSS for Psychologists. A Guide to Data Analysis using SPSS for Windows*. Basingstoke: Macmillan.
- Brem, S. & Boyes, A. J. (2000). Using critical thinking to conduct effektive searches of online resources. *Practical Assessment, Research & Evaluation*, 7 (7). Online verfügbar unter: <http://ericae.net/pare/getvn.asp?v=7&n=7> [14.10.2002].
- Buchner, A., Erdfelder, E. & Faul, F. (1997). How to Use G\*Power [Online]. Verfügbar unter: [http://www.psych.uni-duesseldorf.de/aap/projects/gpower/how\\_to\\_use\\_gpower.html](http://www.psych.uni-duesseldorf.de/aap/projects/gpower/how_to_use_gpower.html).
- Choo, C. W., Detlor, B. & Turnbull, D. (1998). A Behavioral Model of Information Seeking on the Web - Preliminary Results of a Study of How Managers and IT Specialists Use the Web. In C. M. Preston (Hrsg.), *Proceedings of the 61st Annual Meeting of the American Society for Information Science (ASIS) in Pittsburgh, October 24-29, 1998* (S. 290-302): Information Today, Inc.
- Choo, C. W., Detlor, B. & Turnbull, D. (1999). A Behavioral Model of Information Seeking on the Web - An Integrated Model of Browsing and Searching. *First Monday* [Online], 5 (2). Verfügbar unter: [http://www.firstmonday.dk/issues/issue5\\_2/choo/](http://www.firstmonday.dk/issues/issue5_2/choo/).
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. Aufl.). Hillsdale: Lawrence Erlbaum Associates, Inc.
- Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, 112 (1), 155-159.
- Cook, T. D. & Campbell, D. T. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: McNally.
- Cooke, A. (1999). *Authoritative Guide to Evaluating Information on the Internet*. London: Neal-Schuman Publishers, Inc.
- Couper, M. P. (2000). Web Surveys. A Review of Issues and Approaches. *Public Opinion Quarterly*, 64, 464-494.
- Couper, M. P., Traugott, M. W. & Lamias, M. J. (2001). Web Survey Design and Administration. *Public Opinion Quarterly*, 65, 230-253.
- DeMicheli, M. (1998). *Erfolgreiches Suchen und Finden im Internet*. St Gallen: Smartbooks Publishing AG.
- Diaper, D. (Hrsg.). (1989). *Task Analysis for Human-Computer Interaction*. Chichester: Ellis Horwood Limited.
- Dillman, D. A. (2000). *Mail and Internet Surveys: The Tailored Design Method* (2. Aufl.). New York: John Wiley Company.
- Dix, A., Finley, J., Abowd, G. & Beale, R. (1995). *Mensch, Maschine, Methodik*. München: Prentice Hall Verlag.

- Dodge, M. & Kitchin, R. (2001). *Atlas of Cyberspace*: Addison Wesley.
- Drori, O. (2000a). The Benefits of Displaying Additional Internal Document Information on Textual Database Search Result Lists. In *Proceedings of the 4th European Conference on Research and Advanced Technology for Digital Libraries - ECDL2000 in Lisbon, Portugal, September 2000* (S. 69-82). Berlin: Springer.
- Drori, O. (2000b). Using Text Elements by Context to Display Search Results in Information Retrieval Systems. In *Proceedings of the Hypertext 2000 & Digital Libraries 2000 Workshop on Information Doors - Where Information Search and Hypertext Link in San Antonio, Texas, USA, 2000* (S. 17-22). New York: ACM Press.
- Drori, O. (2001). Improving Display of Search Results in Information Retrieval Systems - User's Study. In P. T. Isaias (Hrsg.), *Proceedings of the 1st International Workshop on New Developments in Digital Libraries (NDDL). In conjunction with ICEIS in Setúbal, Portugal, July 6-7, 2001* (S. 20-33): ICEIS Press.
- Dumais, S., Cutrell, E. & Chen, H. (2001). Optimizing Search by Showing Results In Context. In *Proceedings of the SIGCHI conference on Human factors in computing systems in Seattle, Washington, United States, 2001* (S. 277-284). New York: ACM Press.
- Dutke, S. (1994). *Mentale Modelle: Konstrukte des Wissens und Verstehens. Kognitionspsychologische Grundlagen für die Software-Ergonomie*. Göttingen: Hogrefe.
- Eichstaedt, J. (2001). An inaccurate-timing filter for reaction-time measurement by JAVAapplets implementing Internet-based experiments. *Behavior Research Method, Instruments, & Computers*, 33, 179-186.
- Endres-Niggemeyer, B. (1998). *Summarizing Information*. Berlin: Springer.
- Erbach, G., Neumann, G. & Uszkoreit, H. (1997). MULINEX. Multilingual Indexing, Navigation and Editing Extension for the Word-Wide Web. In *Proceedings of the 1997 AAAI Spring Symposium on Cross-Language Text and Speech in Stanford, March 24-26, 1997*.
- Fang, X. & Salvendy, G. (2000). Keyword Comparison: A User-Centered Feature for Improving Web Search Tools. *Human-Computer Studies*, 52, 915-931.
- Faul, F. & Erdfelder, E. (1992). GPOWER: a priori, post-hoc, and compromise power analysis for MS-DOS (Version 2.0) [computer programm]. Bonn: Bonn University, Dep. of Psychology.
- Fidel, R., Davies, R. K., Douglass, M. H., Holder, J. K., Hopkins, C. J., Kushner, E. J., et al. (1999). A Visit to the Information Mall: Web Searching Behavior of High School Students. *Journal of the American Society for Information Science*, 50 (1), 24-37.
- Forschungsgruppe Wahlen Online. (2002). *Internet-Strukturdaten IV. Quartal 2002* [Online]. Verfügbar unter: [http://www.fgw-online.de/Ergebnisse/Strukturdaten\\_Internet/](http://www.fgw-online.de/Ergebnisse/Strukturdaten_Internet/) [20.02.2003].
- Fox, S. (2002). *Pew Internet Project Data Memo* [Online]. Verfügbar unter: <http://www.pewinternet.org/reports/index.asp> [14.10.2002].
- Garber, S. R. & Grunes, M. B. (1992). The Art of Search: A Study of Art Directors. In *Conference proceedings on Human factors in computing systems in Monterey, California, United States, 1992* (S. 157-163). New York: ACM Press.
- Google. (2002a). *Google-Informationen für Webmaster* [Online]. Verfügbar unter: <http://www.google.de/intl/de/webmasters/seo.html> [19.01.2003].

- Google. (2002b). *Warum man Google benutzen sollte* [Online]. Verfügbar unter: [http://www.google.de/intl/de/why\\_use.html](http://www.google.de/intl/de/why_use.html) [20.01.2003].
- Gourmelon, A. (1994). Richtlinien für die gedächtnisfreundliche und altersgerechte Gestaltung von Dialogschnittstellen. Nürnberg: Friedrich-Alexander-Universität.
- Greene, S. L., Devlin, S. J., Cannata, P. E. & Gomez, L. M. (1990). No IFs, ANDs, or ORs: A study of database querying. *International Journal of Man-Machine Studies*, 32 (3), 303-326.
- Grossan, B. (1997). Search Engines. What they Are, How They Work, and Practical Suggestions for Getting the Most Out of Them. *Webreference* [Online]. Verfügbar unter: <http://webreference.com/content/search/how.html> [19.10.2002].
- Grundman, D. & Michalek, A. (2000). Improving Hypertext Presentation by Structuring Information Space. In Proceedings of the Hypertext 2000 & Digital Libraries 2000 Workshop on Information Doors - Where Information Search and Hypertext Link in San Antonio, Texas, USA, 2000. New York: ACM Press.
- Gunn, H. (2002). *Effective Internet Searches: Power Searching versus Surfing* [Online]. Verfügbar unter: <http://www.accesswave.ca/~hgunn/searching/> [02.02.2003].
- Harter, S. P. (1992). Psychological Relevance and Information Science. *Journal of the American Society for Information Science*, 43 (9), 602-615.
- Hartmann, W., Näf, M. & Schäuble, P. (2000). *Informationsbeschaffung im Internet, Grundlegende Konzepte verstehen und umsetzen* (2. Aufl.). Zürich: Orell Füssli Verlag AG.
- Hassenzahl, M. & Prümper, J. (1999). "Benutzererwartung eingebaut": Gestaltungsempfehlungen für Suchfunktionen auf der Basis einer empirischen Benutzerbefragung. In U. Arend, E. Eberleh & K. Pitschke (Hrsg.), *Software-Ergonomie '99, Design von Informationswelten, Gemeinsame Fachtagung des German Chapter of the ACM, der Gesellschaft für Informatik (GI) und der SAP AG in Walldorf, März 8-11, 1999*. Stuttgart: Teubner.
- Hauptmanns, P. (1999). Grenzen und Chancen von quantitativen Befragungen mit Hilfe des Internet. In B. Batinic, A. Werner, L. Gräf & W. Bandilla (Hrsg.), *Online Research. Methoden, Anwendungen und Ergebnisse* (S. 21-38). Göttingen: Hogrefe.
- Heinrich, L. J. (1993). *Wirtschaftsinformatik. Einführung und Grundlegung*. München: Oldenbourg.
- Hock, R. (1998). How to Do Field Searching in Web Search Engines: A Field Trip. *Online*, 22 (3). Online verfügbar unter: <http://infoday.com/online/OL1998/hock5.html>.
- Hoelscher, C. & Strube, G. (1999). Searching on the Web: Two Types of Expertise. In Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval in Berkeley, California, United States, 1999 (S. 305-306). New York: ACM Press.
- Hsieh-Yee, I. (1993). Effects of Search Experience and Subject Knowledge on the Search Tactics of Novice and Experienced Searchers. *Journal of the American Society for Information Science*, 45 (3), 161-174.
- Janetzko, D. (1999). *Statistische Anwendungen im Internet : Daten in Netzumgebungen erheben, auswerten und präsentieren*. München: Addison Wesley.

- Jansen, B. J., Goodrum, A. & Spink, A. (2000). Searching for Multimedia: Video, Audio, and Image Web Queries. *World Wide Web Journal*, 3 (4), 249-254. Online verfügbar unter: <http://jimjansen.tripod.com/academic/pubs/mm00.html> [14.10.2002].
- Jansen, B. J. & Pooch, U. (2000). A Review of Web Searching Studies and a Framework for Future Research. *Journal of the American Society for Information Science and Technology*, 52 (3), 235-246. Online verfügbar unter: <http://jimjansen.tripod.com/academic/pubs/wus.html> [14.10.2002].
- Jansen, B. J., Spink, A. & Pfiff, A. (2000). Linguistic Aspects of Web Queries. *Journal of the American Society for Information Science*, 11-16. Online verfügbar unter: <http://jimjansen.tripod.com/academic/pubs/linuis.html> [14.10.2002].
- Jansen, B. J., Spink, A. & Saracevic, T. (1998). Failure Analysis in Query Construction: Data and Analysis from A Large Sample of Web Queries. In *Proceedings of the third ACM conference on Digital libraries in Pittsburgh, Pennsylvania, United States, 1998* (S. 289-290). New York: ACM Press.
- Jonassen, D. H., Hannum, W. H. & Tessmer, M. (1989). *Handbook of Task Analysis Procedures*. London: Praeger.
- Jungermann, H., Pfister, H.-R. & Fischer, K. (1998). *Die Psychologie der Entscheidung*. Heidelberg: Spektrum Akademischer Verlag.
- Karzauninkat, S. (1998a). *Die Suchfibel* [Online]. Verfügbar unter: <http://www.suchfibel.de> [14.10.2002].
- Karzauninkat, S. (1998b). *Die Suchfibel*. Leipzig: Klett Verlag.
- Khan, K. & Locatis, C. (1998a). Searching through Cyberspace: The Effects of Link Cues and Correspondence on Information Retrieval from Hypertext on the World Wide Web. *Journal of the American Society for Information Science*, 49 (14), 1248-1253.
- Khan, K. & Locatis, C. (1998b). Searching Through Cyberspace: The Effects of Link Display and Link Density on Information Retrieval from Hypertext on the World Wide Web. *Journal of the American Society for Information Science*, 49 (2), 176-182. Online verfügbar unter: [www.asis.org/Publications/JASIS/v49n298.html](http://www.asis.org/Publications/JASIS/v49n298.html) [14.10.2002].
- Kleinberg, J. M. (1998). Authoritative sources in a hyperlinked environment. In *Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms in San Francisco, California, United States, 1998* (S. 668-677). New York: ACM Press.
- Kleinfeldt, S. & Baphna, J. (2000). A Commercial Perspective on Hypertext Search Results. In *Proceedings of the Hypertext 2000 & Digital Libraries 2000 Workshop on Information Doors - Where Information Search and Hypertext Link in San Antonio, Texas, USA, 2000*. New York: ACM Press.
- Konrad, E. (1992). Zur Effektivitätsbewertung von Information-Retrieval-Systemen. In R. Kuhlen (Hrsg.), *Experimentelles und praktisches Information Retrieval* (S. 119-130). Konstanz: Universitätsverlag Konstanz.
- Körber, S. (2000). Suchmuster erfahrener und unerfahrener Suchmaschinennutzer im deutschsprachigen World Wide Web. Ein Experiment. Unveröffentlichte Magisterarbeit, Westfälische Wilhelms-Universität Münster.
- Körndle, H. (1993). *Mensch-Computer-Interaction*. Wiesbaden: Deutscher Universitätsverlag GmbH.

- Koster, M. (2002). *The Web Robots Pages* [Online]. Verfügbar unter: <http://www.robotstxt.org/wc/robots.html> [27.10.2002].
- Kuhlen, R. (1992). Hypertext und Information Retrieval - mehr als Browsing und Suche. In R. Kuhlen (Hrsg.), *Experimentelles und praktisches Information Retrieval* (S. 309-324). Konstanz: Universitätsverlag Konstanz.
- Lawrence, S. & Giles, C. L. (1998a). Context and Page Analysis for Improved Web Search. *IEEE Internet Computing*, 2 (4), 28-46. Online verfügbar unter: <http://computer.org/internet/> [14.10.2002].
- Lawrence, S. & Giles, C. L. (1998b). Searching the World Wide Web. *Science*, 280, 98. Online verfügbar unter: [http://www.ist.psu.edu/faculty\\_pages/giles/publications/Science-98.pdf](http://www.ist.psu.edu/faculty_pages/giles/publications/Science-98.pdf) [14.10.2002].
- Lawrence, S. & Giles, C. L. (1998c). *Tips for Searching the Web. New Study on the Accessibility and Distribution of Information on the Web* [Online]. Verfügbar unter: <http://www.neci.nec.com/~lawrence/searchtips.html> [26.01.2003].
- Lawrence, S. & Giles, C. L. (1999). Searching the Web: General and Scientific Information Access. *IEEE Communications*, 37 (1), 116-122.
- Lennartz, S. (1999). Ich bin wichtig! Promotion-Maßnahmen für suchdienstgerechte Webseiten. *c't*, 23, 180-186.
- Lergier, R. & Resnick, M. (2001). A framework for evaluating user strategies in Internet search and evaluation. In 7th Conference on Human Factors and the Web in Madison, Wisconsin, USA, June 4-6, 2001 [Online]. Verfügbar unter: <http://www.optavia.com/hfweb/7thconferenceproceedings.zip/Resnick.pdf>.
- Lienert, G. A. & Ratz, U. (1994). *Testaufbau und Testanalyse* (5. Aufl.). Weinheim: Psychologie Verlags Union.
- Little, R. J. & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Maaß, S. (1993). Software-Ergonomie. Benutzer- und aufgabenorientierte Systemgestaltung. *Informatik-Spektrum*, 16, 191-205.
- Marchionini, G. (1995). *Information Seeking in Electronic Environments*. Cambridge: Cambridge University Press.
- McArthur, R. & Bruza, P. D. (2000). Ranking of Query Refinements in Interactive Web-based Retrieval. In Proceedings of the Hypertext 2000 & Digital Libraries 2000 Workshop on Information Doors - Where Information Search and Hypertext Link in San Antonio, Texas, USA, 2000. New York: ACM Press.
- Miller, R. B. (1966). Task Description and Analysis. In R. M. Gagné (Hrsg.), *Psychological principles in system development* (S. 187-228). New York: Holt, Rinehart and Winston.
- Münz, S. (2001, 27.10.2001). *SelfHtml. Einführung* [Online]. Verfügbar unter: <http://www.netzwelt.com/selfhtml/intro/technologien/html.htm> [13.01.2003].
- Müsseler, J. & Prinz, W. (Hrsg.). (2002). *Allgemeine Psychologie*. Spektrum Akademischer Verlag.
- Navarro-Prieto, R., Scaife, M. & Rogers, Y. (1999). Cognitive Strategies in Web Searching. In *Human Factors and the Web*.

- Nielsen, J. (1990). *Hypertext and Hypermedia*. San Diego: Academic Press, Inc.
- Nielsen, J. (1993). *Usability Engineering*. Boston: AP Professional.
- Nielsen, J. (1997). Search and You May Find. *Alertbox* [Online], July. Verfügbar unter: <http://www.useit.com/alertbox/9707b.html> [14.10.2002].
- Noller, S. (2000). Mentale Modelle und Webnavigation. Ein Usability-Experiment zur Informationssuche im World-Wide-Web. Unveröffentlichte Diplomarbeit, Universität Köln.
- Notess, G. R. (1997). On The Net: Internet Search Techniques and Strategies. *Online*, 21 (4), 63-66. Online verfügbar unter: <http://www.onlinemag.net/JulOL97/net7.html> [14.10.2002].
- Notess, G. R. (1998). On The Net: More Internet Search Strategies. *Online*, 22 (5), 71-74. Online verfügbar unter: <http://www.onlinemag.net/OL1998/net9.html> [14.10.2002].
- Notess, G. R. (2001). On The Net: Customization Options for Web Searching. *Online*, 25 (1), 55-65. Online verfügbar unter: [http://www.onlinemag.net/OL2001/net1\\_01.html](http://www.onlinemag.net/OL2001/net1_01.html) [14.10.2002].
- o.V. (2000). *General FAQ #25: Handling missing or incomplete data* [Online]. Verfügbar unter: <http://www.utexas.edu/cc/faqs/stat/general/gen25.html> [11.01.2003].
- Ohnemus, K. R. (1997). Web Style Guides: Who, What, Where. In Proceedings of the 15th annual international conference on Computer documentation (SIGDOC 97) in Salt Lake City, Utah, United States, 1997 (S. 189-197). New York: ACM Press.
- Panero, J. C. (1995). Navigation and Mental Models of the World Wide Web: An Exploratory Study: Rice University, Human Factors Psychology.
- Payne, J. W., Bettman, J. R. & Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge: Cambridge University Press.
- Peck, V. A. & John, B. E. (1992). Browser-Soar: A Computational Model of a Highly Interactive Task. In *Conference proceedings on Human factors in computing systems in Monterey, California, United States, 1992* (S. 165-172). New York: ACM Press.
- Pior, R. & Tiedemann, P. (2000). *Internet für Psychologen. Eine praxisorientierte Einführung*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Pirelli, P. & Card, S. K. (1999). *Information Foraging* (UIR Technical Report).
- Pittkow, J. E. & Kehoe, C. M. (1996). Emerging Trends in the WWW User Population. *Communications of the ACM*, 39 (6).
- Pollock, A. & Hockley, A. (1997). What's Wrong with Internet Searching. *D-Lib Magazine* [Online], March 1997. Verfügbar unter: <http://www.dlib.org/dlib/march97/bt/03pollock.html> [19.10.2002].
- Pursche, O. (2002). Brisante Downloads. *PC Professionell*, 5, 72-83.
- Rappoport, A. (1999). *Web Site Searching and the User Experience* [Online]. Verfügbar unter: <http://www.searchtools.com/info/ui/search-HI-0.html> [14.10.2002].
- Reibold, H. (2000). Findigkeit gefragt. *PC Professionell*, 2, 170-185.
- Reips, U.-D. (2000). Das psychologische Experiment im Internet. In B. Batinic (Hrsg.), *Internet für Psychologen* (S. 319-344). Göttingen: Hogrefe.
- Rommel, M. (2000). *CoAn für Windows (Version 2.05)* [Computer-Programm]. Bielefeld: Eigenverlag.

- Rosenfeld, L. & Morville, P. (1998). *Information Architecture for the World Wide Web*. Cambridge: O'Reilly.
- RRZN (Hrsg.). (2000). *Suchen & Finden im Internet oder: "Die Nadel im Heuhaufen"* (5. Aufl.). Hannover: Regionales Rechenzentrum für Niedersachsen / Universität Hannover (RRZN).
- Saito, M. & Ohmura, K. (1998). A Cognitive Model for Searching for Ill-Defined Targets on the Web. The Relationship between Search Strategies and User Satisfaction. In *SIGIR '98: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in Melbourne, Australia* (S. 155-163). New York: ACM Press.
- Schneider, W. (1998). Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten - Grundsätze der Dialoggestaltung. Kommentar zu DIN EN ISO 9241-10. Berlin: Beuth Verlag.
- Schulz, U. (2001). *Usability Kriterien für Suchmaschinen (Search Engine Usability)* [Online]. Verfügbar unter: <http://www.bui.fh-hamburg.de/pers/ursula.schulz/webusability/suchma.html> [14.10.2002].
- Selvidge, P. (2003). Examining Tolerance for Online Delays. *Usability News* [Online], 5 (1). Verfügbar unter: <http://psychology.wichita.edu/surl/usabilitynews/51/delaytime.htm>.
- Shneiderman, B. (1997). Designing Information-Abundant Web Sites: Issues and Recommendations. *International Journal of Human-Computer Studies*, 47, 5-29. Online verfügbar unter: <ftp://ftp.cs.umd.edu/pub/hcil/Reports-Abstracts-Bibliography/3634html/3634.html> [14.10.2002].
- Shneiderman, B., Byrd, D. & Croft, W. B. (1997). Clarifying Search: A User-Interface Framework for Text Searches. *D-Lib Magazine* [Online], January. Verfügbar unter: <http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html>.
- Silverstein, C. & Henzinger, M. (1999). Analysis of a Very Large Web Search Engine Query Log. *SIGIR Forum* [Online], 33 (1). Verfügbar unter: <http://www.acm.org/sigir/forum/F99/Silverstein.pdf>.
- Spink, A., Bateman, J. & Jansen, B. J. (1998). Searching heterogeneous collections on the Web: behaviour of Excite users. *Information Research*, 4 (2). Online verfügbar unter: <http://informationr.net/ir/4-2/paper53.html>.
- Stenmark, D. (1997). *To Search is Great, to Find is Greater: a Study of Visualisation Tools for the Web* [Online]. Verfügbar unter: <http://w3.informatik.gu.se/~dixi/publ/mdi.htm> [04.11.2002].
- Thimbleby, H., Jones, M. & Theng, Y. L. (1997). Is "lost in hyperspace" lost in controversy? In *Hypertext '97 Conference*. Southampton.
- Tillmann, H. N. (2000). Evaluating Quality on the Net [Online]. Verfügbar unter: <http://www.hopetillman.com/findqual.html> [14.10.2002].
- Unz, D. C., Capstick, J. & Erbach, G. (1999). Damit die Suche (schneller) zum Ziel führt: Interface-Design einer Suchmaschine zur Unterstützung von Selektionsentscheidungen. In W. Wirth & W. Schweiger (Hrsg.), *Selektion im Internet. Empirische Analysen zu einem Schlüsselkonzept* (S. 249-270). Opladen: Westdeutscher Verlag.
- Vogt, K. (1999). Verzerrungen in elektronischen Befragungen? In B. Batinic, A. Werner, L. Gräf & W. Bandilla (Hrsg.), *Online Research. Methoden, Anwendungen und Ergebnisse* (S. 127-144). Göttingen: Hogrefe.

- Wandke, H. & Hurtienne, J. (1999). Psychologische Analysen und Gestaltungsvorschläge zur Informationsauswahl im World Wide Web. In W. Wirth & W. Schweiger (Hrsg.), *Selektion im Internet. Empirische Analysen zu einem Schlüsselkonzept* (S. 271-292). Opladen: Westdeutscher Verlag.
- Wanger, J., McDonald, D. & Berger, M. C. (1980). *Evaluation of the online search process: a final report*. Santa Monica, CA: Cuadra Associates.
- Weber, C. & Groner, R. (1999). Suchstrategien im WWW bei Laien und Experten. In W. Wirth & W. Schweiger (Hrsg.), *Selektion im Internet. Empirische Analysen zu einem Schlüsselkonzept* (S. 181-196). Opladen: Westdeutscher Verlag.
- Weiss, R., Vélez, B., Sheldon, M. A., Namprempre, C., Szilagyi, P., Duda, A., et al. (1996). HyPursuit: A Hierarchical Network Search Engine that Exploits Content-Link Hypertext Clustering. In *Proceedings of the the seventh ACM conference on Hypertext (HT 1996) in Bethesda, Maryland, United States, 1996* (S. 180-193). New York: ACM Press.
- Wirth, W. & Brecht, M. (1999). Selektion und Rezeption im WWW: Eine Typologie. In W. Wirth & W. Schweiger (Hrsg.), *Selektion im Internet. Empirische Analysen zu einem Schlüsselkonzept* (S. 149-180). Opladen: Westdeutscher Verlag.
- Wirth, W. & Schweiger, W. (1999a). Selektion neu betrachtet: Auswahlentscheidungen im Internet. In W. Wirth & W. Schweiger (Hrsg.), *Selektion im Internet. Empirische Analysen zu einem Schlüsselkonzept* (S. 43-74). Opladen: Westdeutscher Verlag.
- Wirth, W. & Schweiger, W. (Hrsg.). (1999b). *Selektion im Internet. Empirische Analysen zu einem Schlüsselkonzept*. Opladen: Westdeutscher Verlag.
- Wittmann, W. W. (1985). *Evaluationsforschung: Aufgaben, Probleme und Anwendungen*. Berlin: Springer.
- Wolfram, D. & Dimitroff, A. (1997). Preliminary Findings On Searcher Performance and Perceptions of Performance in a Hypertext Bibliographic Retrieval System. *Journal of the American Society for Information Science*, 48 (2), 1142-1145.



## Anhang A: Mittelwerte der Varianzanalysen

Tabelle 19: Beobachtete Mittelwerte für Genauigkeit.

Aufgabe	Beschreibung	Mittelwert	Standardabweichung	N
<b>spezifisch</b>	Kontext	,9821	,0936	56
	Textanfang	,9700	,1199	50
	Schlüsselwörter	,9857	,0845	35
	MaxInfo	,9634	,1318	41
	Gesamt	,9753	,1087	182
<b>unspezifisch</b>	Kontext	,9048	,2417	56
	Textanfang	,8919	,2126	50
	Schlüsselwörter	,9005	,1667	35
	MaxInfo	,9695	,1204	41
	Gesamt	,9150	,1983	182

Tabelle 20: Beobachtete Mittelwerte für Vollständigkeit.

Aufgabe	Beschreibung	Mittelwert	Standardabweichung	N
<b>spezifisch</b>	Kontext	,7672	,2516	58
	Textanfang	,7647	,2521	51
	Schlüsselwörter	,8421	,2872	38
	MaxInfo	,7727	,2944	44
	Gesamt	,7827	,2688	191
<b>unspezifisch</b>	Kontext	,6897	,2919	58
	Textanfang	,7974	,2589	51
	Schlüsselwörter	,7544	,2649	38
	MaxInfo	,7348	,2649	44
	Gesamt	,7417	,2729	191

**Tabelle 21:** Beobachtete Mittelwerte für Entscheidungssicherheit.

<b>Aufgabe</b>	<b>Beschreibung</b>	<b>Mittelwert</b>	<b>Standardabweichung</b>	<b>N</b>
<b>spezifisch</b>	Kontext	2,2226	,2099	53
	Textanfang	2,1551	,3748	49
	Schlüsselwörter	2,0474	,3585	38
	MaxInfo	2,1762	,3621	42
	Gesamt	2,1571	,3313	182
<b>unspezifisch</b>	Kontext	1,7170	,4299	53
	Textanfang	1,7580	,6708	49
	Schlüsselwörter	1,5038	,4132	38
	MaxInfo	1,7483	,5173	42
	Gesamt	1,6907	,5268	182

**Tabelle 22:** Beobachtete Mittelwerte der Entscheidungszeiten.

<b>Aufgabe</b>	<b>Beschreibung</b>	<b>Mittelwert</b>	<b>Standardabweichung</b>	<b>N</b>
<b>spezifisch</b>	Kontext	6,0790	2,2322	53
	Textanfang	7,2842	8,1716	44
	Schlüsselwörter	5,6451	1,7450	31
	MaxInfo	8,2827	3,4652	37
	Gesamt	6,8130	4,8228	165
<b>unspezifisch</b>	Kontext	7,0825	2,7784	53
	Textanfang	6,9464	3,6888	44
	Schlüsselwörter	5,9532	3,4927	31
	MaxInfo	11,0121	7,5852	37
	Gesamt	7,7152	4,9197	165

Tabelle 23: Beobachtete Mittelwerte Erfolgsrate.

Aufgabe	Beschreibung	Mittelwert	Standardabweichung	N
<b>spezifisch</b>	Kontext	,1428	,0704	51
	Textanfang	,1527	,0966	44
	Schlüsselwörter	,1728	,0685	29
	MaxInfo	,1117	,0621	35
	Gesamt	,1441	,0786	159
<b>unspezifisch</b>	Kontext	,1098	,0570	51
	Textanfang	,1362	,0774	44
	Schlüsselwörter	,1573	,0929	29
	MaxInfo	,0856	,0422	35
	Gesamt	,1205	,0721	159

Tabelle 24: Beobachtete Mittelwerte Zufriedenheit.

Aufgabe	Beschreibung	Mittelwert	Standardabweichung	N
<b>spezifisch</b>	Kontext	,1428	,0704	51
	Textanfang	,1527	,0966	44
	Schlüsselwörter	,1728	,0685	29
	MaxInfo	,1117	,0621	35
	Gesamt	,1441	,0786	159
<b>unspezifisch</b>	Kontext	,1098	,0570	51
	Textanfang	,1362	,0774	44
	Schlüsselwörter	,1573	,0929	29
	MaxInfo	,0856	,0422	35
	Gesamt	,1205	,0721	159

Tabelle 25: Geschätzte Mittelwerte des direkten Vergleichs aller vier Bedingungen im Faktor Beschreibung.

Beschreibung	Treffer	Mittelwert	Standardfehler	N
<b>Kontext</b>	relevant	3,302	,078	179
	irrelevant	2,179	,070	179
<b>Textanfang</b>	relevant	3,330	,088	179
	irrelevant	2,955	,084	179
<b>Schlüsselwörter</b>	relevant	3,117	,082	179
	irrelevant	2,413	,090	179
<b>MaxInfo</b>	relevant	3,397	,096	179
	irrelevant	3,034	,090	179

Tabelle 26: Geschätzte Mittelwerte der Zufriedenheit aller Items.

Beschreibung	Aufgabe	Mittelwert	Standardfehler	N
<b>Kontext</b>	spezifisch	,509	,087	47
	unspezifisch	-,170	,093	47
<b>Textanfang</b>	spezifisch	,491	,093	41
	unspezifisch	,305	,099	41
<b>Schlüsselwörter</b>	spezifisch	,087	,104	33
	unspezifisch	-,128	,111	33
<b>MaxInfo</b>	spezifisch	,502	,098	37
	unspezifisch	-,058	,105	37

## Anhang B: Korrelationen zur Ebene Mensch

### Kodierschema:

WSEA3: Anzahl in der letzten Woche verwendeter Suchmaschinen  
 WSEH: Dauer der Suchmaschinennutzung in Stunden  
 WSEA0: Anzahl der gestellten Suchanfragen  
 WSEA1: Davon einfache Suchanfragen  
 WSEA2: Davon komplexe Suchanfragen  
 WPCH: Anzahl Tage mit Computernutzung  
 WINH: Anzahl Tage mit Internetnutzung  
 WSEL: Geschätztes Wissen über Suchmaschinen  
 WEUL: Geschätztes Wissen über Suchmaschinen  
 WWIL: Geschätztes Wirtschaftswissen  
 SEX: Geschlecht  
 AGE: Alter

LRECAG: Vollständigkeit spezifisch  
 LRECAW: Vollständigkeit unspezifisch  
 LPRECAG: Genauigkeit spezifisch  
 LPRECW: Genauigkeit unspezifisch  
 TIMEGREL: Entscheidungszeit spezifisch  
 TIMEWREL: Entscheidungszeit unspezifisch  
 LSICHG: Entscheidungssicherheit spezifisch  
 LSICHW: Entscheidungssicherheit unspezifisch  
 SUCCGT: Erfolgsrate spezifisch  
 SUCCWT: Erfolgsrate unspezifisch  
 OPGN: Zufriedenheit spezifisch  
 OPWN: Zufriedenheit unspezifisch

Pearson correlation matrix

	LRECAG	LRECAW	LPRECAG	LPRECW	TIMEGREL
WSEA3	-0.071	0.008	-0.241	-0.026	-0.088
WSEH	0.022	-0.012	-0.097	-0.126	-0.051
WSEA0	-0.176	-0.071	-0.011	-0.031	-0.106
WSEA1	-0.037	0.029	-0.107	0.017	-0.135
WSEA2	-0.175	-0.069	0.006	-0.031	-0.064
WPCH	0.015	0.023	0.012	0.027	-0.091
WINH	0.037	-0.018	0.019	0.031	-0.050
WSEL	-0.041	-0.038	0.014	0.032	-0.017
WEUL	-0.012	0.058	-0.075	0.126	-0.054
WWIL	-0.016	-0.102	-0.085	-0.080	0.015
SEX	0.145	0.237	0.041	0.135	0.092
AGE	-0.110	-0.196	-0.237	-0.244	0.173

	TIMEWREL	LSICHG	LSICHW	SUCCGT	SUCCWT
WSEA3	-0.007	-0.070	-0.106	-0.014	-0.048
WSEH	-0.047	-0.193	-0.098	0.109	0.012
WSEA0	-0.039	-0.114	-0.097	0.035	-0.063
WSEA1	0.052	0.052	-0.080	0.129	-0.058
WSEA2	-0.056	-0.142	-0.070	-0.022	-0.036
WPCH	0.059	0.123	0.006	0.014	-0.081
WINH	0.048	0.060	-0.099	0.076	-0.016
WSEL	-0.003	0.062	0.064	0.005	-0.064
WEUL	0.044	0.068	0.094	0.035	-0.021
WWIL	0.039	-0.080	-0.075	-0.012	-0.132
SEX	-0.063	0.109	0.167	0.042	0.135
AGE	0.068	-0.292	-0.249	-0.192	-0.166

	OPGN	OPWN
WSEA3	-0.125	0.060
WSEH	-0.109	-0.066
WSEA0	-0.193	-0.077
WSEA1	0.098	0.137
WSEA2	-0.276	-0.143
WPCH	-0.039	-0.003
WINH	-0.104	-0.093
WSEL	0.137	-0.032
WEUL	0.128	0.139
WWIL	0.123	0.139
SEX	-0.012	0.117
AGE	0.027	0.001

## Matrix of Bonferroni Probabilities

	LRECAG	LRECAW	LPRECG	LPRECW	TIMEGREL
WSEA3	1.000	1.000	0.164	1.000	1.000
WSEH	1.000	1.000	1.000	1.000	1.000
WSEA0	1.000	1.000	1.000	1.000	1.000
WSEA1	1.000	1.000	1.000	1.000	1.000
WSEA2	1.000	1.000	1.000	1.000	1.000
WPCH	1.000	1.000	1.000	1.000	1.000
WINH	1.000	1.000	1.000	1.000	1.000
WSEL	1.000	1.000	1.000	1.000	1.000
WEUL	1.000	1.000	1.000	1.000	1.000
WWIL	1.000	1.000	1.000	1.000	1.000
SEX	1.000	0.144	1.000	1.000	1.000
AGE	1.000	1.000	0.185	0.130	1.000

	TIMEWREL	LSICHG	LSICHW	SUCCGT	SUCCWT
WSEA3	1.000	1.000	1.000	1.000	1.000
WSEH	1.000	1.000	1.000	1.000	1.000
WSEA0	1.000	1.000	1.000	1.000	1.000
WSEA1	1.000	1.000	1.000	1.000	1.000
WSEA2	1.000	1.000	1.000	1.000	1.000
WPCH	1.000	1.000	1.000	1.000	1.000
WINH	1.000	1.000	1.000	1.000	1.000
WSEL	1.000	1.000	1.000	1.000	1.000
WEUL	1.000	1.000	1.000	1.000	1.000
WWIL	1.000	1.000	1.000	1.000	1.000
SEX	1.000	1.000	1.000	1.000	1.000
AGE	1.000	0.009	0.106	1.000	1.000

	OPGN	OPWN
WSEA3	1.000	1.000
WSEH	1.000	1.000
WSEA0	1.000	1.000
WSEA1	1.000	1.000
WSEA2	0.125	1.000
WPCH	1.000	1.000
WINH	1.000	1.000
WSEL	1.000	1.000
WEUL	1.000	1.000
WWIL	1.000	1.000
SEX	1.000	1.000
AGE	1.000	1.000

Pairwise frequency table

	LRECAG	LRECAW	LPRECG	LPRECW	TIMEGREL
WSEA3	183	183	179	178	173
WSEH	160	160	156	155	152
WSEA0	162	162	159	157	153
WSEA1	153	153	150	148	144
WSEA2	149	149	145	144	139
WPCH	189	189	185	184	179
WINH	189	189	185	184	179
WSEL	188	188	184	183	178
WEUL	186	186	182	181	176
WWIL	189	189	185	184	179
SEX	189	189	185	184	179
AGE	186	186	182	182	176

	TIMEWREL	LSICHG	LSICHW	SUCCGT	SUCCWT
WSEA3	162	180	177	170	159
WSEH	141	157	154	149	138
WSEA0	144	159	156	150	141
WSEA1	136	150	147	141	133
WSEA2	131	146	143	136	128
WPCH	166	186	183	176	163
WINH	166	186	183	176	163
WSEL	165	185	182	175	162
WEUL	164	183	180	173	161
WWIL	166	186	183	176	163
SEX	166	186	183	176	163
AGE	163	183	180	173	161

	OPGN	OPWN
WSEA3	173	160
WSEH	151	138
WSEA0	154	143
WSEA1	145	137
WSEA2	142	132
WPCH	178	165
WINH	178	165
WSEL	177	165
WEUL	175	163
WWIL	178	165
SEX	178	165
AGE	175	162

## Anhang C: t-Tests zu systematischen Verzerrungen bei Dropouts

Die notwendige Bonferoni-Korrektur zur Irrtumswahrscheinlichkeit ergibt die Irrtumswahrscheinlichkeit die erreicht oder unterschritten werden muss, damit einer der t-Tests als signifikant bezeichnet werden kann:  $\alpha' = 0,05/13 = 0,0039$  (Bortz, 1999).

Das Kodierschema ist in Anhang B beschrieben. Verglichen werden die Angaben im Anfangsfragebogen (S2) zwischen tatsächlichen Teilnehmern und Nicht-Teilnehmern.

Variable	Varianzen sind	Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit		
		F	Signifikanz	T	df	Sig. (2-seitig)
WSEA3	Gleich	4,412	,037	-1,013	302	,312
	nicht gleich			-1,167	266,715	,244
WSEH	Gleich	5,395	,021	1,271	255	,205
	nicht gleich			,990	96,307	,325
WSEA0	gleich	,348	,556	,178	254	,859
	nicht gleich			,199	251,900	,842
WSEA1	gleich	,794	,374	,933	232	,352
	nicht gleich			,931	162,446	,353
WSEA2	gleich	3,860	,051	-1,109	226	,269
	nicht gleich			-1,460	182,916	,146
WPCH	gleich	4,803	,029	-1,137	306	,256
	nicht gleich			-1,053	191,364	,294
WINH	gleich	6,073	,014	-1,092	304	,276
	nicht gleich			-1,032	203,142	,303
WSEL	gleich	1,965	,162	-3,299	305	,001
	nicht gleich			-3,156	215,190	,002
WEUL	gleich	,653	,420	-2,299	301	,022
	nicht gleich			-2,219	218,315	,028
WWIL	gleich	1,212	,272	-1,815	306	,070
	nicht gleich			-1,769	229,700	,078
BEK	gleich	5,798	,017	1,185	305	,237
	nicht gleich			1,280	301,920	,201
SEX	gleich	,649	,421	-2,071	305	,039
	nicht gleich			-1,914	189,191	,057
AGE	gleich	5,753	,017	2,193	286	,029
	nicht gleich			2,099	183,068	,037



## Anhang D: Zitierte Suchmaschinen und Webangebote

Informationdienst Arbeit und Gesundheit - Schwerpunkt Bildschirmarbeit	<a href="http://www.sozialnetz-hessen.de/Ergo-Online/">http://www.sozialnetz-hessen.de/Ergo-Online/</a>
KommDesign	<a href="http://www.kommdesign.de/texte/din.htm">http://www.kommdesign.de/texte/din.htm</a>
Nedstat, Besucherseitenzähler	<a href="http://www.nedstat.com/">http://www.nedstat.com/</a>
Network Information Center (NIC)	<a href="http://www.nic.de.vu/">http://www.nic.de.vu/</a>
Suchmaschine Abacho	<a href="http://www.abacho.de/">http://www.abacho.de/</a>
Suchmaschine Altavista	<a href="http://de.altavista.com/">http://de.altavista.com/</a>
Suchmaschine Ask Jeeves	<a href="http://www.ask.com/">http://www.ask.com/</a>
Suchmaschine Excite	<a href="http://www.excite.de/">http://www.excite.de/</a>
Suchmaschine Fireball	<a href="http://fireball.de/">http://fireball.de/</a>
Suchmaschine Google	<a href="http://www.google.de/">http://www.google.de/</a>
Suchmaschine Kartoo	<a href="http://www.kartoo.com">http://www.kartoo.com</a>
Suchmaschine Northern Light	<a href="http://www.northernlight.com/">http://www.northernlight.com/</a>
Suchmaschine Vivisimo	<a href="http://vivisimo.com/">http://vivisimo.com/</a>
Verlag der EN ISO 9241	<a href="http://www2.beuth.de/">http://www2.beuth.de/</a>

## Anhang E: Texte zur Teilnehmerrekrutierung

### E-Mail-Text für direkte Werbung

Hallo Freunde!

Was macht gute Ergebnisseiten von Suchmaschinen aus? Um diese Frage zu beantworten benötige ich Teilnehmer.

Ein 20minütiges Onlineexperiment zur Benutzerfreundlichkeit der Ergebnisseiten von Suchmaschinen. Diplomarbeit an der Universität Mannheim. Einfach auf einen der beiden Links klicken und dann gehts auch schon los. Alternativ kannst Du auch den Link in Deinem Browser eingeben, falls das aus dem Emailprogramm heraus nicht funktionieren sollte.

<http://kaczmirek.de.vu>

oder

<http://webrum.uni-mannheim.de/ppe/larska/start.html>

Mit Deiner Teilnahme würdest Du mir sehr helfen. Leite diese Mail bitte auch an andere weiter.

Die Untersuchung ist Online bis zum 10. Juni

Frohe Grüße,

Lars

### E-Mail-Text an Webseitenbetreiber für Bitte um Verlinkung

Sehr geehrter Webmaster <bzw. Name wenn bekannt>!

„Usability bei Ergebnisseiten von Suchmaschinen“ ist das Thema meiner Diplomarbeit am Lehrstuhl Arbeits- und Organisationspsychologie der Universität Mannheim. Für eine 20minütige Onlinestudie suche ich Teilnehmer. Daher meine Bitte, ob Sie die Studie auf Ihrer Domäne <hier folgte der Domainname des Webmasters> verlinken würden? Z.B. mit folgendem Link:

<a href="http://webrum.uni-mannheim.de/ppe/larska/start.html">Was macht gute Ergebnisseiten von Suchmaschinen aus?</a>

Ein 20minütiges Onlineexperiment zur Benutzerfreundlichkeit der Ergebnisseiten von Suchmaschinen. Diplomarbeit an der Universität Mannheim.

Die Untersuchung ist Online bis zum 10. Juni

Mit freundlichen Grüßen, Lars Kaczmirek

## Anhang F: Webangebote, die auf die Untersuchung verwiesen

Im Folgenden finden sich Listen zu Suchmaschinen und anderen Webanbietern, bei denen die Untersuchung angemeldet wurde. Sofern die Startseite der Untersuchung auch tatsächlich von einer Suchmaschine indexiert wurde, erzielt der Suchausdruck „linklisten ergebnisseiten suchmaschinen“ eine Ergebnisliste, in der sie häufig auf dem ersten Platz aufgeführt wird (Stand 23.07.2002).

### **Manuelle Anmeldung in der Suchmaschine mit erfolgreicher Indexierung:**

Altavista	<a href="http://de.altavista.com/">http://de.altavista.com/</a>
Excite	<a href="http://www.excite.com/">http://www.excite.com/</a>
Fireball	<a href="http://fireball.de/">http://fireball.de/</a>
Google	<a href="http://www.google.de/">http://www.google.de/</a>
Web.de	<a href="http://web.de/">http://web.de/</a>
Yahoo	<a href="http://de.yahoo.com/">http://de.yahoo.com/</a>

### **Erfolgreiche automatische Anmeldung in der Suchmaschine mit Indexierung:**

Lycos	<a href="http://www.lycos.de/">http://www.lycos.de/</a>
-------	---

### **Angeblich erfolgreiche automatische Anmeldung in der Suchmaschine ohne tatsächliche Indexierung:**

Abacho	<a href="http://abacho.de/">http://abacho.de/</a>
Acoon	<a href="http://acoon.de/">http://acoon.de/</a>
Alexana	<a href="http://alexana.de/">http://alexana.de/</a>
Ameta	<a href="http://ameta.de/">http://ameta.de/</a>
Bellnet	<a href="http://bellnet.de/">http://bellnet.de/</a>
i01	<a href="http://i01.de/">http://i01.de/</a>
klug-suchen	<a href="http://www.klug-suchen.de/">http://www.klug-suchen.de/</a>
Lotse	<a href="http://lotse.de/">http://lotse.de/</a>
Nathan	<a href="http://www.nathan.de/">http://www.nathan.de/</a>
NetReal	<a href="http://netreal.de/">http://netreal.de/</a>
Netzindex	<a href="http://netzindex.de/">http://netzindex.de/</a>
Online-Favoriten	<a href="http://online-favoriten.de/">http://online-favoriten.de/</a>
OnlinePilot	<a href="http://onlinepilot.de/">http://onlinepilot.de/</a>
Top1000	<a href="http://top1000.de/">http://top1000.de/</a>
WebObserver	<a href="http://webobserver.de/">http://webobserver.de/</a>

### **Webangebote, die einen Link auf die Untersuchungsseite gesetzt hatten:**

Internetkompetenz	<a href="http://www.internet-kompetenz.ch/infosuche/">http://www.internet-kompetenz.ch/infosuche/</a>
Mailingliste auf kommdesign.de	<a href="http://www.kommdesign.de">http://www.kommdesign.de</a>
Onlineforschung.de	<a href="http://forschung-online.de">http://forschung-online.de</a>
Psychologie Onlineforschung.de	<a href="http://psychologie-onlineforschung.de">http://psychologie-onlineforschung.de</a>

Searchcode.de	<a href="http://searchcode.de">http://searchcode.de</a>
Suchfibel	<a href="http://www.suchfibel.de">http://www.suchfibel.de</a>
Webseiten verschiedener Lehrstühle im Fachbereich Psychologie der Universität Mannheim	<a href="http://psychologie.uni-mannheim.de">http://psychologie.uni-mannheim.de</a>

## Anhang G: Teststärkeberechnungen

Für eine Berechnung der Teststärke im mixed ANOVA Design werden die Korrelationen zwischen den Bedingungen im within-subjects Faktor Aufgabe benötigt. Die Freiheitsgrade des F-Tests können dann für jeden zu untersuchenden Effekt einzeln mit G-Power errechnet werden (Faul & Erdfelder, 1992). Eine allgemeine Methode zur Berechnung existiert nicht. Die einzelnen Rechenschritte, die erforderlich sind, um die in G-Power bei „Other F-Tests“ einzutragenden Daten zu erhalten, werden im Folgenden aufgeführt. Berechnungen werden für die beiden Kriterien mit der höchsten und niedrigsten Korrelation im within-subjects Faktor durchgeführt (Erfolgsrate und Vollständigkeit). Die anderen Kriterien liegen somit in ihrer Teststärke zwischen diesen beiden Ausprägungen.

### Powerberechnungen (Buchner, Erdfelder & Faul, 1997)

Faktorstufen Beschreibung = a = 4

Faktorstufen Aufgabe = b = 2

Kleine Effektgröße  $f^2 = 0,01$

### Effektmaße für den Haupteffekt Beschreibung (between-subjects)

Freiheitsgrade im Zähler = a - 1 = 4 - 1 = 3

Freiheitsgrade im Nenner = N - a = 191 - 4 = 187

$f^2$  für G-Power =  $(b / (1 + (b-1) \cdot \text{Korrelation})) \cdot \text{Effektgröße}$

Erfolgsrate  $f^2 = (2 / (1 + (2-1) \cdot 0,456)) \cdot 0,01 = 0,0137$  ; Ergibt: Power = 0,24

Vollständigkeit  $f^2 = (2 / (1 + (2-1) \cdot 0,08)) \cdot 0,01 = 0,0185$  ; Ergibt: Power = 0,31

### Effektmaße für den Haupteffekt Aufgabe (within-subjects)

Freiheitsgrade im Zähler = b - 1 = 2 - 1 = 1

Freiheitsgrade im Nenner = (N - a) \* (b - 1) = (191 - 4) \* (2 - 1) = 187

$f^2$  für G-Power =  $2 \cdot \text{Effektgröße} / (1 - \text{Korrelation})$

Erfolgsrate  $f^2 = 2 \cdot 0,01 / (1 - 0,456) = 0,0368$  ; Ergibt: Power = 0,75

Vollständigkeit  $f^2 = 2 \cdot 0,01 / (1 - 0,08) = 0,0217$  ; Ergibt: Power = 0,53

### Effektmaße für die Interaktion

Freiheitsgrade im Zähler = (a - 1) \* (b - 1) = (4 - 1) \* (2 - 1) = 3

Freiheitsgrade im Nenner = (N - a) \* (b - 1) = (191 - 4) \* (2 - 1) = 187

$f^2$  für G-Power =  $2 \cdot \text{Effektgröße} / (1 - \text{Korrelation})$

Erfolgsrate  $f^2 = 2 \cdot 0,01 / (1 - 0,456) = 0,0368$  ; Ergibt: Power = 0,58

Vollständigkeit  $f^2 = 2 \cdot 0,01 / (1 - 0,08) = 0,0217$  ; Ergibt: Power = 0,36



Ihr Geschlecht?

- männlich  
 weiblich

Wie hoch schätzen Sie ihr Wissen über den Euro und die Währungsumstellung ein?

- |                       |                       |                       |                       |                       |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| sehr gering           | gering                | durch-<br>schnittlich | hoch                  | sehr hoch             |
| -2                    | -1                    | 0                     | 1                     | 2                     |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Wie hoch schätzen Sie ihr Wirtschaftswissen ein?

- |                       |                       |                       |                       |                       |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| sehr gering           | gering                | durch-<br>schnittlich | hoch                  | sehr hoch             |
| -2                    | -1                    | 0                     | 1                     | 2                     |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Wie hoch schätzen Sie ihr Wissen über Suchmaschinen ein?

- |                       |                       |                       |                       |                       |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| sehr gering           | gering                | durch-<br>schnittlich | hoch                  | sehr hoch             |
| -2                    | -1                    | 0                     | 1                     | 2                     |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Wieviele Stunden pro Woche verbringen Sie am Computer?

- |                       |                       |                       |                       |                       |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| bis 8                 | 9-16                  | 17-24                 | 25-32                 | über 33               |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Wieviele Stunden pro Woche surfen Sie aktiv im Internet?

- |                       |                       |                       |                       |                       |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| bis 8                 | 9-16                  | 17-24                 | 25-32                 | über 33               |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Wieviele Stunden pro Woche nutzen oder beschäftigen Sie sich mit Suchmaschinen? Zählen Sie dabei jeweils die komplette Zeit von einer ersten Suchanfrage bis Sie die gewünschte(n) Zielseite(n) gefunden haben. Also auch das Betrachten von unbrauchbaren Webseiten.

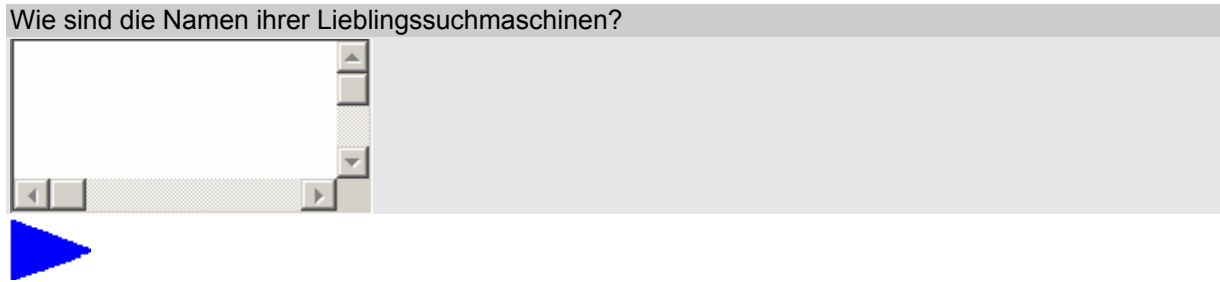
Wieviele Anfragen stellen Sie pro Woche an Suchmaschinen im Internet? Neuformulierungen von Suchanfragen zu einem Sachverhalt zählen hierbei nur einmal.

Wieviele Anfragen sind davon für die Suche...

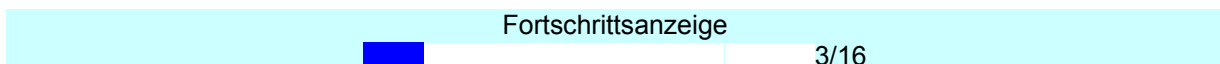
nach einfachen Tatsachen oder Fakten? Beispielsweise die Homepageadresse einer Firma, die Emailadresse einer Person oder das Datum eines Ereignisses.

nach umfassenden Informationen oder zur Beantwortung komplexer Fragestellungen? Beispielsweise Artikel zu einem Thema oder Hintergrundinformationen zur Politik.

Wie viele verschiedene Suchmaschinen verwenden Sie?



### Voruntersuchung Webseite 3, Instruktion zur Übung:

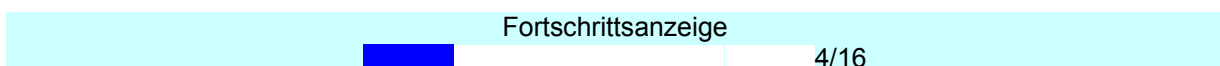


In dieser Untersuchung können die Links auf den Ergebnisseiten nicht angeklickt werden, wie dies ansonsten bei Suchmaschinen üblich ist. Stattdessen gibt es vor jedem Link eine kleine Checkbox () . Möchten Sie einen Link auswählen, so klicken Sie auf die Checkbox. Es erscheint dann ein Haken darin. Die Dokumente, zu denen die Links führen, spielen in dieser Untersuchung keine Rolle. Sie werden daher auch nicht angezeigt.

Auf den nächsten zwei Seiten wird das Auswählen von Links geübt. Sie bekommen dazu eine Seite mit 10 Links. Bei einigen Links steht der Text "Diesen Text auswählen". Setzen Sie jeweils einen Haken in die Checkbox davor. Die anderen Links sollen nicht ausgewählt werden. Klicken Sie dann auf den Weiterbutton. Bei der Auswahl von Links werden die Zeiten gemessen. Achten Sie bei den nächsten zwei Seiten darauf alle richtigen Links zu markieren und arbeiten dabei so schnell wie möglich.



### Voruntersuchung Webseite 4, Übungsdurchgang 1:



Wählen Sie alle Links mit "Diesen Text auswählen" und klicken dann auf den Weiterbutton ganz unten.

- [Diesen Text auswählen](#)
- [Diesen Text auswählen](#)
- [Text nicht auswählen](#)
- [Text nicht auswählen](#)
- [Diesen Text auswählen](#)
- [Text nicht auswählen](#)
- [Text nicht auswählen](#)



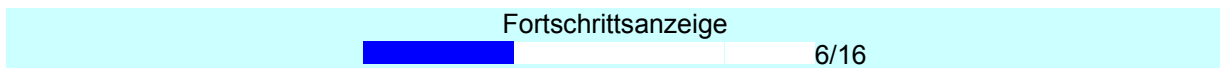
[Text nicht auswählen](#) [Text nicht auswählen](#) [Text nicht auswählen](#)**Voruntersuchung Webseite 5, Übungsdurchgang 2:**

Fortschrittsanzeige

5/16

Ein zweites Mal bitte: Wählen Sie alle Links mit "Diesen Text auswählen" und klicken dann auf den Weiterbutton ganz unten.

 [Text nicht auswählen](#) [Diesen Text auswählen](#) [Text nicht auswählen](#) [Diesen Text auswählen](#) [Text nicht auswählen](#) [Text nicht auswählen](#) [Diesen Text auswählen](#) [Diesen Text auswählen](#) [Text nicht auswählen](#) [Text nicht auswählen](#)

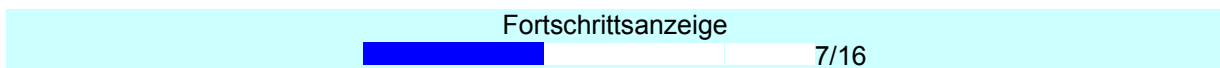
**Voruntersuchung Webseite 6, Instruktion Suchaufgabe (spezifisch):**

Sie wissen nun wie Sie auf den folgenden Seiten Links auswählen können.

Stellen Sie sich bitte folgendes vor:

Sie möchten die Größe aller neuen Euroscheine wissen. Die Wörter "euroscheine größe" haben Sie bereits in eine Suchmaschine eingetippt. Auf der nächsten Seite sehen Sie die Ergebnisse dieser Suchanfrage. Machen Sie einen Haken in die Checkboxen, wenn Sie den Link verfolgen würden.

Gehen Sie beim Auswählen möglichst wie bei einer echten Suche vor. Wieviele Links Sie auswählen, bleibt Ihnen überlassen. Treffen Sie ihre Entscheidungen möglichst schnell mit Hilfe der Titel und Texte.

**Voruntersuchung Webseite 7, Ergebnisseite (spezifisch):**

Die Suchanfrage lautete: Welche Größe haben die Euroscheine?

- [\[PDF\] www.bitzcore.com/pdf/GermanJugend.pdf](#)  
... ist nicht möglich! Bestell-Nr.: TEX\_004\_S Bestell-Nr.: TEX\_001\_S Größe: S Preis: 18 Euro Größe: S Preis: 128 Euro Grössenangaben: Brustumfang ca. 90 cm ...
- [Menu](#)  
... Blau Anzug: Stil: Olympic - Code SW-KL-UO Power Back - Code SW-KL-UP Racer Back - Code SW-KL-UR Größe: 26"-32" Preis: 19 Euro Größe: 34"-42" Preis: 28 Euro. ...
- [redmark.de - Euro Center #EURO-Scheine](#)  
... in Stückelungen von 5, 10, 20, 50, 100, 200 und 500 Euro in Umlauf gesetzt. Größe und Farbe der 7 Banknotenwerte sind unterschiedlich. Die Motive der ...
- [EICHNER TANZSCHUHE - Damenschuhe Latein](#)  
... Schuh anklicken! Click on shoe, Volle Größe und weitere Bestellinformationen, D01, 125.- DM 63,92 Euro. ...
- [Allergate AG](#)  
... Zwirnhandschuhe Lohmann Rauscher Zwirnhandschuh (Euro-Größe 9 = dt. Gr. 7,5) 1 Paar aus weißer Baumwolle Lieferzeit: Versandfertig innerhalb von 10 Tagen ...
- [Freewarepage.de - Programme der Rubrik: Büro/Rechner/Euro](#)  
... Euro-Rechner 0.4, Bewerten User (5), Der Euro-Rechner hilft Ihnen bei der Umstellung der Währung von der DM zum Euro. Win 95/98/Me. Größe: 1474 KB. ...
- [Universität Heidelberg - Gästehäuser](#)  
... Feld [ Voll-Ansicht ]. Wohnungen: 57 Einzimmerwohnungen Größe 25 m ... 2 behindertengerechte Apartments Größe 45 m². Erbaut 1997 ... DM 70,- (Euro 36,-) Zwei Zimmer ...
- [\[PDF\] www.bcl.lu/pdf/euro/bienvenue\\_de.pdf](#)  
... 82 mm Grün Größe 153 x 82 mm Gelblichbraun Größe 160 x 82 mm Lila akzeptiert, und die Euro-Banknoten und - 4. Die Flagge der Europäischen Union. Münzen ...
- [SR-Online/Programm/SR 3/Euro-Gewinnspiel](#)  
... Und so sehen sie aus, die neuen Geldscheine. 5 Euro Größe: 120 x 62 mm Farbe: Grau Dargest. Baustil: Klassik. ...

[Euro Information - Daten - Fakten - Design der Scheine und ...](#)

... 500 Euro. Größe: 160 mm x 82 mm Das Thema des 500 Euro-Scheins ist die. "moderne Architektur des 20. Jahrhunderts". 200 Euro. ...

[Die nächsten 10 Ergebnisse würde ich mir auch ansehen](#)



### Voruntersuchung Webseite 8, Priorisierung der Treffer (spezifisch):

Fortschrittsanzeige

8/16

Unten sehen Sie die selbe Ergebnisseite nochmal. Wie erfolgversprechend finden Sie die Links? Tragen Sie nun Zahlen ein. Beginnen Sie mit 1 für den Link, der am erfolgversprechendsten ist. Machen Sie dann mit 2, 3 usw. weiter. Links, die Sie jetzt nicht auswählen würden benötigen keine Nummer. Sie können sich nun auch für andere Links entscheiden, als auf der vorherigen Seite. Zeit spielt hierbei keine Rolle.

[\[PDF\] www.bitzcore.com/pdf/GermanJugend.pdf](#)

... ist nicht möglich! Bestell-Nr.: TEX\_004\_S Bestell-Nr.: TEX\_001\_S Größe: S Preis: 18 Euro Größe: S Preis: 128 Euro Größenangaben: Brustumfang ca. 90 cm ...

[Menu](#)

... Blau Anzug: Stil: Olympic - Code SW-KL-UO Power Back - Code SW-KL-UP Racer Back - Code SW-KL-UR Größe: 26"-32" Preis: 19 Euro Größe: 34"-42" Preis: 28 Euro. ...

[redmark.de - Euro Center #EURO-Scheine](#)

... in Stückelungen von 5, 10, 20, 50, 100, 200 und 500 Euro in Umlauf gesetzt. Größe und Farbe der 7 Banknotenwerte sind unterschiedlich. Die Motive der ...

[EICHNER TANZSCHUHE - Damenschuhe Latein](#)

... Schuh anklicken! Click on shoe, Volle Größe und weitere Bestellinformationen, D01, 125.- DM 63,92 Euro. ...

[Allergate AG](#)

... Zwirnhandschuhe Lohmann Rauscher Zwirnhandschuh (Euro-Größe 9 = dt. Gr. 7,5) 1 Paar aus weißer Baumwolle Lieferzeit: Versandfertig innerhalb von 10 Tagen ...

[Freewarepage.de - Programme der Rubrik: Büro/Rechner/Euro](#)

... Euro-Rechner 0.4, Bewerten User (5), Der Euro-Rechner hilft Ihnen bei der Umstellung der Währung von der DM zum Euro. Win 95/98/Me. Größe: 1474 KB. ...

[Universität Heidelberg - Gästehäuser](#)

... Feld [ Voll-Ansicht ]. Wohnungen: 57 Einzimmerwohnungen Größe 25 m ... 2 behindertengerechte Appartments Größe 45 m<sup>2</sup>. Erbaut 1997 ... DM 70,- (Euro 36,-) Zwei Zimmer ...

[\[PDF\] www.bcl.lu/pdf/euro/bienvenue\\_de.pdf](#)

... 82 mm Grün Größe 153 x 82 mm Gelblichbraun Größe 160 x 82 mm Lila akzeptiert, und die Euro-Banknoten und - 4. Die Flagge der Europäischen Union. Münzen ...

[SR-Online/Programm/SR 3/Euro-Gewinnspiel](#)

... Und so sehen sie aus, die neuen Geldscheine. 5 Euro Größe: 120 x 62 mm Farbe: Grau Dargest. Baustil: Klassik. ...

[Euro Information - Daten - Fakten - Design der Scheine und ...](#)



Nun geht es um die Links die sie nicht ausgewählt haben.  
Die Texte zu den nicht ausgewählten Links finde ich...

	sehr	eher	teils/ teils	eher	sehr	
	-2	-1	0	1	2	
schlecht	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	gut
unwichtig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	wichtig
unbrauchbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	brauchbar
unverständlich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	verständlich
unpräzise	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	präzise
unübersichtlich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	übersichtlich
unzusammenhängend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	zusammenhängend
unsinnig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sinnvoll

Mit den angebotenen Links war ich zufrieden.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die Ergebnisseite war für die Suche relevant.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die Texte zu den Links sind so gewählt, daß sie Zeit sparen helfen.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die Reihenfolge der Links war sinnvoll.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die Ergebnisseite war benutzerfreundlich.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Es ließ sich erkennen, was die Seiten beinhalten zu denen die Links führten, obwohl diese nicht angezeigt wurden.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Es war schwierig eine Auswahl zu treffen.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Eine der von mir ausgewählten Ergebnisseiten enthält die richtige Antwort.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Es hat mir Spaß gemacht die Links auszuwählen.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die dargebotene Informationsmenge zu den Links reichte für meine Entscheidungen aus.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die Ergebnisseite war frustrierend.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die Ergebnisseite war gut zu lesen.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die Organisation der Informationen war verwirrend.

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



**Voruntersuchung Webseite 10, Instruktion Suchaufgabe (unspezifisch):**

Fortschrittsanzeige

10/16

Die vorherige Suche haben Sie erfolgreich abgeschlossen.

Für die nächste Suchanfrage stellen sie sich bitte folgendes vor:

Sie suchen Informationen zu wirtschaftlichen Folgen der Währungsumstellung. Die Wörter "euro wirtschaftliche folgen" haben Sie bereits in eine Suchmaschine eingetippt. Auf der nächsten Seite sehen sie die Ergebnisse ihrer Suchanfrage. Machen sie einen Haken in die Checkboxen, deren Links Sie anklicken würden.

Gehen Sie beim Auswählen möglichst wie bei einer echten Suche vor. Wieviele Links Sie auswählen bleibt Ihnen überlassen. Treffen Sie ihre Entscheidungen möglichst schnell mit Hilfe der Titel und Texte.

**Voruntersuchung Webseite 11, Ergebnisseite (unspezifisch):**

Fortschrittsanzeige

11/16

Die Suchanfrage lautete: Welche wirtschaftlichen Folgen hat die Währungsumstellung?

[tagesschau: Wirtschaftliche Folgen für Bayer](#)

... Wirtschaftliche Folgen für Bayer, Der wegen des Lipobay-Rückrufs ... es auf einen Umsatz von 636 Millionen Euro - vier Millionen mehr als mit dem Klassiker ...

[Kölnische Rundschau - Wirtschaft](#)

... Kommentar Wirtschaftliche Folgen unabsehbar - Besser besonnen handeln Von Michael Gaßmann Bringt der Terrorakt, zusätzlich zu allem menschlichen Leid, die ...

[Die Einführung des Euro birgt erhebliche wirtschaftliche und ...](#)

... Kapital entgegenzutreten und die wirtschaftliche Entwicklung unter die ... muß der Einführung des Euro in der arbeitenden ... Perspektive und Orientierung folgen. ...

[Yahoo! - Die Europäische Zentralbank](#)

... engl.). -, Organisation für wirtschaftliche Zusammenarbeit und Entwicklung ... Leitzinsen in der Euro-Zone unverändert gelassen ... angesichts möglicher Folgen der US ...

[Yahoo! Schlagzeilen - Verbraucherschützer warnen vor Folgen ...](#)

Zusammenfassung

[Euro-Folgen für Bund und Nationalbank](#)

... Verbraucherschützer warnen vor Folgen des TV-Kabelnetzverkaufs. ... 5,5 Milliarden Euro (10,8 Milliarden Mark). Der wirtschaftliche Übergang erfolgt ...

[FTD - EU-Gipfel: Regierungschefs sehen optimistisch in die ...](#)

... Aktien für 7,50 Euro zu haben Artikel. ... sehen optimistisch in die wirtschaftliche Zukunft Die EUStaats

... führten dazu, dass die Folgen des 11. September für ...

[Alois Bischofberger: Zwischen Zeitgeist und Gewinnstreben](#)

... und des Ausgegrentseins stark. Daraus folgen vereinfachende Schuldzuweisungen. Der Missbrauch des ... nicht nur für die wirtschaftliche, sondern auch für die ...

[liste](#)

... Nedakis. Prognose ueber die wirtschaftliche Entwicklung in der Bundesrepublik ... Der EURO, Vor- und Nachteile, Stimmen. 21. ... Tendenzen und sozial-oekonomische Folgen. ...

[euro](#)

... Die Folgen für die WWU Länder ... Auswirkungen der Einführung des Euro auf die Länder Mittel ... der Abteilung für wirtschaftliche Entwicklung, Ministerium für ...

[Die nächsten 10 Ergebnisse würde ich mir auch ansehen](#)



#### Voruntersuchung Webseite 12, Priorisierung der Treffer (unspezifisch):

Fortschrittsanzeige

12/16

Unten sehen Sie die selbe Ergebnisseite nochmal. Wie erfolgversprechend finden Sie die Links? Tragen Sie nun Zahlen ein. Beginnen Sie mit 1 für den Link, der am erfolgversprechendsten ist. Machen Sie dann mit 2, 3 usw. weiter. Links, die Sie jetzt nicht auswählen würden benötigen keine Nummer. Sie können sich nun auch für andere Links entscheiden, als auf der vorherigen Seite. Zeit spielt hierbei keine Rolle.

[tagesschau: Wirtschaftliche Folgen für Bayer](#)

... Wirtschaftliche Folgen für Bayer, Der wegen des Lipobay-Rückrufs ... es auf einen Umsatz von 636 Millionen Euro - vier Millionen mehr als mit dem Klassiker ...

[Kölnische Rundschau - Wirtschaft](#)

... Kommentar Wirtschaftliche Folgen unabsehbar - Besser besonnen handeln Von Michael Gaßmann Bringt der Terrorakt, zusätzlich zu allem menschlichen Leid, die ...

[Die Einführung des Euro birgt erhebliche wirtschaftliche und ...](#)

... Kapital entgegenzutreten und die wirtschaftliche Entwicklung unter die ... muß der Einführung des Euro in der arbeitenden ... Perspektive und Orientierung folgen. ...



[Yahoo! - Die Europäische Zentralbank](#)

... engl.). -, Organisation für wirtschaftliche Zusammenarbeit und Entwicklung ... Leitzinsen in der Euro-Zone unverändert gelassen ... angesichts möglicher Folgen der US ...

[Yahoo! Schlagzeilen - Verbraucherschützer warnen vor Folgen ...](#)

Zusammenfassung

[Euro-Folgen für Bund und Nationalbank](#)

... Verbraucherschützer warnen vor Folgen des TV-Kabelnetzverkaufs. ... 5,5 Milliarden Euro (10,8 Milliarden Mark). Der wirtschaftliche Übergang erfolgt ...

[FTD - EU-Gipfel: Regierungschefs sehen optimistisch in die ...](#)

... Aktien für 7,50 Euro zu haben Artikel. ... sehen optimistisch in die wirtschaftliche Zukunft Die EUStaats

... führten dazu, dass die Folgen des 11. September für ...

[Alois Bischofberger: Zwischen Zeitgeist und Gewinnstreben](#)

... und des Ausgegrenztseins stark. Daraus folgen vereinfachende Schuldzuweisungen. Der Missbrauch des ... nicht nur für die wirtschaftliche, sondern auch für die ...

[liste](#)

... Nedakis. Prognose ueber die wirtschaftliche Entwicklung in der Bundesrepublik ... Der EURO, Vor- und Nachteile, Stimmen. 21. ... Tendenzen und sozial-oekonomische Folgen. ...

[euro](#)

... Die Folgen für die WWU Länder ... Auswirkungen der Einführung des Euro auf die Länder Mittel ... der Abteilung für wirtschaftliche Entwicklung, Ministerium für ...

[Die nächsten 10 Ergebnisse würde ich mir auch ansehen](#)



**Voruntersuchung Webseite 13, Fragebogen Treatmentcheck, Zufriedenheit (unspezifisch):**

Fortschrittsanzeige

13/16

*Entspricht Seite 9, Fragebogen Treatmentcheck, Zufriedenheit*



**Voruntersuchung Webseite 14, Abschlussfragen:**

Fortschrittsanzeige

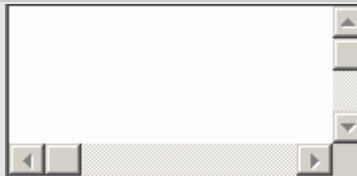
14/16

Nun noch drei letzte Fragen...

Was verstehen Sie unter benutzerfreundlichen Ergebnisseiten?



Welche Informationen nutzen Sie bei der Entscheidung, ob ein Link angeklickt wird oder nicht? Hätten Sie gerne mehr Informationen auf den Ergebnisseiten gehabt? Wenn ja, welche?



Die Texte zu den Links und deren Titel waren Zusammenfassungen der dahinterliegenden Seiten. Wie müssten Zusammenfassungen aussehen, damit Sie daran erkennen können, ob sich der Link lohnt?



**Voruntersuchung Webseite 15, Debriefing 1:**

Fortschrittsanzeige

15/16

Vielen Dank für ihre Teilnahme an der Untersuchung.

Diese Untersuchung möchte mehrere Fragen beantworten:

- Welche Darstellung der Ergebnisseiten ist für die Nutzer optimal?
- Welchen Einfluß haben unterschiedlich komplexe Suchanfragen auf die Suche?
- Wie können die Ergebnisseiten auf die unterschiedlichen Fähigkeiten und Bedürfnisse der Suchenden besser angepasst werden?
- Sind die hier neu entwickelten Untersuchungsmethoden für die ersten zwei Fragestellungen geeignet?

Falls Sie benachrichtigt werden möchten, wenn die Auswertung abgeschlossen ist, können Sie hier ihre Emailadresse angeben. Sie wird unabhängig von den restlichen Daten gespeichert, so daß Ihre Anonymität gesichert ist. *Das ist noch nicht der Fall. Du kannst jedoch gerne Deinen Namen angeben, damit ich Dich von zukünftigen Anfragen verschone \*g\* [Bitte schreibe mir Deine Gedanken zur Untersuchung](#) . Folgende Aussagen sollten im Idealfall alle mit ja beantwortet werden können. Alle Instruktionen sind klar. Du bist in der Lage jemandem zu erklären was die Begriffe Ergebnisseite, Suchanfrage und Checkbox bedeuten. Die Untersuchung war nicht langweilig. Die Fragen waren verständlich und eindeutig. Die Antwortmöglichkeiten waren sinnvoll, es fehlten keine Optionen. Die Fragen zu den Ergebnisseiten konntest Du relativ leicht ohne viel nachzudenken oder durcheinander zu geraten beantworten.*

Email: 

Für Fragen, Anregungen, Diskussionen [schreiben Sie an den Untersuchungsleiter](#), Lars Kaczmirek  
Auf der nächsten Seite erhalten Sie Antworten zu den gestellten Suchanfragen.

**Voruntersuchung Webseite 16, Debriefing 2:**

Fortschrittsanzeige

16/16

Unter den folgenden Links finden Sie einige Antworten zu den gestellten Suchanfragen:

- [Welche Größe haben die Euroscheine?](#)
- [Welche wirtschaftlichen Folgen hat die Währungsumstellung?](#)
- [Eine weitere Seite zu den Folgen.](#)



## Anhang I: Webseiten der Hauptuntersuchung

Im Folgenden sind die 20 in der Hauptuntersuchung verwendeten Webseiten mit allen acht unterschiedlichen Ergebnisseiten abgebildet (vgl. Kap. 3.1 und Abbildung 8). In der Untersuchung waren die Texte farbig unterlegt.

### Startseite (S1):

## Ergebnisseiten von Suchmaschinen eine Untersuchung

Diplomarbeit am Fachbereich Psychologie der Universität Mannheim

Liebe Teilnehmerin, lieber Teilnehmer

**Einführung:** Suchmaschinen liefern bei einer Suchanfrage Linklisten. Der Sucher entscheidet dann, welche Links er weiterverfolgt und welche unbrauchbar sind.

**Fragestellung:** Wie sieht der Link aus, der den Sucher bei seiner Entscheidung optimal unterstützt? Dazu werden beispielhafte Suchanfragen zum Thema Währungsumstellung und Euro verwendet.

**Ablauf:** Die Untersuchung dauert etwa 20 Minuten. Zunächst werden Sie um einige Angaben zu Ihrer Person gebeten. Es folgt eine kurze Übungsphase zur Linkauswahl. Dann bekommen Sie Suchanfragen gestellt. Dabei entscheiden Sie welche Links Sie weiterverfolgen würden. In anschließenden Fragebögen bewerten Sie die Linklisten. Insgesamt bekommen Sie zwei Suchanfragen.

**Bedienungshinweis:** Wenn Sie eine Seite bearbeitet haben, kommen Sie immer mit einem Klick auf den Weiterbutton (▶) zur nächsten Seite. Dieser Button befindet sich immer ganz unten auf jeder Seite. Klicken Sie ihn unten an, um jetzt an der Untersuchung teilzunehmen. Bitte verwenden Sie während der Untersuchung nicht die Backfunktion ihres Browsers. Das würde die Untersuchungsergebnisse negativ beeinflussen.

**Datenschutz:** Falls Sie es wünschen, können Sie am Ende der Untersuchung Ihre Emailadresse angeben, damit Sie über die Ergebnisse informiert werden. Die Untersuchungsdaten sind und bleiben dabei anonym. Alle Daten werden ausschließlich zu wissenschaftlichen Zwecken ausgewertet.



### Eingangsfragebogen (S2):

Fortschrittsanzeige



2/20

Beantworten Sie bitte zunächst einige Angaben zu Ihrer Person. Sie erzeugen damit eindeutigeren Untersuchungsergebnisse. Falls Sie bei einer Antwort unsicher sind, machen Sie bitte eine Angabe, die am ehesten zutrifft.

Welche Suchmaschine verwendeten Sie letzte Woche...

am häufigsten?

- Altavista  
 Fireball  
 Google  
 Lycos  
 Metager  
 Web.de  
 Yahoo

Andere:

keine

am zweithäufigsten?

- Altavista  
 Fireball  
 Google  
 Lycos  
 Metager  
 Web.de  
 Yahoo

Andere:

keine

Wie viele verschiedene Suchmaschinen verwendeten Sie letzte Woche insgesamt?

weiß nicht

Wie viele Stunden nutzten Sie letzte Woche Suchmaschinen? Zählen Sie die komplette Zeit von einer ersten Suchanfrage bis Sie die gewünschte(n) Information(en) gefunden hatten.

weiß nicht

Wie viele Anfragen haben Sie letzte Woche in Suchmaschinen eingegeben? Jede Anfrage, mit der die Sie eine neue Linkseite bekommen haben zählt.

weiß nicht

Wie viele Anfragen waren davon für die Suche...

nach einfachen Tatsachen oder Fakten? Beispielsweise die Homepageadresse einer Firma, die Emailadresse einer Person oder das Datum eines Ereignisses.

weiß nicht

nach umfassenden Informationen oder zur Beantwortung komplexer Fragestellungen? Beispielsweise Artikel zu einem Thema oder Hintergrundinformationen zur Politik.

weiß nicht

An wie vielen Tagen saßen Sie letzte Woche an einem Computer?

1

2

3

4

5

6

7

weiß nicht

An wie vielen Tagen surfen Sie letzte Woche im Internet?

1

2

3

4

5

6

7

weiß nicht

Wie hoch schätzen Sie ihr Wissen über Suchmaschinen ein?

sehr gering

gering

durch-  
schnittlich

hoch

sehr hoch

-2

-1

0

1

2

weiß nicht

Wie hoch schätzen Sie ihr Wissen über die mit dem Euro verbundene Währungsumstellung ein?

sehr gering    gering    durchschnittlich    hoch    sehr hoch

-2    -1    0    1    2

weiß nicht

Wie hoch schätzen Sie ihr Wirtschaftswissen ein?

sehr gering    gering    durchschnittlich    hoch    sehr hoch

-2    -1    0    1    2

weiß nicht

Haben Sie bereits in diesem Jahr an einer Untersuchung mit Linklisten von Suchmaschinen teilgenommen (z.B. der Voruntersuchung)?

- ja  
 nein

Ihr Geschlecht?

- männlich  
 weiblich  
 keine Angabe

Ihr Alter?

keine Angabe



### Instruktion zur Übung (S3):

Fortschrittsanzeige

3/20

In dieser Untersuchung können die Links der Linkliste nicht angeklickt werden, wie dies ansonsten bei Suchmaschinen üblich ist. Die Dokumente, zu denen die Links führen, spielen in dieser Untersuchung keine Rolle. Sie werden daher auch nicht angezeigt. Stattdessen gibt es vor jedem Link die Auswahl

ja nein

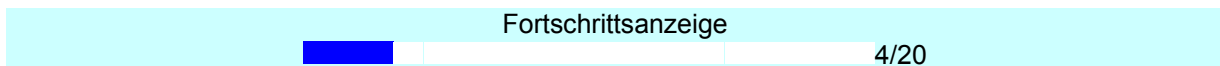


Geben Sie für jeden Link an, ob Sie ihn auswählen würden oder nicht.

Auf den nächsten beiden Seiten wird das Auswählen von Links geübt. Bei einigen Links steht der Text "Diesen Link auswählen". Wählen Sie für diese Links "ja", für alle anderen "nein". Klicken Sie dann auf den Weiterbutton.

Achten Sie bei den nächsten zwei Seiten auf korrekte Angaben und arbeiten dabei so schnell wie möglich.



**Übungsdurchgang 1 (S4):**

Wählen Sie "ja" bei den Links, die Sie auswählen würden, "nein" bei allen anderen.  
Klicken Sie dann auf den Weiterbutton ganz unten.

ja nein

[Diesen Link auswählen](#)

ja nein

[Diesen Link auswählen](#)

ja nein

[Link nicht auswählen](#)

ja nein

[Link nicht auswählen](#)

ja nein

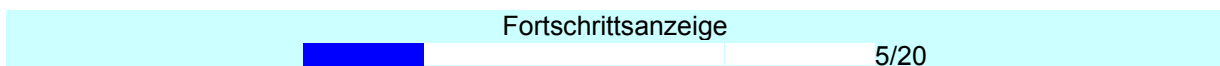
[Diesen Link auswählen](#)

ja nein

[Link nicht auswählen](#)

ja nein

[Diesen Link auswählen](#)

**Übungsdurchgang 2 (S5):**

Ein zweites Mal bitte: Wählen Sie "ja" bei den Links, die Sie auswählen würden, "nein" bei allen anderen.

ja nein

[Link nicht auswählen](#)

ja nein

[Diesen Link auswählen](#)

ja nein

[Link nicht auswählen](#)

ja nein

[Diesen Link auswählen](#)

ja nein

[Link nicht auswählen](#)

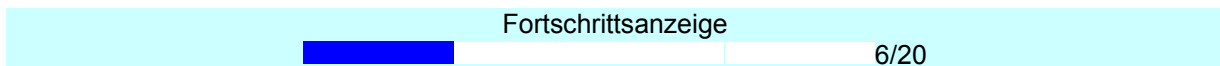
ja nein

[Link nicht auswählen](#)

ja nein

[Diesen Link auswählen](#)



**Instruktion Suchaufgabe (exemplarisch für die Bedingung spezifisch-Kontext, S6):**

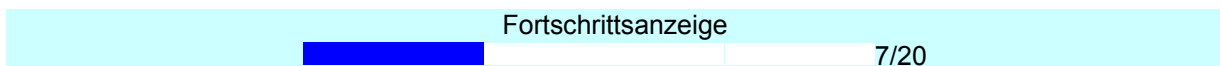
Sie wissen nun wie Sie auf den folgenden Seiten Links auswählen.

Stellen Sie sich bitte Folgendes vor: Sie suchen Antworten auf die Frage

**"Welche Größe haben die deutschen Euroscheine in Millimeter?"**

Auf der nächsten Seite sehen Sie das Ergebnis dieser Suchanfrage. Es ist eine Linkliste, wie sie eine Suchmaschine liefern könnte. Bei jedem Link ist für das dahinterliegende Dokument folgendes angegeben: **Titel und Zusammenhang in dem die gesuchten Wörter vorkommen**. Bei Links, die Sie auswählen würden klicken Sie "ja" an, ansonsten "nein".

Gehen Sie beim Auswählen möglichst wie bei einer echten Suche vor. Wieviele Links Sie auswählen, bleibt Ihnen überlassen. Treffen Sie ihre Entscheidungen **zügig** mit Hilfe der Titel und Texte.

**Ergebnisseite (Bedingung spezifisch-Kontext, S7a):**

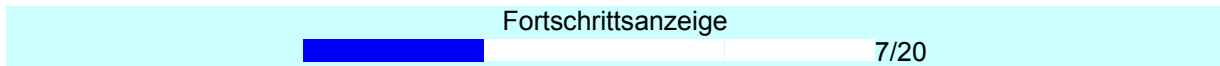
Die Suchanfrage lautet: Welche Größe haben die Euroscheine? Verwendete Suchwörter sind "Euroscheine Größe".

- ja  nein  [Gelton - Online-Banking Brokerage und mehr](#)  
  ... 5 Euro Größe: 120x62 Millimeter Farbe: grau Abbildung: Epoche "Klassik", 10 Euro  
 Größe: 127x67 Millimeter Farbe: rot Abbildung: Epoche "Romantik", ...
- ja  nein  [Allergate AG](#)  
  ... Zwirnhandschuhe Lohmann Rauscher Zwirnhandschuh (Euro-Größe 9 = dt. Gr. 7,5) 1  
 Paar aus weißer Baumwolle Lieferzeit: Versandfertig innerhalb von 10 Tagen ...
- ja  nein  [redmark.de - Euro Center](#)  
  ... in Stückelungen von 5, 10, 20, 50, 100, 200 und 500 Euro in Umlauf gesetzt. Größe und  
 Farbe der 7 Banknotenwerte sind unterschiedlich. Die Motive der ...
- ja  nein  [Orthopädie und Reha-Technik](#)  
  ... 0)3 41/9 41 15 16 Preis pro Zelt: 2,05 Euro (Größe 1), 3,07 Euro (Größe 2),  
 Erwachsene: 2,56 Euro, Ermäßigt: 1,79 Euro (Kinder u. Jugendliche 6-14 Jahre ...
- ja  nein  [Fünf Tester berichten über Ihre Praxis-Erfahrungen](#)  
  ... Bei meinen bisherigen Schuhen war die US-Größe jeweils 10 1/2, die der Euro-Größe  
 44 1/2 gleichzusetzen war. Bei diesem Schuh ist die US-Größe 10 1/2 ...
- ja  nein  [Willkommen auf "www.unser-burgenland.at"](#)  
  ... 5 Euro - Größe 120 mm x 62 mm. 10 Euro - Größe 127 mm x 67 mm. 20 Euro - Größe  
 133 mm x 72 mm. ...
- ja  nein  [Euro-Münzen](#)  
  ... und Zwei-Euro-Münzen vor. Designer: Die Entwürfe für die Münzen stammen von dem  
 Belgier Luc Luycx. Zwei-Euro-Münze, Größe: 25,25 mm x 1,95 mm (8,5 Gramm ...





### Ergebnisseite (Bedingung spezifisch-Textanfang, S7b):



Die Suchanfrage lautet: Welche Größe haben die Euroscheine? Verwendete Suchwörter sind "Euroscheine Größe".

- ja  nein  [Gelon - Online-Banking Brokerage und mehr](#)  
  5 Euro Größe: 120x62 Millimeter Farbe: grau Abbildung: Epoche "Klassik" 10 Euro Größe: 127x67 Millimeter Farbe: rot Abbildung: Epoche "Romantik" 20 Euro Größe: 133x72 Millimeter Farbe: blau Abbildung: Epoche "Gotik" 50 Euro Größe: 140x77 Millimeter Farbe: orange Abbildung: Epoche "Renaissance" 100 Euro Größe: 147x82 Millimeter Farbe: grün Abbildung: Epoche "Barock und Rokoko" 200 Euro Größe: 153x82...
- ja  nein  [Allergate AG](#)  
  Suche nach "handschuh": Spezialprodukte für Neurodermitiker Lotties Neurodermitis Handschuhe bis 2 Jahre Größe M Größe M Weiche Baumwollfäustlinge als Kratzschutz in der Nacht und am Tage, mit Hilfe eines Bändchens am Handgelenk zu fixieren. Lieferzeit: Versandfertig innerhalb von 10 Tagen. 3.53 Spezialprodukte für Neurodermitiker Lotties Neurodermitis Handschuhe 2-6 Jahre Größe L Größe L...
- ja  nein  [redmark.de - Euro Center](#)  
  EURO-Banknoten Am 1. Januar 2002 werden in 12 Mitgliedstaaten der Europäischen Union Banknoten in Stückelungen von 5, 10, 20, 50, 100, 200 und 500 Euro in Umlauf gesetzt. Größe und Farbe der 7 Banknotenwerte sind unterschiedlich. Die Motive der Banknoten sind europaweit einheitlich: Auf der Vorderseite der Banknoten symbolisieren Fenster und Tore den Geist der Offenheit und Zusammenarbeit in...
- ja  nein  [Orthopädie und Reha-Technik](#)  
  ORTHOPÄDIE + REHA-TECHNIK 11. Internationale Fachmesse und Weltkongress, 8. - 11. Mai 2002 11th International Trade Show and World Congress, May 8th to 11th, 2002 Besonders preiswerte Unterkünfte / Very reasonable accommodations Hostel Sleepy Lion Käthe-Kollwitz-Straße 3, D-04109 Leipzig Tel. +49(0)341/9 93 94 80, e-Mail info@hostelleipzig.de, Internet www.hostel-leipzig.de 50 Betten EZ 23 Euro...
- ja  nein  [Fünf Tester berichten über Ihre Praxis-Erfahrungen](#)  
  Fila FGT 340 Klaus Spielbücher: 1. Erfahrung (10. 4. 2001) Tester gesucht Leserbericht: Nach kurzer Wartezeit konnte ich schon den neuen FILA FGT 340 in Händen halten, und noch am selben Abend der Lieferung (Donnerstag, 5. April) konnte ich sie bei einem leichten Dauerlauf von 12 km Länge ausprobieren. Vom Anblick her, war ich positiv überrascht, durch die geradlinigen Streifen wirkt er optisch...
- ja  nein  [Willkommen auf "www.unser-burgenland.at"](#)  
  Banknoten Am 13.12.1996, wurden auf dem EU-Gipfel in Dublin die Muster der künftigen Euro-Banknoten veröffentlicht. Gleichzeitig wurden sie vom Europäischen Währungsinstitut (EWI) in Frankfurt/M. der Presse vorgestellt. Diese Banknoten sollen zum 01.01.2002 in Umlauf gebracht werden. Die nationalen Währungen werden noch längstens bis zum

30.06.2002 - in Österreich bis zum 28. Februar 2002...

ja  nein  [Euro-Münzen](#)

- Münzen Motiv: Die Vorderseiten der Euro-Münzen werden europaweit einheitlich gestaltet. Sie stellen - neben dem Wert in großen, klar lesbaren Zahlen - die Europäische Union auf vielfältige Art und Weise dar, vor dem dynamischen Hintergrund von europäischen Symbolen, nämlich Sternen. Die Ein- und Zwei-Euro-Münzen zeigen die Europäische Union ohne Grenzen. Die Gestaltung der Münz-Rückseiten blieb...



### Ergebnisseite (Bedingung spezifisch-Schlüsselwörter, S7c):

Fortschrittsanzeige

7/20

Die Suchanfrage lautet: Welche Größe haben die Euroscheine? Verwendete Suchwörter sind "Euroscheine Größe".

Ja  nein  [Gelon - Online-Banking Brokerage und mehr](#)

- Schlüsselwörter: Euro Größe Farbe Abbildung News Anbieter Vergleich

Ja  nein  [Allergate AG](#)

- Schlüsselwörter: Spezialprodukte Neurodermitiker Neurodermitis Handschuhe Latexfreie Gummihandschuhe Zwirnhandschuh

Ja  nein  [redmark.de - Euro Center](#)

- Schlüsselwörter: Euro Banknoten Europa EUR Größe Farbe Baustil

Ja  nein  [Orthopädie und Reha-Technik](#)

- Schlüsselwörter: Euro Größe Tel Leipzig Preis Zelt Betten

Ja  nein  [Fünf Tester berichten über Ihre Praxis-Erfahrungen](#)

- Schlüsselwörter: Euro Größe Schuhe FILA FGT Zehen laufen

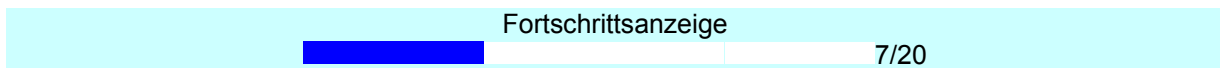
Ja  nein  [Willkommen auf "www.unser-burgenland.at"](#)

- Schlüsselwörter: mm Euro Größe Banknoten EU Umlauf europäisch

Ja  nein  [Euro-Münzen](#)

- Schlüsselwörter: mm Münze Euro Cent Größe Form Gramm



**Ergebnisseite (Bedingung spezifisch-MaxInfo, S7d):**

Die Suchanfrage lautet: Welche Größe haben die Euroscheine? Verwendete

Suchwörter sind "Euroscheine Größe".

Ja nein [Gelon - Online-Banking Brokerage und mehr](#)

- Euro Größe Farbe Abbildung News Anbieter Vergleich  
 ... 5 Euro Größe: 120x62 Millimeter Farbe: grau Abbildung: Epoche "Klassik", 10 Euro  
 Größe: 127x67 Millimeter Farbe: rot Abbildung: Epoche "Romantik", ...  
 5 Euro Größe: 120x62 Millimeter Farbe: grau Abbildung: Epoche "Klassik" 10 Euro Größe:  
 127x67 Millimeter Farbe: rot Abbildung: Epoche "Romantik" 20 Euro Größe: 133x72  
 Millimeter Farbe: blau Abbildung: Epoche "Gotik" 50 Euro Größe: 140x77 Millimeter Farbe:  
 orange Abbildung: Epoche "Renaissance" 100 Euro Größe: 147x82 Millimeter Farbe: grün  
 Abbildung: Epoche "Barock und Rokoko" 200 Euro Größe: 153x82...

Ja nein [Allergate AG](#)

- Spezialprodukte Neurodermitiker Neurodermitis Handschuhe Latexfreie  
 Gummihandschuhe Zwirnhandschuh  
 ... Zwirnhandschuhe Lohmann Rauscher Zwirnhandschuh (Euro-Größe 9 = dt. Gr. 7,5) 1  
 Paar aus weißer Baumwolle Lieferzeit: Versandfertig innerhalb von 10 Tagen ...  
 Suche nach "handschuh": Spezialprodukte für Neurodermitiker Lotties Neurodermitis  
 Handschuhe bis 2 Jahre Größe M Größe M Weiche Baumwollfäustlinge als Kratzschutz in  
 der Nacht und am Tage, mit Hilfe eines Bändchens am Handgelenk zu fixieren. Lieferzeit:  
 Versandfertig innerhalb von 10 Tagen. 3.53 Spezialprodukte für Neurodermitiker Lotties  
 Neurodermitis Handschuhe 2-6 Jahre Größe L Größe L...

Ja nein [redmark.de - Euro Center](#)

- Euro Banknoten Europa EUR Größe Farbe Baustil  
 ... in Stückelungen von 5, 10, 20, 50, 100, 200 und 500 Euro in Umlauf gesetzt. Größe und  
 Farbe der 7 Banknotenwerte sind unterschiedlich. Die Motive der ...  
 EURO-Banknoten Am 1. Januar 2002 werden in 12 Mitgliedstaaten der Europäischen  
 Union Banknoten in Stückelungen von 5, 10, 20, 50, 100, 200 und 500 Euro in Umlauf  
 gesetzt. Größe und Farbe der 7 Banknotenwerte sind unterschiedlich. Die Motive der  
 Banknoten sind europaweit einheitlich: Auf der Vorderseite der Banknoten symbolisieren  
 Fenster und Tore den Geist der Offenheit und Zusammenarbeit in...

Ja nein [Orthopädie und Reha-Technik](#)

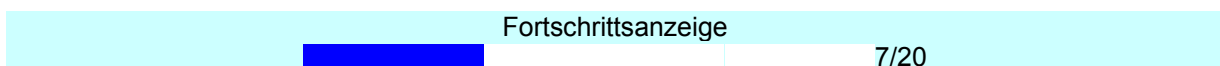
- Euro Größe Tel Leipzig Preis Zelt Betten  
 ... 0)3 41/9 41 15 16 Preis pro Zelt: 2,05 Euro (Größe 1), 3,07 Euro (Größe 2),  
 Erwachsene: 2,56 Euro, Ermäßigt: 1,79 Euro (Kinder u. Jugendliche 6-14 Jahre ...  
 ORTHOPÄDIE + REHA-TECHNIK 11. Internationale Fachmesse und Weltkongress, 8. -  
 11. Mai 2002 11th International Trade Show and World Congress, May 8th to 11th, 2002  
 Besonders preiswerte Unterkünfte / Very reasonable accommodations Hostel Sleepy Lion  
 Käthe-Kollwitz-Straße 3, D-04109 Leipzig Tel. +49(0)341/9 93 94 80, e-Mail  
 info@hostelleipzig.de, Internet www.hostel-leipzig.de 50 Betten EZ 23 Euro...

Ja nein [Fünf Tester berichten über Ihre Praxis-Erfahrungen](#)

- Euro Größe Schuhe FILA FGT Zehen laufen  
 ... Bei meinen bisherigen Schuhen war die US-Größe jeweils 10 1/2, die der Euro-Größe 44 1/2 gleichzusetzen war. Bei diesem Schuh ist die US-Größe 10 1/2 ...  
 Fila FGT 340 Klaus Spielbüchler: 1. Erfahrung (10. 4. 2001) Tester gesucht Leserbericht:  
 Nach kurzer Wartezeit konnte ich schon den neuen FILA FGT 340 in Händen halten, und noch am selben Abend der Lieferung (Donnerstag, 5. April) konnte ich sie bei einem leichten Dauerlauf von 12 km Länge ausprobieren. Vom Anblick her, war ich positiv überrascht, durch die geradlinigen Streifen wirkt er optisch...
- Ja nein [Willkommen auf "www.unser-burgenland.at"](#)
- mm Euro Größe Banknoten EU Umlauf europäisch  
 ... 5 Euro - Größe 120 mm x 62 mm. 10 Euro - Größe 127 mm x 67 mm. 20 Euro - Größe 133 mm x 72 mm. ...  
 Banknoten Am 13.12.1996, wurden auf dem EU-Gipfel in Dublin die Muster der künftigen Euro-Banknoten veröffentlicht. Gleichzeitig wurden sie vom Europäischen Währungsinstitut (EWI) in Frankfurt/M. der Presse vorgestellt. Diese Banknoten sollen zum 01.01.2002 in Umlauf gebracht werden. Die nationalen Währungen werden noch längstens bis zum 30.06.2002 - in Österreich bis zum 28. Februar 2002...
- Ja nein [Euro-Münzen](#)
- mm Münze Euro Cent Größe Form Gramm  
 ... und Zwei-Euro-Münzen vor. Designer: Die Entwürfe für die Münzen stammen von dem Belgier Luc Luycx. Zwei-Euro-Münze, Größe: 25,25 mm x 1,95 mm (8,5 Gramm ...  
 Münzen Motiv: Die Vorderseiten der Euro-Münzen werden europaweit einheitlich gestaltet. Sie stellen - neben dem Wert in großen, klar lesbaren Zahlen - die Europäische Union auf vielfältige Art und Weise dar, vor dem dynamischen Hintergrund von europäischen Symbolen, nämlich Sternen. Die Ein- und Zwei-Euro-Münzen zeigen die Europäische Union ohne Grenzen. Die Gestaltung der Münz-Rückseiten blieb...



### Ergebnisseite (Bedingung unspezifisch-Kontext, S7e):



Die Suchanfrage lautet: Welche wirtschaftlichen Folgen hat die Währungsumstellung? Verwendete Suchwörter sind "Euro wirtschaftliche Folgen".

- Ja nein [World Socialist Web Site](#)
- ... Kapital entgegenzutreten und die wirtschaftliche Entwicklung unter die ... muß der Einführung des Euro in der arbeitenden ... Perspektive und Orientierung folgen. ...
- Ja nein [Finanzpark - Capital & Feri AG](#)
- ... 2001 - 17:06] Capital & Feri AG: wirtschaftliche Folgen des Terroraktes. ... von 22 Dollar, einen auf 80 Euro-Cent erstarkten Dollar und eine Trotzreaktion ...
- Ja nein [Die Grünen](#)
- ... DER EURO-BARGELDEINFÜHRUNG: DIE FAKTEN. ... Bundesministerium für wirtschaftliche Angelegenheiten wurde ... Einführung des Euro zu keinen ... Mögliche Folgen für die ...

Ja  nein  [Mediko: Maul- und Klauenseuche](#)

... Der Ausbruch der Maul- und Klauenseuche (MKS ... Union verheerende wirtschaftliche Folgen: Allein in ... den MKS-Folgen betroffen sind ... 6 Milliarden Euro) geschätzt. Die ...

Ja  nein  [Wirtschaft - Spiegel Online](#)

... 08.08.2001). Rückruf-Aktion: Dramatische wirtschaftliche Folgen (08.08.2001). ... mit dem Arzneimittel 636 Millionen Euro Umsatz (plus 82 Prozent) erlöst ...

Ja  nein  [Contraste Hamburg](#)

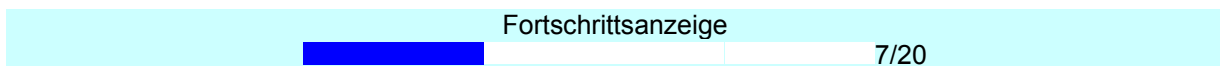
... Zu den Thesen der Befürworter - ... oben geschilderten Folgen. Die Hamburger ... kurzfristig vom EURO profitieren, Arbeitsplätze ... auf die wirtschaftliche Entwicklung ...

Ja  nein  [Rheinhochwasser](#)

... Folgen der Extremhochwasser. ... Wirtschaftliche Schäden des Dezemberhochwassers 1993. Es gibt drei Bereiche zu ... RP : 250 Mio. Euro. NRW : <100 Mio. Euro. BW : 70 Mio ...



#### Ergebnisseite (Bedingung unspezifisch-Textanfang, S7f):



Die Suchanfrage lautet: Welche wirtschaftlichen Folgen hat die Währungsumstellung? Verwendete Suchwörter sind "Euro wirtschaftliche Folgen".

Ja  nein  [World Socialist Web Site](#)

Die Einführung des Euro birgt erhebliche wirtschaftliche und soziale Konflikte Von Chris Talbot und Chris Marsden 23. Januar 1999 aus dem Englischen (21. Januar 1999) Die Einführung des Euro zu Beginn diesen Monats, mit der elf nationale Wirtschaften zu einer gemeinsamen Währungszone zusammengeschlossen werden, schafft einen Markt mit 300 Millionen Menschen, dessen Umfang nur noch von den...

Ja  nein  [Finanzpark - Capital & Feri AG](#)

[22.09.2001 - 17:06] Capital & Feri AG: wirtschaftliche Folgen des Terroraktes >DRUCKVERSION Die Terrorakte in den USA sind fuer die Weltwirtschaft weit weniger dramatisch als von vielen Experten befuerchtet. Vorausgesetzt, der Konflikt bleibt unter Kontrolle und eskaliert nicht in einen Weltkrieg, droht der Weltwirtschaft selbst im schlimmsten Fall nicht einmal ein Jahr Stillstand. Das impliziert...

Ja  nein  [Die Grünen](#)

GELD SORGEN! Die Euro-Bargeldeinführung und ihre Folgen: Ab 1.Jänner 2002 werden wir mit dem Euro-Bargeld bezahlen und ab 1.März 2002 wird der Schilling dann entgültig verschwinden. Der Euro wird alleiniges gesetzliches Zahlungsmittel. Umso näher der Tag der Euro-Bargeldeinführung rückt, desto lauter werden die Stimmen jener, die vor negativen Folgen sowohl für KonsumentInnen als auch Klein- und...

- Ja  nein  [Mediko: Maul- und Klauenseuche](#)  
  Maul- und Klauenseuche hat verheerende wirtschaftliche Folgen - Neben Landwirtschaft leidet auch britischer Tourismus 20.06.2001 - Der Ausbruch der Maul- und Klauenseuche (MKS) hat für mehrere Länder der Europäischen Union verheerende wirtschaftliche Folgen: Allein in Großbritannien, wo die Seuche am schlimmsten wütete, wurden mehr als 3,2 Millionen Tiere getötet, darunter 2,6 Millionen Schafe...
- Ja  nein  [Wirtschaft - Spiegel Online](#)  
  GEFÄHRLICHES MEDIKAMENT Bayer droht Millionenklage Spastische Zuckungen, Krämpfe, Nierenversagen: Der Pharmakonzern Bayer muss eine cholesterinsenkende Arznei wegen möglicherweise lebensgefährlicher Nebenwirkungen vom Markt nehmen. Nun wollen Geschädigte in den USA den Konzern offenbar verklagen. DPA Der Cholesterinsenker Baycol/Lipobay: Möglicherweise mit tödlichen Nebenwirkungen Leverkusen -...
- Ja  nein  [Contraste Hamburg](#)  
  Die Folgen des Euro Zu den Thesen der Befürworter - Als nächster Schritt der Weiterentwicklung der EU kommt die gemeinsame Währung, der EURO. Welche Wirkungen der EURO auf die politische und ökonomische Situation hat, ist nicht nur in Deutschland umstritten. Von den Befürwortern der schnellen Einführung des EURO wird behauptet, die Währungsumstellung sei keine Währungsreform und auch mit der...
- Ja  nein  [Rheinhochwasser](#)  
  Rheinhochwasser - Folgen der Extremhochwasser - Während Niedrigwasser so gut wie keine Schäden anrichten, ist der Schaden durch ein Hochwasser immens. Besonders für das Extremhochwasser 1993/1994 liegen Bilanzen vor. Im Rheingebiet brachten Tiefdruckgebiete schon Anfang Dezember viel Regen, so daß die Nebenflüsse und auch der Rheinpegel deutlich anstiegen. Schon in den ersten Dezembertagen...



### Ergebnisseite (Bedingung unspezifisch-Schlüsselwörter, S7g):



Die Suchanfrage lautet: Welche wirtschaftlichen Folgen hat die Währungsumstellung? Verwendete Suchwörter sind "Euro wirtschaftliche Folgen".

- Ja  nein  [World Socialist Web Site](#)  
  Euro Dollar Folgen Prozent Frage Einführung Wirtschaft
- Ja  nein  [Finanzpark - Capital & Feri AG](#)  
  USA Terrorakte Jahr Prozent Weltwirtschaft Zinssenkungen Oelpreis
- Ja  nein  [Die Grünen](#)  
  Euro Folgen Grünen Bargeldeinführung Schilling Zeitplan Preis
- Ja  nein  [Mediko: Maul- und Klauenseuche](#)  
  Maul Klauenseuche Millionen Milliarden Euro Niederlanden Großbritannien

Ja  nein  [Wirtschaft - Spiegel Online](#)

Euro Patienten Bayer Nebenwirkungen Jahr Arznei Millionen

Ja  nein  [Contraste Hamburg](#)

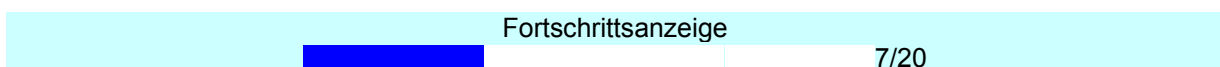
Euro Wahrung Kriterien Wechselkurs Devisen Wirkung Konvergenzkriterien

Ja  nein  [Rheinhochwasser](#)

Euro Bereich Mio Extremhochwasser Jahre Hochwasser Wassermenge



### Ergebnisseite (Bedingung unspezifisch-MaxInfo, S7h):



Die Suchanfrage lautet: Welche wirtschaftlichen Folgen hat die Wahrungsumstellung? Verwendete Suchwortersind "Euro wirtschaftliche Folgen".

Ja  nein  [World Socialist Web Site](#)

Euro Dollar Folgen Prozent Frage Einfuhrung Wirtschaft  
 ... Kapital entgegenzutreten und die wirtschaftliche Entwicklung unter die ... mu der Einfuhrung des Euro in der arbeitenden ... Perspektive und Orientierung folgen. ... Die Einfuhrung des Euro birgt erhebliche wirtschaftliche und soziale Konflikte Von Chris Talbot und Chris Marsden 23. Januar 1999 aus dem Englischen (21. Januar 1999) Die Einfuhrung des Euro zu Beginn diesen Monats, mit der elf nationale Wirtschaften zu einer gemeinsamen Wahrungszone zusammengeschlossen werden, schafft einen Markt mit 300 Millionen Menschen, dessen Umfang nur noch von den...

Ja  nein  [Finanzpark - Capital & Feri AG](#)

USA Terrorakte Jahr Prozent Weltwirtschaft Zinssenkungen Oelpreis  
 ... 2001 - 17:06] Capital & Feri AG: wirtschaftliche Folgen des Terroraktes. ... von 22 Dollar, einen auf 80 Euro-Cent erstarkten Dollar und eine Trotzreaktion ... [22.09.2001 - 17:06] Capital & Feri AG: wirtschaftliche Folgen des Terroraktes  
 >DRUCKVERSION Die Terrorakte in den USA sind fuer die Weltwirtschaft weit weniger dramatisch als von vielen Experten befuerchtet. Vorausgesetzt, der Konflikt bleibt unter Kontrolle und eskaliert nicht in einen Weltkrieg, droht der Weltwirtschaft selbst im schlimmsten Fall nicht einmal ein Jahr Stillstand. Das impliziert...

Ja  nein  [Die Grunen](#)

Euro Folgen Grunen Bargeldeinfuhrung Schilling Zeitplan Preis  
 ... DER EURO-BARGELDEINFUHRUNG: DIE FAKTEN. ... Bundesministerium fur wirtschaftliche Angelegenheiten wurde ... Einfuhrung des Euro zu keinen ... Mogliche Folgen fur die ...  
 GELD SORGEN! Die Euro-Bargeldeinfuhrung und ihre Folgen: Ab 1.Janner 2002 werden wir mit dem Euro-Bargeld bezahlen und ab 1.Marz 2002 wird der Schilling dann entgultig verschwinden. Der Euro wird alleiniges gesetzliches Zahlungsmittel. Umso naher der Tag der Euro-Bargeldeinfuhrung ruckt, desto lauter werden die Stimmen jener, die vor negativen Folgen sowohl fur KonsumentInnen als auch Klein- und...



Ja  nein  [Mediko: Maul- und Klauenseuche](#)

- Maul Klauenseuche Millionen Milliarden Euro Niederlanden Großbritannien  
 ... Der Ausbruch der Maul- und Klauenseuche (MKS ... Union verheerende wirtschaftliche Folgen: Allein in ... den MKS-Folgen betroffen sind ... 6 Milliarden Euro) geschätzt. Die ... Maul- und Klauenseuche hat verheerende wirtschaftliche Folgen - Neben Landwirtschaft leidet auch britischer Tourismus 20.06.2001 - Der Ausbruch der Maul- und Klauenseuche (MKS) hat für mehrere Länder der Europäischen Union verheerende wirtschaftliche Folgen: Allein in Großbritannien, wo die Seuche am schlimmsten wütete, wurden mehr als 3,2 Millionen Tiere getötet, darunter 2,6 Millionen Schafe...

Ja  nein  [Wirtschaft - Spiegel Online](#)

- Euro Patienten Bayer Nebenwirkungen Jahr Arznei Millionen  
 ... 08.08.2001). Rückruf-Aktion: Dramatische wirtschaftliche Folgen (08.08.2001). ... mit dem Arzneimittel 636 Millionen Euro Umsatz (plus 82 Prozent) erlöst ... GEFÄHRLICHES MEDIKAMENT Bayer droht Millionenklage Spastische Zuckungen, Krämpfe, Nierenversagen: Der Pharmakonzern Bayer muss eine cholesterinsenkende Arznei wegen möglicherweise lebensgefährlicher Nebenwirkungen vom Markt nehmen. Nun wollen Geschädigte in den USA den Konzern offenbar verklagen. DPA Der Cholesterinsenker Baycol/Lipobay: Möglicherweise mit tödlichen Nebenwirkungen Leverkusen -...

Ja  nein  [Contraste Hamburg](#)

- Euro Währung Kriterien Wechselkurs Devisen Wirkung Konvergenzkriterien  
 ... Zu den Thesen der Befürworter - ... oben geschilderten Folgen. Die Hamburger ... kurzfristig vom EURO profitieren, Arbeitsplätze ... auf die wirtschaftliche Entwicklung ... Die Folgen des Euro Zu den Thesen der Befürworter - Als nächster Schritt der Weiterentwicklung der EU kommt die gemeinsame Währung, der EURO. Welche Wirkungen der EURO auf die politische und ökonomische Situation hat, ist nicht nur in Deutschland umstritten. Von den Befürwortern der schnellen Einführung des EURO wird behauptet, die Währungsumstellung sei keine Währungsreform und auch mit der...

Ja  nein  [Rheinhochwasser](#)

- Euro Bereich Mio Extremhochwasser Jahre Hochwasser Wassermenge  
 ... Folgen der Extremhochwasser. ... Wirtschaftliche Schäden des Dezemberhochwassers 1993. Es gibt drei Bereiche zu ... RP : 250 Mio. Euro. NRW : <100 Mio. Euro. BW : 70 Mio ...  
 Rheinhochwasser - Folgen der Extremhochwasser - Während Niedrigwasser so gut wie keine Schäden anrichten, ist der Schaden durch ein Hochwasser immens. Besonders für das Extremhochwasser 1993/1994 liegen Bilanzen vor. Im Rheingebiet brachten Tiefdruckgebiete schon Anfang Dezember viel Regen, so daß die Nebenflüsse und auch der Rheinpegel deutlich anstiegen. Schon in den ersten Dezembertagen...



**Erhebung Entscheidungssicherheit (exemplarisch für die Bedingung spezifisch-Kontext, S8):**



Unten sehen Sie die selbe Linkliste wie eben zur Frage "Welche Größe haben die Euroscheine". Bitte geben Sie nun Ihre Einschätzung für jeden Link ab: Wie wahrscheinlich führt dieser Link zu



den gesuchten Informationen?

#### [Gelon - Online-Banking Brokerage und mehr](#)

... 5 Euro Größe: 120x62 Millimeter Farbe: grau Abbildung: Epoche "Klassik", 10 Euro Größe: 127x67 Millimeter Farbe: rot Abbildung: Epoche "Romantik", ...

sehr wahr scheinlich	wahr- scheinlich	unklar/ unsicher	unwahr- scheinlich	sehr unwahr- scheinlich
2	1	0	-1	-2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

#### [Allergate AG](#)

... Zwirnhandschuhe Lohmann Rauscher Zwirnhandschuh (Euro-Größe 9 = dt. Gr. 7,5) 1 Paar aus weißer Baumwolle Lieferzeit: Versandfertig innerhalb von 10 Tagen ...

sehr wahr scheinlich	wahr- scheinlich	unklar/ unsicher	unwahr- scheinlich	sehr unwahr- scheinlich
2	1	0	-1	-2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

#### [redmark.de - Euro Center](#)

... in Stückelungen von 5, 10, 20, 50, 100, 200 und 500 Euro in Umlauf gesetzt. Größe und Farbe der 7 Banknotenwerte sind unterschiedlich. Die Motive der ...

sehr wahr scheinlich	wahr- scheinlich	unklar/ unsicher	unwahr- scheinlich	sehr unwahr- scheinlich
2	1	0	-1	-2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

#### [Orthopädie und Reha-Technik](#)

... 0)3 41/9 41 15 16 Preis pro Zelt: 2,05 Euro (Größe 1), 3,07 Euro (Größe 2), Erwachsene: 2,56 Euro, Ermäßigt: 1,79 Euro (Kinder u. Jugendliche 6-14 Jahre ...

sehr wahr scheinlich	wahr- scheinlich	unklar/ unsicher	unwahr- scheinlich	sehr unwahr- scheinlich
2	1	0	-1	-2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

#### [Fünf Tester berichten über Ihre Praxis-Erfahrungen](#)

... Bei meinen bisherigen Schuhen war die US-Größe jeweils 10 1/2, die der Euro-Größe 44 1/2 gleichzusetzen war. Bei diesem Schuh ist die US-Größe 10 1/2 ...

sehr wahr scheinlich	wahr- scheinlich	unklar/ unsicher	unwahr- scheinlich	sehr unwahr- scheinlich
2	1	0	-1	-2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Willkommen auf "www.unser-burgenland.at"](#)

... 5 Euro - Größe 120 mm x 62 mm. 10 Euro - Größe 127 mm x 67 mm. 20 Euro - Größe 133 mm x 72 mm. ...

sehr wahr scheinlich	wahr- scheinlich	unklar/ unsicher	unwahr- scheinlich	sehr unwahr- scheinlich
2	1	0	-1	-2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Euro-Münzen](#)

... und Zwei-Euro-Münzen vor. Designer: Die Entwürfe für die Münzen stammen von dem Belgier Luc Luycx. Zwei-Euro-Münze, Größe: 25,25 mm x 1,95 mm (8,5 Gramm ...

sehr wahr scheinlich	wahr- scheinlich	unklar/ unsicher	unwahr- scheinlich	sehr unwahr- scheinlich
2	1	0	-1	-2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



**Fragebogen 1, Treatmentcheck, Zufriedenheit (S9):**

Fortschrittsanzeige

9/20

Bei den folgenden Fragen geht es ausschließlich um Ihre persönliche Meinung. Es gibt keine richtigen oder falschen Antworten. Nur Ihre Meinung ist wichtig. Bitte geben Sie an wie weit Sie den folgenden Aussagen zustimmen oder ablehnen.

Wie ist Ihre Meinung zur eben durchgeführten Suche?

Die Suche finde ich...

	sehr	eher	teils/ teils	eher	sehr	
	-2	-1	0	1	2	
langweilig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	interessant
schwierig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	leicht
komplex	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	einfach

Bei einer echten Suche, hätte ich mir noch weitere Links anzeigen lassen

- ja
- nein

Hätten Sie andere Suchwörter eingegeben? Falls ja, welche?

Die Reihenfolge (das Ranking) der Links war schlecht

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die Linkliste ließ sich leicht lesen

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die Linkliste war für die Suche relevant

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Es war schwierig zwischen den Links auszuwählen

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die Linkliste war frustrierend

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Mit den angebotenen Links war ich zufrieden

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Fragebogen 2, Zufriedenheit (S10):

Fortschrittsanzeige

10/20

Bitte geben Sie an wie weit Sie den folgenden Aussagen zustimmen oder ablehnen.

Die Texte waren unübersichtlich

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Die Texte waren unverständlich

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Die Texte waren präzise

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Die Texte unterstützten schnelle Entscheidungen

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Die Texte waren unbrauchbar

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Der Zusammenhang der Suchbegriffe mit den Links war nicht erkennbar

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Die Texte waren zusammenhängend

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Es hat Spaß gemacht Links auszuwählen

lehne stark ab	lehne ab	weder noch	stimme zu	stimme stark zu
-2	-1	0	1	2
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



**Instruktion Suchaufgabe (exemplarisch für die Bedingung unspezifisch-Kontext, S11):**

Fortschrittsanzeige

11/20

Die vorherige Suche haben Sie erfolgreich abgeschlossen.

Für die nächste Suchanfrage stellen Sie sich bitte Folgendes vor: Sie suchen Antworten auf die Frage **"Welche wirtschaftlichen Folgen hat die Währungsumstellung?"**

Auf der nächsten Seite sehen Sie das Ergebnis dieser Suchanfrage. Es ist eine Linkliste, wie sie eine Suchmaschine liefern könnte. Bei jedem Link ist für das dahinterliegende Dokument folgendes angegeben: **Titel und Zusammenhang in dem die gesuchten Wörter vorkommen**. Bei Links, die Sie auswählen würden klicken Sie "ja" an, ansonsten "nein".

Gehen Sie beim Auswählen möglichst wie bei einer echten Suche vor. Wieviele Links Sie auswählen, bleibt Ihnen überlassen. Treffen Sie ihre Entscheidungen **zügig** mit Hilfe der Titel und Texte.



**Ergebnisseite (S12): vergleiche S7.**

**Erhebung Entscheidungssicherheit (S13): vergleiche S8.**

**Fragebogen 1, Treatmentcheck, Zufriedenheit (S14): siehe S9.**

**Fragebogen 2, Zufriedenheit (S15): siehe S10.**

**Direkter Vergleich aller Beschreibungsvarianten bei unspezifischer Suchanfrage und relevantem Treffer (S16):**

Fortschrittsanzeige

16/20

Im Folgenden sehen Sie Links zur Suchanfrage "Welche wirtschaftlichen Folgen hat die Währungsumstellung". Die vier Links verweisen auf das gleiche Dokument, in dem es Antworten zur Frage gibt. Jeder Link zeigt jedoch unterschiedliche Informationen des Dokuments an. Welche Darstellung würden Sie bei einer Suchmaschine bevorzugen? Beantworten Sie bitte:

**Wie gut oder schlecht ist die jeweilige Darstellung des Links?**

**Titel und Zusammenhang in dem die gesuchten Wörter vorkommen:**

[Contraste Hamburg](#)

... Zu den Thesen der Befürworter -. ... oben geschilderten Folgen. Die Hamburger ... kurzfristig vom EURO profitieren, Arbeitsplätze ... auf die wirtschaftliche Entwicklung ...

sehr gut      gut      teils/teils      schlecht      sehr schlecht

1              2              3              4              5

**Titel und die ersten 400 Zeichen:**[Contraste Hamburg](#)

Die Folgen des Euro Zu den Thesen der Befürworter - Als nächster Schritt der Weiterentwicklung der EU kommt die gemeinsame Währung, der EURO. Welche Wirkungen der EURO auf die politische und ökonomische Situation hat, ist nicht nur in Deutschland umstritten. Von den Befürwortern der schnellen Einführung des EURO wird behauptet, die Währungsumstellung sei keine Währungsreform und auch mit der...

sehr gut	gut	teils/teils	schlecht	sehr schlecht
1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Titel und die 7 wichtigsten Schlüsselwörter:**[Contraste Hamburg](#)

Euro Währung Kriterien Wechselkurs Devisen Wirkung Konvergenzkriterien

sehr gut	gut	teils/teils	schlecht	sehr schlecht
1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Titel, Zusammenhang in dem die gesuchten Wörter vorkommen, die 7 wichtigsten Schlüsselwörter und die ersten 400 Zeichen:**[Contraste Hamburg](#)

Euro Währung Kriterien Wechselkurs Devisen Wirkung Konvergenzkriterien

... Zu den Thesen der Befürworter - ... oben geschilderten Folgen. Die Hamburger ... kurzfristig vom EURO profitieren, Arbeitsplätze ... auf die wirtschaftliche Entwicklung ...

Die Folgen des Euro Zu den Thesen der Befürworter - Als nächster Schritt der Weiterentwicklung der EU kommt die gemeinsame Währung, der EURO. Welche Wirkungen der EURO auf die politische und ökonomische Situation hat, ist nicht nur in Deutschland umstritten. Von den Befürwortern der schnellen Einführung des EURO wird behauptet, die Währungsumstellung sei keine Währungsreform und auch mit der...

sehr gut	gut	teils/teils	schlecht	sehr schlecht
1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Direkter Vergleich aller Beschreibungsvarianten bei unspezifischer Suchanfrage und irrelevantem Treffer (S17):**

Fortschrittsanzeige

17/20

Vergleichen Sie bitte auch die folgenden vier Links zur selben Suchanfrage "Welche wirtschaftlichen Folgen hat die Währungsumstellung". Die vier Links verweisen auf das gleiche Dokument, in dem es **keine** Antworten zur Frage gibt. Welche Darstellung würden Sie bei einer Suchmaschine bevorzugen? Beantworten Sie bitte:

**Wie gut oder schlecht ist die jeweilige Darstellung des Links?**

**Titel und Zusammenhang in dem die gesuchten Wörter vorkommen:**[Wirtschaft - Spiegel Online](#)

... 08.08.2001). Rückruf-Aktion: Dramatische wirtschaftliche Folgen (08.08.2001). ... mit dem Arzneimittel 636 Millionen Euro Umsatz (plus 82 Prozent) erlöst ...

sehr gut	gut	teils/teils	schlecht	sehr schlecht
1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Titel und die ersten 400 Zeichen:**[Wirtschaft - Spiegel Online](#)

GEFÄHRLICHES MEDIKAMENT Bayer droht Millionenklage Spastische Zuckungen, Krämpfe, Nierenversagen: Der Pharmakonzern Bayer muss eine cholesterinsenkende Arznei wegen möglicherweise lebensgefährlicher Nebenwirkungen vom Markt nehmen. Nun wollen Geschädigte in den USA den Konzern offenbar verklagen. DPA Der Cholesterinsenker Baycol/Lipobay: Möglicherweise mit tödlichen Nebenwirkungen Leverkusen -...

sehr gut	gut	teils/teils	schlecht	sehr schlecht
1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Titel und die 7 wichtigsten Schlüsselwörter:**[Wirtschaft - Spiegel Online](#)

Euro Patienten Bayer Nebenwirkungen Jahr Arznei Millionen

sehr gut	gut	teils/teils	schlecht	sehr schlecht
1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Titel, Zusammenhang in dem die gesuchten Wörter vorkommen, die 7 wichtigsten**[Schlüsselwörter und die ersten 400 Zeichen:](#)

Wirtschaft - Spiegel Online

Euro Patienten Bayer Nebenwirkungen Jahr Arznei Millionen

... 08.08.2001). Rückruf-Aktion: Dramatische wirtschaftliche Folgen (08.08.2001). ... mit dem Arzneimittel 636 Millionen Euro Umsatz (plus 82 Prozent) erlöst ...

GEFÄHRLICHES MEDIKAMENT Bayer droht Millionenklage Spastische Zuckungen, Krämpfe, Nierenversagen: Der Pharmakonzern Bayer muss eine cholesterinsenkende Arznei wegen möglicherweise lebensgefährlicher Nebenwirkungen vom Markt nehmen. Nun wollen Geschädigte in den USA den Konzern offenbar verklagen. DPA Der Cholesterinsenker Baycol/Lipobay: Möglicherweise mit tödlichen Nebenwirkungen Leverkusen -...

sehr gut	gut	teils/teils	schlecht	sehr schlecht
1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



**Abschlussfragen, Anmerkungen (S18):**

Fortschrittsanzeige

18/20

Bei Onlineuntersuchungen gibt es viele "Durchklicker", die unbrauchbare Daten erzeugen. Wählen Sie deshalb bitte

Ich habe ernsthaft an der Untersuchung teilgenommen.

Falls Sie Ihre bisherige Mitarbeit zurückziehen möchten, geben Sie folgendes an

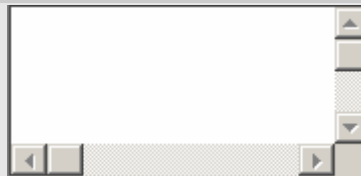
Meine Daten sollen nicht ausgewertet werden.

Nun noch vier letzte Fragen...

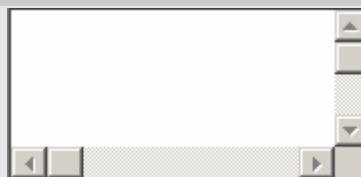
Wie sind Sie als erstes auf die Untersuchung aufmerksam geworden?

- Uni Mannheim Webseiten
- Über eine Suchmaschine
- anderer Link auf Webseiten
- Mailingliste
- Forum oder Newsgroup
- persönliche Email
- mündlich
- Aushang
- Anderes

Die Texte zu den Links und deren Titel waren Zusammenfassungen der dahinterliegenden Seiten. Wie sehen Zusammenfassungen aus, an denen Sie erkennen können, ob sich der Link lohnt?



Welche weiteren Informationen nutzen Sie bei der Entscheidung, ob ein Link angeklickt wird oder nicht?



Haben Sie weitere Anmerkungen zur Untersuchung?





**Debriefing (S19):**

Fortschrittsanzeige

19/20

Vielen Dank für ihre Teilnahme an der Untersuchung.

Diese Untersuchung möchte mehrere Fragen beantworten:

- Welche Darstellung der Linkliste ist für die Nutzer optimal?
- Welchen Einfluss haben unterschiedlich komplexe Suchanfragen (Euroscheingröße vs. wirtschaftliche Folgen) auf die Suche?
- Welche Rolle spielen unterschiedliche Fähigkeiten und Bedürfnisse der Suchenden?
- Wie gut eignen sich die hier entwickelten Untersuchungsmethoden?

Falls Sie benachrichtigt werden möchten, sobald die Auswertung abgeschlossen ist, können Sie hier ihre Emailadresse angeben. Sie wird unabhängig von den restlichen Daten gespeichert, so daß Ihre Anonymität gesichert ist.

Email: 

Für Fragen, Anregungen, Diskussionen [schreiben Sie an den Untersuchungsleiter](#), Lars Kaczmirek. Auf der nächsten Seite erhalten Sie Antworten zu den gestellten Suchanfragen.

**Linkliste mit Antworten zu den Suchaufgaben (S20):**

Fortschrittsanzeige

20/20

Unter den folgenden Links finden Sie ausgewählte Antworten zu den gestellten Suchanfragen ohne Anspruch auf Vollständigkeit.

**Welche Größe haben die Euroscheine?**

Bilder der Euroscheine mit Größenangaben hinter den Links: [1](#) [2](#) [3](#)

**Welche wirtschaftlichen Folgen hat die Währungsumstellung?**

Verschiedene Artikel zu den Folgen hinter den Links: [1](#) [2](#) [3](#)



## Glossar

Bit	Kleinste binäre Informationseinheit
Browser	Computerprogramm mit dem
CGI	Common Gateway Interface. Ein Serverseitiges Programm das bestimmte Aufgaben in der Interaktion mit Internetsurfern erledigt, z.B. Abspeichern von Formulardaten.
Cookie	Eine kleine Datei, die von einem Internetanbieter auf dem Computer eines Internetsurfers abgelegt wurde, um über seinen Besuch oder ihn selbst Informationen zu speichern.
Crawler	siehe Robot.
Dienst	Ein Unterbereich des Internet, z.B. WWW, Newsgroup, E-Mail, FTP.
FTP	File Transfer Protocol. Ein Dienst des Internet. Protokoll zur Datenübertragung.
HTML	HyperText Markup Language. Eine Beschreibungssprache, um die strukturellen Bestandteile eines Texts festlegen zu können. Vergleiche Tag.
Internet	Die Verbindung zahlreicher Computer weltweit, die verschiedene Dienste unterstützen.
IR	Information Retrieval. Methoden und Werkzeuge zum Auffinden von Informationen in großen Datenmengen.
Klick	Auswahl eines Elements mit dem Mauszeiger.
Newsgroup	Diskussionsforum. Ein Dienst des Internet.
OPAC	Online Public Access Catalogue. Bibliothekssystem an Universitäten.
PageRank-Technologie™	Patentierter Ranking Technologie von Google für die Sortierung der Treffer auf Ergebnisseiten.
Robot	Computerprogramm, das automatisch Webseiten besuchen und Informationen von diesen sammeln kann.
Spider	siehe Robot.
Tag	Strukturelement von HTML, das zur Beschreibung von Text dient, z.B. <h1>Überschrift</1>.
URL	Uniform Resource Locator. Eindeutige Adresse eines Webdokuments.
WWW	World Wide Web. Ein Dienst des Internet. Webseiten des Internet die mit Browsern dargestellt werden können.

## Autorenverzeichnis

- Abowd, G. 24, 91  
Amitay, E. 12, 14, 90  
Auble, R. S. 33, 90  
Baeza-Yates, R. 2, 8, 10, 11, 18, 22, 23, 35, 90  
Baker, R. 16, 90  
Banke, K. 49, 90  
Baphna, J. 17, 18, 94  
Bateman, J. 32, 97  
Bates, M. J. 22, 90  
Batinic, B. 43, 71, 90, 91, 93, 96, 97  
Beale, R. 24, 91  
Bekavac, B. 8, 10, 11, 90  
Berger, M. C. 48, 98  
Bernard, M. 15, 16, 90  
Bettman, J. R. 28, 30, 96  
Bloom, B. S. 20, 90  
Borgman, C. L. 31, 32, 90  
Bortz, J. 48, 55, 60, 61, 63, 67, 79, 90  
Bosnjak, M. 43, 53, 65, 67, 71, 90, 91  
Boyes, A. J. 33, 91  
Brace, N. 60, 61, 91  
Brecht, M. 54, 98  
Brem, S. 33, 91  
Bruza, P. D. 18, 23, 35, 95  
Buchner, A. 85, 91  
Byrd, D. 97  
Campbell, D. T. 62, 66, 67, 86, 91  
Cannata, P. E. 10, 33, 93  
Capstick, J. 2, 33, 39, 40, 43, 51, 97  
Card, S. K. 22, 96  
Chaparro, B. 15, 90  
Chen, H. 14, 92  
Choo, C. W. 2, 22, 91  
Cohen, J. 43, 84, 85, 91  
Cook, T. D. 62, 66, 67, 86, 91  
Cooke, A. 2, 91  
Couper, M. P. 46, 86, 91  
Croft, W. B. 22, 97  
Cutrell, E. 14, 92  
Davies, R. K. 33, 92  
DeLongchamp, D. L. 33, 90  
DeMicheli, M. 21, 91  
Detlor, B. 2, 22, 91  
Devlin, S. J. 10, 33, 93  
Diaper, D. 24, 91  
Dillman, D. A. 46, 54, 91  
Dimitroff, A. 58, 98  
Dix, A. 24, 91  
Dodge, M. 11, 92  
Döring, N. 48, 55, 90  
Douglass, M. H. 33, 92  
Drori, O. 2, 10, 11, 15, 33, 40, 51, 58, 87, 92  
Duda, A. 14, 98  
Dumais, S. 14, 92  
Dutke, S. 31, 92  
Eichstaedt, J. 52, 92  
Endres-Niggemeyer, B. 12, 92  
Erbach, G. 2, 33, 39, 40, 43, 51, 92, 97  
Erdfelder, E. 39, 84, 85, 91, 92  
Fang, X. 79, 92  
Faul, F. 39, 84, 85, 91, 92  
Fernandez, M. 16, 90  
Fidel, R. 33, 92  
Finley, J. 24, 91  
Fischer, K. 29, 30, 94  
Fox, S. 1, 92  
Garber, S. R. 8, 92  
Giles, C. L. 1, 21, 32, 95  
Gomez, L. M. 10, 33, 93  
Goodrum, A. 8, 94  
Gourmelon, A. 37, 93  
Greene, S. L. 10, 33, 93  
Groner, R. 22, 31, 32, 98  
Grossan, B. 10, 32, 93  
Grundman, D. 14, 93  
Grunes, M. B. 8, 92  
Gunn, H. 33, 93  
Hannum, W. H. 24, 94  
Harter, S. P. 56, 62, 93

- Hartmann, W. 33, 93  
Hassenzahl, M. 33, 48, 93  
Hauptmanns, P. 70, 93  
Heinrich, L. J. 6, 93  
Henzinger, M. 8, 97  
Hock, R. 43, 93  
Hockley, A. 32, 33, 96  
Hoelscher, C. 22, 32, 93  
Holder, J. K. 33, 92  
Hopkins, C. J. 33, 92  
Hsieh-Yee, I. 33, 48, 93  
Hurtienne, J. 28, 98  
Janetzko, D. 52, 93  
Jansen, B. J. 2, 8, 32, 94, 97  
John, B. E. 22, 91, 96  
Johnson, E. J. 28, 30, 96  
Jonassen, D. H. 24, 94  
Jones, M. 23, 97  
Jungermann, H. 29, 30, 94  
Karzauninkat, S. 1, 2, 9, 10, 18, 72, 94  
Kehoe, C. M. 71, 96  
Kemp, R. 60, 61, 91  
Khan, K. 16, 28, 79, 94  
Kitchin, R. 11, 92  
Kleinberg, J. M. 2, 14, 32, 94  
Kleinfeldt, S. 17, 18, 94  
Konrad, E. 56, 94  
Körber, S. 44, 94  
Körndle, H. 49, 94  
Koster, M. 10, 95  
Kuhlen, R. 23, 94, 95  
Kushner, E. J. 33, 92  
Lamias, M. J. 46, 91  
Lawrence, S. 1, 21, 32, 91, 95  
Lennartz, S. 13, 95  
Lergier, R. 2, 19, 25, 29, 30, 40, 41, 43, 49, 51, 58, 95  
Lienert, G. A. 46, 75, 95  
Little, R. J. 64, 65, 95  
Locatis, C. 16, 28, 79, 94  
Maaß, S. 6, 95  
Marchionini, G. 19, 20, 21, 33, 95  
McArthur, R. 18, 23, 35, 95  
McDonald, D. 48, 98  
Michalek, A. 14, 93  
Miller, R. B. 24, 95  
Morville, P. 11, 21, 97  
Münz, S. 12, 13, 95  
Müsseler, J. 33, 95  
Näf, M. 33, 93  
Namprempre, C. 14, 98  
Navarro-Prieto, R. 6, 20, 21, 22, 33, 95  
Neumann, G. 39, 92  
Nielsen, J. 1, 32, 33, 37, 96  
Noller, S. 31, 32, 96  
Notess, G. R. 9, 33, 37, 96  
Ohmura, K. 28, 97  
Ohnemus, K. R. 34, 37, 96  
Panero, J. C. 32, 96  
Paris, C. 12, 90  
Payne, J. W. 28, 30, 96  
Peck, V. A. 22, 96  
Pfiff, A. 32, 94  
Pfister, H.-R. 29, 30, 94  
Pior, R. 33, 96  
Pirelli, P. 22, 96  
Pittkow, J. E. 96  
Pollock, A. 32, 33, 96  
Pooch, U. 2, 94  
Prinz, W. 33, 95  
Prümper, J. 33, 48, 93  
Pursche, O. 8, 96  
Rappoport, A. 34, 96  
Ratz, U. 46, 75, 95  
Reibold, H. 10, 96  
Reips, U.-D. 43, 96  
Resnick, M. 2, 19, 25, 29, 30, 40, 41, 43, 49, 51, 58, 95  
Ribeiro-Neto, B. 2, 8, 10, 11, 18, 22, 23, 35, 90  
Rogers, Y. 6, 20, 21, 22, 33, 95  
Romppel, M. 45, 96  
Rosenfeld, L. 11, 21, 97  
Rubin, D. B. 64, 95  
Saito, M. 28, 97  
Salvendy, G. 79, 92

- 
- Saracevic, T. 33, 94  
Scaife, M. 6, 20, 21, 22, 33, 95  
Schäuble, P. 33, 93  
Schneider, W. 34, 38, 97  
Schulz, U. 16, 34, 97  
Schweiger, W. 2, 27, 28, 30, 31, 97, 98  
Selvidge, P. 16, 97  
Sheldon, M. A. 14, 98  
Shneiderman, B. 21, 22, 97  
Silverstein, C. 8, 97  
Snelgar, R. 60, 61, 91  
Spink, A. 8, 32, 94, 97  
Stenmark, D. 11, 97  
Strube, G. 22, 32, 93  
Szilagyi, P. 14, 98  
Tessmer, M. 24, 94  
Theng, Y. L. 23, 97  
Thimbleby, H. 23, 97  
Thomasson, R. 15, 90  
Tiedemann, P. 33, 96  
Tillmann, H. N. 1, 97  
Traugott, M. W. 46, 91  
Turnbull, D. 2, 22, 91  
Tuten, T. L. 53, 67, 91  
Unz, D. C. 2, 33, 39, 40, 43, 51, 97  
Uszkoreit, H. 39, 92  
Vélez, B. 14, 98  
Vogt, K. 71, 97  
Wandke, H. 28, 98  
Wanger, J. 48, 98  
Weber, C. 22, 31, 32, 98  
Weiss, R. 14, 98  
Wirth, W. 2, 27, 28, 30, 31, 54, 97, 98  
Wittmann, W. W. 49, 98  
Wolfram, D. 58, 98